THE UNIVERSITY OF CHICAGO

LEARNING-BY-DOING AND PREFERENCE DISCOVERY IN VIDEO GAME PLAY

A DISSERTATION SUBMITTED TO
THE FACULTY OF THE UNIVERSITY OF CHICAGO
BOOTH SCHOOL OF BUSINESS
IN CANDIDACY FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

BY
CHARLES ZOU

CHICAGO, ILLINOIS
JUNE 2019

TABLE OF CONTENTS

LIST OF FIGURES . . . vi

LIST OF TABLES . . . viii

ACKNOWLEDGMENTS . . . ix

ABSTRACT . . . x

1 INTRODUCTION . . . 1
  1.1 Paper Organization . . . 6

2 DATA . . . 8
  2.1 Data introduction . . . 8
  2.2 Data cleaning . . . 10
    2.2.1 Partitioning modes into ordinal levels . . . 10

3 DESCRIPTIVE SUMMARY . . . 13
  3.1 Adoption Time . . . 13
    3.1.1 Hazard models for adoption . . . 14
  3.2 Propensity to play . . . 15
  3.3 Conditioning on play . . . 16
    3.3.1 Play duration | play . . . 17
    3.3.2 Max level played (MLP) | play . . . 18
    3.3.3 Shares | play . . . 19
  3.4 Cross generational patterns . . . 20
  3.5 Covariates: cluster analysis . . . 22
  3.6 Stylized facts . . . 28

4 MODEL . . . 30
  4.1 Mapping stylized facts to model . . . 30
  4.2 General econometric model . . . 31
    4.2.1 Bayesian learning specification . . . 33
    4.2.2 Solution to the constrained optimization . . . 35
    4.2.3 Likelihood . . . 35
  4.3 Specific model . . . 36
    4.3.1 Restricted utility specification . . . 37
    4.3.2 Covariance specification for learning components . . . 37
    4.3.3 Latent segmentation . . . 39
  4.4 Identification . . . 40
    4.4.1 Bias introduced by incorrect risk aversion parameter . . . 43
    4.4.2 Bias introduced by incorrect prior variance . . . 44
    4.4.3 Bias introduced by incorrect discretization parameter . . . 45
  4.5 Model discussion . . . 45

5 ESTIMATION . . . 48
5.1 Estimation preliminaries . . . 48
5.2 Model I: pure behavioral . . . 50
5.2.1 Recovery of data patterns . . . 54
5.3 Model II: advertising effects . . . 56
5.3.1 Advertising data . . . 57
5.3.2 Advertising specification . . . 61
5.3.3 Interpretation of advertising coefficient . . . 61
5.3.4 Discussion of parameters . . . 63
5.3.5 Segment summary statistics . . . 68

6 COUNTERFACTUALS . . . 72
6.1 Increased advertising . . . 72
6.2 Console switching . . . 74
6.3 Summary of counterfactuals . . . 83
6.4 Firm relevance of engagement . . . 85

7 CONCLUSION . . . 91

REFERENCES . . . 93

A DATA . . . 98
A.1 Coefficients in ZI-exponential adoption model . . . 98
A.2 Additional figures/tables for level vs experience groupings . . . 99
A.3 Conditional play patterns with alternate slicings . . . 100
A.4 Conditional play patterns grouped by previous level . . . 102
A.5 Correlation between (log) duration and (log) no. choices . . . 104
A.6 Evolution of high level shares across all generations . . . 105
A.7 Cluster analysis of $K = 2, 3$ clusters . . . 106
A.8 Regressions of outcome variables for each of $K = 5$ clusters and overall . . . 111
A.9 Advertising: evidence of geographic targeting . . . 114
A.10 Advertising: evolution of advertising stock, using different depreciation κ . . . 116
A.11 Regression of outcomes on console switching . . . 117
A.12 Counterfactual simulation visuals for New player console switching . . . 119
A.13 Regressions of adoption/in-game purchases on play . . . 123
A.14 Varying ad stock specifications . . . 126

B THEORY . . . 128
B.1 Derivation of closed-form posterior distribution for Bayesian learning specification . . . 128
B.2 Alternate learning specifications . . . 130
B.3 Non-identification of the γ-model in the presence of learning . . . 134
B.4 Additional properties of the SDCEV model . . . 136
B.4.1 Closed-form characterization of inside consumption . . . 140
B.5 Likelihood of the nested extreme value (NEV) model . . . 143

B.6 Differentiation from existing projects 145

LIST OF FIGURES

3.1 Adoption time by type. 14
3.2 Predictions from hazard models 16
3.3 Propensity to play (extensive margin) by type. 17
3.4 Conditional play duration (intensive margin) by type. 18
3.5 Conditional max level by type. 18
3.6 Conditional high level shares and HHI by type. 20
3.7 Evolution of individual level shares by type. 20
3.8 Cross generational high level shares and HHI by type. 22
3.9 Evolution of extensive margin, intensive margin, and high level shares for each cluster. 25

5.1 Model I: Empirical vs simulated data moments. 57
5.2 Model I: Underlying parameter evolution. 58
5.3 Weekly advertising impressions and expenditures. 60
5.4 Model II: Empirical vs simulated data moments. 70
5.5 Model II: Underlying parameter evolution. 71

6.1 Simulated moments under advertising counterfactuals. 75
6.2 Evolution of beliefs under advertising counterfactuals. 76
6.3 Simulated moments under PlayStation-switching counterfactual. 79
6.4 Evolution of beliefs under PlayStation-switching counterfactual. 80
6.5 Simulated moments under console-upgrading counterfactual. 81
6.6 Evolution of beliefs under console-upgrading counterfactual. 82
6.7 Various distributions for retained vs. churned players. 87
6.8 Number of in-game purchases against various engagement metrics. 90

A.1 Total play under alternate grouping. 99
A.2 No. of weeks played under alternate grouping. 99
A.3 Conditional moments, additional robustness checks. 101
A.4 Conditional play and max level, by past level with robustness check. 102
A.5 Conditional HHI and high level shares, by past level with robustness check. 103
A.6 Time played and number of sessions. 104
A.7 Evolution of high level shares across all generations. 105
A.8 Data moments with $K = 2$ clusters. 106
A.9 Data moments with $K = 3$ clusters. 108
A.10 Distribution of local advertising over time. 115
A.11 Evolution of advertising stock under various κ. 116
A.12 Simulated moments for PlayStation-switching counterfactual, new players. 119
A.13 Evolution of beliefs for PlayStation-switching counterfactual, new players. 120
A.14 Simulated moments for console-upgrading counterfactual, new players. 121
A.15 Evolution of beliefs for console-upgrading counterfactual, new players. 122
A.16 Predicted purchasing behavior. 125
A.17 Simulated moments under alternate advertising specifications (κ). 127

B.1 Non-identification of $\gamma-$ SDCEV model under Bayesian learning. 134
B.2 Non-identification of $\gamma-$ MDCEV model under Bayesian learning. 135
B.3 Simulated distribution of first-week play, SDCEV. 138
B.4 SMM estimates of satiation and utility. 138
B.5 Crossover point property of SDCEV. 142

LIST OF TABLES

2.1 Summary statistics for each game. 9

3.1 Adoption statistics by type. 14
3.2 Clustered behavioral and demographic means, $K = 5$. 22
3.3 OLS of log(Total play G) for each cluster and overall. 27

4.1 Simulation results varying no. of simulation draws S 42
4.2 Simulation results mis-specifying risk aversion parameter r. 43
4.3 Simulation results mis-specifying prior variance σ_{0j}^2. 44
4.4 Simulation results mis-specifying discretization parameter Δ. 46

5.1 Model I: Average pessimism by segment and level. 50
5.2 Model I: Estimated parameters. 52
5.3 Model I: Sample statistics by predicted segment identity. 55
5.4 Model I: Aggregate moments, data vs simulation. 56
5.5 Model II: Average prior bias by segment and level 64
5.6 Model II: Estimation results. 65
5.7 Model II: Sample statistics by predicted segment identity 67
5.8 Model II: Aggregate moments, data vs simulation. 68

6.1 Console switching. For brand (left), N = 13804; for generations (right), N = 12732. 77
6.2 Data moments relative to baseline (as percentage), various counterfactuals. . . . 84
6.3 Adoption vs. play, with interactions. 88
6.4 In-game purchases vs. play, with interactions. 89

A.1 ZI-exponential adoption hazard model. 98
A.2 Adoption summaries, by player past level. 99
A.3 Clustered behavioral and demographic means, $K = 2$. 106
A.4 Clustered behavioral and demographic means, $K = 3$. 108
A.5 OLS of no. wks. played G for each cluster and overall. 111
A.6 OLS of last wk. played G for each cluster and overall. 112
A.7 OLS of overall high shares G for each cluster and overall. 113
A.8 Outcomes on console upgrade dummy and controls 117
A.10 Outcomes on PlayStation switch dummy and controls 118
A.12 Adoption vs play, main effects only. 123
A.13 In-game purchases vs. play, main effects only. 124
A.14 Ad stock coefficients by κ. 126

B.1 First week statistics for early adopters, by week of adoption. 136

ACKNOWLEDGMENTS

I thank my dissertation committee: Professors Pradeep Chintagunta, Ruey Tsay, Bradley Shapiro, and Anita Rao for their comments, insights, and support on this project. I thank the Wharton Consumer Analytics Initiative (WCAI) and the corporate sponsor for granting me this opportunity to work with such a rich dataset. This work was additionally supported by the Kilts Center for Marketing. My analyses were calculated based in part on data from the Nielsen Company (US), LLC and marketing databases provided through the Nielsen Datasets at the Kilts Center for Marketing Data Center at the University of Chicago Booth School of Business. The conclusions drawn from the Nielsen data are mine and do not reflect the views of Nielsen. Nielsen is not responsible for, had no role in, and was not involved in analyzing and preparing the results reported herein.

ABSTRACT

In this paper I consider the effect of learning-by-doing and preference discovery on engagement for users of a popular franchise video game. An important data aspect is competition–players must budget their game time between competitive and non-competitive modes. I observe that player behavior is consistent with learning and initial competition aversion. For example, shares of time spent in competitive levels tend to drop upon new game adoption and then rise with time played. Within this rich dataset, I also discover significant heterogeneity in usage patterns. Thus, a one-size-fits-all approach is insufficient. To study how the firm can increase player engagement, and to understand the relative values of competitive and non-competitive play, I propose a novel structural model nesting Bayesian learning within a multiple-discrete continuous framework. This allows me to jointly explain usage along the extensive and intensive margins. With this model, I find that consumers can, broadly speaking, be categorized as high types ("hardcore") or low types ("casual"). High types tend to be competition-seeking and more naturally engaged, while low types tend to be competition-averse and drop out quickly before learning their true match values. I consider actions the firm can undertake to improve consumer engagement–in particular, I perform counterfactual analysis on advertising and console switching (i.e. bundling) policies. I find that low types tend to be more responsive to both policies and primarily respond by increasing consumption along the extensive margin. I find that console switching has a significant positive effect on total play, but cause players to substitute away from competitive levels. This is consistent with the learning framework, where players pay a skill or familiarity cost when switching consoles. Finally, I discuss the value of engagement to the firm, both qualitatively and with respect to revenue-related outcomes.

CHAPTER 1

INTRODUCTION

The video game industry is both an important contributor to the US economy as well as a significant supplier of leisure goods to the US consumer. In 2016, consumers spent $23.5 billion on video games, hardware, and accessories. 63% of households contain at least one person who plays over 3 hours per week and 48% own a dedicated game console (Entertainment Software Association). Gamers spend over half of their non-social leisure time on games (American Time Use Survey, 2016). Of course, as digital and online consumption continues rising gaming is increasingly a channel, rather than a substitute, for socialization. 53% of frequent gamers feel video games help them connect with friends and 42% feel they help spend time with family (ESA, 2016). Further, the demographic of video gamers is shifting: dedicated gaming console ownership now favors women (Pew Research Center, 2015) and the average gamer is now 35 (ESA, 2016). Underlying these trends are the evolving, unobserved motivations for play. For example, survey-based literature suggests that competition is particularly important to younger or male gamers while strategy is relatively age-invariant (Quantic Foundry, 2016). At the same time, video games are not monolithic. Unlike the traditional marketing case of CPG's, games are experiential and often have multiple modalities that satisfy different motivations (i.e. for socialization, competition, or immersion). Thus, from the perspective of the video game firm, it is crucial to deeply understand the interplay between consumer preferences and game modalities that generate observed behavior.

The marketing literature has traditionally focused on the adoption or purchase stage (e.g. Norton and Bass, 1987). Additionally, there has been work considering formation of consumer consideration and search sets (e.g. Mehta et al., 2003; Honka et al., 2017). On the other hand, research on post-adoption behavior has been relatively scant due to the dearth of revealed preference data (Nevskaya and Albuquerque, 2019). This is certainly a rich field and filling in the gaps between repeat-purchase or product upgrade with usage data allows for a deeper understanding of consumer preferences. For example, Huh and Kim (2008) study repeat purchasing among smartphone users and find that sticky behavior is driven by post-adoptive usage of brand-specific innovations. Conclusions of this type require a type of data that is often simply unavailable. Usage behavior also tends to be leveraged in the literature on churn in retention (see, e.g. Ascarza and Hardie, 2013; Braun and Schweidel, 2011 for applications to the warehouse retail and telecommunications industries).

In this paper I aim to structurally understand the drivers of product usage and its consequences for consumer engagement. My primary research interest is discovering how firms can utilize observed usage data to understand the evolving preferences of consumers and what "levers" may best be called upon to improve user engagement. In particular I posit a structural model of usage framed within consumers learning-by-doing, resolving uncertainty about the product at hand–this is especially relevant in the case of new products or inexperienced consumers. Moreover, I consider the scenario where usage is a non-binary vector, i.e. multiple discrete-continuous[1]. This allows me to understand usage not only along the extensive (use-or-not) and intensive (how much to use) margins, but also the composition of modalities. The empirical application at hand is a proprietary dataset detailing users of a major franchise console video game with sessions-level play over many game modes. Thus, I am further interested in understanding motivations for play, and specifically, the role of appetite for competition in driving usage and long-term engagement. Through this lens, a consumer resolving uncertainty over game modes ordinally indexed by level of competition can alternately be viewed as increasing her appetite for competition or potentially game skill. To my knowledge, this is the first paper to consider: (1) learning in the framework of multiple discrete-continuous models and (2) the nature of the relationship between competitive appetite and play from an economically-founded point of view. Applying my model

1. In other words, this case is a generalization of the single discrete, multiple discrete, and single discrete-continuous decisions oft considered in marketing literature.

to the play data, I find that there are two latent segments of consumers, which I define to be "hardcore" (high types) and "casuals" (low types). Relative to hardcore gamers, casuals are characterized by a large degree of aversion towards the competitive modes and a high degree of satiation. For all consumers, I find evidence of "upward" learning that is eventually overwhelmed by declining novelty–on average there is a significant amount of learning left on the table, particularly for casuals. Turning to policy, I find that advertising has a 0.02-0.06% elasticity (of play) and console switching (i.e. upgrading generations) has a markedly larger effect on net play (up to 7%) but lead to a large degree of substitution away from the competitive modes and have an empirically low uptake. Finally, I consider next-game adoption and in-game purchasing and find a marginal but positive indirect effect (on the order of 0.05%).

This work primarily draws from several streams of literature. First, it builds on research in quantitative marketing/empirical industrial organization on understanding the relationship between consumer behavior (specifically micro-level usage) and future outcomes. For example, Hartmann and Viard (2008) look at frequency rewards programs for users of a golf course. They find that high (low) frequency users place a high (low) ex ante value on both the product and the reward, and that it is this valuation rather than switching costs that drives observed loyalty patterns. In a similar application to this paper, Nevskaya and Albuquerque (2019) use popular MMO game *World of Warcraft* player behavior to estimate a dynamic structural model of experiential good consumption. One particular novelty in their paper is the observation of player content progression. Using this, they find that a significant driver of usage is pace of content updates and that save for a small contingent of "hardcore" users, most users do not consume content quickly enough to justify current update frequency.

Similar to this literature is consumer learning, usually framed as Bayesian learning. For a comprehensive survey of learning models in marketing, please refer to Ching et al., 2013.

For example, Narayanan et al., 2007 analyze consumers pre-committing to fixed or metered telephone plans and find that those on metered plans learn their true usage more quickly than those on fixed plans. In a similar vein I consider evolving consumer behavior through the lens of uncertainty resolution and show that several prominent data patterns can be considered as players solving static budget problems with differential learning rates. Further, the patterns I observe are similar to that in Huang, 2019, who studies the camera upgrade behavior for users of the popular image-hosting website *Flickr*: we both observe an immediate drop in the indexing variable (photograph ratings and competitive level, respectively) following adoption which slowly rises again over time. Huang, 2019 attributes this pattern to forward-looking consumers "investing" in more complex technologies with a higher quality ceiling while I attribute it to the novelty-risk aversion tradeoff.

Methodologically, I also build upon work in the environmental economics/transportation operations literature on multiple discrete-continuity. In particular, in this paper I extend Bhat's (2008) multiple discrete-continuous nested extreme value (MDCNEV) model to accommodate consumer learning. Similar models without learning have been previously introduced in the marketing literature (cf. Satomura et al., 2011). Through a simulation study, I present necessary conditions for model identification and show it has sufficient flexibility to explain several prominent patterns in the data. At the same time the model parameters retain a simple, intuitive interpretation that lends naturally to answering the policy questions I am interested in: whether shifting consumer prior beliefs or improving flow utility can increase user engagement.

Additionally, I draw from the psychology and sociological literature on playful consumption and competition to develop my modeling framework. In particular, this literature posits that consumer utility is driven by intangibles (e.g. fantasy, emotion, novelty) in addition to tangible attributes (e.g. Holbrook and Hirschman, 1982; Berlyne, 1970). Holbrook et al.

(1984) show that personality-game compatibility and mastery are important drivers of enjoyment in game play. Another stream of literature focuses on the effect of competition on players (both in sports and games). Vorderer et al. (2003) show that competition and interactivity are key elements to explain the entertainment experience in videogames. Frey et al. (2003) find that in competitive sports, there is significantly greater recruitment of mental skills in competitive vs. practice settings. Boudreau et al. (2016) find that in a tournament setting, increasing competition has a J-shaped effect on performance with respect to underlying player skill. A similar line of inquiry is the effect of social play on outcomes. For example, Jansz and Tanis (2007) find that the social interaction motive was the strongest predictor of time played among *Counterstrike* players and Weibel et al. (2008) find that playing vs. human (compared to computer) opponents leads to greater enjoyment and the affect of flow. Gu et al. (2016) find that highly social players have higher short-term engagement but lower long-term retention while less social, casual players are the opposite. They caution that firms cannot expect social users to purchase more, nor expect increasing social play to increase engagement: in fact the best predictor of future purchase is past purchasing. In summary, this literature provides both a theoretic and empiric starting point in developing my utility formulation that considers learning and competition as key drivers in usage.

While I specialize my model to the context of video game engagement, there are many applications where joint estimation of extensive and intensive margins alongside consumer learning is natural. Within the extant marketing literature, for example, researchers have examined consumer packaged goods (CPGs) such as beverages (Satomura et al., 2011; Kifer, 2015) or yogurt (Kim et al., 2002). In these applications, it is important to allow both interior and boundary solutions within a utility maximization framework that enables consumer welfare measurement. Here learning accommodates the introduction of new brands–perhaps consumers are uncertain about new products but form partially informed priors based on experiences with the umbrella brand. In fact, in the same way that omission of the MDC component can lead to biased price elasticities (cf. Bonnet and de Mouzon, 2014), the omission of learning where present also biases estimated parameters. A second stream of diverse literature ranging from resource economics to transportation examines time-use decisions. For example, researchers have studied recreation time budgeting (Bhat, 2005; Luo et al., 2013), mobile application usage (Han et al., 2016), or motor vehicle usage (Fang, 2008). These applications highlight the ability of the MDC model to recover structural parameters and perform policy experiments. The machinery introduced in this paper is naturally applicable to time use with preference evolution. While Luo et al. (2013) model dynamics through skill accumulation, I explain similar data patterns through Bayesian learning. More generally, Sanders (2016) shows that with sufficiently rich data, both human capital accumulation and learning can be jointly identified. Finally, it is important to caveat that there are scenarios where this extension is unnatural and thus not recommended. For example, physician learning over patient-drug matches is well-studied in marketing (cf. Chintagunta et al., 2009; Ching and Lim, 2016; Coscelli and Shum, 2004). In these scenarios it is difficult to imagine physicians are solving a budget problem, either over drug costs or dosages. Thus while discrete-continuity may be a feature of the data, one should carefully examine the context before blindly adopting the framework. In summary, the model I propose is general and appropriate for both the standard utility maximization as well as usage budgeting frameworks when there exists uncertainty about alternatives that is then resolved with usage. With all tools, however, it is up to the user to properly assess the circumstances and suitability.

1.1 Paper Organization

The remainder of this paper is structured as follows: in Section 2 I give a brief overview of the data and how it is cleaned. In Section 3 I present descriptive summaries for notable data patterns alongside potential explanations and stylized facts. Following that in Section 4 I present a structural model that has sufficient flexibility to accommodate these data patterns, then provide simulation results to establish identification. In Section 5 I provide estimation results of my model and show that it can recover important data patterns. In Section 6 I employ the model augmented with ad effects to perform counterfactuals and discuss actions the firm can undertake to improve player engagement. Finally in Section 7 I discuss potential model extensions and conclude.

CHAPTER 2

DATA

2.1 Data introduction

The data contains purchase and play behavior for consumers (players) of a franchise console video game with annual releases over five years. While the central premise, game play, and overarching feel of the game remain consistent, with each new release there may be substantial mechanical and graphical changes. For example, individual game sessions are played upon selection of a game mode. The first game I observe in the sample consists of 20 modes while the final has 34–while there are certain core modes that consistently return, quite a few modes are added or dropped each year. Further, there are often core mechanical tweaks and qualify-of-life improvements year-over-year.

I observe play across 5 games in the franchise (and consequently 5 years of play data) from a "left-censored" sample of players. By construction, all players from previous generations that adopt the next game are retained. In addition roughly 10,000 new players are sampled each year. In Table 2.1 I present some basic summary statistics for each game. Due to the accumulating nature of the sampling process, the number of players I observe playing each game grows by around 15% annually. The mean number of game sessions also increases rather significantly the first few games before stabilizing in the final two games[1]. On the other hand, the average session length is about one hour across all games. I observe that about half of the players in each game are retained each year. Unsurprisingly, new players churn out significantly more than returning players.

1. A session is bookended by a player initializing and exiting the game. This is a somewhat nebulous entity as it can range from a few seconds to several hours (e.g. the longest session I observe is almost 24 hours long and the 99% quantile is 5.5 hours). The mapping from session length to number of actual in-game matches played, which is perhaps of greater interest, is unfortunately unobserved and difficult to back out in the data.

	Players	Modes	Sessions	Min/ses.	Retention rate	Ret.|Returning	Ret.|New
Game 1	15462	20	21.04	63.75			
Game 2	19651	30	36.96	60.17	0.50	0.73	0.34
Game 3	20963	28	40.92	63.44	0.47	0.66	0.29
Game 4	25455	29	49.80	63.47	0.52	0.68	0.35
Game 5	26812	34	49.77	56.44	0.57	0.71	0.38

Table 2.1: Summary statistics for each game.

In this paper I will solely focus on explaining player behavior for the final game in the sample (Game 5). Because my policy question relates to player engagement, I only consider a single game and condition on adoption. While I show that more engaged players are likelier to adopt the next game and it is certainly reasonable to ask whether increased engagement has a causal effect on future adoption, that is outside the scope of the current paper. Conditioning on adoption means my interest is moreso measuring a kind of "average treatment effect on treated" (ATT) than the general ATE. I believe that from the perspective of the firm this conditional treatment effect is important.

The inclusion of the adoption decision would certainly allow me to tackle a greater range of questions, but there are both data and modeling challenges that impede satisfactory treatment of adoption. From a modeling perspective, a joint model of adoption and usage necessitates the introduction of dynamics, as I must consider that consumers choose when to upgrade. Unfortunately as I show in Section 3 for a majority of the players for whom I have useful historical information the observed adoption times lack variation[2], and for new players for whom there is large variation in adoption times, there is zero information that can be used to explain adoption–this includes any price information. Finally, as I show in Section 4 the proposed model is highly complex and identification is non-trivial. Typical dynamic problems in the marketing literature are simple in the sense that the decision variable is scalar discrete. Because I am modeling a multiple discrete-continuous outcome (vector continuous), the computational and analytic burden would be tremendous. As I

2. For example, about 85% of highly experienced players adopt by the second week.

show later, the proposed static model can sufficiently capture important data moments in the post-adoptive framework.

2.2 Data cleaning

After basic cleaning, I observe 26,715 players from adoption of the final game until end of sample for a possible maximum of 304 days. Due to a high degree of sparsity in daily sessions data, I aggregate the data to a weekly level. After cleaning, I have around 898,000 player-weeks and have net play duration for each of 30 remaining game modes. Even after the weekly aggregation, about 70% of player-weeks have no play. For player-weeks with play, the median play duration is 2.7 hours (mean = 4.5 hours).

As I am interested in the consumer's evolution of appetite for competition, I then map each of the 30 game modes to one of four **levels** of competition, detailed below. This additionally serves the benefit of reducing the dimensionality and sparsity of the play data.

1. Solo: single-player, played against AI. Includes training modes.
2. Friendly: multi-player, played with others from user's friends list, can be cooperative or competitive. Includes "couch co-op" mode.
3. Competitive: multi-player, played against strangers in a matchmade environment, purely competitive.
4. Ranked: same as competitive but with an additional "ranking score" attached to the account.

2.2.1 Partitioning modes into ordinal levels

While there is a possibility that grouping several modes into one level throws away useful information, I believe that the simplicity and gains to interpretability (as well as dimension reduction) more than compensate for the cost. Below I detail the reasoning used to form these levels:

The distinction between Solo and other levels is fairly straightforward: single- vs multi-player and is a common characterization used in industry. While it partitions modes that offer no competition vs those that do, it also partitions modes that have no social element from those that do. In the literature there is significant support both for the social effect of gaming (e.g. Tyack et al., 2016; Yee, 2007) and the effect of playing against computer vs human opponents (e.g. Williams and Clippinger, 2002). Empirically, a potential concern then is that the social/network and competitive effects may be conflated with this partitioning. I believe this may be resolved by construction of the Friendly level, which is the only level to include a cooperative mode. Intuitively, the elasticity of substitution away from Solo to Friendly compared to that of Solo to (Competitive + Ranked) can help pin down the competitive relative to the social effect.

The distinction between Friendly and the next two levels is that games are purely competitive and played against strangers. Again, there is significant support in the psychology literature that players fundamentally view cooperation and competition, even within the same game, differently (see, e.g. Schmierbach, 2010; Ewoldsen et al., 2012). Additionally, evidence suggests that playing against friends (even virtual) can be both differently motivated and experienced than against complete strangers (see e.g. Tyack et al., 2016; Pollmann and Krahmer, 2017).

Finally, the distinction between Competitive and Ranked is largely empirical: looking over online discussion boards[3], the general consensus appears to be that the ranked modes

3. Sources include the official game forums maintained by the firm, and various unofficial enthusiast discussion boards

are the most "hardcore". Consistent with competitive aversion, I observe that almost all players have played a solo mode in their career (>95%), while over half have played a friendly (65%) or competitive (50%) mode, while less than half (<40%) have ever touched the ranked mode. Within the first four weeks of purchase, the numbers are 94%/51%/40%/24%, which provides further justification this definition.

CHAPTER 3

DESCRIPTIVE SUMMARY

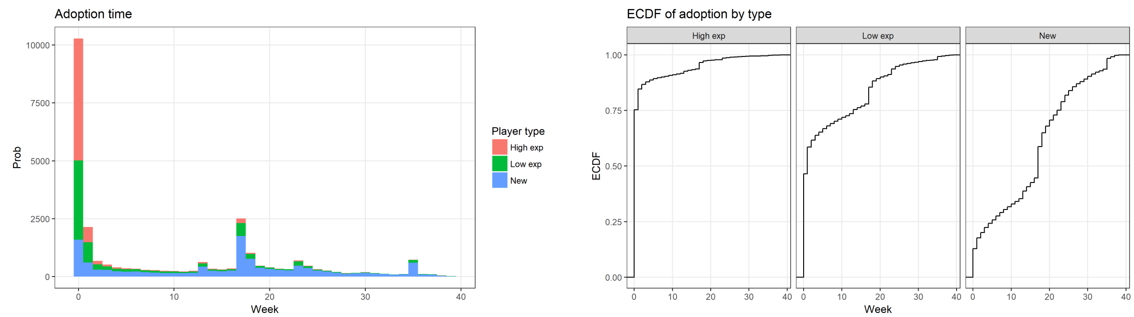
Below I give a high-level description of the data, and show patterns that motivate my model and the empirical questions I plan to tackle. I visualize the data by partitioning the sample into three groups: new players who are introduced to the franchise with the current game (46%, denoted **new**), returning players who played less than the median duration in the previous franchise game (28%, **low exp**), and returning players who were above the median (26%, **high exp**). I consider the extensive margin to be number of players who play in any given week and the intensive margin the conditional duration for those who decide to play. I show that the low exp and new players behave rather similarly, while the high exp players are significantly different in both intensive and extensive margins, and propose mechanisms for the observed behavior.

3.1 Adoption Time

For completeness I present summary statistics for adoption here. There is significant heterogeneity in adoption time for the game. As seen in the table below and Fig. 3.1, high exp players purchase the game far earlier than new players, with low exp players somewhere in between. In fact, by the second week 85% of high exp players have purchased (compared to less than 20% of new players). Additionally note that around week 17 there is a spike in purchases corresponding to a major US holiday. In that particular week, 14% of all new players adopt compared to only 3% for high exp players. These numbers may suggest that high exp (and to some extent low exp) players have already decided to deliberate, perhaps requiring an exogenous holiday shock.

Type	Mean adoption	Median adopt	3rd quartile adopt	% adoption in first 2 weeks
New	15.27	17.00	22.00	0.18
Low exp	6.46	1.00	13.00	0.59
High exp	2.06	0.00	0.00	0.85

Table 3.1: Adoption statistics by type.



(a) (b)

Figure 3.1: Adoption time by type.

3.1.1 Hazard models for adoption

Next I use two simple hazard models to gain some intuition for the adoption process. As I do not observe demographics, I have scant information to predict adoption times for new players. For returning players, I use play in the previous game[1]. Below I present selected results:

First, I consider the virtual Bass adoption model (VBM; Jiang et al., 2006). Recall it is parameterized by the set (p, q, τ) where the CDF of adoption $F_1(A_t)$ is defined:

$$F_1(A_{it}) = \frac{1 - e^{-(p+q)(\tau+t)})}{(q/p)e^{-(p+q)(\tau+t)} + 1}$$

where $\tau > 0, q > p$. Here p, q represent the coefficients of innovation and imitation, respectively, and τ represents the initial "release" (i.e. when consumers begin forming mental purchase commitments). I estimate $p = 0.035, q = 0.054, \tau = 9.8$ which is reasonable given

1. Specifically, I restrict the previous game data to usage before the current game's release. This simplifies treatment of the models by abstracting away from "port-back" behavior (see e.g. Tao and Sweeting, 2019). Among those who do adopt the current game (the sample of interest), I observe that there is very little play in the previous game after release of the current game.

that the typical product lifespan here is a single year. The game was officially announced roughly 16 weeks before release, so the τ estimate suggests that players may not commit to purchase until several weeks later.

Next, I incorporate history variables using a zero-inflated exponential (which provides a significantly superior fit than ZI-Poisson despite fitting discrete data). Here I model density $f_2(A_t)$ as:

$$f_2(A_{it}) = \pi_i \mathbb{I}(A_{it} = 0) + (1 - \pi_i)\lambda_i e^{-\lambda_i A_{it}}$$

where $\lambda_i = e^{X_{i0}^1 \delta^1}$, $\pi_i = (1 + e^{X_{i0}^2 \delta^2})^{-1}$, so that covariates that affect the likelihood of purchase on release and purchase after are allowed to differ. In the Appendix I present results for estimates on $\delta = (\delta^1, \delta^2)$ using a dummy for new player (i.e. no prior history), log(previous game experience), and average level played in previous game. Roughly speaking, a returning player is about 4 times more likely to purchase immediately, and it is strongly increasing in both previous experience and average level played. Further, the rate associated with the exponential function is decreasing in both experience variables as well as for new players. The fit is significantly better than the vanilla VBM. Predictions are given in Fig. 3.2.

3.2 Propensity to play

The literature on experiential product usage suggests that consumer utility depends not only on product tangibles but also intangibles such as social and novelty effects (Holbrook and Hirschman, 1982). Similar to Nevskaya and Albuquerque, 2019, I observe declining play across all groups, consistent with diminishing novelty and reductions in social interactions as other players stop playing as well. In another study relating to a popular game in the multiplayer online battle arena (MOBA) genre, Tyack et al. (2016) find that many

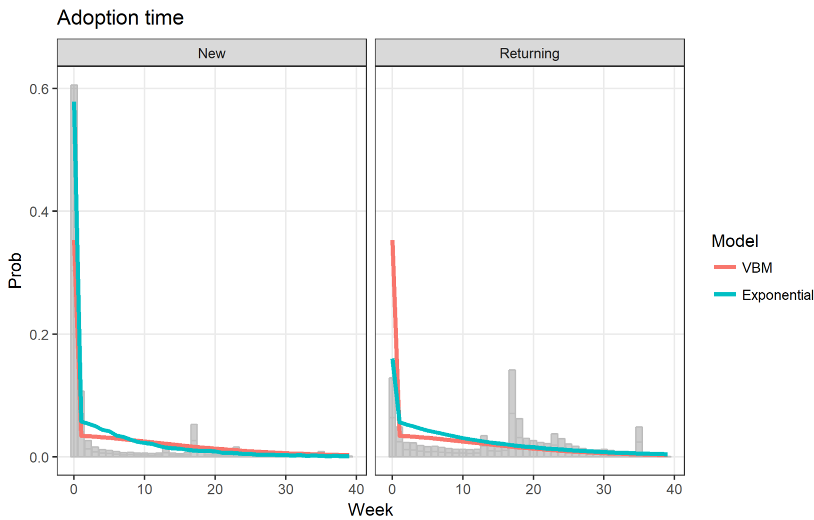


Figure 3.2: Predictions from hazard models

players quit primarily due to time constraints, competing interests in other games, and loss of enjoyment; the primary reason for returning is game updates. Within a single game, I do not observe a quit decision, and cannot determine whether a sequence of observed zeros from time t until end of sample are structural (quit) or sampling (budget-allocated) zeros.

Below I present probability of play per week post-purchase by experience in the exact way as Fig. 3.1. Note that low exp players are more similar to new players than high exp players. The median number weeks played for new players is 3 (mean = 6.1) and returning players is 12 (mean = 14.1). For high exp players, the median is 19 (mean = 19.2) and for low exp players, it is 7 (mean = 9.2). In the Appendix, I show the pattern persists when the outcome is total weeks with non-zero play.

3.3 Conditioning on play

The declining probability of play empirically dominates all other unconditional data moments. In other words declining engagement along the extensive margin (from e.g. declining

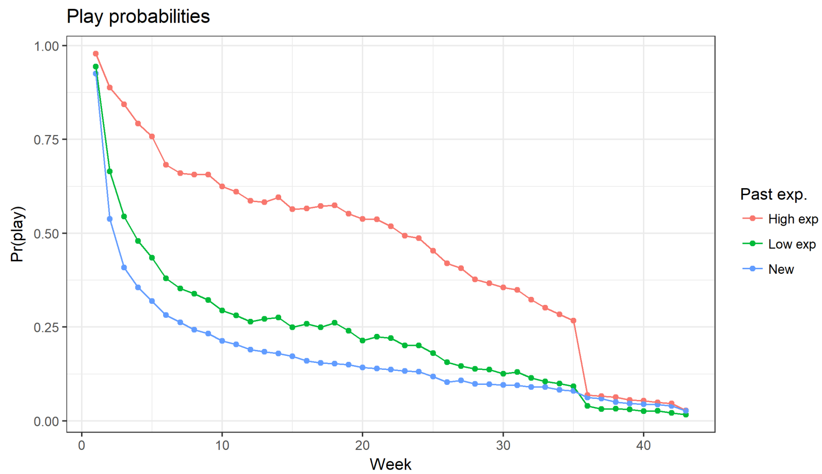


Figure 3.3: Propensity to play (extensive margin) by type.

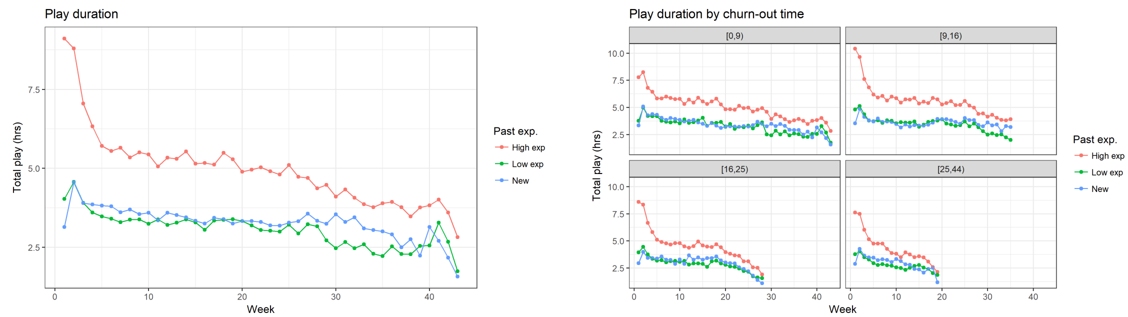
novelty, time constraints, or other games) masks trends in total play, competitive play, etc. This leads to rather uninteresting patterns in the unconditional data moments. Thus I will consider conditional data moments (conditioning on play). This will inevitably introduce selection bias since the subsample that plays in each period shifts over time (note that these subsamples **are not** necessarily decreasing in time). I account for this bias by conditioning on churn proxied by last week played. Specifically, I provide visualizations for static subsamples of players who "churn-out" within similar bins[2]. By binning players with similar attrition times, I believe structural group differences in play propensities can be approximately accounted for. As a robustness check, in the Appendix I provide visualizations using alternate conditioning variables.

3.3.1 Play duration / play

In Fig. 3.4a I present conditional play durations for each of the three groups (by experience). Note that even conditional play decreases over time for all groups, but high level players tend to play more than low level players, who are similar to new players. This pattern is generally preserved conditioning on churn-out quartiles (Fig. 3.4b). Again, this suggests that even

2. Here I use sample quartiles of last week played: 9, 16, 25, and 44 weeks.

among hardcore fans of the game, the novelty of the game wears off.

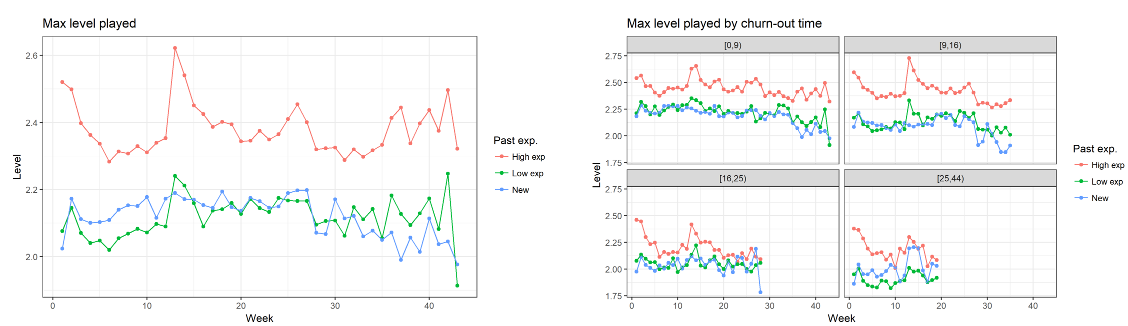


(a) (b)

Figure 3.4: Conditional play duration (intensive margin) by type.

3.3.2 Max level played (MLP) / play

For each user-time, I observe the play vector $(x_1, \ldots, x_4)$ where 1 denotes Solo and 4 denotes Ranked. Here I consider max level played (MLP) defined as MLP $= \text{argmax}_j x_j$, shown in Fig. 3.5. It tends to be rather volatile but largely flat over time, with clear separation once again between high exp and other players. Sliced by churn-out time there appears to be a slightly declining trend for long-playing users. While a declining trend may be consistent with experimentation, where consumers are trying more competitive levels than their current tastes may otherwise dictate and dropping down afterwards (Sanders 2016), a flat curve would rule it out.



(a) (b)

Figure 3.5: Conditional max level by type.

3.3.3 Shares / play

Finally, I consider shares $s_j = x_j / \sum_k x_k$ which normalizes the declining intensive margin and allows me to focus on how play behavior shifts between levels over time. First in Fig. 3.6a I consider the shares of High levels, $s_{34} = s_3 + s_4$, which serves as a measure of competitive appetite. Note that it is increasing for all players, suggesting players are increasingly seeking out competition in play. High exp players appear to have greater competitive appetite than the low exp and new players. Next in Fig. 3.6c I compute HHI, defined as $\sum_j s_j^2$. Recall that for J alternatives, HHI has an upper bound of 1 when all consumption is allocated to a single good and a lower bound of $1/J$ when consumption is evenly divided amongst all alternatives. It is thus a measure of increasing concentration. Empirically, I observe increasing HHI over time for all groups, indicating that all player groups are becoming specialists and that, on average, players are far from uniform play allocation (with $J = 4$, uniform play would correspond to an HHI of 0.25). Note that new and low exp players tend to concentrate their play more than high exp players, possibly indicating an aversion to certain levels. In Figs. 3.6b and 3.6d I show this effect persists for players conditioning on churn-out time. Finally, in Fig. 3.7 I show the individual evolution of shares for each group. Through the entire sampling period, the dominant level is Solo, although it is decreasing over time (moreso for high exp players). High exp players are substituting away from Solo into all three other levels, which generates the increasing High level shares. Low exp and new players are substituting away from Solo and Friendly to Competitive, while Ranked play remains low. This also generates increasing High level shares, but in a different manner. These visualizations suggest that, in general, players are increasingly competitive and specialize away from the Solo levels. These patterns are suggestive of the Bayesian learning scenario where users are ex ante more pessimistic or uncertain about the good with higher true match value. An alternate explanation is that players are ex ante uncertain about their own competitive appetite and resolve it through play. While shares of High level play are increasing, low exp and new players do not necessarily end up substituting to the highest level (Ranked), indicating that there may be a plateau in the learning effect.

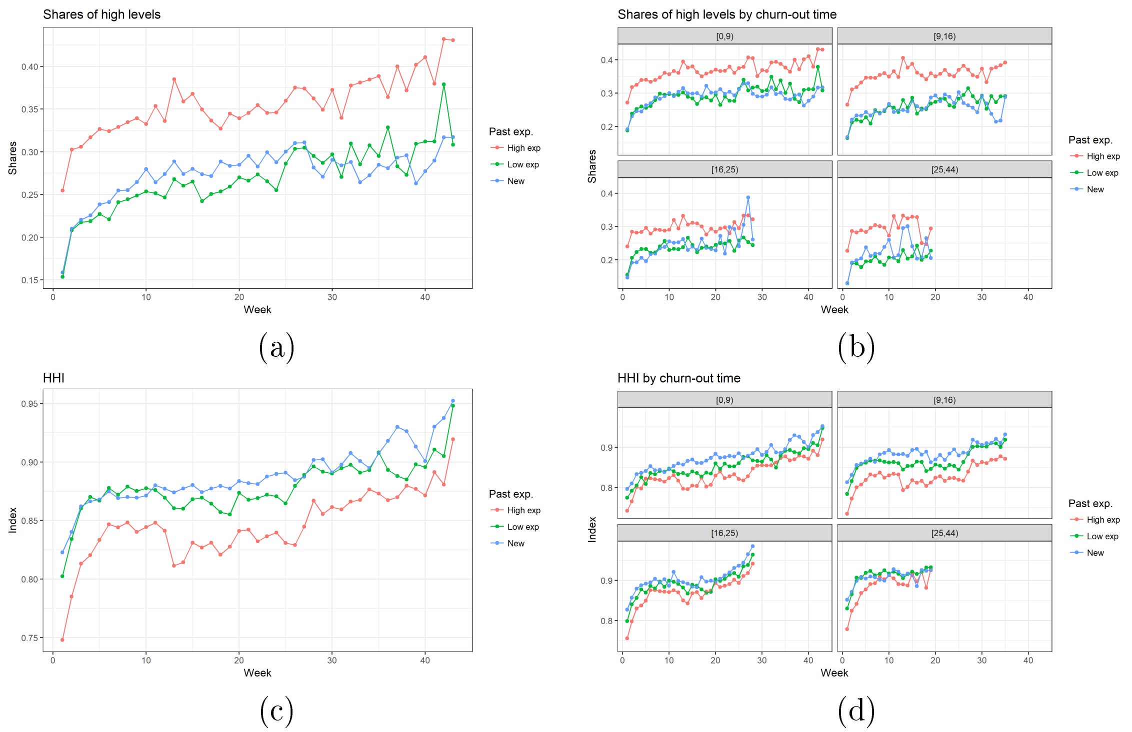


Figure 3.6: Conditional high level shares and HHI by type.

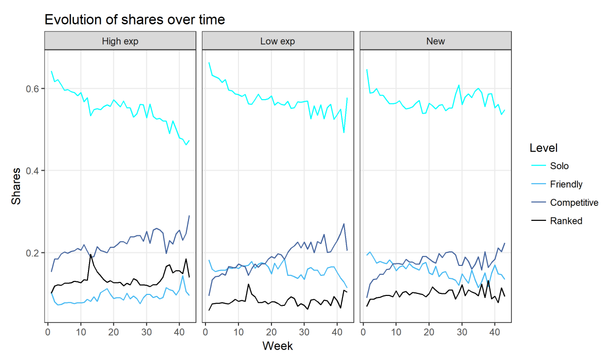


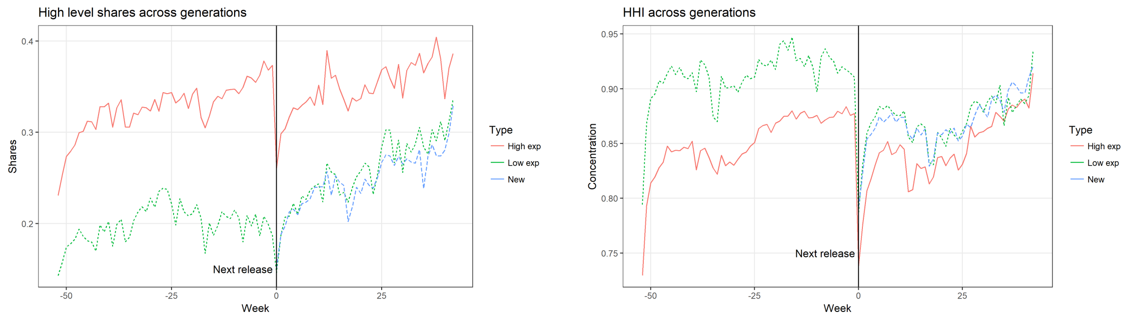
Figure 3.7: Evolution of individual level shares by type.

3.4 Cross generational patterns

Although the focus of my paper is solely current generation play patterns, I believe it is useful to understand how players behave between game versions. In Fig. 3.8 note that for returning players, there are significant drops in both high level shares and HHI immediately after upgrading to the newer game[3] This would once again be consistent with a learning framework, where players are uncertain about their match values for game G's levels. This is consistent with the fact that each franchise game introduces new core mechanics and game modes. For example, a player's favorite game mode may even have been removed entirely between game $G-1$ and G. In the same vein, the drop in HHI can be interpreted as attempts to experiment in order to learn about the new game's modes and resolve uncertainty. A final note is that the low exp and new players are very similar in Fig. 3.8. Given that the threshold for low exp was defined to be roughly 40 hours (<1 hour/week), it makes sense that from an uncertainty and experience perspective, low exp players behave similarly to new ones.

These patterns are seen in Huang (2019), who attributes the immediate quality drop when *Flickr* users purchase new cameras to a mental switching cost as users must accustom themselves to the new technology. Thus, he rationalizes adoption of more advanced cameras through the tradeoff between current skill and a higher future skill ceiling. On the other hand, I argue that these patterns can be fully rationalized even without dynamics: for example, consider pessimism over level j defined $pess_j = q_j^* - \bar{q}_{0j}$, i.e. difference between true and initial valuation. Consider just the Competitive ($j=3$) and Solo ($j=1$) levels. When $pess_3 > pess_1$ but $q_3^* > q_1^*$, in other words a player is more pessimistic about the Competitive level but it has better true match value, the observed high level shares can be rationalized. The drop in HHI can similarly be explained by a jump in uncertainty over game G relative to $G-1$.

3. Here the plots are conditional on observing a play session, which leads to the obvious criticism of self-selection. In the Appendix I replicate these plots for the subsample of high utilization players who recorded at least one session in the final month and show this pattern remains. Without further assumptions I cannot conclude anything about those low-utilization players who stopped playing game $G-1$ much earlier–this is a standard treatment effects issue (i.e. Rubin's potential outcomes).



(a) (b)

Figure 3.8: Cross generational high level shares and HHI by type.

3.5 Covariates: cluster analysis

In this section I work backwards from outcome data to gain intuition on consumer "types". In particular, I cluster the sample using ex post summary behavioral data such as play duration, max level played, etc.[4] and examine the ex ante (game $G-1$) covariates as well as current-game play evolution for each cluster. While previous game covariates cannot be generally considered exogenous, in my model I will condition on adoption, allowing treatment of pre-adoption behavioral variables as given. Below I present results for $K \in \{2, 3, 5\}$ clusters, focusing on $K = 5$ clusters. In Fig. 3.9 I visualize the evolution of three data moments–extensive margin, intensive margin, and high level shares for each cluster, and in Table 3.2 I provide summary "demographics" (i.e. including previous game behavior) for each cluster.

Table 3.2: Clustered behavioral and demographic means, $K = 5$.

	Clust. 1	Clust. 2	Clust. 3	Clust. 4	Clust. 5
No. games owned	1.58	3.58	2.95	2.88	3.54
No. yrs in samp.	0.34	1.84	1.61	1.44	2.01
Adopt day G	159.21	27.91	23.28	66.63	36.90

4. I use: no. weeks played, no. weeks owned, percent of weeks with play, last week played, overall max level played, first {1, 4} weeks play duration/max level, total play in each level as well as overall, overall high shares, and overall HHI for the current game.

Table 3.2: Clustered behavioral and demographic means, $K = 5$, continued.

Owns Xbox 360	0.11	0.06	0.07	0.07	0.06
Owns Xbox One	0.50	0.47	0.41	0.43	0.43
Owns PS3	0.05	0.03	0.07	0.08	0.05
Owns PS4	0.35	0.49	0.47	0.46	0.49
Owns Xbox	0.60	0.51	0.48	0.48	0.49
Owns PS	0.40	0.51	0.53	0.53	0.53
Owns older cons.	0.16	0.09	0.13	0.15	0.11
Owns multiple cons.	0.01	0.05	0.02	0.03	0.04
High exp	0.02	0.52	0.11	0.27	0.40
Low exp	0.14	0.23	0.42	0.25	0.27
Returning	0.16	0.75	0.53	0.52	0.67
New	0.84	0.25	0.47	0.48	0.33
Adopt day $G-1$	131.34	59.19	61.61	63.32	54.12
No. wks. played $G-1$	6.36	21.44	8.18	15.53	17.91
No. wks. owned $G-1$	33.69	43.87	43.56	43.29	44.58
Pct. wks. played $G-1$	0.21	0.48	0.20	0.36	0.40
Last wk. played $G-1$	34.61	42.63	28.75	37.18	38.93
Overall max lvl $G-1$	1.89	2.77	1.84	3.23	1.76
First wk. play $G-1$	2.44	8.35	4.29	6.36	6.60
First 4 wks. play $G-1$	6.95	25.80	11.36	19.28	20.29
First wk. max lvl $G-1$	1.13	1.09	1.09	1.23	1.06
First 4 wks. max lvl $G-1$	1.46	1.56	1.34	2.34	1.22
Total solo play $G-1$	11.67	62.74	20.65	13.88	72.03
Total friendly play $G-1$	3.83	10.35	4.02	5.53	6.92
Total competitive play $G-1$	2.28	38.34	3.66	13.75	6.26
Total ranked play $G-1$	1.98	17.78	2.19	41.39	1.97
Total play $G-1$	19.76	126.78	30.52	74.57	86.67
Overall high shares $G-1$	0.17	0.34	0.13	0.59	0.08
Overall HHI $G-1$	0.74	0.59	0.76	0.65	0.79
Above med. high lvls, $G-1$	0.14	0.67	0.17	0.70	0.24
Above med. low lvls, $G-1$	0.33	0.77	0.47	0.45	0.79
Above med. ranked play $G-1$	0.07	0.33	0.09	0.57	0.08
Above med. competitive play $G-1$	0.08	0.56	0.11	0.33	0.20
Above med. friendly play $G-1$	0.23	0.51	0.27	0.34	0.36
Above med. solo play $G-1$	0.16	0.63	0.30	0.21	0.69
Above med. last wk. played $G-1$	0.37	0.65	0.28	0.48	0.53
Above med. HHI $G-1$	0.75	0.96	0.76	0.93	0.83
Above med. no. wks. played $G-1$	0.14	0.68	0.21	0.50	0.60
Above med. high shares $G-1$	0.36	0.75	0.32	0.85	0.27
Above med. max lvl played $G-1$	0.25	0.65	0.22	0.77	0.21

Table 3.2: Clustered behavioral and demographic means, $K = 5$, continued.

Adopted during weekend $G-1$	0.40	0.18	0.26	0.22	0.19
Adopted 1st 2 weeks $G-1$	0.15	0.54	0.54	0.52	0.57
N	6950	5162	4949	3430	6224

Cluster 1 represents the low-play, low-duration, and low-level players. They are overwhelmingly likely to be new players. Additionally, they have by far the latest adoption times. Returning players in this cluster also tended to repeat these patterns in the previous game. Note they are the most likely to adopt on a weekend. I characterize this cluster as the "newcomers."

Cluster 2 is comprised of the greatest proportion of returning, as well as High Exp players. They are among the oldest players, and adopt the earliest. In almost every period, they are both the most likely to play and play the longest. In the previous game, they also tended to play the most. Similar to the newcomer group, they spent about half their previous-game play time in the Solo level. However, they spent a far greater share of time in the higher levels. They stand out as specialists in the Competitive level. This cluster can be characterized as the "hardcore" players.

Clusters 3 & 4 are both comprised of around half returning players but exhibit drastically different characteristics. Cluster 3 has similar play patterns to the newcomers while Cluster 4 has medium-level play but play at the highest levels of any group. Interestingly, Cluster 3 differs from newcomers in their very early adoption (in both current and previous game) and somewhat higher previous-game play. I consider this group the "casual fans" who exhibit relatively high loyalty but low utilization. On the other hand, Cluster 4 are the "ranked level specialists", and among returning players have the greatest share of Ranked playtime. They are not the highest utilization players in either the current or previous game. In the

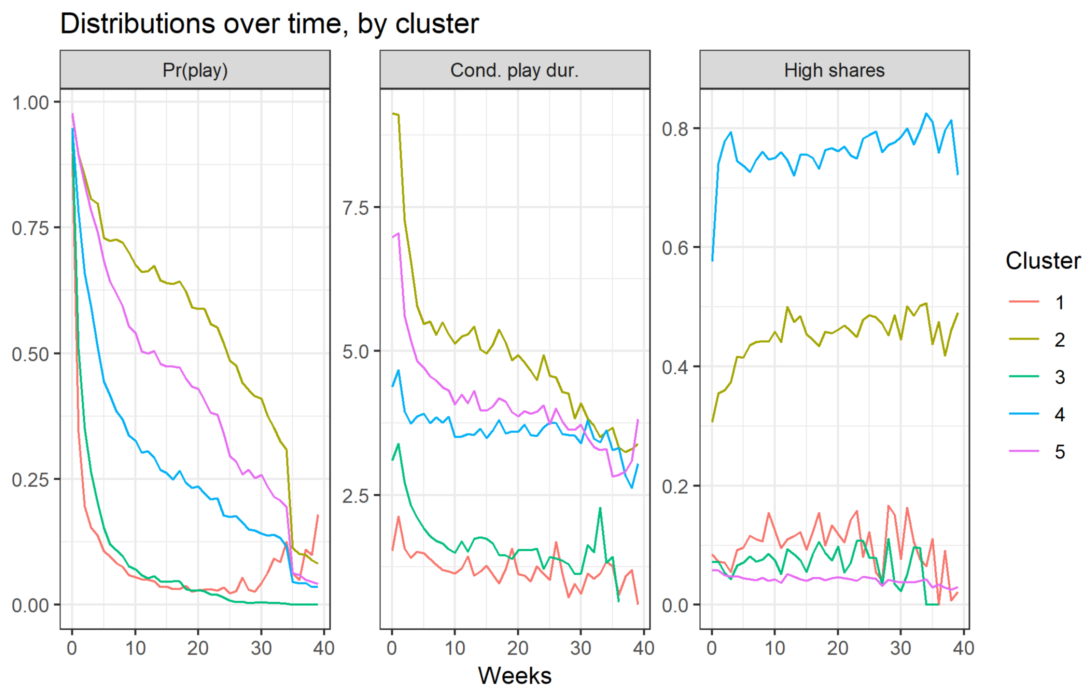


Figure 3.9: Evolution of extensive margin, intensive margin, and high level shares for each cluster.

first week of the previous game, their average max level is on par with the other clusters, but in the first four weeks their average max level is far greater than any cluster, perhaps suggestive of experimentation or quick learning rate.

Finally, Cluster 5 is roughly two-thirds returning players and about 40% High Exp players. They play nearly as much as Cluster 2, but at consistently the lowest levels. This pattern repeats for the previous game, where these players spend a far greater share of time in the Solo level than even the newcomers. This cluster can best be characterized as "solo specialists".

Finally I note that with $K = 2$ the clustering seems to be new vs advanced (high exp) players with low exp players roughly equally likely to belong to either cluster. Further, the advanced players are more likely to own a PlayStation. With $K = 3$ the primary intuition is that there is a group of players who play at a similar intensity to the advanced players but at a competitive level more like new players. In essence, this cluster is similar to the solo specialists in the full 5 cluster specification. These results are further detailed in the Appendix.

To more precisely understand the relationship between "true" demographics, game $G-1$ behavior, and game G behavior, I consider a set of regressions using various summary variables for the present game on these covariates. In Table 3.3 I regress log(total play)[5] on a selection of variables including console ownership and previous-game behavior for returning users. First, note the significant heterogeneity in cluster intercepts. One interpretation is that the covariate dummies alone are insufficient in teasing out behavioral differences between players in different clusters. Nonetheless these variables clearly affect playtime. For example, players with PlayStation, multiple, or newer consoles play more in total. In most groups, low exp and new players have similar total play while high exp players have significantly greater play. For a cluster with a relatively larger proportion of returning players, above median previous play in any level (except interestingly, Friendly) is associated with increased current total play. Finally, weekend adoption is negatively associated with playtime. Econometrically, it is important to note that the power of each within-cluster regression is highly dependent on the proportion of returning players with historical data to draw from. This is of course unsurprising, but serves to highlight the necessity of being able to differentiate new players.

I briefly summarize the above findings now. First, there is significant heterogeneity in play patterns, and broad categorizations such as high exp, low exp, and new provide useful but limited predictions for next-game play. Second, previous game behavior is a strong predictor of current game behavior. However, this information is completely missing for new players. Further, new players do not neatly fall into any one group, and in fact exhibit great

5. At the end of sample, so each user represents one data point, i.e. a cross-sectional regression.

	Dependent variable:					
	log(Total play G)					
	Cl. 1	Cl. 2	Cl. 3	Cl. 4	Cl. 5	All
(Intercept)	0.07	3.94***	0.99***	1.98***	3.64***	1.29***
	(0.05)	(0.03)	(0.05)	(0.06)	(0.03)	(0.03)
Owns PS	0.38***	0.06**	0.15***	0.21***	0.04**	0.37***
	(0.04)	(0.02)	(0.04)	(0.05)	(0.02)	(0.02)
Owns older cons.	−0.22***	−0.13***	−0.23***	−0.17**	0.002	−0.33***
	(0.05)	(0.05)	(0.06)	(0.07)	(0.03)	(0.03)
Owns multiple cons.	0.81***	0.22***	0.49***	0.31**	0.06	0.87***
	(0.22)	(0.06)	(0.18)	(0.14)	(0.05)	(0.07)
No. games owned	0.15***	−0.002	0.03	0.03	0.01	0.11***
	(0.04)	(0.02)	(0.03)	(0.04)	(0.01)	(0.02)
No. yrs in samp.	−0.05	0.02	0.06**	0.04	0.02**	0.02
	(0.03)	(0.01)	(0.02)	(0.03)	(0.01)	(0.01)
Low exp	−0.03	−0.12*	0.23***	0.20	−0.01	0.35***
	(0.12)	(0.07)	(0.09)	(0.14)	(0.04)	(0.05)
High exp	0.27	0.32***	0.28**	0.86***	0.44***	1.01***
	(0.27)	(0.09)	(0.14)	(0.17)	(0.06)	(0.07)
Above med. ranked play $G-1$	−0.24	0.15***	−0.06	0.64***	0.09*	0.21***
	(0.21)	(0.03)	(0.12)	(0.09)	(0.05)	(0.05)
Above med. competitive play $G-1$	−0.23	0.29***	0.02	−0.03	0.39***	0.32***
	(0.21)	(0.04)	(0.12)	(0.08)	(0.04)	(0.04)
Above med. friendly play $G-1$	0.03	0.04	0.08	−0.14*	−0.07***	0.05
	(0.12)	(0.03)	(0.07)	(0.07)	(0.03)	(0.03)
Above med. solo play $G-1$	0.27	0.11***	0.58***	−0.11	0.15***	0.47***
	(0.18)	(0.04)	(0.09)	(0.09)	(0.04)	(0.04)
Above med. last wk. played $G-1$	0.14	0.18***	0.19***	0.25***	0.13***	0.40***
	(0.11)	(0.03)	(0.07)	(0.07)	(0.02)	(0.03)
Above med. HHI $G-1$	0.16	0.003	0.05	0.10	−0.05	0.10**
	(0.13)	(0.07)	(0.08)	(0.15)	(0.03)	(0.05)
Above med. no. wks. played $G-1$	0.11	−0.05	−0.18*	0.20**	−0.02	0.27***
	(0.20)	(0.04)	(0.10)	(0.10)	(0.03)	(0.05)
Above med. high shares $G-1$	−0.01	0.03	0.01	−0.16	−0.06	0.06
	(0.14)	(0.04)	(0.09)	(0.13)	(0.04)	(0.05)
Above med. max lvl played $G-1$	−0.09	−0.02	−0.02	0.004	−0.17***	0.01
	(0.15)	(0.03)	(0.09)	(0.11)	(0.04)	(0.04)
Adopted during weekend $G-1$	−0.04	−0.19***	−0.24***	−0.18**	−0.13***	−0.33***
	(0.10)	(0.03)	(0.07)	(0.08)	(0.03)	(0.04)
Adopted 1st 2 weeks $G-1$	−0.03	−0.08***	−0.07	0.02	−0.06**	−0.19***
	(0.14)	(0.03)	(0.06)	(0.08)	(0.03)	(0.03)
Observations	6,950	5,162	4,949	3,430	6,224	26,715
R^2	0.03	0.21	0.07	0.25	0.18	0.30
Note:						*p<0.1; **p<0.05; ***p<0.01

Table 3.3: OLS of log(Total play G) for each cluster and overall.

variation in play patterns. Third, adoption time does not appear to have an unconditionally strong (or even consistently-directional) effect on play patterns. For example, on average, the casual fans adopt the earliest but play very little while the ranked specialists are the second-latest adopting group and play at the highest levels. Fourth, while the data contains limited true demographic information[6].

3.6 Stylized facts

So far I have given an exploratory description of the data. Using observations in the previous sections, I will try to formulate some stylized facts about player behavior:

First, there is evidence in support of player learning-by-doing. Despite significant heterogeneity in behavior, on average players tend to shift to competitive levels and specialize. Furthermore, at the moment of adoption players tend to reduce their level of competition and experiment across a greater variety of levels. These behaviors are consistent with players learning about their true valuations for each level as they play. Additionally, I find evidence that suggests it is not the adoption timing or player age, but rather amount of prior experience that influences player behavior–returning low utilization players play very similarly to new ones.

Second, I find interesting patterns along both the extensive and intensive margins. In particular, using both latent and observed segmentation I find that over time players are simultaneously less likely to play and play for shorter durations. Both patterns are consistent with findings in the literature. For example, Nevskaya and Albuquerque (2019) find that declining novelty is a significant driver of usage among *World of Warcraft* players. Tyack et al. (2016) find that alongside this loss of enjoyment, an increasingly dominant outside option

6. The only true demographic I observe is console information. However, as I show console ownership does have explanatory power in behavior. This is supported by, e.g. branding, social, and performance differences between consoles studied.

(e.g. time constraints, competing games) is important in determining player behavior. It is clear that these factors should affect player behavior along both margins. While there is a plethora of research considering each in turn, a model that can jointly predict along both margins would be highly useful.

Third, even though I observe a rich set of historical data (at least for returning players), there is still significant unobserved heterogeneity in behavior. This heterogeneity does not appear to readily map to any observed characteristics, and is especially difficult to identify for new players. At the same time, it may be of significant interest to the firm to use this rich data to target players for interventions that may, for example, improve engagement.

CHAPTER 4

MODEL

4.1 Mapping stylized facts to model

The workhorse model I employ belongs in the class of multiple discrete-continuous models, which allows for corner solutions. Intuitively, the probability of an interior solution translates to the extensive margin and the conditional distribution of the strictly positive consumption is the intensive margin. I further choose a specification with a basis in microeconomic theory and utility maximization, allowing recovery of deep structural parameters that I use to perform policy simulations as well as avoid the endogeneity of conditioning on play. I explain evolving play patterns using a combination of decreasing outside satiation, declining novelty, and Bayesian learning, where uncertainty and risk aversion are key drivers of behavior. I consider learning in the standard marketing sense: players have beliefs over their true match value for each level and through play obtain signals for that match value. In this sense, my model is similar to the one used by Narayanan et al. (2007), who utilize Bayesian learning in a discrete-continuous model to explain phone plan-choice and usage under uncertainty. As reasoned previously, I will **not** explicitly model adoption and simply condition on it. However, the Bayesian framework can be highly useful to account for whatever endogeneity is induced by conditioning on adoption. In particular, consider initial play after adoption as dependent on priors derived from past play. Accounting for selection is then a simple matter of modeling priors as functions of past play[1]. This learning framework naturally both lends itself to explaining the different play patterns observed between player types and similarities in their evolution. I begin by introducing a fully general econometric model and subsequently specialize that model to the one I will use.

1. This is supported by the observation that low exp players (recall they are defined as those who have played, on average, less than 1 hr/week) behave similarly to new players. In a Bayesian learning framework, we can say that their priors are still weak.

4.2 General econometric model

Consider a game release at time t_0. Individual i adopts the game at time $t_i > t_0$ and may costlessly access the game in any period $t = t_i, t_i + 1, \ldots, T$. The game is a quintessential experience good. For the remainder of the exposition I will suppress the i subscript. During each period the player allocates his/her luxury time to either the game or other recreational activities (i.e. an outside good). Additionally, the game itself has several "levels" $1, \ldots, J$. Wlog I assume these levels are ordered in increasing competition.

I assume players are uncertain about their true match values over the game's levels and that they learn in a Bayesian manner over these match values as they play. Further, I assume they are static optimizers which rules out experimentation in the typical marketing sense (i.e. forward-looking consumers). In the Appendix I provide alternate mechanisms under myopic consumers that can nonetheless explain data patterns resembling dynamics.

I will consider a variation on the model formulated in Bhat (2008) and assume consumers solve the following problem at time t:

$$
\begin{aligned}
&\max \mathbb{E}_{\tilde{q}_t}\left[u_0(x_{0t}) + \sum_{j>0} u_j(x_{jt}; q_{jt})\right] \; s.t. \; x_{0t} + \sum_{j>0} x_{jt} = E \\
&u_0(x_{0t}) = \frac{1}{\alpha_0}\Psi_{0t} x_{0t}^{\alpha_0} \\
&u_j(x_{jt}) = \frac{\gamma_j}{\alpha_j}\Psi_{jt}(q_{jt})\left[(\frac{x_{jt}}{\gamma_j} + 1)^{\alpha_j} - 1\right]
\end{aligned}
\tag{4.1}
$$

Here E represents the recreational time budget, x_{0t} is time allocated to non-game activities (i.e. "numeraire" good). Ψ_{jt} takes on the interpretation of baseline marginal utility at the point of 0 consumption, $\alpha_j < 1$ are satiation parameters, and $\gamma_j > 0$ are translation parameters that allow corner solutions. As in Bhat (2005) I introduce a RUM by parameterizing $\Psi_{0t} = \exp(\epsilon_{0t})$ and define:

$$\Psi_{jt} : \mathbb{E}_{\tilde{q}_t}[\Psi_j(q_{jt})] = \exp(\mu_{jt} - r\sigma_{jt}^2 + Z_{jt}\beta + \epsilon_{jt}) \tag{4.2}$$

I assume that $q_{jt} \sim N(\mu_{jt}, \sigma_{jt}^2)$, r is risk aversion, Z_{jt} are observed characteristics, and ϵ_t are demand shocks unobserved to the econometrician. The primary difference between my model and Bhat's (2008) is the expectation taken over q_{jt}.

Note that the utility specification assumes additive separability, which as outlined in Bhat (2008) immediately implies:

- None of the goods are a priori inferior. Considering "income" to be luxury time, I believe this is an acceptable assumption.
- All goods are strictly Hicksian substitutes. Note that in the present context prices do not have a readily available definition. Abstractly, I consider a price increase to be anything that increases the time spent in a match holding utility fixed (recall only total duration is observed, not individual matches, so this can only be an abstract argument). Perhaps loading or matchmaking times are longer. Under the assumption that these factors do not directly affect (dis)utility, the Hicksian substitute assumption roughly states that if, say matchmaking times for level j increase time spent in mode j' will increase, compensating for utility.
- Marginal utility with respect to any good is independent of consumption levels of all other goods. This is perhaps a problematic assumption in that it rules out cross-satiation effects. However, I believe there are several accommodations I can make to alleviate potential violations. First, I aggregate play sessions up to the weekly level. While intra-day cross-satiation seems quite likely, abstracting away from a session- or daily-level budget problem can also abstract away from the satiation effects on those

scales. Second, because I treat the budget problem over all leisure activities, inside consumption will generally represent a small percentage of total allocation. Thus, it may be reasonable that *relative to total budget*, cross-satiation is not a significant effect. Finally, to induce correlations for marginal utilities for the inside goods, I impose correlation structures in the error terms. In particular, I assume errors are distributed nested extreme value with all inside goods (game levels) in one nest and the outside good in another.

Additionally, I assume multiplicative separability between q_j and x_j so that discoveries in preferences q_j act solely as multipliers to baseline marginal utility.

4.2.1 Bayesian learning specification

For the present I consider a general case where the consumer has prior preferences $\tilde{q}_0 = (q_{10}, ..., q_{J0})$ distributed multivariate normal with mean $(\bar{q}_{10}, ..., \bar{q}_{J0})$ and covariance matrix Σ_0 with $(\Sigma_0)_{j,j'} \equiv \sigma^2_{jj'}$. Wlog, we can let any or all of these parameters depend on player characteristics and history. In the simulation section and the Appendix I discuss alternate specifications as well as identification.

Next, I assume that consumers are learning not only over choice occasions (extensive margin) but the duration decision as well (intensive margin). As I explained in the stylized facts, both margins contain information. Intuitively, a consumer who spent more time with level j should resolve a greater amount of uncertainty than one with less, *ceteris paribus*. To model this I discretize the learning process. Consider some small unit for time Δ. I assume that when a player spends x_{jt} time playing level j they receive $k_{jt} \equiv \lceil \frac{x_{jt}}{\Delta} \rceil$ signals informing them about their true preferences. Of course as $\Delta \to 0$ this approaches the continuous learning process. For larger Δ preference uncertainty and learning lose explanatory power over variation in x_{jt}. Δ is not identified in the model and must be calibrated ex post. Empirically, an alternative to choosing a small Δ is to consider the abstraction of "game matches" which are in fact discrete. Individual game sessions are not observed in the data but it would seem that setting Δ to an average session length may serve as a useful approximation. A third solution would be to discrete the choice space entirely and solve a multiple discrete-count problem (see, e.g. von Haefen and Phaneuf, 2003). I do not consider that particular approach in this paper.

In the current specification let R_{jts} be one signal observed by the consumer about level j at time t ($s = 1, ..., k_{jt}$). I assume signals are informative and unbiased for true level-specific preferences. Let $q_i^* = (q_{i1}^*, ..., q_{iJ}^*)$ denote true preferences. For exposition I consider the general case where signals are serially uncorrelated but potentially correlated contemporaneously. Let $K_t \equiv \sum_j k_{jt}$ be the total number of discretized signals observed by the consumer at time t across all levels. Then the $K_t \times 1$ vector $\vec{R}_t \equiv \{R_{jts}\}_{j=1,...,J;s=1,...,k_{jt}}$ will have $K_t \times K_t$ covariance matrix Σ_{Rt}. I now define the $J \times K$ collapsing matrix M_t whose j^{th} row is comprised of 1's in the $\sum_{j'<j} k_{j'} + 1, ..., \sum_{j'<j} k_{j'} + k_j$ positions and 0 elsewhere. For example, if $J = 3$ and $k_1 = 1, k_2 = 3, k_3 = 2$, we have:

$$M_t = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 \end{bmatrix}$$

Even without explicit assumptions over the structure of Σ_{Rt} we can still derive posteriors. In particular, let the consumer's period t prior preferences be denoted $\tilde{q}_{it} \sim N(\bar{q}_{it}, \Sigma_{qt})$. Then it can be shown[2] that the posterior preferences (given K_t total signals) has distribution $\tilde{q}_{i,t+1} \equiv \tilde{q}_{it}|\vec{R}_t \sim N(\bar{q}_{i,t+1}, \Sigma_{q,t+1})$ with:

2. See Appendix.

$$\begin{aligned}\Sigma_{q,t+1}^{-1} &= M_t \Sigma_{Rt}^{-1} M_t' + \Sigma_{qt}^{-1} \\ \bar{q}_{i,t+1} &= \Sigma_{q,t+1}\left(\Sigma_{qt}^{-1}\bar{q}_{it} + M_t \Sigma_{Rt}^{-1} \vec{R}_t\right)\end{aligned} \tag{4.3}$$

4.2.2 Solution to the constrained optimization

Equipped with a closed form expression for $(\mu_{jt}, \sigma_{jt}^2)$, I revisit Eq. (4.1). Note that under my assumptions for $\Psi_{jt}(\cdot)$, the expected utility maximization problem here takes on the exact same form as in Bhat (2005). In particular, the KKT conditions will be:

$$\begin{aligned}&V_{jt} + \epsilon_{jt} = V_{0t} + \epsilon_{0t}, \quad x_{jt}^* > 0 \\ &V_{jt} + \epsilon_{jt} < V_{0t} + \epsilon_{0t}, \quad x_{jt}^* = 0 \\ &V_{0t} = (\alpha_0 - 1)\log(x_{0t}^*), \quad V_{jt} = \mu_{jt} - r\sigma_{jt}^2 + Z_{jt}\beta + (\alpha_j - 1)\log\left(\frac{x_{jt}^*}{\gamma_j} + 1\right)\end{aligned} \tag{4.4}$$

Note that it is always possible to compare V_{jt} to the outside option V_{0t} because $x_{0t}^* > 0$ w.p. 1. This is due to the lack of a translation parameter for the outside option[3]. Intuitively, consumers should not spend all their leisure time on playing a single video game. Of course, there are extreme outliers in the data (e.g. over 40 hours/week played) but they occur so rarely that either dropping or censoring them seem reasonable enough.

4.2.3 Likelihood

In my empirical application I consider the case when errors are correlated–this assumption is crucial as it allows cross-satiation for game levels. In this section I present a simpler likelihood with independent errors $\epsilon \overset{i.i.d.}{\sim}$ Gumbel. Under this specification an elegant, closed-form

3. A warning to this approach, particularly in the model without prices, is that the problem is no longer scale invariant. In particular, scaling budget and consumption down by a positive factor will tend to distort the model, with an extreme result that consumers are predicted to consume the outside good more as their outside good satiation **increases**. More details are provided in the Appendix.

expression is available for the likelihood of the solution vector $(x_{0t}^*, x_{1t}^*, x_{2t}^*, ..., x_{mt}^*, 0, ..., 0)$ for $m \leq J$, where $m \geq 0$ is the number of distinct levels played (Bhat 2005):

$$Pr(x_{0t}^*, x_{1t}^*, x_{2t}^*, ..., x_{mt}^*, 0, ..., 0) = \det(T_{V_t}) \left[\frac{\prod_{i=0}^{m} \exp V_{it}}{(\sum_{i=0}^{J} \exp V_{it})^{(m+1)}} \right] \cdot m! \tag{4.5}$$

where T_{V_t} is the Jacobian of the mapping $\epsilon_{jt} \to V_{j0} - V_{jt}$ and has closed form:

$$\det(T_{V_t}) = \left(\sum_{i=0}^{m} \frac{x_{it}^* + \gamma_i}{1 - \alpha_j} \right) \left(\prod_{i=0}^{m} \frac{1 - \alpha_j}{x_{it}^* + \gamma_i} \right)$$

I denote the probability in Eq. (4.5) by $p(x_{it}^*, \tilde{q}_{it}; \theta)$ to make explicit its dependence on not just the choice vector but the current belief vector as well. In my current application signals are unobserved so that $\tilde{q}_{it}$ is not precisely known. To resolve this I follow standard marketing practice and simulate the unobserved learning process. If, for each consumer, I draw S sequences of signals I can derive $\tilde{q}_{it}^s, s = 1, ..., S$ and approximate a consumer's data likelihood as:

$$L_i(\theta) \approx \frac{1}{S} \sum_s p(x_{it_i}^*, \tilde{q}_0) \prod_{t=t_i+1}^{T} p(x_{it}^*, \tilde{q}_{it}^s; \theta) \tag{4.6}$$

When NEV errors are used, Eq. 4.5 must be modified, but the remaining likelihood components are the same. Bhat (2008) provides a constructive general derivation for consumption probabilities and I present the expression for my particular nesting structure, which despite its unwieldiness remains analytic, in the Appendix.

4.3 Specific model

In the previous section I presented a fully general model nesting Bayesian learning in a MDCEV framework. Here I describe the particular specification I use in my empirical application. In particular, I restrict the utility specification given in Eq. (4.1) due to identification concerns, specify the assumed generative process for the covariance matrices in the learning framework, and allow for discrete consumer heterogeneity via latent segmentation.

4.3.1 Restricted utility specification

Bhat (2008) shows that the full utility specification given in Eq. (4.1) is typically non-identified in empirical settings. In particular, α-models that restrict $\gamma_j = 1 \; \forall j$ and γ-models that restrict $\alpha_j = 0 \; \forall j$ tend to fit data equally well. In the Appendix I show using simulation that the γ-model is not identified in the presence of learning and give some intuition for that result. On the other hand, the α-model **is** identified. In this case I reformulate the consumer's problem as:

$$\begin{aligned}
&\max \mathbb{E}_{\tilde{q}_t}\left[u_0(x_{0t}) + \sum_{j>0} u_j(x_{jt}; q_{jt})\right] \; s.t. \; x_{0t} + \sum_{j>0} x_{jt} = E \\
&u_0(x_{0t}) = \frac{1}{\alpha_0}\Psi_{0t} x_{0t}^{\alpha_0} \\
&u_j(x_{jt}) = \frac{1}{\alpha_j}\Psi_{jt}(q_{jt})\left[(x_{jt}+1)^{\alpha_j} - 1\right]
\end{aligned} \tag{4.7}$$

This is similar to Bhat's (2008) first utility form with an outside good. Consequently, the KKT conditions are:

$$\begin{aligned}
&V_{jt} + \epsilon_{jt} = V_{0t} + \epsilon_{0t}, \quad x_{jt}^* > 0 \\
&V_{jt} + \epsilon_{jt} < V_{0t} + \epsilon_{0t}, \quad x_{jt}^* = 0 \\
&V_{0t} = (\alpha_0 - 1)\log(x_{0t}^*), \quad V_{jt} = \mu_{jt} - r\sigma_{jt}^2 + Z_{jt}\beta + (\alpha_j - 1)\log\left(x_{jt}^* + 1\right)
\end{aligned} \tag{4.8}$$

The likelihood remains unchanged except in the computation of the Jacobian term there is the restriction $\gamma_i = 1 \; \forall i > 0$.

4.3.2 Covariance specification for learning components

For the present I consider the simplest specification: uncorrelated learning. Specifically, I define the prior covariance to be diagonal: $(\Sigma_0)_{j,j'} = \mathbb{I}(j = j') \cdot \sigma_{j0}^2$. Further, I assume

$\text{Var}(R_{jts}) = \sigma^2_{Rj}$, $\text{Cov}(R_{jts}, R_{j'ts'}) = 0 \; \forall j \neq j', s \neq s'$. The signal covariance matrix is thus also diagonal. Note that in the Bayes update, this implies the posterior precision can be written:

$$\Sigma^{-1}_{q,t+1} = \Sigma^{-1}_{qt} + \Lambda_{Rt}$$
$$(\Lambda_{Rt})_{j,j'} = \mathbb{I}(j = j') \cdot \frac{k_{jt}}{\sigma^2_{Rj}}$$

Simple algebra reveals that posteriors can be computed without recursion:

$$\sigma^2_{tj} = (\frac{1}{\sigma^2_{0j}} + \frac{N_{jt}}{\sigma^2_{Rj}})^{-1}$$
$$\mu_{tj} = \sigma^2_{jt} \left[\frac{q_{0j}}{\sigma^2_{0j}} + \frac{N_{jt-1} \sum_{t'<t} R_{jt'}}{\sigma^2_{Rj}} \right]$$

where $N_{jt} \equiv \sum_{t'<t} k_{jt}$ is the total number of signals observed before time t for good j and $\sum_{t'<t} R_{jt'}$ is the cumulative signal for good j observed before time t with distribution $N(N_{jt}q^*_j, N^2_{jt}\sigma^2_{Rj})$ under the assumption of i.i.d. unbiased signals[4]. This non-recursive formulation also has significant implications for computation time, which is a non-trivial burden for this model.

I discuss alternate learning specifications that incorporate correlated learning in the Appendix.

4. Note this formulation is very similar to that in the single-discreteness model in Jiang et al., 2006. The treatment of discretized learning in continuous models is in fact identical to a multi-period single-discreteness or single-period multiple-discreteness models.

4.3.3 Latent segmentation

As described in Section 3, there is evidence of significant consumer heterogeneity in the sample. To account for this heterogeneity I will employ a latent segmentation approach[5]. Following standard marketing protocol, I assume there are G homogeneous but unobserved segments in the population. At best, the researcher can probabilistically assess individual segment identities using observed individual characteristics W_i and segment identity loadings δ_g:

$$Pr(i \in g) \equiv \pi_g = \frac{\exp(W_i \delta_g)}{1 + \sum_{g'=2}^{G} \exp(W_i \delta_{g'})}$$

There are two empirical considerations here. First, a well-documented limitation of latent segmentation models is their poor performance even under a moderate number of specified segments. In particular, if a non-segmentation model specifies P parameters, the segmentation model must solve for at least $GP + (G-1)|W|$ parameters. In addition, the number of observations used in identifying segment-specific parameters decreases in G. As a result most studies consider only $G \in \{2, 3\}$[6]. Second, in my current application I am heavily constrained in the set of exogenous observed characteristics W_i. As a result I instead estimate the intercept-only formulation:

$$\pi_g = \frac{\exp(\delta_g)}{1 + \sum_{g'=2}^{G} \exp(\delta_{g'})}$$

This presents the additional advantage of amenability to direct likelihood maximization: due to stability issues with quasi-Newton techniques in latent segmentation models (cf. Bhat 1997), expectation-maximization (EM) is often employed instead (for an example of latent segmentation in a MDCEV model estimated with EM, see Sobhani et al., 2013). In

5. For a survey of methods dealing with consumer heterogeneity in marketing, see Allenby and Rossi (1998).

6. See Bhat (1997) for a more complete discussion of issues commonly encountered in latent segmentation models.

contrast, an ex ante constant prior with $G = 2$ segments only requires estimation of a single θ parameter. Under latent segmentation, the consumer-level likelihood from Eq. (4.6) is expanded to:

$$L_i(\theta) = \sum_g \pi_g L_i(\theta | i \in g) = \sum_g \pi_g L_i(\theta_g)$$

4.4 Identification

In this section I discuss identification for this model. Identification of the multiple discrete-continuous model with generalized extreme value (GEV) with random coefficients (continuous unobserved heterogeneity) is established in Bhat (2008). Separately, identification of Bayesian learning in a single discrete model with multiple signals per period is discussed in Coscelli and Shum (2004). As noted previously, when a discrete-continuous outcome is approximated by discretized signals, the results from Coscelli and Shum (2004) will apply. However, it is unclear whether these results extend to any new parameters introduced by MDC. For example, as described earlier and in the Appendix, the linear satiation ($\lambda-$) formulation is not compatible with learning while the power satiation ($\alpha-$) formulation is. In the remainder of this section I extend my study to the rest of the model, seeking to provide intuition for identification, describe the critical role of tuning parameters in estimation, and provide heuristics for estimating this class of models.

In the following simulation, I consider $N = 1,000$ consumers who solve Eq. (4.7) with $J = 3$ inside goods and budget $E = 100$ over $T = 50$ time periods. I assume each good has 3 time-varying attributes, so $M \equiv |Z_{ijt}| = 3$. Consumers are risk averse with $r = 0.10$ and satiation parameter $\alpha_0 = 0.8$ for the outside option and $\alpha_1 = \ldots = \alpha_J = 0.35$ for each inside option.

For the first simulation I assume individuals have common true preferences $q_1^* = \ldots =$

$q_J^* = 0$ and vary priors: $\bar{q}_{01} = -1, \bar{q}_{02} = 0, \bar{q}_{03} = 1$, which correspond to pessimistic, rational, and optimistic priors, respectively. I assume prior variances are common as well as signal variances: $\sigma_{10}^2 = \ldots = \sigma_{J0}^2 = 20, \sigma_{\nu_1}^2 = \ldots = \sigma_{\nu_J}^2 = 4.5$. The prior and signal variances are chosen so that by the end of simulation most consumers have selected an inside good at least once, and ones that have multiple choices have updated their beliefs to be somewhat close to the truth. Finally I assume learning is discretized over $\Delta = 12.5$, which roughly corresponds to one signal per day if the budget were uniformly distributed over the week.

Simulation is performed solving the KKT conditions from Eq. (4.8) using the method of moving asymptotes algorithm in the R package `nloptr`. Estimation is performed using the sequential quadratic programming algorithm (SQP) implemented by the Knitro solver (called from Matlab[7]). To ensure constraints on certain parameters are satisfied, I re-parameterize $\alpha_j = \frac{1}{1+\exp(-\delta_j)}$ as recommended in Bhat (2005) and all variance parameters as $\sigma_s^2 = \exp(\gamma_s)$. Further, it is known that to achieve asymptotic consistency and efficiency, we must have $\lim_{N\to\infty} \sqrt{N}/S \to 0$ (see, e.g. Lee, 1995). To gain some intuition for the relative simulation draws needed in my model, in Table 4.1 I provide estimation results varying number of simulation draws S.

First note the omission of certain parameters. It is well-known in the marketing literature that risk aversion is difficult to pin down in learning models. In fact, Coscelli and Shum (2004) analytically show that $(\bar{q}_{0j}, r)$ are not jointly identified in the single-good case. They further provide simulation-based evidence that the two are also not jointly identifiable with $J > 1$. This result naturally extends here. I fix risk aversion at its true value and note that bias in fixing risk aversion is given by the following relation: $C = \bar{q}_{0j} - r\sigma_{0j}^2$ for a constant C. I further find that $(\sigma_{0j}^2, \sigma_{\nu_j}^2)$ (i.e. prior and signal variances) are difficult to jointly identify so I fix one (prior) at its true value. By jointly fixing these parameters (regardless of

7. Documentation for the solver can be found on the Artelys website.

		$S=100$		1000		2000		5000	
	True	Est	SE	Est	SE	Est	SE	Est	SE
β_1	1	0.95	0.02	0.97	0.02	0.97	0.02	0.97	0.02
β_2	-1	-0.96	0.02	-0.98	0.02	-0.98	0.02	-0.98	0.02
β_3	0.5	0.5	0.01	0.51	0.01	0.51	0.01	0.51	0.01
α_1	0.35	0.37	0.00	0.36	0.00	0.36	0.00	0.36	0.00
α_2	0.35	0.37	0.00	0.36	0.00	0.36	0.00	0.36	0.00
α_3	0.35	0.36	0.00	0.36	0.00	0.35	0.00	0.35	0.00
α_0	0.8	0.79	0.00	0.79	0.00	0.79	0.00	0.79	0.00
$\bar{q}_1$	-1	-0.98	0.03	-0.97	0.03	-0.97	0.03	-0.97	0.03
$\bar{q}_2$	0	-0.08	0.00	-0.07	0.00	-0.07	0.00	-0.07	0.00
$\bar{q}_3$	1	0.96	0.04	0.97	0.04	0.98	0.04	0.98	0.04
q_1^*	0	-0.08	0.00	-0.05	0.00	-0.04	0.00	-0.04	0.00
q_2^*	0	-0.07	0.00	-0.05	0.00	-0.04	0.00	-0.04	0.00
q_3^*	0	-0.08	0.00	-0.05	0.00	-0.05	0.00	-0.05	0.00
σ^2_{u1}	4.5	3.09	0.14	3.89	0.19	3.87	0.19	4.15	0.21
σ^2_{u2}	4.5	3.41	0.16	4.07	0.21	4.26	0.20	4.33	0.22
σ^2_{u3}	4.5	3.07	0.15	3.87	0.21	3.89	0.21	4.1	0.22

Table 4.1: Simulation results varying no. of simulation draws S

correctness) I am able to recover: (1) ratio of signal to prior variance, (2) prior means up to a constant multiple of prior variances.

Second note that under my assumptions even using a small number of draws (e.g. $S = 100$) I am able to recover almost all parameters except perhaps the posterior variance[8] which tends to be biased downwards. As I increase the number of draws this bias continues to diminish (and uncertainty increases), until the true posterior variance is contained in the confidence interval.

In the follow sections I will study the result of selecting an incorrect risk aversion parameter r, incorrect prior variance σ^2_{0j}, and incorrect discretization parameter Δ.

8. I note that coverage, in particular for true means and satiation parameters, is troublesome at this number of draws. However, an eye test shows at $S = 100$ I am still able to recover useful estimates.

4.4.1 *Bias introduced by incorrect risk aversion parameter*

In Table. 4.2 I give simulation results for the model varying r, the risk aversion parameter fixing $S = 1000$. In the first set I assume the researcher knows the true risk aversion parameter, which gives the same results as Table. 4.1, $S = 1000$. In the next two panels I assume the researcher has under- and over-guessed r, respectively by a factor of 2. Because my learning framework is purely within the realm of standard Bayesian models, I find the results from Coscelli & Shum hold here and the bias is easily characterizable as:

$$\text{Bias}(\hat{\bar{q}}_{0j}) = (r_{guess} - r)\sigma_{0j}^2$$

Note that in general, mis-specifying r still allows successful recovery of all other parameters.

		$\tilde{r} = r$		$\tilde{r} = 0.5r$		$\tilde{r} = 2$	
	True	Est	SE	Est	SE	Est	SE
β_1	1	0.97	0.02	0.97	0.02	0.98	0.02
β_2	-1	-0.98	0.02	-0.98	0.02	-0.98	0.02
β_3	0.5	0.51	0.01	0.51	0.01	0.51	0.01
α_1	0.35	0.36	0.00	0.36	0.00	0.36	0.00
α_2	0.35	0.36	0.00	0.36	0.00	0.36	0.00
α_3	0.35	0.35	0.00	0.35	0.00	0.35	0.00
α_0	0.8	0.79	0.00	0.79	0.00	0.79	0.00
$\bar{q}_1$	-1	-0.97	0.03	-1.98	0.07	1.03	0.04
$\bar{q}_2$	0	-0.07	0.00	-1.07	0.04	1.93	0.07
$\bar{q}_3$	1	0.98	0.04	-0.02	0.00	2.98	0.12
q_1^*	0	-0.04	0.00	-0.04	0.00	-0.04	0.00
q_2^*	0	-0.04	0.00	-0.04	0.00	-0.04	0.00
q_3^*	0	-0.05	0.00	-0.05	0.00	-0.05	0.00
σ_{u1}^2	4.5	4.15	0.21	4.11	0.20	4.19	0.21
σ_{u2}^2	4.5	4.33	0.22	4.42	0.23	4.43	0.23
σ_{u3}^2	4.5	4.1	0.22	4.18	0.22	4.22	0.23

Table 4.2: Simulation results mis-specifying risk aversion parameter r with $\tilde{r}$.

4.4.2 Bias introduced by incorrect prior variance

In Table 4.3 I give simulation results for the model varying $\sigma^2_{\nu_j}$, the prior variance fixing $S = 1000$. For alternative $j = 1$ I set the prior variance to the correct value (20), $j = 2$ I assume it is mis-specified at half its true value (10) and for $j = 3$ I assume it is mis-specified at twice its true value (40). From the previous section, it is clear that mis-specifying prior variance will also bias prior mean, even in the presence of a properly specified risk aversion parameter. Further, a mis-specified risk aversion parameter will only serve to further amplify the bias with a mis-specific prior variance.

Note here that signal variances can also be biased with mis-specified prior variances. In particular, mis-specifying the prior variance upward (downward) seems to bias the corresponding signal variance downward (upward). However, the magnitude of the bias is unclear given that it tends to be a difficult parameter to pin down (even with 1000 simulation draws).

	True	Est	SE
β_1	1	0.97	0.02
β_2	-1	-0.98	0.02
β_3	0.5	0.51	0.01
α_1	0.35	0.36	0.00
α_2	0.35	0.36	0.00
α_3	0.35	0.35	0.00
α_0	0.8	0.79	0.00
$\bar{q}_1$	-1	-1.96	0.07
$\bar{q}_2$	0	-0.07	0.00
$\bar{q}_3$	1	2.98	0.12
q^*_1	0	-0.03	0.00
q^*_2	0	-0.04	0.00
q^*_3	0	-0.05	0.00
σ^2_{u1}	4.5	4.12	0.21
σ^2_{u2}	4.5	4.33	0.22
σ^2_{u3}	4.5	3.87	0.20

Table 4.3: Simulation results mis-specifying prior variance σ^2_{0j}: $j = 1$ is halved, $j = 2$ is correct, $j = 3$ is doubled.

4.4.3 Bias introduced by incorrect discretization parameter

In Table 4.4 I give simulation results for the model varying Δ, the learning signal discretization parameter fixing $S = 1000$. In the first set I assume the researcher knows the true value, which gives the same results as Table. 4.1, $S = 1000$. In the next two panels I assume the researcher has under- and over-guessed Δ, respectively by a factor of 2. Intuitively, $(\Delta, \sigma^2_{\nu_j})$ characterize rate of learning with the former determining number of signals and the latter the (lack of) informativeness of each signal. Thus, when the researcher mis-specifies Δ there is a compensating effect on $\sigma^2_{\nu_j}$. For example, if Δ is under-guessed (second panel), then the model is estimating learning with too many signals and $\sigma^2_{\nu_j}$ must be smaller and vice versa (third panel).

Specifically, in vanilla Bayesian learning after K signals the variance is:

$$\sigma^2_K = \left(\frac{1}{\sigma^2_0} + \frac{K}{\sigma^2_\nu}\right)^{-1}$$

If K were instead constructive and mapped from total duration X and discretization parameter Δ, i.e. $K = \lceil \frac{X}{\Delta} \rceil$, then we would have:

$$\sigma^2_X(\Delta) = \left(\frac{1}{\sigma^2_0} + \left\lceil \frac{X}{\Delta} \right\rceil \frac{1}{\sigma^2_\nu}\right)^{-1}$$

For small enough Δ (i.e. so that the step function for K is roughly smooth), we would expect $(\Delta, \sigma^2_{\nu_j})$ to in fact be inversely proportional. Finally, note that mis-specification of Δ only biases estimation on $\sigma^2_{\nu_j}$.

4.5 Model discussion

In this section I discuss the simulation results, proposed model, and merits and challenges to its application to my present data:

		$\tilde{\Delta} = \Delta$		$\tilde{\Delta} = 0.5\Delta$		$\tilde{\Delta} = 2\Delta$	
	True	Est	SE	Est	SE	Est	SE
β_1	1	0.97	0.02	0.97	0.02	0.98	0.02
β_2	-1	-0.98	0.02	-0.98	0.02	-0.98	0.02
β_3	0.5	0.51	0.01	0.51	0.01	0.51	0.01
α_1	0.35	0.36	0.00	0.36	0.00	0.35	0.00
α_2	0.35	0.36	0.00	0.36	0.00	0.36	0.00
α_3	0.35	0.35	0.00	0.35	0.00	0.35	0.00
α_0	0.8	0.79	0.00	0.79	0.00	0.79	0.00
$\bar{q}_1$	-1	-0.97	0.03	-0.97	0.03	-0.97	0.03
$\bar{q}_2$	0	-0.07	0.00	-0.06	0.00	-0.07	0.00
$\bar{q}_3$	1	0.98	0.04	0.98	0.04	0.98	0.04
q_1^*	0	-0.04	0.00	-0.05	0.00	-0.06	0.00
q_2^*	0	-0.04	0.00	-0.04	0.00	-0.05	0.00
q_3^*	0	-0.05	0.00	-0.04	0.00	-0.05	0.00
σ^2_{u1}	4.5	4.15	0.21	5.77	0.32	3.44	0.17
σ^2_{u2}	4.5	4.33	0.22	6.22	0.35	3.5	0.17
σ^2_{u3}	4.5	4.1	0.22	6.31	0.37	3.07	0.16

Table 4.4: Simulation results mis-specifying discretization parameter Δ using $\tilde{\Delta}$.

Thus far, I have shown, via simulation, conditions under which a static multiple discrete-continuous model with Bayesian learning can be consistently estimated. A brief discussion of the empirical feasibility and implications of these conditions follows:

1. *Simple, uncorrelated learning where a continuous "experience" is discretized into discrete signals.* In the Appendix I discuss alternate learning specifications that can incorporate correlated learning. This would be an important extension to accommodate the fact that players' behaviors across levels are clearly related and would directly allow for patterns such as increasing play of higher levels after moderate experience with lower levels. However, theoretic properties of correlated learning in MDC are currently unknown. Unfortunately, initial results from simulation experiments suggest they are incredibly difficult to pin down empirically. That said, all the patterns described in Section 3 can be fully accommodated in an uncorrelated learning framework. Even if players only learn about the levels they've played, they are still solving

a budget problem. For example, if players have relatively optimistic priors over the non-competitive levels that would produce the "upward learning" observed in the data. What uncorrelated learning does rule out, however, is the skill accumulation explanation: it makes little sense to consider learning as a proxy for skill improvement when it is non-transferable across levels.

2. *Risk aversion is a priori known.* In the case that risk aversion is mis-specified, I show that prior mean estimates are biased. In the special case that players have the same prior uncertainty over all goods, the order of prior mean estimates is preserved. To fully accommodate patterns observed in the data, it is too restrictive of an assumption. However, in my application I am solely interested in recovering the prior-to-signal ratio and by setting prior variances the same can recover prior preference ordinality.

3. *Prior uncertainty is a priori known.* Mis-specification of prior uncertainty in the presence of mis-specified risk aversion both biases signal variance and compounds the bias in prior preferences. As such, it will be a crucial empirical task to properly tune the prior uncertainty. As discussed previously, one potential starting point is setting prior variances equal to, at the very least, recover the correct ordering on prior preferences. A more data-oriented approach may be to tune prior preferences using historical play data and covariates such as adoption time.

4. *Discretization parameter is a priori known.* Note that mis-specification of discretization parameter biases only signal variances. Here it is useful to consider the underlying learning process. In general, individual sessions tend to last 40-60 minutes and so discretization by hours may be intuitively appealing. In the data, I also observe total play duration by day and contiguous play sessions (e.g. from turn-on to turn-off of the console). Both of these can also serve as proxies for signals. Finally, in the case that learning is truly continuous, setting the discretization parameter $\Delta \to 0$ may be reasonable.

CHAPTER 5

ESTIMATION

5.1 Estimation preliminaries

In my empirical application I use nested extreme value (NEV) errors alongside $G = 2$ latent segments. I also collapse the inside goods to two, $x_{\text{Low}} = x_{\text{Solo}} + x_{\text{Friendly}}$ and $x_{\text{High}} = x_{\text{Competitive}} + x_{\text{Ranked}}$, so that players directly budget their game-time between non-competitive (Low) and competitive (High) modes. I use the following parameterizations: satiation is parameterized as $1 - \exp(x)$ to fall in $(-\infty, 1)$, signal-prior ratio is parameterized as $\exp(x)$, segment probabilities are multinomial logistic: $\frac{\exp(x)}{1+\sum_{g<G}\exp(g)}$, and all inside goods are assumed to belong to one nest with nesting parameter parameterized $\theta = \frac{1}{1+\exp(-x)}$. Recall this implies errors have Kendall rank-correlation $\tau = 1 - \theta$.

I present results with the set of tuning parameters as described below. I let budget $E = 40$: in this case, all users with a weekly play over 40 hours must be removed from the sample. This represents a minuscule, non-representative fraction of the users ($< 1\%$) that have abnormally high play patterns[1]. I note here that the American Time Use Survey for the relevant year gives average leisure as around 36 hours/week. I let risk aversion $r = 1$. This term interacts with belief variance as an uncertainty penalty. It is not jointly identified with prior variance. I let discretization $\Delta = 1$. In sample, the average session length is slightly less than one hour, so I assume users learn discretely from each session (on average). Finally, I use $R = 50$ draws due to the size of my data. As noted above, this should not affect consistency except possibly through signal-prior ratio.

Covariates as selected as follows. In baseline utility (β) I include a linear time trend to capture declining novelty and console ownership demographics, which are taken to be exoge-

1. I consider a separate model with budget parameterized as $E = 40 + \exp(x)$ and find the implied $\hat{E} \approx 40$.

nous. I estimate inside satiation (α_j) as a constant and outside satiation (α_0) with a linear time trend–this directly captures the competing effect of other games as noted in Tyack et al. (2016) and allows an alternate pathway for loss of enjoyment to enter into the model. Empirically I find that assigning novelty decline to only one of β or α_0 was insufficient to explain the particular shape of the extensive margin evolution. In the Appendix I provide a closed-form expression for play propensity of the single discrete-continuous model and show that the two enter the play decision in different places. For the prior means (q_0) I incorporate historical information in the form of a suite of dummies, including whether a player is returning, adopted previous game early, played at a high level, etc. Under the assumption that unobserved heterogeneity is adequately captured by segmentation and within a segment users only differ with respect to their histories and experience[2], true means (q^*) are constant. Signal-prior ratio (σ_ν^2/σ_0^2) includes a dummy for whether a player had above-median experience in the associated level in game $G-1$: this reflects the inverse relationship between experience and uncertainty in the Bayesian framework. Finally, all parameters except the error nesting parameter θ are allowed to vary between segments.

Estimation is performed in Matlab using the Knitro solver and bootstrapped standard errors are provided. All estimates are presented in unrestricted parameter space. Below I present estimation results and show the proposed model can adequately recover data patterns. I present estimates in Table 5.2 and sample demographics for each cluster in Table 5.3.

2. In other words, conditional on being in the same segment, at time t user i differs from k only in their current beliefs $\tilde{Q}(\mathcal{H}_{it}), \tilde{Q}(\mathcal{H}_{kt})$ which then differs only because of their differing histories $\mathcal{H}_{\cdot t}$. In this formulation history includes unobserved signals from playing so is still stochastic from the researcher's point of view.

5.2 Model I: pure behavioral

We can roughly characterize the latent segments into 1: high usage, majority returning players ("hardcore"), and 2: low usage, majority new players ("casual"). The implied split is around 55-45. Using K-means I previously showed that the sample can roughly be described by a similar behavioral split[3]. Note that history-based parameters tend to have greater precision in Segment 1, because the effective sample size (returning users) is greater there. Log-MU is dominated by a declining novelty effect, more strongly for Segment 1, and as discussed previously, there is some evidence for a "PlayStation" effect: positive coefficient, with a positive interaction with level for high types.

Across the board, prior means are most strongly positively associated with high play amounts in the previous game and negatively associated with weekend adoption of the previous game. Above-average high shares in the previous games is positively associated with priors in the high level (and vice versa). In general, priors seem to be most shaped by play amounts and only moderately by play type. Relative to the intercept, heterogeneity in priors is limited but still significant. Further, because play is jointly determined by satiation, priors, and novelty, it is difficult to directly interpret raw intercepts. Instead I consider ex ante pessimism (prior bias) about a level $j : \Delta_j = q_j^* - q_{0j}$: In Table 5.1 above I compute *average*

Seg	Lvl	Prior	True	Pessimism
1	Low	-10.64	-6.20	-4.43
1	High	-11.94	-7.48	-4.46
2	Low	-0.84	6.86	-7.70
2	High	-2.49	3.67	-6.16

Table 5.1: Model I: Average pessimism by segment and level.

priors over each segment (using posterior segment probabilities as assignment). Note that Segment 2 is much more pessimistic, on average, than Segment 1, and slightly less pessimistic

3. In fact, the latent segment demographics I compute here are very similar to those from the K-means cluster demographics with $K = 2$ provided in the Appendix.

about High levels while Segment 1 is equally pessimistic about each level. All differences are statistically significant. While all players are ex ante pessimistic, note that true means are higher for the Low levels (vs High levels), potentially indicating competitive aversion.

Baseline signal-to-prior variance ratio is around 13-35% higher for Segment 1, which suggests that newer players learn relatively quicker. In the Gaussian-conjugacy Bayesian learning paradigm, recall that the lion's share of learning occurs in the initial few signals. Thus, the coefficient estimate is consistent with the hypothesis that these players have greater ex ante uncertainty[4]. Note that playing High levels is significant more informative (3.5x for Segment 1 and 4.2x for Segment 2). One hypothesis for this result is that competitive levels require greater attention and/or engagement from the player, leading to greater information acquisition. Finally, the nesting parameter implies a roughly 0.26 correlation (Kendall's tau) for the inside goods. The inclusion of this term is thus necessary and I find evidence for a NEV vs MDC model (i.e. I reject the assumption that MU w.r.t. to any good is independent of consumption levels of all other goods).

In Table 5.3 I present average pre-adoption (previous game) statistics for each latent segment, where users are assigned to a segment based on their posterior segment probability:

$$i \in g^* \Leftrightarrow \pi_{g^*} L_i(\theta_{g^*}) = \max_g \pi_g L_i(\theta_g)$$

For a more apples-to-apples comparison, I compare only returning players for all statistics except the final block of statistics (experience type). In general, returning players in Segment 1 adopt earlier, play significantly more and at a higher level. Further note that these players are likelier to adopt during a weekend. The significance of this is twofold: (1) the

4. Note, however, that as I only estimate signal-to-prior variance ratio and **not** prior variance, I cannot conclusively rule out the alternate story: newer players both have lower uncertainty and learn faster. However, this narrative seems to contrast basic intuition.

game traditionally releases on a weekday so weekend adoption can signal not adopting upon release[5], and (2) intuitively, weekday adoption may imply either a more uniform budget distribution (w.r.t to days of week) or simply one shifted out. Finally, as noted in the previous section, the unconditional segment probabilities are 0.55 and 0.45, respectively, and Segment 1 consists of the majority of High Exp players while Segment 2 comprises mostly New players (with Low Exp players roughly splitting between the two).

Table 5.2: Model I: Estimated parameters.

Group	Par	Seg 1	Seg 2
log-MU, low levels	Weeks	-8.474	-17.224
		(0.326)	(0.735)
	Owns PS	0.059	0.127
		(0.014)	(0.064)
	Owns older cons.	-0.008	0.028
		(0.023)	(0.044)
log-MU, high levels	Weeks	-8.018	-14.582
		(0.343)	(0.607)
	Owns PS	0.099	0.095
		(0.017)	(0.125)
	Owns older cons.	0.027	0.104
		(0.033)	(0.076)
Prior mean, low levels	Constant	-11.096	-1.045
		(0.345)	(0.868)
	Returning	0.251	0.363
		(0.064)	(0.06)
	Adopted 1st 2 weeks, $G-1$	-0.087	0.162
		(0.032)	(0.056)
	Adopted during weekend, $G-1$	-0.143	-0.15
		(0.029)	(0.075)
	Above med. max lvl played, $G-1$	-0.246	-0.201
		(0.085)	(0.114)
	Above med. high shares, $G-1$	-0.039	-0.057
		(0.059)	(0.085)
	Above med. no. wks. played, $G-1$	0.398	0.335
		(0.081)	(0.112)
	Above med. last wk. played, $G-1$	0.251	0.131

5. Among returning players, 40% of those in Segment 1 adopted in the initial weekdays compared to 33% in Segment 2.

Table 5.2: Model I: Estimated parameters, continued.

		(0.022)	(0.117)
	High exp	0.501	0.466
		(0.075)	(0.093)
Prior mean, high levels	Constant	-12.417	-2.653
		(0.304)	(0.911)
	Returning	-0.308	-0.422
		(0.1)	(0.072)
	Adopted 1st 2 weeks, $G-1$	-0.131	0.032
		(0.032)	(0.101)
	Adopted during weekend, $G-1$	-0.135	-0.088
		(0.055)	(0.123)
	Above med. max lvl played, $G-1$	0.309	0.398
		(0.098)	(0.087)
	Above med. high shares, $G-1$	0.801	1.04
		(0.08)	(0.108)
	Above med. no. wks. played, $G-1$	0.218	0.233
		(0.118)	(0.063)
	Above med. last wk. played, $G-1$	0.248	0.136
		(0.033)	(0.164)
	High exp	0.562	0.616
		(0.107)	(0.114)
True means	q_1^*	-6.204	6.86
		(0.386)	(0.27)
	q_2^*	-7.475	3.67
		(0.353)	(0.309)
Signal-prior ratio, low levels	Constant	3.975	3.853
		(0.04)	(0.058)
	Above med. low lvls, $G-1$	0.081	-0.094
		(0.035)	(0.063)
Signal-prior ratio, high levels	Constant	2.709	2.411
		(0.029)	(0.107)
	Above med. high lvls, $G-1$	0.306	0.163
		(0.034)	(0.126)
Satiation	α_1	0.123	0.313
		(0.015)	(0.011)
	α_2	0.106	0.213
		(0.011)	(0.023)
	α_{01}	1.191	-0.602
		(0.029)	(0.665)
	α_{0Weeks}	0.182	-5.124
		(0.022)	(0.844)

Table 5.2: Model I: Estimated parameters, continued.

Addl pars	π_1	0.228
		(0.044)
	θ_1	1.024
		(0.034)
-LL		1235175
N		26438

5.2.1 Recovery of data patterns

In this section I validate my model by simulating data at the estimated parameter values and comparing the evolution of three key moments: (1) extensive margin, (2) intensive margin, and (3) shares of high levels. In Table 5.4 I show that these three moments can be captured on the whole (unconditional, over both users and time periods). Next, in Fig. 5.1a I compare these data moments over time and in Fig. 5.1b I condition on recovered latent segment. It appears that my model can recover the extensive margin very well for both segments but has some trouble with the intensive margin, particularly for Segment 2. Recall this is the segment consisting mainly of low utilization players, reflected in the rapid drop in play probabilities. Past the initial third or so of the sampling period simulated moments become quite noisy since I predict so few players will play. On the other hand, the intensive margin is captured significantly better for the high utilization Segment 1. Finally, while I can capture the increasing competitiveness of play over time, it appears that my model overshoots this growth. This is partly driven by the few remaining users in Segment 1 with a streak of high draws who quickly converge to their true valuation[6]. However, due to the shape of the outside satiation in Segment 2 (Fig. 5.2b) even these high draw players are predicted to stop playing before the end of sample.

6. Note that the high level shares in Segment 2 are very noisy in the second of the sample as well. Although the timing is off, I do in fact manage to recover this noise.

Var	Seg. 1	Seg. 2	Signif
Adopt day, $G-1$	60.318	74.190	*
No. wks. played, $G-1$	18.500	9.643	*
No. wks. owned, $G-1$	43.717	41.759	*
Pct. wks. played, $G-1$	0.421	0.244	*
Last wk. played, $G-1$	40.291	31.228	*
Overall max lvl, $G-1$	2.347	2.072	*
First wk. play, $G-1$	6.510	5.350	*
First 4 wks. play, $G-1$	19.868	15.336	*
First wk. max lvl, $G-1$	1.101	1.104	
First 4 wks. max lvl, $G-1$	1.508	1.494	
Total solo play, $G-1$	54.598	27.072	*
Total friendly play, $G-1$	7.839	4.588	*
Total competitive play, $G-1$	17.725	7.535	*
Total ranked play, $G-1$	13.173	6.841	*
Total play, $G-1$	93.335	46.036	*
Overall high shares, $G-1$	0.245	0.224	*
Overall HHI, $G-1$	0.693	0.730	*
Above med. high lvls, $G-1$	0.444	0.293	*
Above med. low lvls, $G-1$	0.710	0.492	*
Above med. ranked play, $G-1$	0.234	0.168	*
Above med. competitive play, $G-1$	0.333	0.187	*
Above med. friendly play, $G-1$	0.413	0.275	*
Above med. solo play, $G-1$	0.556	0.337	*
Above med. last wk. played, $G-1$	0.573	0.326	*
Above med. HHI, $G-1$	0.890	0.800	*
Above med. no. wks. played, $G-1$	0.590	0.286	*
Above med. high shares, $G-1$	0.517	0.450	*
Above med. max lvl played, $G-1$	0.449	0.332	*
Adopted during weekend, $G-1$	0.208	0.255	*
Adopted 1st 2 weeks, $G-1$	0.535	0.467	*
High exp	0.372	0.114	*
Low exp	0.274	0.233	*
Returning	0.646	0.347	*
New	0.354	0.653	*
N	13790	12648	

Table 5.3: Model I: Sample statistics by predicted segment identity.

In Fig. 5.2a I present the evolution of simulated beliefs for each segment. Note that players in Segment 1 resolve a greater proportion of uncertainty (up to around 45% by end

Var	Data	Sim
Pr(play)	0.31	0.35
E(dur.|play)	4.30	4.21
Pr(high lvl)	0.28	0.30

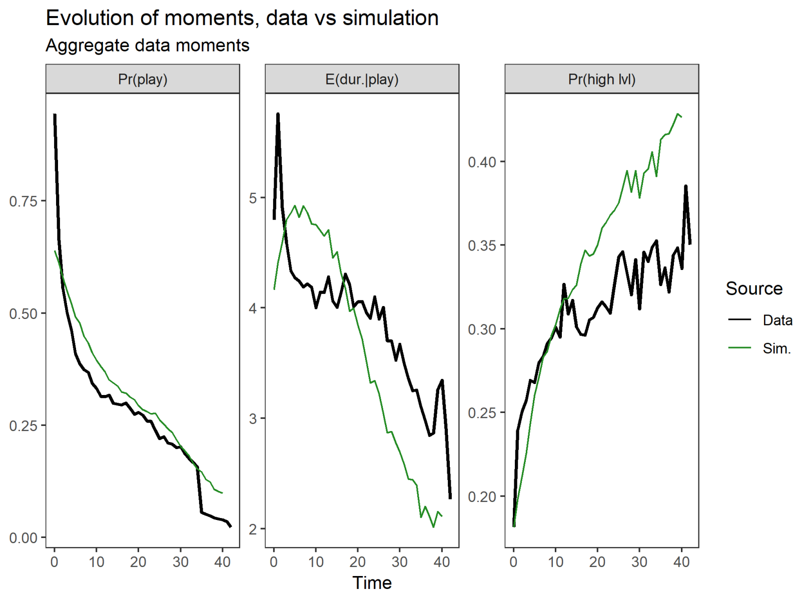
Table 5.4: Model I: Aggregate moments, data vs simulation.

of sample compared to 20% in Segment 2). Uncertainty about the High levels is always comparatively greater than the Low levels. In the end, players in Segment 1 are much closer to learning their true valuation[7]. Note that inside satiation is similar for both goods and segments: slightly higher satiation rate than log-utility. On the other hand, the outside good satiation is significantly lower and converging towards 1 for Segment 2, implying the outside good approaches linear utility. This is one of the primary drivers for the declining play propensity observed in Fig. 5.1. On the other hand outside satiation initializes and remains rather high for Segment 1.

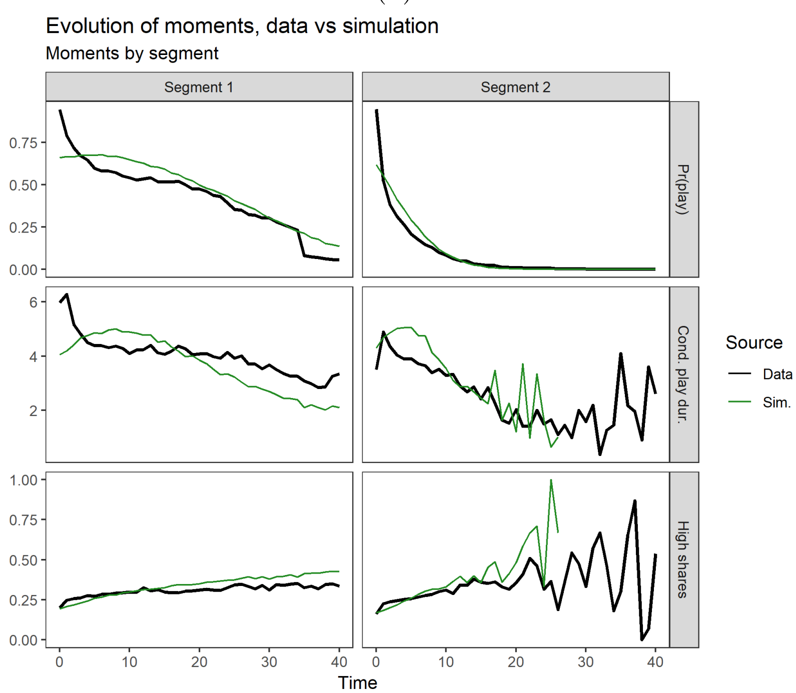
5.3 Model II: advertising effects

In Section 5 I showed that my model can adequately recover important data moments and illustrated observed player behavior as an function of the underlying competing forces of learning, satiation, and novelty loss. While I am able to explain player behavior it is difficult to directly translate the estimated parameters into concrete actions the firm can undertake to improve engagement, at least without a large degree of abstraction. To explore this area of my research question, I turn to a source of non-behavioral variation: advertising. In this section I re-estimate my model with advertising effects. To begin I introduce the advertising data, discuss its significance in the context of my data/model, and present estimation results.

7. Final pessimism at end of sample is around 2 for Segment 1 and 4 to 5 for Segment 2.



(a)

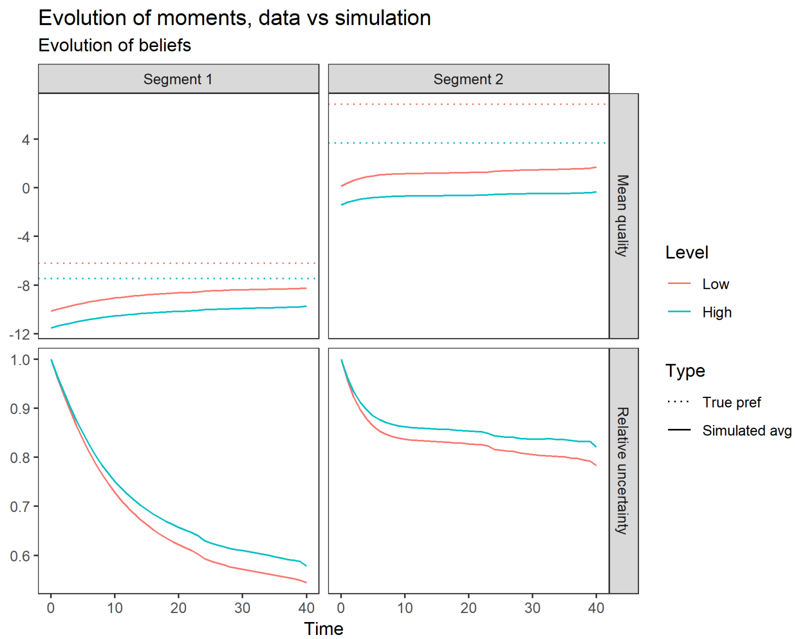


(b)

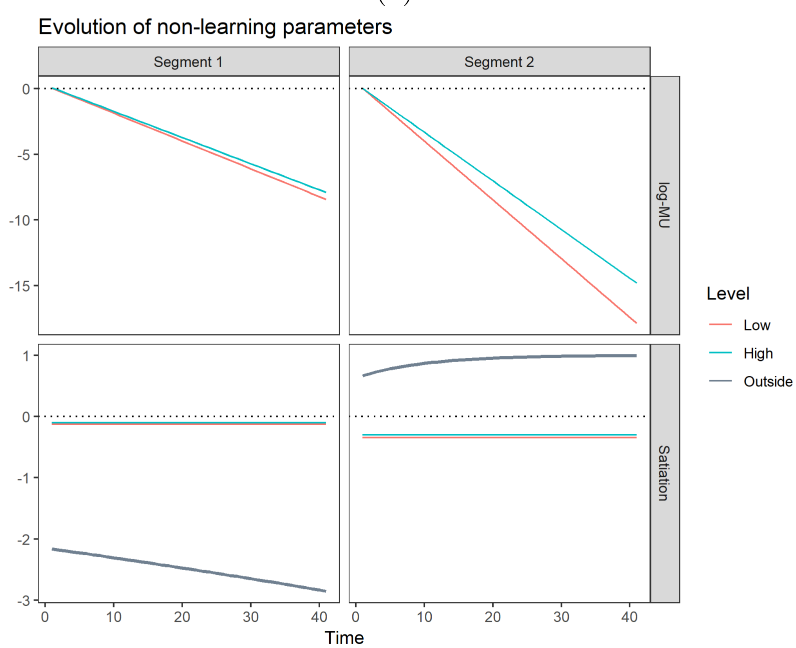
Figure 5.1: The model captures the extensive margin very well, as well as the noisiness in conditional play behavior for Segment 2. There are however some shortcomings.

5.3.1 Advertising data

I use data obtained from the Nielsen Ad Intel database. I observe advertising through national (television and Internet) and local (television and radio) sources over a period ranging



(a)



(b)

Figure 5.2: Segment 1 resolves a greater proportion of uncertainty. Note that Segment 2 both faces stronger novelty loss as well as outside satiation increase.

from around 100 days prior to around 120 days after game G release[8]. In Fig. 5.3a I present

8. Although the firm release a new version of the game annually, within my sample period there is no advertising observed for game $G-1$ or $G+1$, so I can safely ignore issues that arise from within-brand cross-product advertising (i.e. Shapiro et al., 2019).

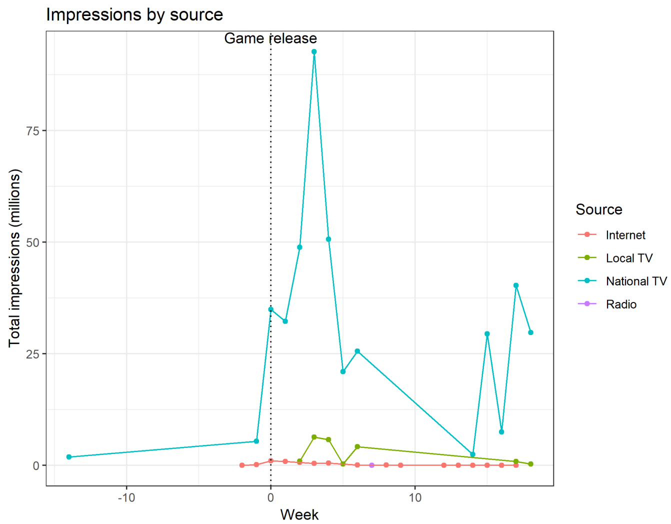
the distribution of advertisements (by net impressions) over time from each source. It is clear that exposure is dominated by national television advertising[9]. I drop radio advertising as it comprises a trivial proportion of total advertising expenditure. Next I consider local television advertising. It is clear that there is geographic targeting from the firm–in particular the firm appears to concentrate advertising in major metropolitan areas[10]. Unfortunately because I do not observe player locations in my data, I cannot measure the effect of this targeting. This leads to two options: (1) aggregate national and local television advertising together or (2) discard local television advertising. Both options abstract away from targeting and make strong assumptions about the firm's profit function. In particular, for the relevant outcome y (adoption or engagement), I must assume that at any given pair of current advertising expenditures $(A_{local,t}, A_{national,t})$ the next unit of advertising must have the same return for both sources: $\frac{\partial y_t/\partial A_{local,t}}{MC(A_{local,t})} = \frac{\partial y_t/\partial A_{national,t}}{MC(A_{national,t})}$. This is a very strong assumption but in the context of my model may have some validity. First, it may be reasonable to assume that adoptions enter directly into the firm's profit equation and engagement only indirectly (through increased likelihood of purchasing in-game goods and future adoption). Because I condition on adoption in my analysis, I can then treat advertising as indirect targeting. Second, I find that less than 5% of total ratings come from local television advertising. Further, the correlation of daily (log) gross ratings computed with and without local advertising is over 0.995 (see Appendix). Thus, I drop local television advertising going forward.

I aggregate the remaining national-level advertising and compute total impressions[11].

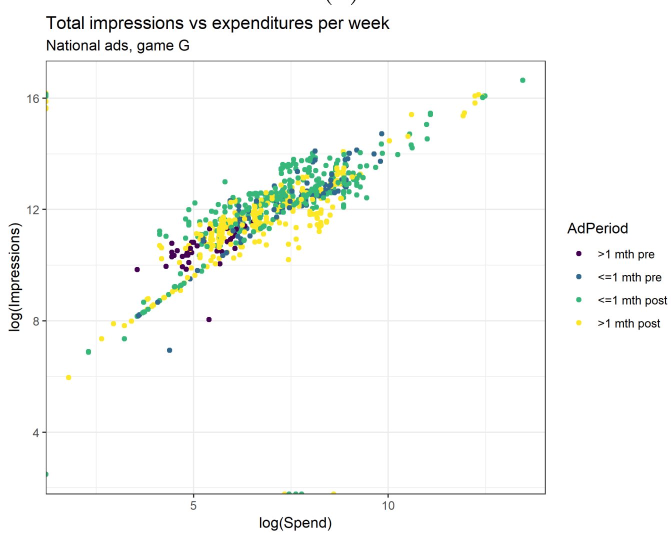
9. I surmise that expenditure follows a similar pattern, but unfortunately do not observe it for the second largest advertising source: local advertising. In Fig. 5.3b I show that for media types that contain advertising expenditures, it is reasonable to proxy impressions (or gross rating) for expenditures. One possible argument is that the firm is buying views at an ex-ante known price.

10. As I show in the Appendix, the four most-targeted DMAs are New York, Los Angeles, Philadelphia, and Chicago and they account for over a quarter of total local advertising exposure.

11. From Fig. 5.3a it is clear that national TV ads dominate Internet ads by exposure. I could have also dropped Internet ads and computed a pure ratings variable for national TV ads. Including Internet ads gives



(a)



(b)

Figure 5.3: Weekly impressions by media source, and impressions vs expenditures for national advertising. The heaviest expenditures are within a month of game release.

Then I define advertising stock as $S_t^A = \kappa A_t + (1-\kappa)S_{t-1}^A$, with $\kappa = \{1, 0.90, 0.75, 0.50\}$. Recall that $\kappa = 1$ corresponds to zero carryover. Due to the temporal concentration of advertising (immediately after release and roughly three months later) the specifications all give a similar shape to the stock variable. In the Appendix I present the evolution of advertising stock over the various levels of depreciation κ.

---

me slightly more variation in the advertising variable but means I can only compute total impressions.

5.3.2 Advertising specification

I assume advertising stock enters into the consumer's utility through their prior mean. Note this is similar to the standard marketing operationalization of the persuasive effect (see e.g. Narayanan and Manchanda, 2009) but differs slightly in that the persuasive effect is specified within the prior mean instead of the deterministic utility component, and leads to a slightly different learning shape. This operationalization implies that advertising affects play behavior, but **not** true valuations. Moreover, I do not include post-adoption advertising stock or flow in the utility specification. This rules out two effects: information and reminder. The former is empirically ruled out because a sample of observed product advertisements reveals little product information. In other words, after adoption learning-by-doing trivializes the information effect. Although the latter may be economically significant, I am unable to identify it with my current model specification–specifically because I look at play aggregated over weeks. First, with cumulative play it is impossible to remove lookahead bias and recover causal estimates. Second, the advertising reminder effect is often very short-lived (e.g. He and Klein (2018) find that the reminder effect on lottery ticket sales is significant for up to 4 hours post-exposure). Since I face the additional constraint of not observing who observes each ad, I believe estimates derived from the inclusion of advertising flow would be essentially uninterpretable. Ultimately, this means the advertising effect I back out is solely the persuasive effect of advertising stock at the moment of purchase conditional on adoption.

5.3.3 Interpretation of advertising coefficient

In this section I discuss the interpretation for the advertising coefficient. There are several model and data limitations that prevent treatment of the coefficient as causal or one of the standard (assumed unbiased) treatment effects typically recovered. I elaborate on each in turn:

The first source of bias occurs when the adoption decision in unobserved (or here, unmodeled). For example, it might be reasonable to expect that users with higher beliefs are both more likely to adopt earlier and play more in any given week. Because adoption is a dynamic program (i.e. optimal stopping problem) associating the advertising coefficient purely with first-period prior beliefs creates an upward bias: in the "true" DGP, advertising raises net discounted value of all future play and not just the first-period. On the other hand, it is also possible that advertising causes the marginal consumer to adopt: she will shift the distribution of priors downward. Holding adoption fixed, the advertising coefficient will be biased downward. Finally, consider the case that advertising has no effect on consumer beliefs. In that case the advertising coefficient simply captures the degree to which early adoptions (where players tend to play significantly more) correlates with a glut of advertising (which does indeed happen to be concentrated in the early parts of the game's release cycle).

The second source of bias is ubiquitous to the advertising literature: the timing and quantity of advertising cannot be considered truly exogenous. It should be expected that firms are targeting in some way with advertisements. In my dataset I observe advertising from a variety of media, including national TV, local TV, and Internet, along with estimated ratings. Alongside temporal targeting, I find evidence of geographic targeting in local TV. However, I find that local TV represents a significantly low proportion of total ratings and use only national level advertising. Therefore I abstract away from spatial endogeneity but cannot separately identify using current data, for example, whether the uptick in adoptions around the US holiday season is caused by the increase in advertising intensity. Finally, because I cannot match users in-sample with viewed advertisements I can at best associate each user-time observation with the national average at the time. Thus, I can at best estimate a homogeneous intent-to-treat.

Given the above issues, it is clear that the resulting advertising coefficient will be limited in what it can reasonably measure. Specifically, I cannot measure the impact of actual exposure to advertising, its effect on adoption, or effects beyond the first period (week). Instead I can only recover the effect of advertising holding the adopt decision and supply-side decision fixed. With this result I am limited in the set of counterfactuals I can consider. This rules out the following scenarios: (1) the firm targets users for advertisements or the behavioral effect of viewing an advertisement, (2) the firm considers a different advertising schedule, (3) the advertising effect on a marginal consumer (w.r.t. adoption **and** adopt time). Nonetheless, the advertising coefficient gives interpretation to a "unit" of prior mean. Even if it is only among a subset of consumers, I can now attach some firm-side cost to improving prior beliefs. In particular, if δ_A is the estimated advertising stock coefficient then an increase in advertising stock from S_t^A to $(1+p)S_t^A$ can be approximately formulated as a prior mean shift of $p\delta_A$ (in the log model).

5.3.4 Discussion of parameters

I present parameter estimates for Model II in Table 5.6. Note the log-likelihood is significantly improved over Model I (around 0.68% better). In general, the estimated parameters are robust between the two models. However, certain parameters (such as novelty decline) do change, representing information now captured by the advertising coefficient. Below I discuss a selection of the parameters and in particular interpret the advertising coefficient in light of the previous discussion.

First, note that the two latent segments can still be characterized as 1: high usage, majority returning players ("hardcores") and 2: low usage, majority new players ("casuals") with an implied split of 65-35. The primary difference here is that prior experience types is less useful in discriminating latent segments, indicating that the inclusion of advertising data is valuable in determining latent types. Next, for Segment 2 both novelty decline and estimated true means are significantly reduced in magnitude. Recall these are offsetting effects. As Segment 2 represents low types who do not in general play much (and thus learn poorly over their preferences), it is unsurprising that true match values are poorly identified. However, there are implications for the estimated pessimism values, presented in Table 5.1. As expected, pessimism is now estimated to be smaller across the board. However, the

Seg	Lvl	Prior	True	Prior bias	Signif
1	Low	-9.02	-6.51	-2.51	*
1	High	-11.87	-7.90	-3.97	*
2	Low	1.26	1.95	-0.69	
2	High	-3.41	1.06	-4.47	*

Table 5.5: Model II: Average prior bias by segment and level

results appear quite robust for Segment 1. The major difference is in Segment 1: for these users rational expectations can no longer be ruled out for the Low level (on average), while pessimism over the High level is very strong. This result is consistent with data patterns described in Section 3.

Finally, note that all advertising coefficients are significant and positive, and that coefficients for the Low levels are much larger than over High levels. From the previous discussion, the rigorous interpretation is that holding fixed the adopt decision, a unit of advertising exposure is positively associated with prior means, particularly the Low level. As periods of high advertising are associated with a large amount of adoptions, one argument is that marginal consumers select into adopting after viewing advertisements. These are the consumers with high ex ante uncertainty and low ex ante value on competition in the game, and as such the ad coefficient simply captures this selection. A second argument is that the advertisement contains a prestige effect on the Low level, and those who adopt shortly after viewing the advertisement retain this effect through their beliefs. Within the constraint of my data I cannot differentiate the two effects, and can only recover the conditional causal effect (second argument) assuming both advertisement targeting and the adoption schedule is unchanged.

Table 5.6: Model II: Estimation results.

log-MU, low levels	Weeks	-6.124	-4.856
		(1.028)	(0.562)
	Owns PS	0.055	0.122
		(0.02)	(0.023)
	Owns older cons.	-0.009	-0.067
		(0.026)	(0.03)
log-MU, high levels	Weeks	-5.493	-4.477
		(1.053)	(0.518)
	Owns PS	0.113	0.127
		(0.013)	(0.024)
	Owns older cons.	0.014	-0.01
		(0.042)	(0.032)
Prior mean, low levels	Constant	-11.833	-2.606
		(0.425)	(0.272)
	Returning	0.296	0.488
		(0.05)	(0.04)
	Adopted 1st 2 weeks, $G-1$	-0.059	0.063
		(0.023)	(0.025)
	Adopted during weekend, $G-1$	-0.19	-0.217
		(0.031)	(0.041)
	Above med. max lvl played, $G-1$	-0.379	-0.214
		(0.034)	(0.06)
	Above med. high shares, $G-1$	0.034	-0.133
		(0.023)	(0.046)
	Above med. no. wks. played, $G-1$	0.235	0.359
		(0.033)	(0.078)
	Above med. last wk. played, $G-1$	0.234	0.287
		(0.033)	(0.034)
	High exp	0.619	0.502
		(0.04)	(0.053)
	Ad stock ($\kappa = 1$)	2.288	3.45
		(0.415)	(1.292)
Prior mean, high levels	Constant	-13.183	-4.195
		(0.41)	(0.23)
	Returning	-0.256	-0.224
		(0.049)	(0.053)
	Adopted 1st 2 weeks, $G-1$	-0.042	-0.044
		(0.044)	(0.027)
	Adopted during weekend, $G-1$	-0.165	-0.156
		(0.043)	(0.061)
	Above med. max lvl played, $G-1$	0.255	0.445

Group	Par	Seg 1	Seg 2
		(0.023)	(0.07)
	Above med. high shares, $G-1$	0.861	0.941
		(0.029)	(0.071)
	Above med. no. wks. played, $G-1$	-0.008	0.253
		(0.033)	(0.071)
	Above med. last wk. played, $G-1$	0.213	0.299
		(0.035)	(0.036)
	High exp	0.666	0.563
		(0.037)	(0.058)
	Ad stock ($\kappa = 1$)	0.695	0.488
		(0.186)	(0.366)
True means	q_1^*	-6.665	1.963
		(0.448)	(0.301)
	q_2^*	-8.06	1.076
		(0.39)	(0.2)
Signal-prior ratio,	Constant	3.887	2.85
low levels		(0.036)	(0.053)
	Above med. low lvls, $G-1$	0.054	-0.187
		(0.021)	(0.02)
Signal-prior ratio,	Constant	2.671	1.639
high levels		(0.021)	(0.077)
	Above med. low lvls, $G-1$	0.251	0.022
		(0.024)	(0.03)
Satiation	α_{Low}	0.215	0.283
		(0.011)	(0.007)
	α_{High}	0.183	0.223
		(0.011)	(0.009)
	$\alpha_{0,Constant}$	1.267	0.176
		(0.032)	(0.043)
	$\alpha_{0,Weeks}$	-0.018	-14.776
		(0.083)	(2.276)
Addl pars	π_1	0.608	
		(0.021)	
	θ_1	1.293	
		(0.024)	
-LL		1226732	
N		26438	

Var	Seg. 1	Seg. 2	Signif
Adopt day, $G-1$	62.811	66.651	*
No. wks. played, $G-1$	20.220	11.678	*
No. wks. owned, $G-1$	43.364	42.824	*
Pct. wks. played, $G-1$	0.460	0.281	*
Last wk. played, $G-1$	41.177	34.044	*
Overall max lvl, $G-1$	2.355	2.174	*
First wk. play, $G-1$	6.422	5.880	*
First 4 wks. play, $G-1$	20.115	16.909	*
First wk. max lvl, $G-1$	1.109	1.097	
First 4 wks. max lvl, $G-1$	1.505	1.501	
Total solo play, $G-1$	61.069	32.454	*
Total friendly play, $G-1$	8.143	5.609	*
Total competitive play, $G-1$	20.937	8.843	*
Total ranked play, $G-1$	15.027	7.772	*
Total play, $G-1$	105.177	54.678	*
Overall high shares, $G-1$	0.248	0.230	*
Overall HHI, $G-1$	0.694	0.714	*
Above med. high lvls, $G-1$	0.456	0.342	*
Above med. low lvls, $G-1$	0.719	0.570	*
Above med. ranked play, $G-1$	0.238	0.191	*
Above med. competitive play, $G-1$	0.351	0.229	*
Above med. friendly play, $G-1$	0.421	0.323	*
Above med. solo play, $G-1$	0.569	0.412	*
Above med. last wk. played, $G-1$	0.602	0.399	*
Above med. HHI, $G-1$	0.892	0.834	*
Above med. no. wks. played, $G-1$	0.626	0.376	*
Above med. high shares, $G-1$	0.518	0.476	*
Above med. max lvl played, $G-1$	0.455	0.373	*
Adopted during weekend, $G-1$	0.215	0.230	*
Adopted 1st 2 weeks, $G-1$	0.521	0.505	
High exp	0.375	0.176	*
Low exp	0.256	0.254	
Returning	0.631	0.430	*
New	0.369	0.570	*
N	9614	16824	

Table 5.7: Model II: Sample statistics by predicted segment identity

5.3.5 Segment summary statistics

In this section I first present average pre-adoption (previous game, $G-1$) statistics for each latent segment (Table 5.7). Next, I show that the inclusion of advertising data improves qualitative aggregate data fit.

In Table 5.7 I compare only returning players for all statistics except the final group (experience type). In general, I find that returning players in Segment 1 adopt earlier, play significantly more and at a higher level. Further note that these players are likelier to adopt during a weekend. The significance of this is twofold: (1) the game traditionally releases on a weekday so weekend adoption can signal not adopting upon release[12], and (2) intuitively, weekday adoption may imply either a more uniform budget distribution (w.r.t to days of week) or simply one shifted out. Finally, as noted in the previous section, the unconditional segment probabilities are 0.35 and 0.65, respectively, and Segment 1 consists of the majority of High Exp players while Segment 2 comprises mostly New players (with Low Exp players roughly splitting between the two).

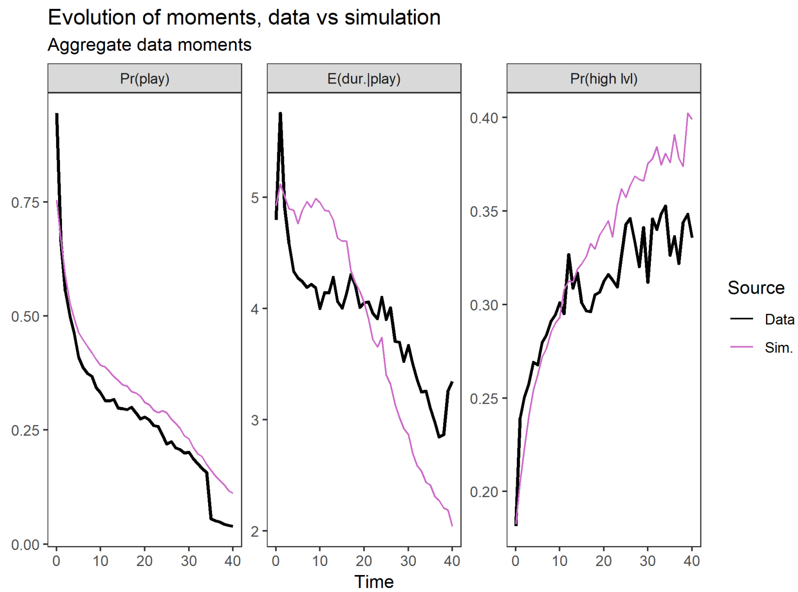
Var	Data	Sim
Pr(play)	0.31	0.36
E(dur.|play)	4.30	4.40
Pr(high lvl)	0.28	0.29

Table 5.8: Model II: Aggregate moments, data vs simulation.

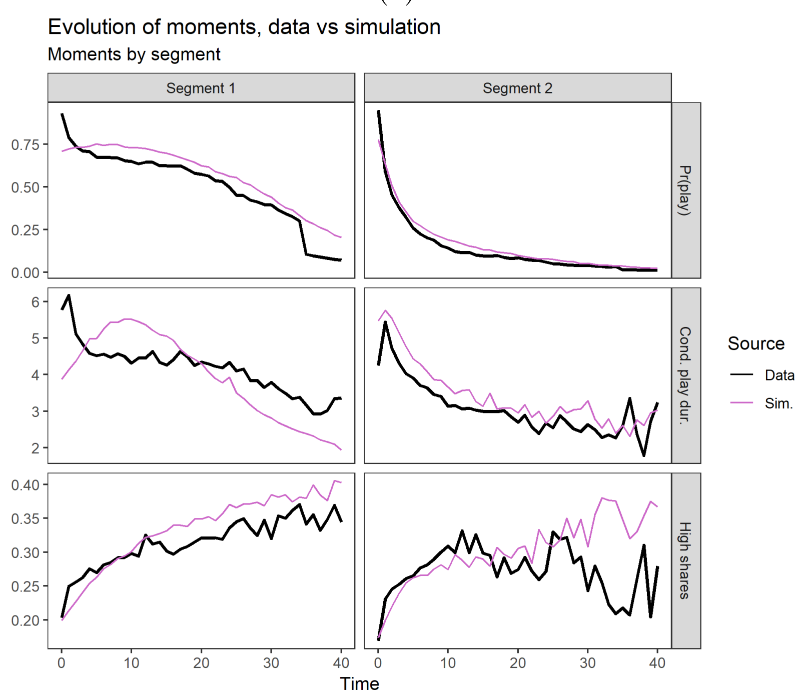
In Fig. 5.4 I once again visualize the evolution of the three key data moments (extensive margin, intensive margin, and high level shares), comparing simulated to raw data. I find that I can fit play patterns in Segment 2 much better, partly due to the reduced novelty decline parameter. However, the model still has some trouble properly recovering the intensive

12. Among returning players, 40% of those in Segment 1 adopted in the initial weekdays compared to 33% in Segment 2.

margin for Segment 1 (i.e. initial "hump" in play). Aggregate moments provided in Table 5.8 show that the model slightly overpredicts both the extensive and intensive margins overall. In Fig. 5.5 I visualize the evolution of the latent model parameters over time. Note that compared to Model I, I find that for Segment 2, novelty decline is less severe and initial outside satiation is similar to the inside goods. Finally, I find that under these results Segment 2 learns significantly better over the Low levels than previously implied.

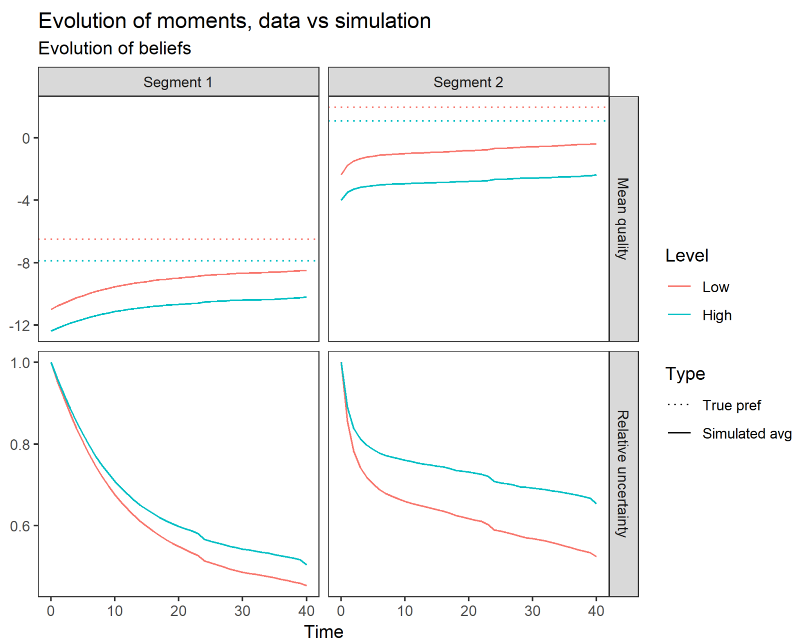


(a)

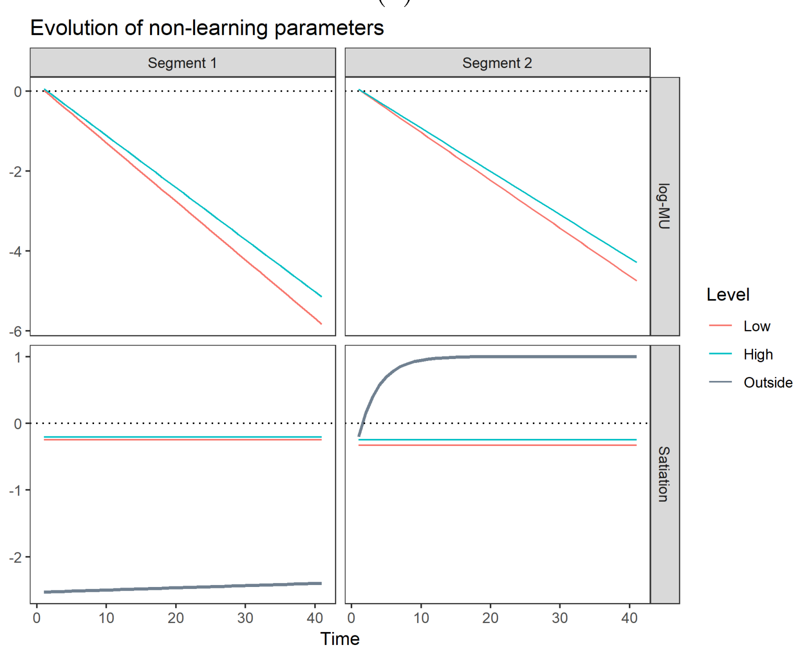


(b)

Figure 5.4: Actual vs simulated data, aggregated and by estimated segment. All major data patterns, except perhaps the tail end of High level shares for Segment 2, are adequately captured.



(a)



(b)

Figure 5.5: Implied evolution of parameters, learning and non-learning. Segment 1 learns "better" than Segment 2. A major driver of Segment 2's patterns in the shape of their outside satiation.

CHAPTER 6

COUNTERFACTUALS

Here I provide several counterfactuals based on parameters from Model II. In particular, I will consider the effect of increased advertising, console upgrading, and console brand-switching. I then conclude with a brief summary of the value to the firm of each policy.

6.1 Increased advertising

First I consider the effect of increasing national-level advertising by a percentage following the same advertising schedule. As discussed previously, I can only back out the effect of advertising conditional on fixing temporal and spatial targeting, and further assuming any changes in advertising do not affect the adopt decision, either in the binary yes-no or timing. Thus, I only consider the counterfactual where a firm holds its advertising distribution fixed but increases it globally by a certain percentage. Among other things, this maintains the advertising ratio between any two points in time.

To reiterate, I observe the sequence of daily national-level advertising $\{A_t\}_{t=-100}^{120}$ where the game is released at time $t = 0$. I assume all users are exposed to a homogeneous advertising stock $S_t^A = \kappa A_t + (1-\kappa)S_{t-1}^A$, which affects their prior mean belief at adoption. Post-adoption I assume there is no persuasive effect of advertising w.r.t. play and that the informative effect is strongly dominated by learning-by-doing. With the current specification of only immediate advertising effect ($\delta = 1$), advertising before release is constrained to have no behavioral effect[1].

In Fig. 6.1 I present counterfactuals with advertising increased from a range of 10% to

1. Note this **does not** rule out the possibility that pre-release advertising affects adoption. However, as I am modeling behavior conditional on adoption I have no way to capture it. That said, I believe it is far more plausible to take $\delta = 1$ in a conditional usage context than in an adoption (and possibly joint) one.

30% (policies Ad10, Ad20, Ad30). The advertising effect on total play across both segments is convexly increasing, ranging from 0.20% to 1.72% with increasing advertising. It ranges from 0.02% to 0.78% for Segment 1 (hardcores) and from 0.53% to 3.45% for Segment 2 (casuals). At low levels of increased advertising, this increase in play is primarily driven along the extensive margin while at high levels the intensive margin begins to be more significant. For Segment 1 the extensive margin is further driven by an increase in high level play at the expense of low level play, while for Segment 2 it is generally balanced (except Ad10, when it is driven by an increase in low play). From Fig. 6.1 it appears that the increased advertising slightly raises high level shares for Segment 1 and (noisily) lowers for Segment 2, with a positive aggregate effect. With the exception of Segment 2 resolving slightly more uncertainty, learning curves do not perceptibly improve.

I posit that the conditional advertising effect is stronger for Segment 2 because they are more likely to be low usage, newer players. Given that they do not have much prior experience with the game, advertising could serve as a substitute for experience (potentially informative effect[2]). As these players are significantly more pessimistic about their match values for the High level, the increase in play is manifested through the Low level. On the other hand, Segment 1 is more balanced in their pessimism between levels. This may partially explain why increasing advertising causes them to substitute away from the Low to High level. Note that increasing advertising has two effects: (1) increases the likelihood that a given player will have seen any ad by adoption, (2) for a user that has seen an ad, increases the expected number of ad exposures. I posit that low increases in advertising (e.g. AdLow) primarily increase the likelihood of any ad exposure (effect 1), which serve as a sort of reminder effect in the sense that it may increase the likelihood of playing the game entering into future consideration sets. As the firm further increases advertising ($\rightarrow$ AdHigh),

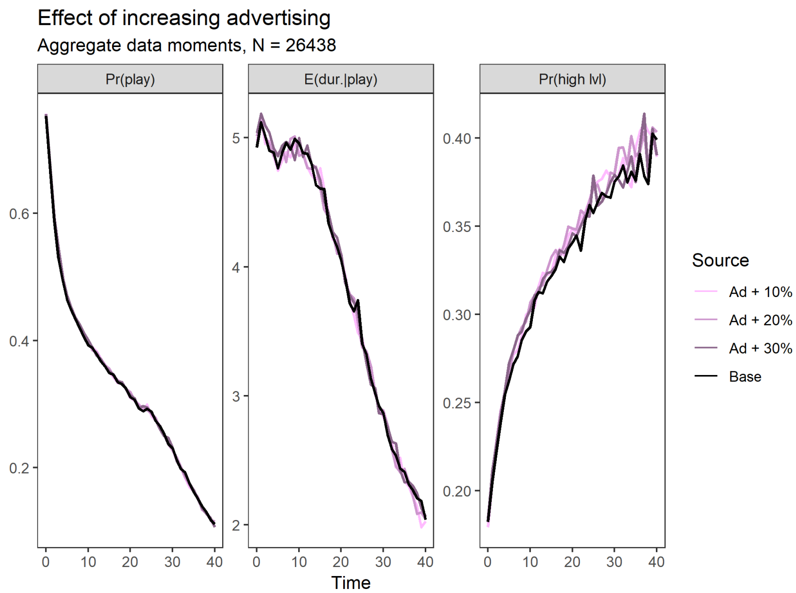
2. As I described previously, the observed advertisements for this game appear to contain little informative content. As such I believe the informative effect to be very weak for newer players and practically non-existent once these players have actually played the game.

effect 1 reaches some saturation level. This is supported by the fact that for Segment 1 the final increase in advertising to 30% does not increase play probability. On the other hand, increasing advertising can affect the intensive margin of play via effect 2: *ceteris paribus*, players who have seen more ads may associate it with greater prestige[3].

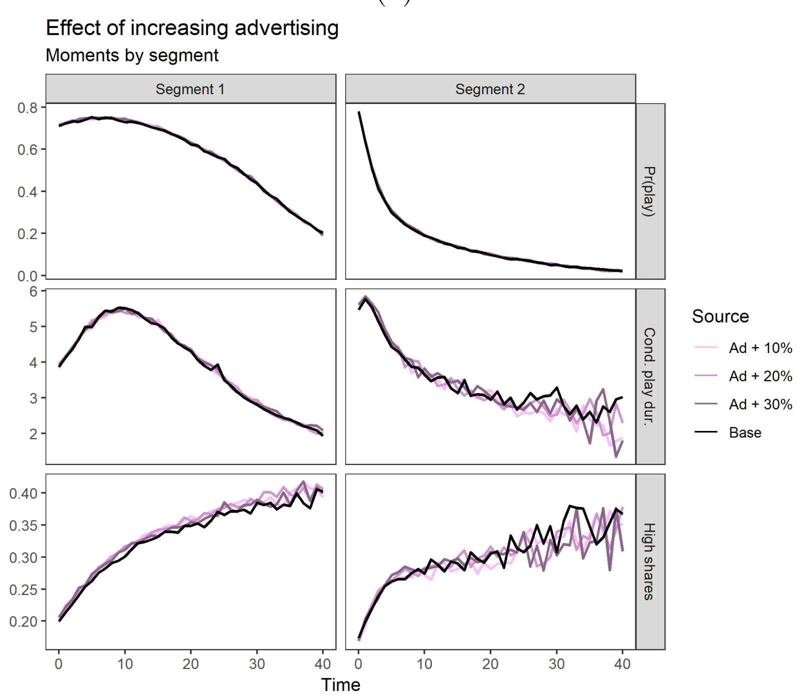
6.2 Console switching

In this section I consider the console effect. In Table 5.6 I showed the existence of a strong, positive PlayStation effect as well as weaker positive console generation effect. There are several possible sources of this PlayStation effect. One potential explanation is that due to intrinsic platform differences, more "hardcore" gamers self-select into PlayStation ownership. For example, Gilbert (2018) notes that while hardware and prices are similar between the PlayStation 4 and Xbox One, the PlayStation boasts a greater selection of console-exclusive titles. Steiner et al. (2016) catalog motivational differences between hardcore and casual gamers. They find that when considering platform adoption, hardcore gamers' valuation is dominated by expected software quality while casual and social gamers pay more mind to price, individual title availability, and socialization. A second explanation occurs at the game level. While games are quintessentially identical between platforms, there are e.g. graphical differences due to hardware differences. In fact a large review aggregator maintains separate ratings for the same game on each console. At the margin, there is therefore an argument for the potential causal effect of inducing a console switch. On the other hand, the effect of upgrading a console to the next generation appears to be more obvious. Selection can be framed through the lens of "innovators" versus "imitators" (e.g. the Bass adoption model). This effect can be bolstered by network effects, as players must often be using the same console version to play with friends. The canonical example of a platform-switching policy is bundling–this is a rather common marketing strategy for video games (and other digital

3. Up until a satiation point, which is commonly cited as around 3, cf. Deighton et al., 1994; Pedrick and Zufryden, 1991.



(a)

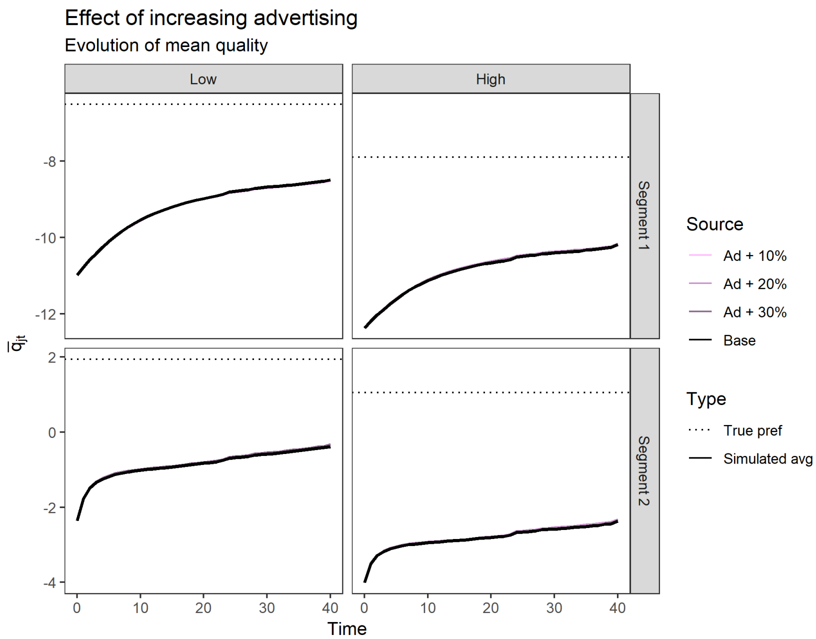


(b)

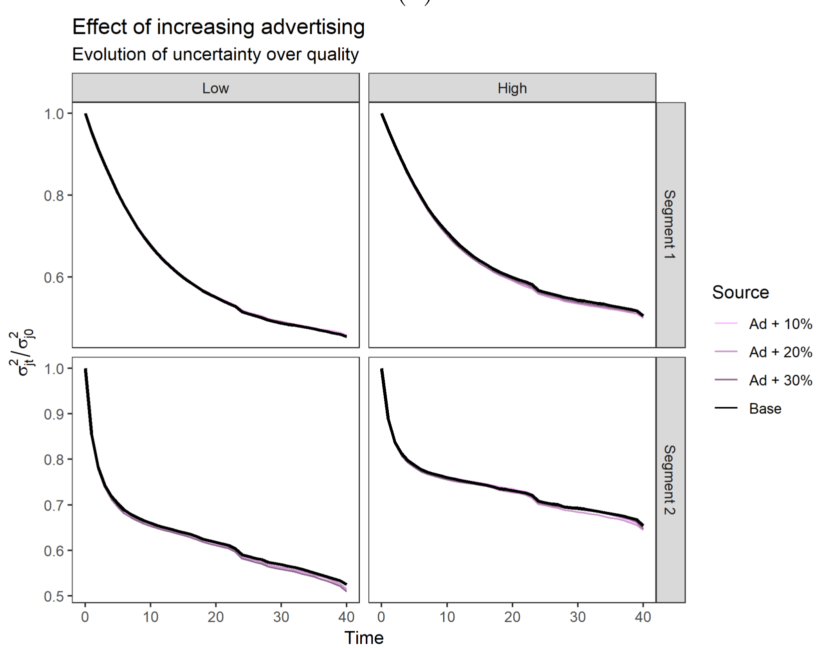
Figure 6.1: Increasing advertising from 10% through 30% increases net engagement by a range of 0.20% to 1.72%, holding all else constant.

goods), due to their low marginal costs[4]. In the context of direct upgrading, additional

4. For example, Derdenger and Vineet (2013) show that bundling for Nintendo games can increase sales by (1) causing infra-marginal consumers to adopt earlier and (2) capture consumers who have ex ante low valuation for the game. A key driver of their results is the power of bundles to segment consumers.



(a)



(b)

Figure 6.2: Evolution of beliefs, increased advertising. Segment 2 appears to resolve slightly more uncertainty over both levels.

policies such as progress carryover or backward-compatibility become relevant as well.

In my model I have included behavioral controls that I believe account for the selection in

To:	Xbox	PS		Prev.	Curr.
From: Xbox	0.506	0.016	Prev. gen.	0.091	0.088
PlayStation	0.008	0.470	Curr. gen.	0.001	0.820

Table 6.1: Console switching. For brand (left), N = 13804; for generations (right), N = 12732.

the first explanation, e.g. play and adoption behavior in the previous game. I thus take the view that the PlayStation and console generation coefficients in Table 5.6 represent players' (log) marginal utility over the console-specific attributes, such as network effects (from the installed base), graphical differences, and performance improvements (in the case of newer console). In Table 6.1 I present summary console statistics for returning players[5]. I observe that there is a marginally positive net switching to PlayStation between games (0.478 to 0.486, $p > 0.10$) and a significant net switching to current generation consoles (0.821 to 0.908, $p << 0.05$). The regression of (log) game G duration on the switching dummies gives coefficients of 0.15 ($p = 0.13$) and 0.50 ($p = 0.07$), respectively[6]. Finally, I find that returning users who do not switch consoles are most likely to belong in Segment 1 (44.6-46.2%) while Xbox to PlayStation switchers are the least frequent (37%, while PlayStation to Xbox is 40%). With these descriptives I can pose the first set of console-related policy questions: what is the console-switching effect on engagement for (1) previous-game Xbox users and (2) non-console upgrading users.

In Fig. 6.3 I present counterfactuals for returning Xbox users switching to PlayStation and in Fig. 6.4 the evolution of beliefs. I find that this switch increases total play by 8%, with Segment 2 significantly more affected. The play increase appears to be primarily driven by

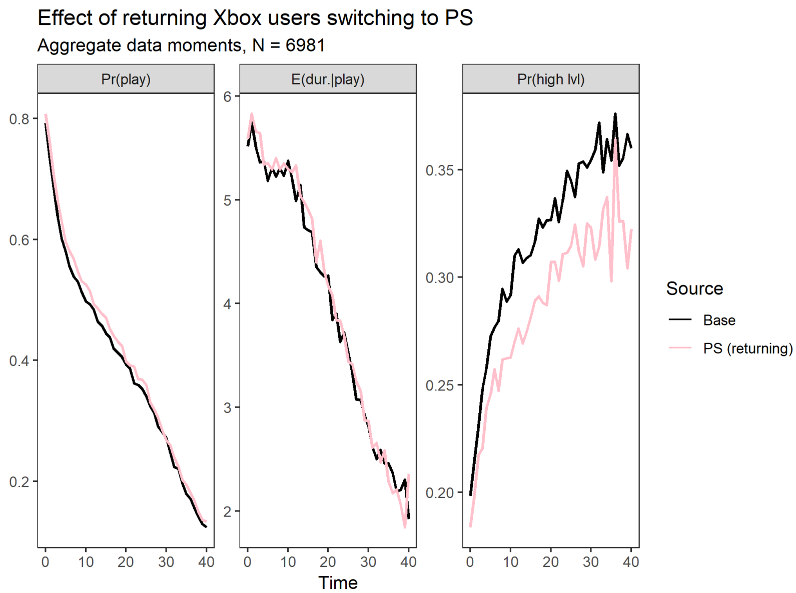
5. I recover console ownership from observed play sessions. Thus, I do not know whether a switching player still allocates significant playtime to other games on the previous platform. I further filter out users with observed sessions on multiple platforms ($< 1\%$ of the sample).

6. While the PlayStation switching coefficient is not significant, I believe it has economic significance. One issue with this regression is a lack of power: only around 200 players switched from Xbox to PlayStation while over 1000 switched from a previous to current generation console. Regression tables are provided in the Appendix.

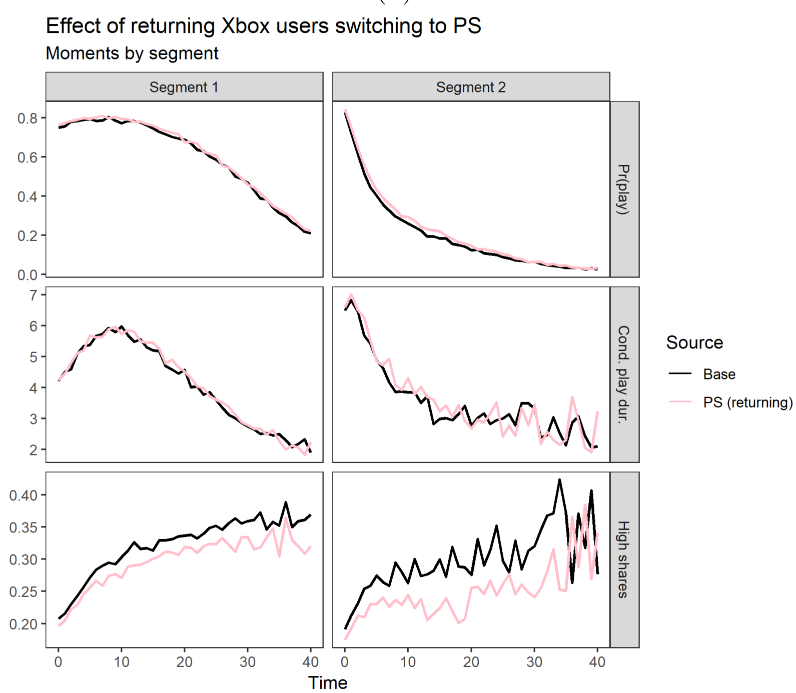
the extensive margin and through Low level play. One potential explanation is adjustment costs: players who purchase a new game concurrently with the new platform pay a mental cost in both game and platform familiarity. This is especially true for Segment 2 players, who decrease their High level shares significantly more than Segment 1 players. While the net increase of 8% appears rather significant, this result is tempered by the low conversion rate–recall that in-sample only around 3% of Xbox players switch console brands. Finally, note that with this counterfactual, players are predicted to learn better over the Low levels (especially Segment 2) and worse over the High ones.

Next in Fig. 6.5 I present results for console upgrade (holding fixed the brand). Note that the predictions are noisy because the sample of players who remained using a previous generation console is small ($N = 1160$). I find a much more modest effect to upgrading: Segment 1 players play 0.90% more while Segment 2 plays 5.6% more, with an average effect of 3.9%. The primary driver here appears to again be the extensive margin and High level play is reduced for both segments, potentially due to a switching cost. Note that similar to the brand switching, Segment 2 is significantly more sensitive to the policy than Segment 1. Additionally, these players are far likelier to *not* upgrade their console in the first place. This is consistent with the hypothesis that Segment 2 comprises the casual players described in Steiner et al. (2016). Compared to the PlayStation switching policy, the console upgrade policy appears to have a much more muted individual effect. However, it has significantly greater uptake: almost 50% of previous generation owners upgrade their console in sample. Unfortunately, as I do not explicitly model the console-switching decision I cannot evaluate which policy has a greater unconditional effect on engagement[7]. That said, it does appear the firm must trade off the value of converting an Xbox user to its relative infrequency, and vice versa for console upgrades.

7. In other words, rather than an average treatment effect I can only recover average treatment on untreated.



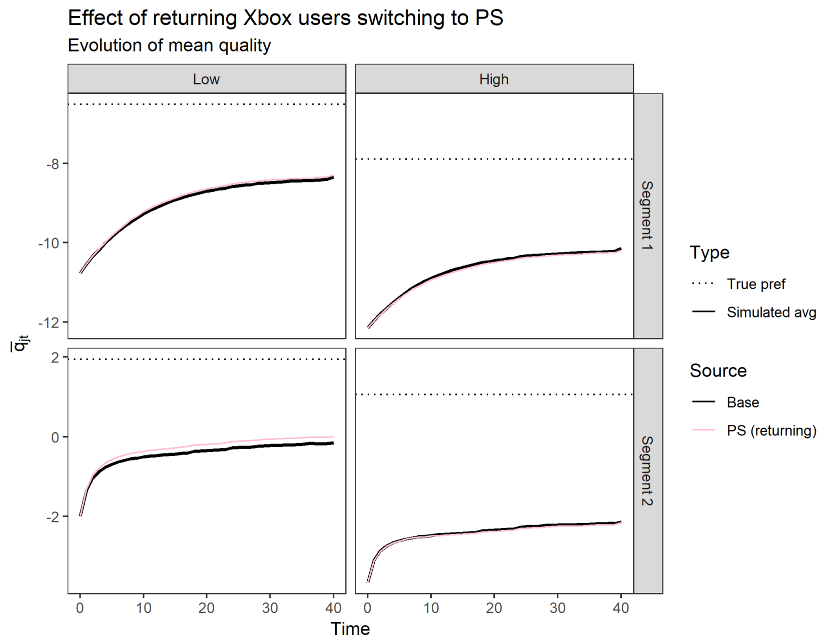
(a)



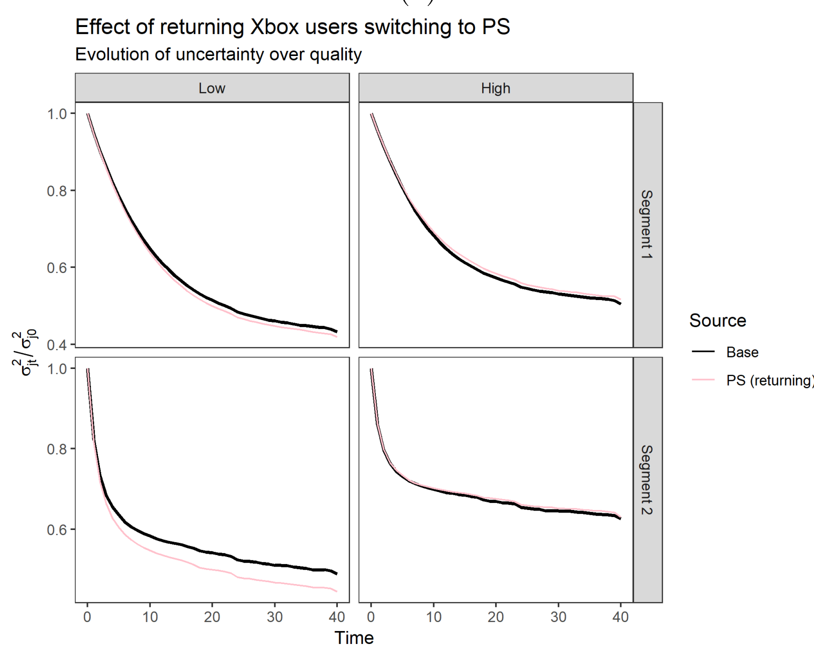
(b)

Figure 6.3: The average effect on play is 8%, and increases Low level play at the cost of High level play. However, in-sample the conversion rate is only 3%.

Finally, I consider the effect of console switching for New players. Without historical information, I have no measure for switching probabilities and can only describe the pol-



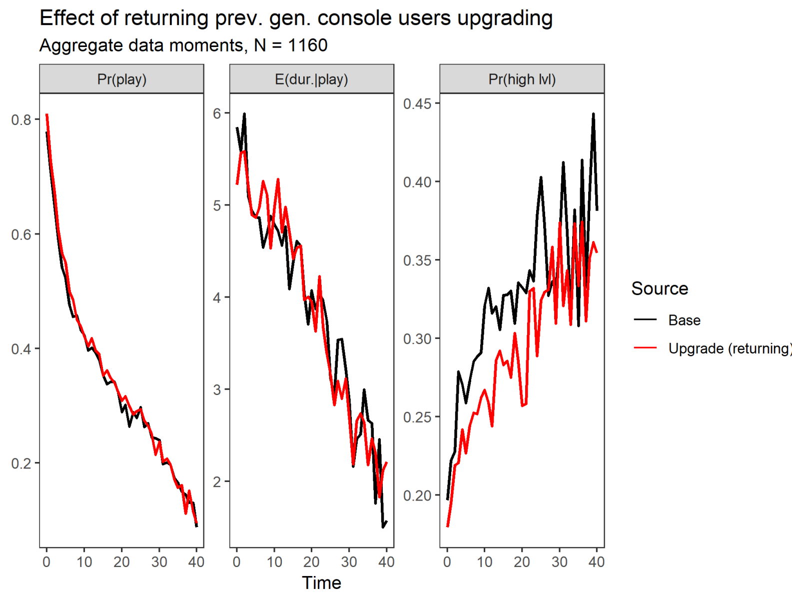
(a)



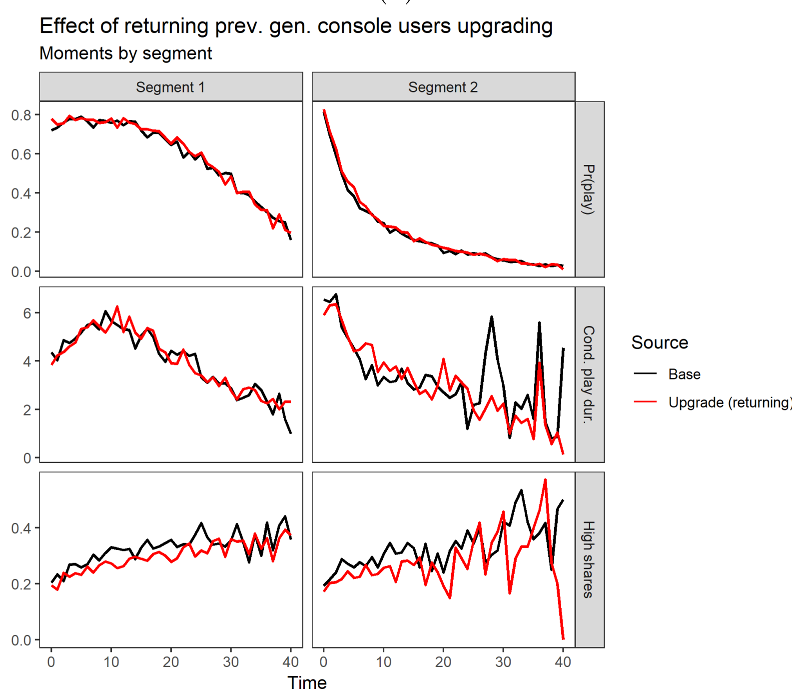
(b)

Figure 6.4: Players learn better over the Low levels (particularly Segment 2), and slightly worse over the High levels.

icy effect conditional on uptake. As the baseline, I observe that 48.6% of New players use PlayStation and 84.6% are using a current generation console, which is qualitatively con-



(a)

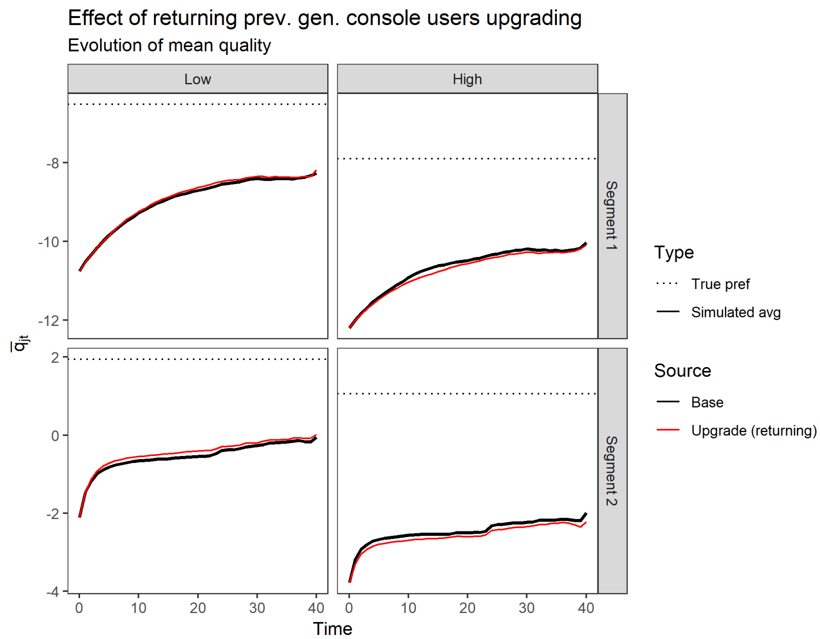


(b)

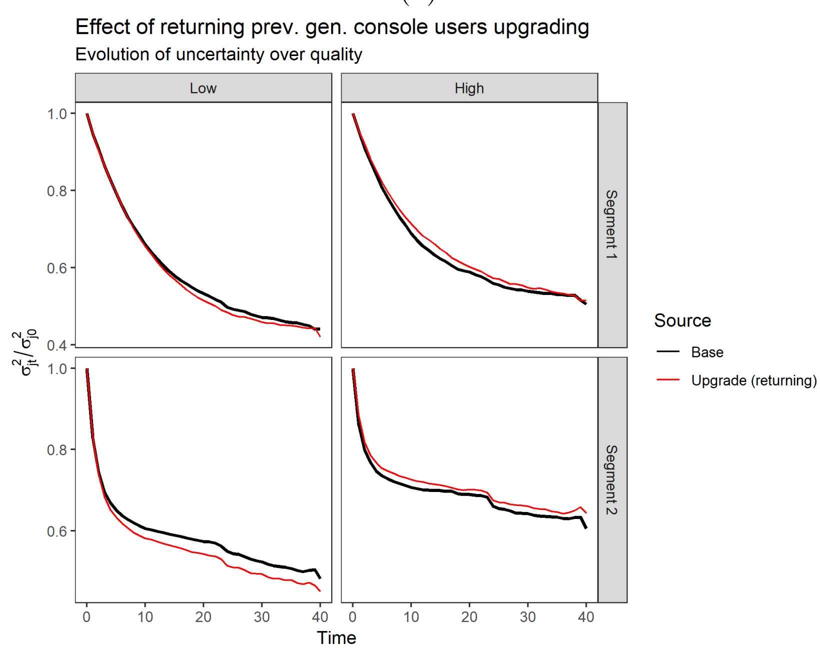
Figure 6.5: The average effect on play is 4%, primarily driven through the extensive margin. Almost 50% of users upgrade between games.

sistent with the reasoning that these players are more casual[8]. I present Figures in the

8. In other words, it is more appropriate to think of New users as qualitatively different than returning users than as a mixture of Low and High types. This is compatible with beliefs under a Bayesian learning framework.



(a)



(b)

Figure 6.6: Learning curves are similar for the console upgrading policy.

Appendix. As a summary, I find that the PlayStation policy for new players is qualitatively similar to that for returning players, with a net effect of 9% primarily driven through the extensive margin. I find that it causes Segment 1 players to actually reduce their intensive margin. I find the generation upgrade policy to have a modest effect for Segment 1 (4.2%)

but actually decrease total play for Segment 2 (-0.2%). This decrease is driven by a reduction in the intensive margin. Unfortunately without historical console information it is difficult to attribute this to any kind of skill or familiarity cost, and thus difficult to make policy recommendations for New players with regards to console switching.

6.3 Summary of counterfactuals

In Table 6.2 I present aggregate counterfactual moments. To briefly summarize, the conditional advertising elasticity is estimated to be around 0.02%. However, it is estimated to be an increasing function, as an increase in advertising of 30% leads to a roughly 1.72% increase in total play (elasticity of 0.06%). I find that at low levels of increased advertising the increase is primarily driven by the extensive margin, and at high levels of advertising the intensive margin growth begins to overtake the extensive margin growth. Segment 1 shifts play towards the High level while Segment 2 has a somewhat balanced response.

I then analyze the response among different consumer groups to policies promoting console switching. In particular, there is evidence from both the data and the literature that type of platform use can affect player utility through tangibles such as networks, graphics, or performance as well as intangibles such as prestige or familiarity. I find that a policy such as bundling that causes consumers to switch from Xbox to PlayStation may increase net engagement. Unfortunately, as I do not model platform switching I cannot assess the unconditional effectiveness of said policies. Using observed switching patterns in the data, I can however recover policy effects conditional on switching. Empirically, the baseline conversion is around 3%. On the other hand, I find that a policy promoting console upgrading has a much smaller unconditional effect but may be significantly easier to achieve greater conversion (in sample, there is an almost 50% rate). I find all these switching policies exact a "skill" cost as consumers significantly substitute away from the High levels afterward.

	Segment 1					Segment 2				
Policy	Play	Pr(play)	E(dur.|play)	Pr(high lvl)	N	Play	Pr(play)	E(dur.|play)	Pr(high lvl)	N
Ad + 10%	0.02	0.14	-0.12	2.86	9614	0.53	0.68	-0.15	-1.57	16824
Ad + 20%	0.53	0.36	0.17	2.83		0.87	0.68	0.19	0.78	
Ad + 30%	0.78	0.30	0.48	1.97		3.45	1.65	1.77	0.40	
PS (returning)	3.16	1.82	1.31	-7.27	3099	11.89	9.07	2.58	-15.29	3752
Upgrade (returning)	0.90	1.12	-0.22	-10.13	398	5.55	4.72	0.79	-14.95	753
PS (new)	1.95	3.36	-1.36	-3.58	1756	11.74	10.04	1.54	-11.39	5124
Upgrade (new)	4.21	3.10	1.08	-14.21	565	-0.24	0.96	-1.18	-10.15	1459

Table 6.2: Data moments relative to baseline (as percentage), various counterfactuals.

6.4 Firm relevance of engagement

Thus far I have evaluated the impact of several potential firm-side actions on player engagement. It is difficult to directly quantify the value of this increased engagement to the firm's bottom line but several qualitative arguments can be made, in the context of my application, for the value of engagement as the focal metric. First, increased engagement can create value through network and word-of-mouth effects. For players seeking a competitive experience, utility is a function of quality of game matches, which itself depends on the current distribution of other players. Additionally high engagement users are more likely to innovate in a way that benefits the firm (Bogers et al., 2010; Nielsen, 2006). In my application, users participate in on-line discussion boards and create a burgeoning media collection relating to the game (i.e. written guides, recorded video). Users also provide direct feedback on their satisfaction and critiques of the game through surveys. In the context of a franchise game with annual releases, each generation's users can be thought of as ideal "beta testers" for the next game. A case can be made, then, that increased play leads to better feedback (both direct and behavioral) for the next iteration of the game.

That said it is not unreasonable to consider that increased engagement may also lead to increased probability of future adoption or in-game purchasing–these are the outcomes that may enter directly into the firm's profit equation. In Figs. 6.7 and 6.8 I visualize the distribution of various engagement metrics, (1) between next-game adopters and non-adopters and (2) wrt number of concurrent in-game purchases. It is clear that users who adopt behave very different than those who do not, while users with increased in-game purchases appears= to be best identified using play time. However, because I explicitly model neither adoption nor in-game purchasing, I cannot recover the average treatment effects. To formalize for the case of adoptions, let A_i be the event that player i adopts the next game, Z_i be a set of exogenous covariates and X_i a set of behavioral covariates. Consider the observed adoption decision as $A_i = A(X_i, Z_i, \epsilon_i)$. Under some binary policy $P \in \{0, 1\}$ we might consider the counterfactual $A_i^P = A(X_i^P, Z_i, P, \epsilon_i)$. At the moment I have only estimated $X_i = X(Z_i|\hat{\theta})$ and can only estimate the impact on A of policy P through its effect on X, i.e. $\hat{A}_i^P(X_i^P, Z_i, 0, \epsilon_i)$. In other words, while a policy may directly increase adoption probability and indirectly through increased engagement, I am limited to the latter effect. For instance, this limits my counterfactuals to the case that advertising and bundling does not shift the adopt decision for either the current or future games. To estimate this indirect adoption function I consider the set of users in the previous game $G - 1$ and perform logistic regression of game G adoption on engagement. The crucial assumption here is latent ignorability (i.e. conditional independence in the Rubin (1974) sense) which is non-trivial. To justify it I include an exhaustive set of behavioral controls.

I present results in Table 6.3[9]. In Regression (1) I consider the effect of total play on next-game adoption while in Regression (2) I decompose total play into Low + High levels. Note that Regression (1) actually provides superior fit (individual level plays are not significant), indicating that total play is more relevant. In Table 6.4 I present quasi-Poisson regression results for total number of game $G - 1$ in-game purchases. Here I find, conversely, that it is the individual level play that is important in predicting purchase, and Low level play in particular[10]. The role of in-game purchases is limited to a subset to modes, both competitive and non-competitive. In these modes, players earn in-game currency by playing and winning, and invest this currency to improve their in-game power. In-game purchases directly provide in-game currency, so are a direct substitute for playing. Players more effective at winning can thus be considered to have more efficient production functions for the in-game currency.

9. To more completely control for confounding effects I include all second-order interactions (with full interactions I still have over 20,000 df). However I omit them here for parsimony. I present results with only main effects in the Appendix. With only main effects an interesting result is that whether a player is new is not predictive of future adoption (although number of previous games is, and that coefficient incorporates this information).

10. The number of in-game purchases is highly skewed, with an in-sample maximum of almost 500 and 99% percentile of 90. To reasonably estimate an effect I limit myself to users within this 99% percentile. My findings are robust to different percentile thresholds.

One possible explanation then is that lower skilled players tend to play non-competitive levels but are also less efficient at producing in-game currency so substitute by making in-game purchases. Thus, from the perspective of the firm both total play and in each level are important levers. Using these results I estimate that an increase in total play of 1% leads to a 0.21% increase in adoptions. With a 50-50 split in Low and High level play, this corresponds to 0.51% increase in in-game purchases, while a 100-0 split corresponds to 0.92% increase in in-game purchases. To quantify this in the context of the considered counterfactuals, an increase in advertising of 10% would indirectly increase adoptions by 0.07% and in-game purchasing by 0.28%. An increase of 10% in the number of users switching from Xbox to PlayStation would increase overall adoptions by 0.003% and in-game purchasing by 0.02%. An increase in 10% in the number of users upgrading to the current generation console would increase overall adoptions by 0.007% and in-game purchasing by 0.02%.

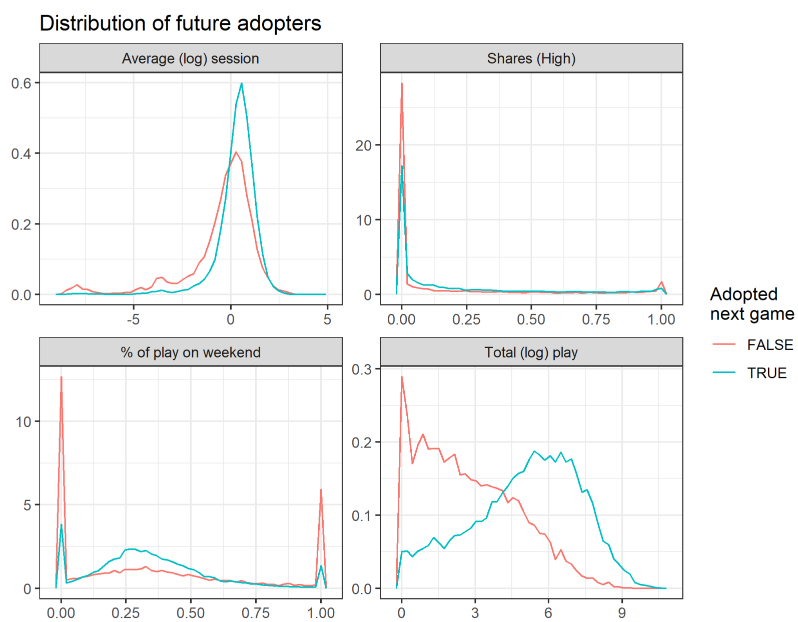


Figure 6.7: Retained and churned consumers have different behavior distributions.

	Dependent variable:	
	Game G adoption	
	(1)	(2)
(Intercept)	−1.208 (1.190)	−1.162 (1.217)
log(Total play)	0.371^{**} (0.155)	
log(Low lvl. play)		0.188 (0.118)
log(High lvl. play)		0.096 (0.116)
New player	−0.520 (0.593)	−0.534 (0.592)
Num. prev. games	0.221 (0.291)	0.220 (0.290)
Adoption time	0.004 (0.003)	0.004^{*} (0.003)
Is weekend adoption	−0.284 (0.437)	−0.284 (0.436)
PlayStation	-1.186^{***} (0.383)	-1.197^{***} (0.382)
Curr. gen. console	−0.580 (0.400)	−0.565 (0.400)
Num. sessions	−0.018 (0.016)	−0.012 (0.015)
log(Weekend play)	−0.648 (0.436)	−0.674 (0.446)
Avg. session length	−1.297 (0.826)	−1.294 (0.830)
Time from purch. to 1st play	−0.008 (0.016)	−0.008 (0.016)
Time from 1st to last play	0.006^{*} (0.003)	0.006^{*} (0.003)
High lvl. shares	−0.715 (0.756)	−0.776 (0.792)
HHI shares	−0.867 (1.128)	−0.919 (1.156)
log(First month play)	0.602 (0.378)	0.664^{*} (0.379)
log(Last month play)	-1.939^{*} (1.128)	-1.938^{*} (1.128)
Avg. date played	−0.001 (0.008)	−0.001 (0.008)
% of time played on wkend	0.900 (1.050)	1.110 (1.060)
Played any High lvl.	2.274^{*} (1.265)	2.166^{*} (1.283)
Played any mult. lvls.	−1.847 (1.238)	−1.729 (1.243)
Observations	22,181	22,181
Log Likelihood	−10,058.740	−10,060.160
Akaike Inf. Crit.	20,489.470	20,494.310
Note:		*p<0.1; **p<0.05; ***p<0.01

Table 6.3: Adoption as function of play, with interactions (main effects presented).

	Dependent variable:	
	In-game purchases	
	(1)	(2)
(Intercept)	-6.244^{***} (2.043)	-6.420^{***} (2.313)
log(Total play)	0.795^{**} (0.333)	
log(Low lvl. play)		0.925^{***} (0.274)
log(High lvl. play)		0.094 (0.097)
New player	2.390^{**} (1.028)	2.377^{**} (1.154)
Num. prev. games	0.977^{***} (0.356)	0.932^{**} (0.399)
Adoption time	0.002 (0.006)	0.003 (0.006)
Is weekend adoption	−0.441 (0.703)	−0.642 (0.793)
PlayStation	0.290 (0.587)	0.243 (0.658)
Curr. gen. console	−0.129 (0.802)	−0.180 (0.901)
Num. sessions	0.008 (0.013)	0.008 (0.013)
log(Weekend play)	0.573 (0.774)	0.310 (0.865)
Avg. session length	−1.234 (1.604)	−1.634 (1.797)
Time from purch. to 1st play	0.046^{**} (0.021)	0.051^{**} (0.023)
Time from 1st to last play	0.009^{**} (0.005)	0.010^{*} (0.005)
High lvl. shares	1.978 (1.231)	2.292 (1.442)
HHI shares	0.063 (1.630)	0.221 (1.873)
log(First month play)	−0.263 (0.492)	−0.354 (0.536)
log(Last month play)	-1.720^{**} (0.851)	-1.745^{*} (0.946)
Avg. date played	−0.004 (0.012)	−0.005 (0.013)
% of time played on wkend	1.960 (2.745)	2.921 (3.088)
Played any High lvl.	68.371 (713.119)	70.202 (822.162)
Played any mult. lvls.	−65.709 (713.118)	−67.743 (822.162)
Observations	21,960	21,960
Note:		*p<0.1; **p<0.05; ***p<0.01

Table 6.4: In-game purchases vs. play, with interactions (main effects presented, for purchases ≤ 90).

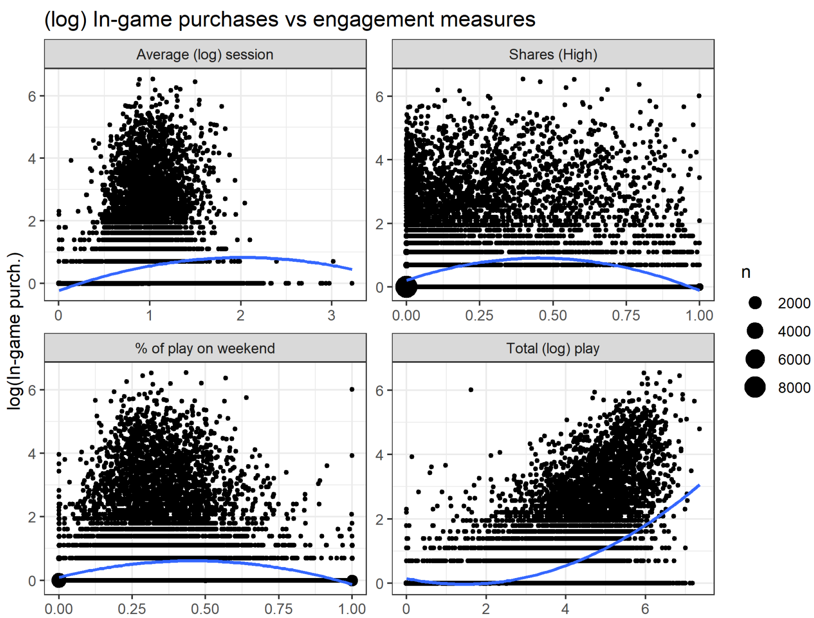


Figure 6.8: The strongest predictor of number of in-game purchases appears to be total play.

CHAPTER 7

CONCLUSION

In this paper I analyzed a novel dataset containing session-level information for a panel of players of a popular franchise video game. I focus on post-adoptive behavior for the most recent release, and observe patterns consistent within a micro-economic framework of consumers solving a budget problem over their leisure time. In particular I find that there is a tension between players learning their true match values, which tend to be more positive than their prior beliefs, and declining engagement due to novelty loss and outside interests. An important feature of the game is that it is comprised of competitive and non-competitive modes–I observe that players tend to substitute away from competition upon adoption of a new game and slowly switch back with experience. This finding is consistent with Bayesian learning. Differential learning rates are captured by a multiple discrete-continuous model that relates play intensity with learning intensity. An important advantage of this class of models is their natural decomposition of consumption into the extensive and intensive margins, allowing for a deeper layer of policy analysis. I contribute to the methodology literature by giving identification results for the MDCNEV model under learning.

In my empirical application I find the data can be characterized by latent segments corresponding to low ("casual") and high ("hardcore") types. Both tend to be ex ante pessimistic about match values, and the low types are additionally more sensitive to satiation effects. On average, neither player segment learns their true match values, and beyond the first few weeks casual players become drastically difficult to retain. These results suggest that increasing player engagement would be beneficial from the consumer welfare perspective. The question then turns to what actions the firm can do to increase engagement, and whether play in the different levels matters? Specifically, I consider counterfactuals where the firm (1) increases advertising and (2) promotes console upgrading (i.e. bundling). I find that advertising has a small effect on engagement while bundling has a much stronger effect but faces the issue of low compliance. Casual players are more responsive to both policies and primarily respond by playing more often. When players do switch consoles, they substitute away from competitive levels. This finding is in line with the hypothesis that players pay a psychic cost when adopting a new platform. For the firm, it is total play that drives future adoption while non-competitive play drives in-game purchasing. The indirect effect (i.e. increased purchasing due to increased play) of both policies is small, and unfortunately assessing the direct policy effect is outside the scope of my model.

There are several extensions that can be considered for my model. First, I do not model the adopt decision. A joint model would allow for direct measurement of more general outcomes such as consumer lifetime value but unfortunately (1) would dramatically complicate a model that is already highly complex, and (2) requires price data which I do not observe. For a similar reason I do not consider dynamics. Second, I only consider the simplest form of learning in my model. A joint learning specification, for example, would allow the learning to be naturally interpreted as skill accumulation. Finally, I do not consider the case of continuous unobserved heterogeneity. While two latent segments captures important data moments I show there is certainly an argument to be made for greater player heterogeneity.

REFERENCES

Allenby, G. and P. Rossi (1998). Marketing models of consumer heterogeneity. *Journal of Econometrics 89*(1–2), 57–78.

Ascarza, E. and B. Hardie (2013). A joint model of usage and churn in contractual settings. *Marketing Science 32*(4), 570–590.

Berlyne, D. (1970). Novelty, complexity, and hedonic value. *Perception and Psychophysics 8*(5), 279–286.

Bhat, C. (1997). Endogenous segmentation mode choice model with an application to intercity travel. *Transportation Science 31*(1), 34–48.

Bhat, C. (2005). A multiple-discrete continuous extreme value model: formulation and application to discretionary time-use decisions. *Transportation Research Part B 39*(8), 679–707.

Bhat, C. (2008). The multiple discrete-continuous extreme value (MDCEV) model: role of utility function parameters, identification considerations, and model extensions. *Transportation Research Part B 42*(3), 274–303.

Bogers, M., A. Afuah, and B. Bastian (2010). Users as innovators: a review, critique, and future research directions. *Journal of Management 36*(4), 857–875.

Bonnet, C. and O. de Mouzon (2014). Analysis of multiple discrete-continuous choices: empirical evidence of biased price elasticities under standard discrete choice models on the soft drink market. Technical Report TSE-520, Toulouse School of Economics.

Boudreau, K., K. Lakhani, and M. Menietti (2016). Performance responses to competition across skill levels in rank-order tournaments: field evidence and implications for tournament design. *The RAND Journal of Economics 47*(1), 140–165.

Braun, M. and D. Schweidel (2011). Modeling consumer lifetimes with multiple causes of churn. *Marketing Science 30*(5), 881–902.

Ching, A., T. Erdem, and M. Keane (2013). Learning models: an assessment of progress, challenges and new developments. *Marketing Science 32*(6), 913–938.

Ching, A. and H. Lim (2016). A structural model of correlated learning and late-mover advantages: the case of statins. Technical Report 2662286, Rotman School of Management Working Paper.

Chintagunta, P., R. Jiang, and G. Jin (2009). Information, learning and drug diffusion: the case of Cox-2 inhibitors. *Quantitative Marketing and Economics 7*(4), 399–443.

Coscelli, A. and M. Shum (2004). An empirical model of learning and patient spillovers in new drug entry. *Journal of Econometrics 122*(2), 213–246.

Cruz, C., M. Hanus, and J. Fox (2017). The need to achieve: players' perceptions and uses of extrinsic meta-game reward systems for video game consoles. *Computers in Human Behavior 71*, 516–524.

Deighton, J., C. Henderson, and S. Neslin (1994). The effects of advertising on brand switching and repeat purchasing. *Journal of Marketing Research 31*(1), 28–43.

Derdenger, T. and K. Vineet (2013). The dynamic effects of bundling as product strategy. *Marketing Science 32*(6), 827–859.

Dominici, F. and G. Parmigiani (2000). Conjugate analysis of multivariate normal data with incomplete observations. *The Canadian Journal of Statistics 28*(3), 533–550.

Ewoldsen, D., C. Eno, B. Okdie, J. Velez, R. Guadagno, and J. DeCoster (2012). Effect of playing violent video games cooperative or competitively on subsequent cooperative behavior. *Cyberpsychology, Behavior, and Social Networking 15*(5), 277–280.

Fang, H. A. (2008). A discrete–continuous model of households' vehicle choice and usage, with an application to the effects of residential density. *Transportation Research Part B: Methodological 42*(9), 736–758.

Frey, M., P. Laguna, and K. Ravizza (2003). Collegiate athletes' mental skill use and perceptions of success: an exploration of the practice and competition settings. *Journal of Applied Sport Psychology 15*(2), 115–128.

Gilbert, B. (2018). Five years after launch, we compared the PlayStation 4 and the Xbox One - and it's a harder decision than ever. *Business Insider*.

Gu, L., I. Ryzhov, S. Mankad, and B. Han (2016). Social behavior and user engagement in competitive online gaming: an empirical analysis.

Han, S. P., S. Park, and W. Oh (2016). Mobile app analytics: a multiple discrete-continuous choice framework. *Management information systems quarterly 40*(4), 983–1008.

Hartmann, W. and V. B. Viard (2008). Do frequency reward programs create switching costs? A dynamic structural analysis of demand in a reward program. *Quantitative Marketing and Economics 6*(2), 109–137.

He, C. and T. Klein (2018). Advertising as reminder: evidence from the Dutch State Lottery. Technical Report 2018-018, TILEC Discussion Paper.

Holbrook, M., R. Chestnut, T. Oliva, and E. Greenleaf (1984). Play as a consumption experience: the roles of emotions, performance, and personality in the enjoyment of games. *Journal of consumer research 11*(2), 728–739.

Holbrook, M. and E. Hirschman (1982). The experiential aspects of consumption: consumer fantasies, feelings, and fun. *Journal of Consumer Research 9*(2), 132–140.

Honka, E., A. Hortacsu, and M. A. Vitorino (2017). Advertising, consumer awareness, and choice: evidence from the u.s. banking industry. *The RAND Journal of Economics 48*(3), 611–646.

Huang, Y. (2019). Learning by doing and the demand for advanced products. *Marketing Science*, Revision requested.

Huh, Y. E. and S.-H. Kim (2008). Do early adopters upgrade early? Role of post-adoption behavior in the purchase of next-generation products. *Journal of Business Research 61*(1), 40–46.

Jansz, J. and M. Tanis (2007). Appeal of playing online first person shooter games. *CyberPsychology and Behavior 10*(1), 133–136.

Jiang, Z., F. Bass, and P. Bass (2006). Virtual Bass Model and the left-hand data-truncation bias in diffusion of innovation studies. *International Journal of Research in Marketing 23*(1), 93–106.

Kifer, A. (2015). The incidence of a soda tax, in pennies and pounds. Technical Report 1-017, Kilts Booth Marketing series.

Kim, J., G. Allenby, and P. Rossi (2002). Modeling consumer demand for variety. *Marketing Science 21*(3), 229–250.

Lee, L.-f. (1995). Asymptotic bias in simulated maximum likelihood estimation of discrete choice models. *Econometric Theory 11*(3), 437–483.

Lee, S., S. Thomas, and G. Allenby (2018). An economic analysis of demand of the very poor.

Luo, L., B. Ratchford, and B. Yang (2013). Why do we do what we do? A model of activity consumption. *Journal of Marketing Research 50*(1), 24–43.

Mehta, N., S. Rajiv, and K. Srinivasan (2003). Price uncertainty and consumer search: a structural model of consideration set formation. *Marketing Science 22*(1), 58–84.

Narayanan, S., P. Chintagunta, and E. Miravete (2007). The role of self selection, usage uncertainty and learning in the demand for local telephone service. *Quantitative Marketing and Economics 5*(1), 1–34.

Narayanan, S. and P. Manchanda (2009). Heterogeneous learning and the targeting of marketing communications for new products. *Marketing Science 28*(3), 424–441.

Nevskaya, Y. and P. Albuquerque (2019). The impact of innovation on product usage: a dynamic model with progression in content consumption. Technical report.

Nielsen, J. (2006). The 90-9-1 rule for participation inequality in social media and online communities. *Nielsen Norman Group*.

Norton, J. and F. Bass (1987). A diffusion theory model of adoption and substitution for successive generations of high-technology products. *Management Science 33*(9), 1069–1086.

Pedrick, J. and F. Zufryden (1991). Evaluating the impact of advertising media plans: a model of consumer purchase dynamics using single-source data. *Marketing Science 10*(2), 111–130.

Pollmann, M. and E. Krahmer (2017). How do friends and strangers play the game Taboo? a study of accuracy, efficiency, motivation, and the use of shared knowledge. *Journal of Language and Social Psychology 37*(4), 497–517.

Rubin, D. (1974). Estimating causal effects of treatments in randomized and nonrandomized studies. *Journal of Educational Psychology 66*(5), 688–701.

Sanders, C. (2016). Skill accumulation, skill uncertainty, and occupational choice. *Under review*.

Satomura, T., J. Kim, and G. Allenby (2011). Multiple constraint choice models with corner and interior solutions. *Marketing Science 30*(3), 481–490.

Schmierbach, M. (2010). Killing spree: exploring the connection between competitive game play and aggressive cognition. *Communication Research 37*(2), 256–274.

Shapiro, B., G. Hitsch, and A. Tuchman (2019). Generalizable and robust TV advertising effects. Technical report, University of Chicago Booth School of Business.

Sobhani, A., N. Eluru, and A. Faghih-Imani (2013). A latent segmentation based multiple discrete continuous extreme value model. *Transportation Research Part B: Methodological 58*(4), 154–169.

Steiner, M., N. Wiegand, A. Eggert, and K. Backhaus (2016). Platform adoption in system markets: the roles of preference heterogeneity and consumer expectations. *International Journal of Research in Marketing*.

Sunada, T. (2019). Customer learning and revenue maximizing trial provision. Technical report, Department of Economics, University of Pennsylvania.

Tao, X. and A. Sweeting (2019). Capital accumulation and preference formation: an application to the video game port-back. Technical report, Department of Economics, University of Maryland.

Tyack, A., P. Wyeth, and D. Johnson (2016). The appeal of MOBA games: what makes people start, stay, stop. In *Proceedings of the 2016 Annual Symposium on Computer-Human Interaction in Play.*

von Haefen, R. H. and D. J. Phaneuf (2003). Estimating preferences for outdoor recreation: a comparison of continuous and count data demand system frameworks. *Journal of Environmental Economics and Management 45*(3), 612–630.

Vorderer, P., T. Hartmann, and C. Klimmt (2003). Explaining the enjoyment of playing video games: the role of competition. In *International Conference on Entertainment Computing*.

Weibel, D., B. Wissmath, S. Habegger, Y. Steiner, and R. Groner (2008). Playing online games against computer- vs. human-controlled opponents: effects on presence, flow, and enjoyment. *Computers in Human Behavior 24*(5), 2274–2291.

Williams, R. and C. Clippinger (2002). Aggression, competition and computer games: computer and human opponents. *Computers in Human Behavior 18*(5), 495–506.

Yee, N. (2007). Motivations for play in online games. *Cyberpsychology and Behavior 9*(6), 772–775.

APPENDIX A

DATA

A.1 Coefficients in ZI-exponential adoption model

X_i	$\hat{\delta}^1$	$\hat{\delta}^2$
Intercept	-4.227***	2.815***
Is new player	1.518***	0.422***
log(Previous game exp.)	1.107***	-1.829***
Average level played in prev. game	0.177***	0.292***

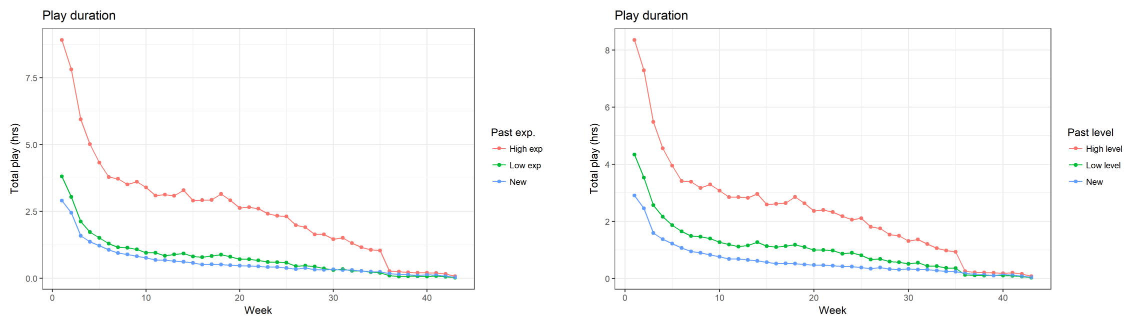
Table A.1: Zero-inflated exponential adoption hazard model. Note that all coefficients are significant, but in particular new player and the intercept have relatively large size.

A.2 Additional figures/tables for level vs experience groupings

Type	Mean adoption	Median adopt	3rd quartile adopt	% adoption in first 2 weeks
New	15.3	17.00	22.00	0.177
Low level	4.24	0.00	3.00	0.717
High level	4.39	0.00	3.00	0.708

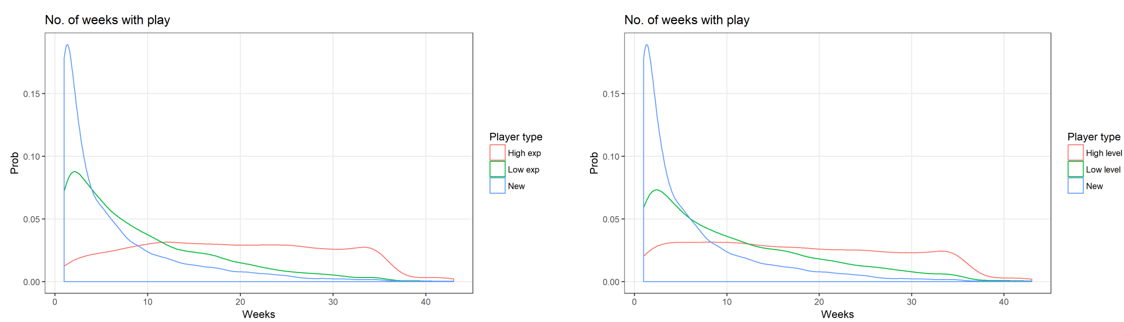
Table A.2: Adoption summaries, by player past level.

Adoption statistics grouped by median split on level played rather than experience.



(a) (b)

Figure A.1: Total play by different groupings. Qualitatively similar findings to Fig. 3.3



(a) (b)

Figure A.2: No. of weeks played by different groupings. Note that grouped by level, returning players look rather similar while grouped by experience, there is a clear difference between low and high exp players

A.3 Conditional play patterns with alternate slicings

Below I present conditional play patterns sliced on (a) adoption date, and (b) number of weeks played. I use sample quartiles to create slices. In general observe that except for noisy-by-construction groups (i.e. 4th quartile adopters, or 1st quartile number of weeks played), the same trends observed in aggregate appear to hold here as well.

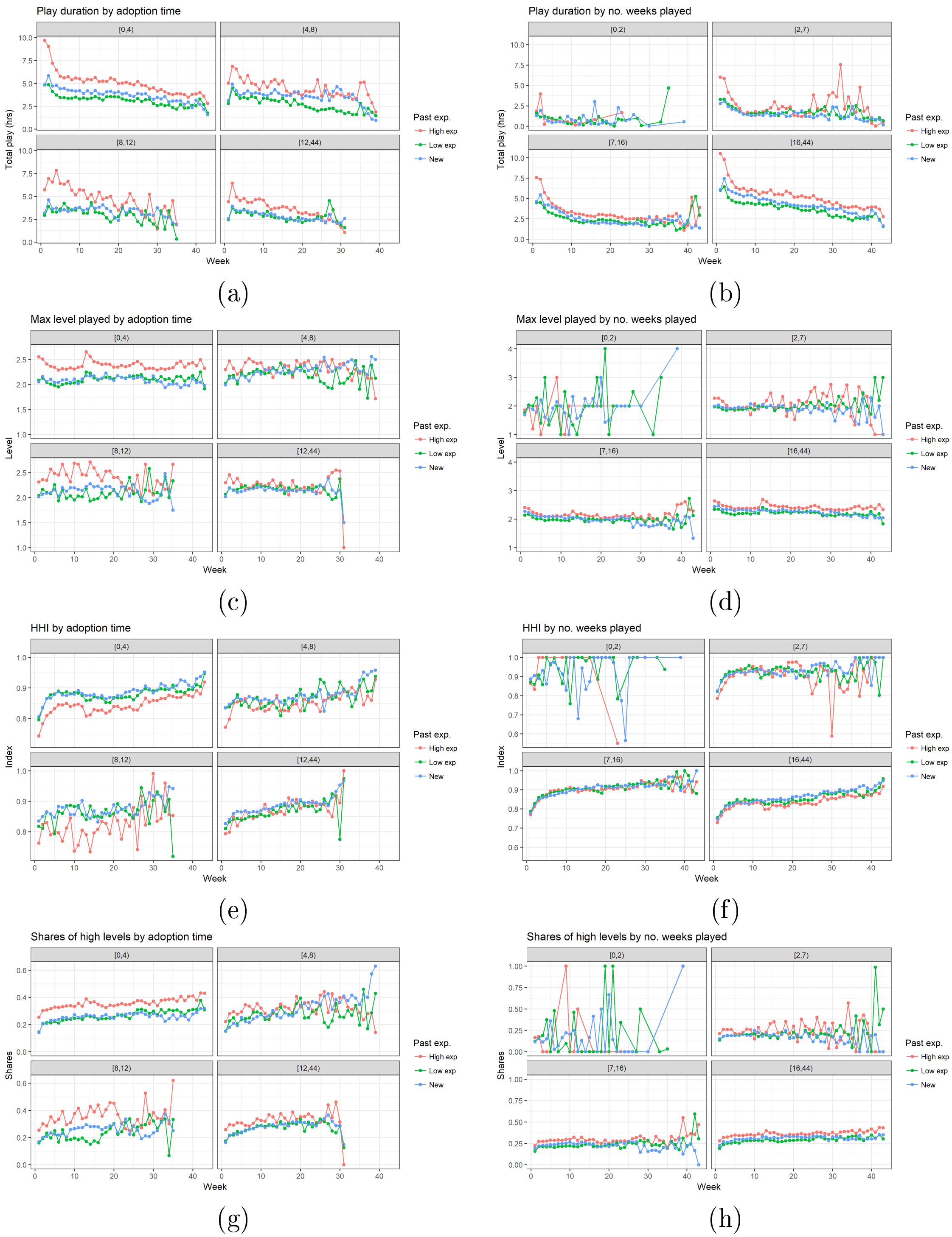


Figure A.3: Conditional total duration, max level played, HHI, and high level shares.

A.4 Conditional play patterns grouped by previous level

Instead of a median split on duration played here I group by median split on past level. Here there is more separation between the low types and new players, except in play probabilities (extensive margin) and high shares, which is unsurprising. As with the experience split, it is important to note that conditional on an outcome variable gives rise to endogeneity and the results can at best be considered illustrative.

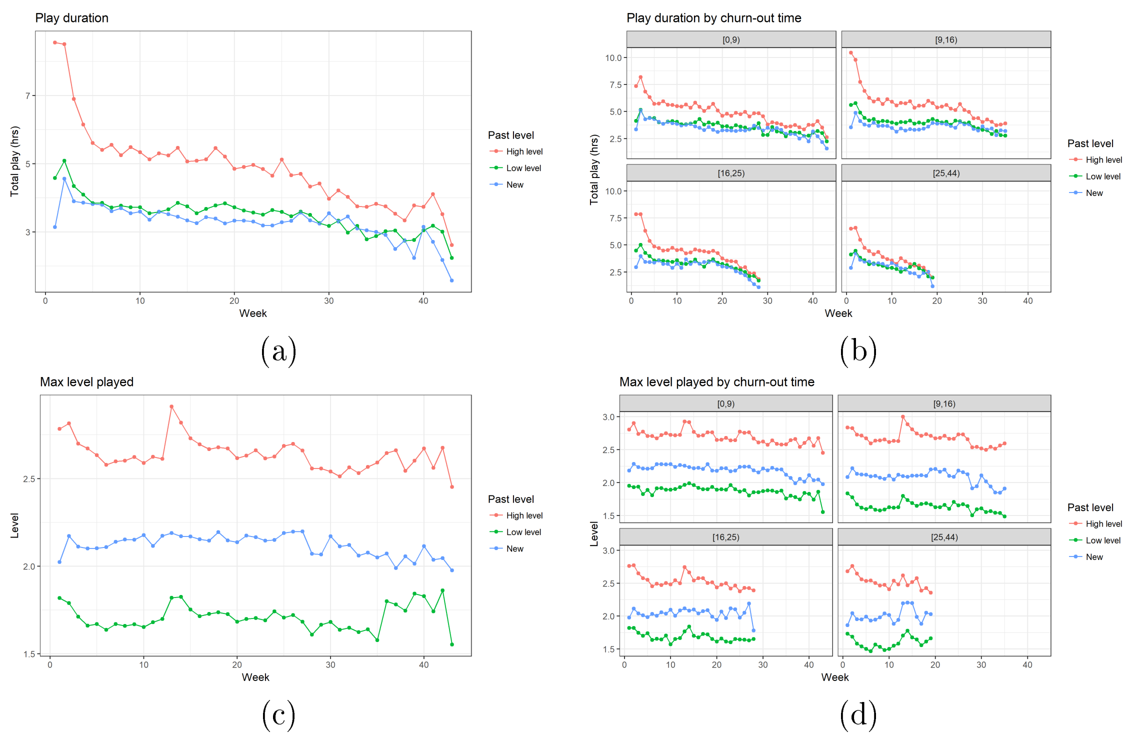
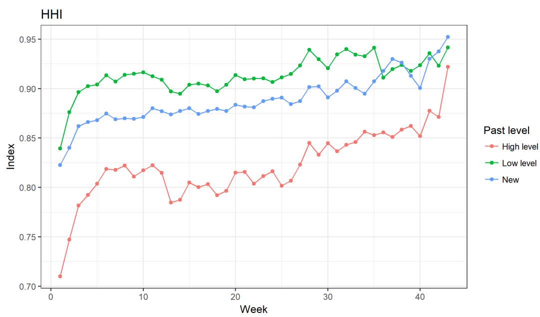
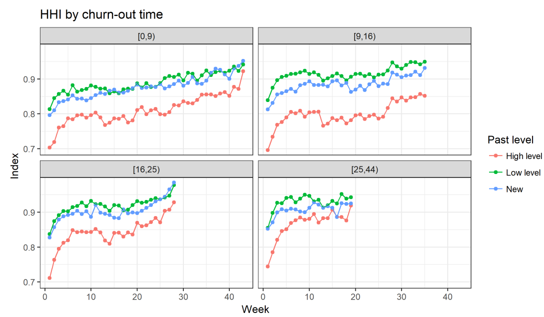


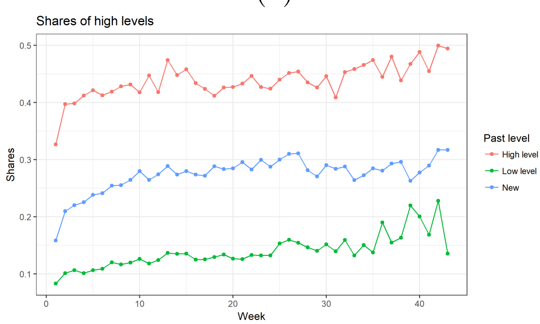
Figure A.4: Conditional play duration and max level by previous level. Note that compared to Fig. 3.4 the monotonicity between groups at each slice is significantly blurrier. Further, max level may be flat or slightly declining across all groups.



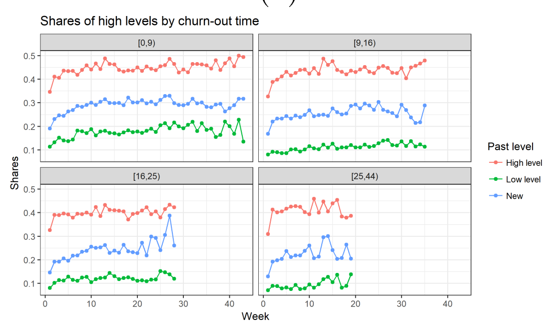
(a)



(b)



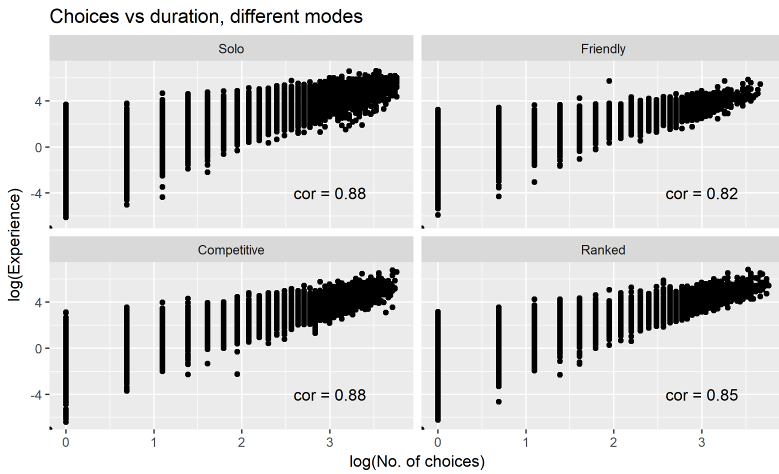
(c)



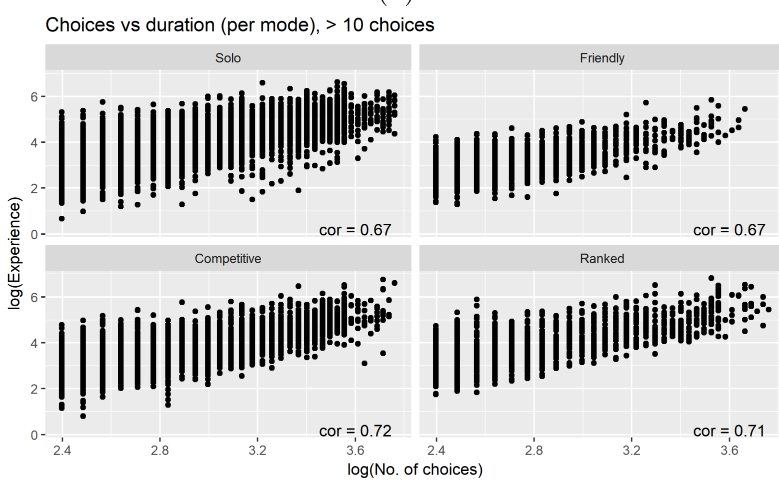
(d)

Figure A.5: Conditional HHI and high level shares by previous level. Unsurprisingly past high levels is predictive of current high levels.

A.5 Correlation between (log) duration and (log) no. choices



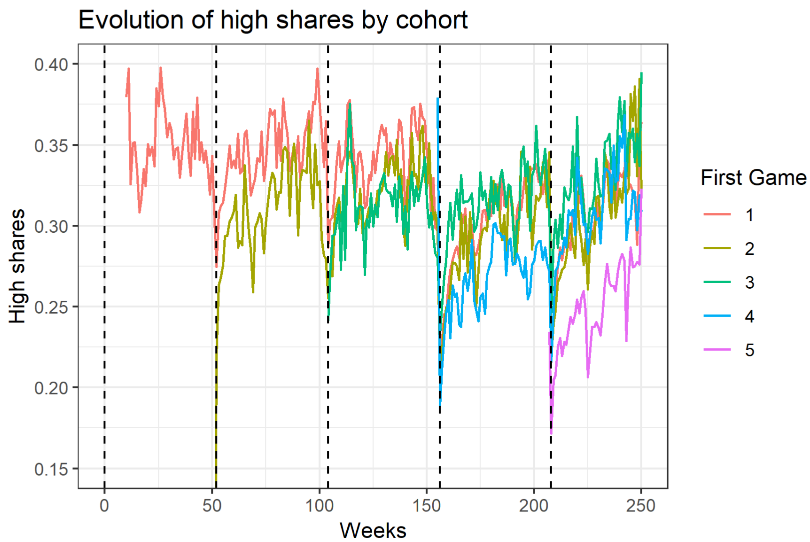
(a)



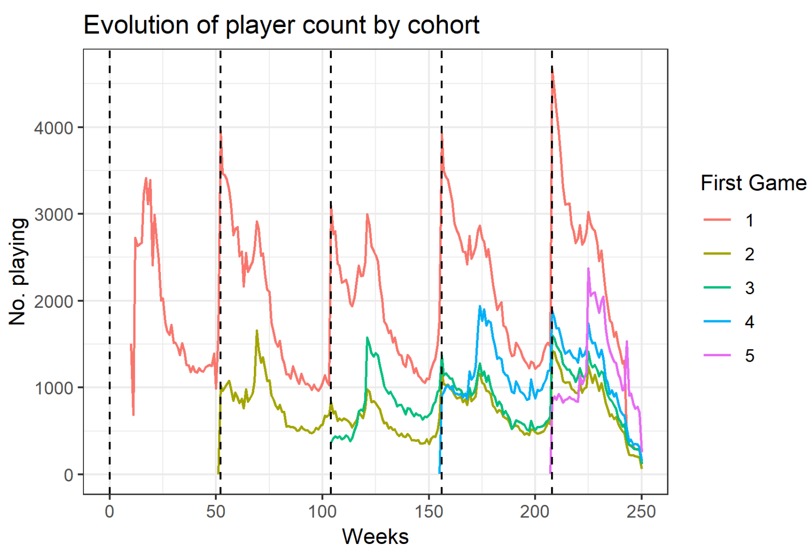
(b)

Figure A.6: Note: only players with positive no. choices in a mode are considered. Unsurprisingly, correlation (a) is rather high. Conditional on high utilization (b), correlation is significantly lower. Here we start to see that the two variables contain different information.

A.6 Evolution of high level shares across all generations



(a)



(b)

Figure A.7: The pattern of increasing high level shares over the course of the game followed by a sharp drop after the next game holds by cohort. However, selection cannot be ruled out.

A.7 Cluster analysis of $K = 2, 3$ clusters

Figure A.8: Data moments with $K = 2$ clusters.

Table A.3: Clustered behavioral and demographic means, $K = 2$.

	Clust. 1	Clust. 2
No. games owned	2.26	3.52
No. yrs in samp.	0.97	1.90
Adopt day G	97.07	34.48
Owns Xbox 360	0.09	0.06
Owns Xbox One	0.46	0.45
Owns PS3	0.06	0.05
Owns PS4	0.41	0.49
Owns Xbox	0.54	0.49
Owns PS	0.46	0.52
Owns older cons.	0.15	0.11
Owns multiple cons.	0.01	0.05
High exp	0.08	0.45
Low exp	0.26	0.24
Returning	0.34	0.70
New	0.66	0.30
Adopt day 15	78.56	56.61
No. wks. played $G-1$	8.72	19.83
No. wks. owned $G-1$	41.15	44.23
Pct. wks. played $G-1$	0.23	0.45

Table A.3: Clustered behavioral and demographic means, $K = 2$, continued.

Last wk. played $G-1$	31.57	40.81
Overall max lvl $G-1$	1.81	2.53
First wk. play $G-1$	4.03	7.57
First 4 wks. play $G-1$	11.05	23.42
First wk. max lvl $G-1$	1.09	1.11
First 4 wks. max lvl $G-1$	1.34	1.60
Total solo play $G-1$	23.46	60.27
Total friendly play $G-1$	3.69	8.75
Total competitive play $G-1$	3.02	22.78
Total ranked play $G-1$	2.06	17.26
Total play $G-1$	32.23	107.71
Overall high shares $G-1$	0.13	0.30
Overall HHI $G-1$	0.77	0.67
Above med. high lvls, $G-1$	0.15	0.54
Above med. low lvls, $G-1$	0.48	0.73
Above med. ranked play $G-1$	0.07	0.30
Above med. competitive play $G-1$	0.10	0.40
Above med. friendly play $G-1$	0.25	0.44
Above med. solo play $G-1$	0.33	0.58
Above med. last wk. played $G-1$	0.33	0.59
Above med. HHI $G-1$	0.75	0.93
Above med. no. wks. played $G-1$	0.24	0.64
Above med. high shares $G-1$	0.31	0.61
Above med. max lvl played $G-1$	0.22	0.53
Adopted during weekend $G-1$	0.29	0.18
Adopted 1st 2 weeks $G-1$	0.44	0.56
N	14425.00	12290.00

First, I use $K = 2$ clusters. Here the main separation appears to be between newer and advanced players: the average player in the second cluster is about one year younger and much more likely to be new. Notably, low exp players are just as likely to belong to either cluster. For returning players all the previous game play statistics indicate the first cluster consists of more engaged players. Unsurprisingly, these players are more likely to both own multiple consoles and newer consoles. What is surprising is that these "hardcore" players are also more likely to own a PlayStation than an Xbox[1].

1. $p << 0.05$. Gilbert (2018) concludes that the Xbox One and PlayStation 4 are ultimately very similar

Figure A.9: Data moments with $K = 3$ clusters.

Table A.4: Clustered behavioral and demographic means, $K = 3$.

	Clust. 1	Clust. 2	Clust. 3
No. games owned	3.57	3.24	2.03
No. yrs in samp.	2.01	1.64	0.77
Adopt day G	32.96	46.65	108.87
Owns Xbox 360	0.06	0.06	0.09
Owns Xbox One	0.45	0.44	0.47
Owns PS3	0.05	0.05	0.06
Owns PS4	0.48	0.48	0.39
Owns Xbox	0.50	0.49	0.56
Owns PS	0.52	0.53	0.45
Owns older cons.	0.11	0.12	0.15
Owns multiple cons.	0.04	0.04	0.01
High exp	0.41	0.39	0.05
Low exp	0.28	0.25	0.24
Returning	0.69	0.64	0.28

except in one major dimension: the PS4 has a much larger library of system-exclusive games. Cruz et al. (2017) note that trophy or achievement points (which accrue by user account across games but **not** consoles) are an important driver of motivation, enjoyment, and engagement. I hypothesize then, that in addition to brand loyalty, whether a player can retain their points contributes to the brand "stickiness". In my data I note that across the latest two generations of games, 2% of users on a current-generation console switch brands across games, while for those who upgrade consoles from previous- to current-generation alongside games, brand-switching is 10%.

Table A.4: Clustered behavioral and demographic means, $K = 3$, continued.

New	0.31	0.36	0.72
Adopt day $G-1$	54.79	60.88	88.24
No. wks. played $G-1$	18.03	18.79	7.17
No. wks. owned $G-1$	44.49	43.63	39.79
Pct. wks. played $G-1$	0.41	0.43	0.20
Last wk. played $G-1$	39.06	40.14	30.49
Overall max lvl $G-1$	1.89	3.04	1.85
First wk. play $G-1$	6.91	7.50	3.33
First 4 wks. play $G-1$	21.08	22.99	8.92
First wk. max lvl $G-1$	1.05	1.16	1.11
First 4 wks. max lvl $G-1$	1.24	1.93	1.38
Total solo play $G-1$	71.78	36.67	15.52
Total friendly play $G-1$	7.86	7.88	3.69
Total competitive play $G-1$	11.12	29.80	2.98
Total ranked play $G-1$	2.72	29.68	2.08
Total play $G-1$	92.32	102.94	24.27
Overall high shares $G-1$	0.11	0.47	0.15
Overall HHI $G-1$	0.76	0.60	0.75
Above med. high lvls, $G-1$	0.30	0.70	0.15
Above med. low lvls, $G-1$	0.79	0.63	0.40
Above med. ranked play $G-1$	0.10	0.46	0.08
Above med. competitive play $G-1$	0.25	0.48	0.09
Above med. friendly play $G-1$	0.38	0.44	0.25
Above med. solo play $G-1$	0.68	0.43	0.22
Above med. last wk. played $G-1$	0.54	0.57	0.30
Above med. HHI $G-1$	0.85	0.95	0.75
Above med. no. wks. played $G-1$	0.59	0.60	0.17
Above med. high shares $G-1$	0.34	0.82	0.33
Above med. max lvl played $G-1$	0.27	0.73	0.23
Adopted during weekend $G-1$	0.18	0.20	0.32
Adopted 1st 2 weeks $G-1$	0.57	0.54	0.39
N	8382.00	7203.00	11130.00

Next I use $K = 3$ clusters, and observe that clusters 1 and 3 are exceedingly similar (and seem to largely reflect returning players) with one major behavioral difference: cluster 1 primarily stays in the low levels while cluster 3 plays the high levels. These patterns persist from the previous game. Possible explanations for this pattern include preference heterogeneity and state-dependent, largely non-transferred learning. Finally, note that cluster 3 has play frequency and duration patterns similar to cluster 1, but high level shares similar to cluster 2.

A.8 Regressions of outcome variables for each of $K = 5$ clusters and overall

	Dependent variable:					
	No. wks. played G					
	Cl. 1	Cl. 2	Cl. 3	Cl. 4	Cl. 5	All
(Intercept)	2.00***	16.27***	2.64***	6.03***	12.38***	4.87***
	(0.06)	(0.33)	(0.08)	(0.30)	(0.25)	(0.12)
Owns PS	0.33***	0.74***	0.01	1.06***	0.04	1.04***
	(0.05)	(0.23)	(0.06)	(0.24)	(0.19)	(0.10)
Owns older cons.	0.09	−0.93**	−0.11	−0.67*	−0.38	−1.20***
	(0.06)	(0.46)	(0.10)	(0.35)	(0.32)	(0.15)
Owns multiple cons.	1.96***	2.34***	1.26***	1.88***	1.99***	4.58***
	(0.25)	(0.59)	(0.27)	(0.73)	(0.53)	(0.30)
No. games owned	0.25***	0.20	0.14***	0.36*	0.31**	0.68***
	(0.04)	(0.15)	(0.04)	(0.18)	(0.13)	(0.07)
No. yrs in samp.	−0.10***	−0.07	−0.005	−0.05	−0.04	−0.19***
	(0.04)	(0.14)	(0.04)	(0.16)	(0.11)	(0.06)
Low exp	−0.21	−0.92	0.47***	−0.67	0.72*	0.23
	(0.14)	(0.73)	(0.13)	(0.72)	(0.41)	(0.23)
High exp	−0.19	1.15	0.09	2.81***	2.97***	3.44***
	(0.31)	(0.88)	(0.21)	(0.88)	(0.56)	(0.33)
Above med. ranked play $G-1$	−0.27	0.93***	−0.13	3.08***	0.17	1.08***
	(0.25)	(0.32)	(0.19)	(0.45)	(0.48)	(0.21)
Above med. competitive play $G-1$	−0.06	1.78***	−0.32*	−0.33	2.90***	2.54***
	(0.25)	(0.37)	(0.18)	(0.40)	(0.38)	(0.20)
Above med. friendly play $G-1$	0.07	0.39	0.32***	−1.24***	−0.42	0.29*
	(0.15)	(0.28)	(0.11)	(0.36)	(0.26)	(0.15)
Above med. solo play $G-1$	−0.22	0.27	0.40***	−1.72***	−0.09	1.46***
	(0.21)	(0.38)	(0.13)	(0.45)	(0.36)	(0.19)
Above med. last wk. played $G-1$	0.50***	2.79***	0.36***	2.43***	2.58***	3.42***
	(0.12)	(0.29)	(0.10)	(0.36)	(0.25)	(0.15)
Above med. HHI $G-1$	0.40***	0.52	0.23*	0.56	−0.46	0.36
	(0.15)	(0.74)	(0.12)	(0.75)	(0.34)	(0.22)
Above med. no. wks. played $G-1$	0.44*	2.41***	0.76***	3.51***	3.11***	4.00***
	(0.23)	(0.43)	(0.15)	(0.50)	(0.34)	(0.21)
Above med. high shares $G-1$	−0.40**	−0.60	−0.19	−0.91	−1.01***	−0.05
	(0.17)	(0.43)	(0.14)	(0.65)	(0.38)	(0.21)
Above med. max lvl played $G-1$	−0.12	−0.02	0.07	0.66	−1.52***	0.09
	(0.18)	(0.34)	(0.14)	(0.55)	(0.35)	(0.20)
Adopted during weekend $G-1$	0.16	−1.37***	−0.05	−0.45	−1.02***	−1.51***
	(0.12)	(0.34)	(0.10)	(0.40)	(0.30)	(0.16)
Adopted 1st 2 weeks $G-1$	−0.09	−0.79**	−0.24**	−0.48	−0.82***	−1.20***
	(0.17)	(0.31)	(0.10)	(0.38)	(0.27)	(0.16)
Observations	6,950	5,162	4,949	3,430	6,224	26,715
R^2	0.04	0.19	0.09	0.30	0.20	0.38
Note:						*p<0.1; **p<0.05; ***p<0.01

Table A.5: OLS of no. wks. played G for each cluster and overall.

	Dependent variable:					
	Last wk. played G					
	Cl. 1	Cl. 2	Cl. 3	Cl. 4	Cl. 5	All
(Intercept)	27.87***	34.78***	9.78***	26.48***	31.98***	25.62***
	(0.24)	(0.29)	(0.27)	(0.46)	(0.27)	(0.16)
Owns PS	−1.08***	−0.48**	−0.40*	−0.29	−0.85***	−0.93***
	(0.19)	(0.20)	(0.22)	(0.36)	(0.20)	(0.13)
Owns older cons.	−1.90***	−0.46	0.49	−1.42***	−1.01***	−1.77***
	(0.26)	(0.41)	(0.33)	(0.53)	(0.34)	(0.21)
Owns multiple cons.	6.12***	2.04***	9.68***	5.27***	2.33***	6.18***
	(1.01)	(0.52)	(0.91)	(1.10)	(0.57)	(0.42)
No. games owned	0.52***	0.25*	0.48***	1.12***	0.29**	0.95***
	(0.18)	(0.14)	(0.15)	(0.28)	(0.13)	(0.10)
No. yrs in samp.	−0.39**	−0.69***	−0.55***	−1.37***	−0.71***	−1.30***
	(0.16)	(0.12)	(0.13)	(0.25)	(0.12)	(0.09)
Low exp	−1.37**	−4.27***	−0.69	−6.32***	−3.14***	−5.35***
	(0.55)	(0.65)	(0.44)	(1.09)	(0.44)	(0.31)
High exp	−0.33	−4.33***	−2.59***	−5.04***	−3.28***	−3.91***
	(1.25)	(0.78)	(0.72)	(1.34)	(0.60)	(0.45)
Above med. ranked play $G-1$	−0.23	0.37	−0.39	1.27*	−0.49	0.38
	(0.98)	(0.28)	(0.64)	(0.69)	(0.52)	(0.28)
Above med. competitive play $G-1$	−3.25***	−0.10	−0.76	0.07	−0.15	0.43
	(0.99)	(0.33)	(0.61)	(0.61)	(0.41)	(0.28)
Above med. friendly play $G-1$	0.95	0.58**	1.56***	−0.10	−0.01	0.68***
	(0.58)	(0.24)	(0.37)	(0.55)	(0.28)	(0.21)
Above med. solo play $G-1$	−0.26	−0.07	−0.28	−0.94	−0.51	0.74***
	(0.85)	(0.34)	(0.45)	(0.69)	(0.39)	(0.27)
Above med. last wk. played $G-1$	1.52***	2.05***	1.17***	3.25***	2.88***	3.74***
	(0.49)	(0.26)	(0.36)	(0.54)	(0.26)	(0.20)
Above med. HHI $G-1$	−0.08	−0.02	0.73*	0.37	−0.63*	0.46
	(0.59)	(0.66)	(0.41)	(1.14)	(0.36)	(0.30)
Above med. no. wks. played $G-1$	2.32**	1.64***	1.50***	3.27***	2.15***	3.89***
	(0.92)	(0.38)	(0.50)	(0.76)	(0.37)	(0.29)
Above med. high shares $G-1$	0.34	−0.68*	−0.21	−2.50**	−0.92**	−0.17
	(0.67)	(0.38)	(0.48)	(0.98)	(0.41)	(0.30)
Above med. max lvl played $G-1$	−0.45	0.05	0.07	1.58*	−0.32	0.69**
	(0.71)	(0.30)	(0.49)	(0.84)	(0.37)	(0.27)
Adopted during weekend $G-1$	0.26	0.17	0.17	1.74***	0.97***	0.16
	(0.46)	(0.30)	(0.35)	(0.60)	(0.32)	(0.23)
Adopted 1st 2 weeks $G-1$	−0.15	−1.54***	−1.44***	−2.82***	−1.99***	−3.33***
	(0.66)	(0.27)	(0.34)	(0.58)	(0.29)	(0.22)
Observations	6,950	5,162	4,949	3,430	6,224	26,715
R^2	0.02	0.10	0.05	0.09	0.10	0.10
Note:						*p<0.1; **p<0.05; ***p<0.01

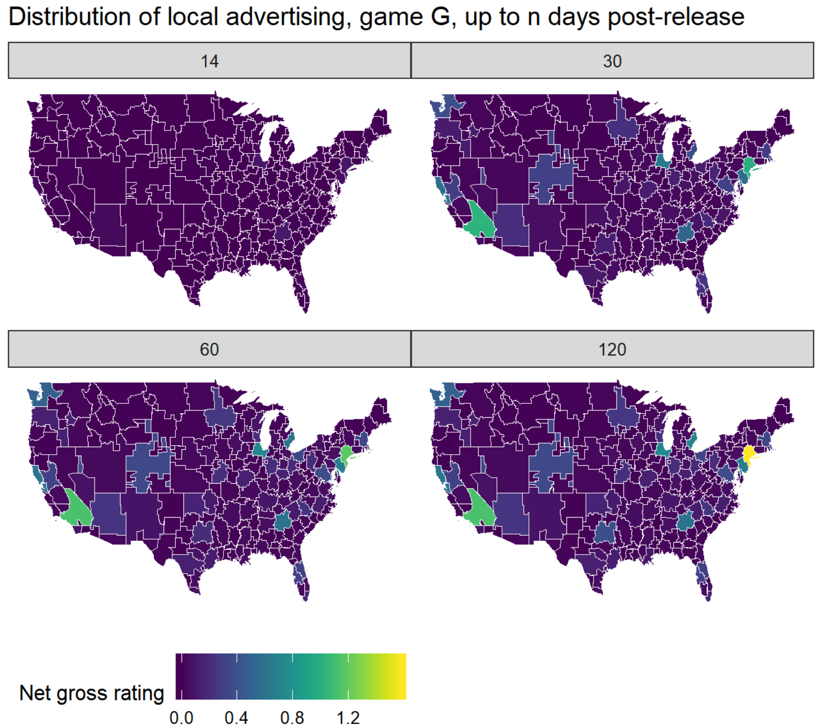
Table A.6: OLS of last wk. played G for each cluster and overall.

	Dependent variable:					
	Overall high shares G					
	Cl. 1	Cl. 2	Cl. 3	Cl. 4	Cl. 5	All
(Intercept)	0.06***	0.40***	0.07***	0.70***	0.05***	0.16***
	(0.01)	(0.01)	(0.01)	(0.01)	(0.003)	(0.004)
Owns PS	0.02***	0.002	0.01	−0.01	−0.003	0.02***
	(0.01)	(0.01)	(0.01)	(0.01)	(0.002)	(0.003)
Owns older cons.	0.03***	−0.003	0.01	0.04***	−0.01*	0.01**
	(0.01)	(0.01)	(0.01)	(0.01)	(0.003)	(0.01)
Owns multiple cons.	0.003	0.01	−0.004	−0.13***	0.01*	0.01
	(0.03)	(0.02)	(0.02)	(0.03)	(0.01)	(0.01)
No. games owned	0.01***	−0.01	0.01***	−0.01**	0.002	0.01***
	(0.005)	(0.004)	(0.004)	(0.01)	(0.001)	(0.002)
No. yrs in samp.	−0.01*	0.005	−0.01***	0.01*	−0.003***	−0.01***
	(0.004)	(0.004)	(0.003)	(0.01)	(0.001)	(0.002)
Low exp	0.005	−0.03	−0.05***	−0.09***	−0.02***	−0.07***
	(0.02)	(0.02)	(0.01)	(0.03)	(0.004)	(0.01)
High exp	0.02	0.05*	−0.04**	0.02	−0.03***	0.02
	(0.03)	(0.02)	(0.02)	(0.03)	(0.01)	(0.01)
Above med. ranked play $G-1$	−0.02	0.10***	0.02	0.13***	0.01	0.18***
	(0.03)	(0.01)	(0.01)	(0.02)	(0.01)	(0.01)
Above med. competitive play $G-1$	0.08***	0.08***	0.05***	0.02	0.04***	0.07***
	(0.03)	(0.01)	(0.01)	(0.02)	(0.004)	(0.01)
Above med. friendly play $G-1$	−0.003	−0.03***	−0.02**	−0.07***	−0.01***	−0.03***
	(0.02)	(0.01)	(0.01)	(0.01)	(0.003)	(0.01)
Above med. solo play $G-1$	−0.04	−0.12***	−0.02*	−0.11***	0.001	−0.15***
	(0.02)	(0.01)	(0.01)	(0.02)	(0.004)	(0.01)
Above med. last wk. played $G-1$	−0.02	−0.0002	0.01	0.02	0.0002	0.01**
	(0.01)	(0.01)	(0.01)	(0.01)	(0.003)	(0.01)
Above med. HHI $G-1$	−0.03*	−0.06***	0.02**	−0.08***	0.01***	0.01
	(0.02)	(0.02)	(0.01)	(0.03)	(0.004)	(0.01)
Above med. no. wks. played $G-1$	0.001	−0.03***	−0.01	−0.04**	−0.01***	−0.01
	(0.03)	(0.01)	(0.01)	(0.02)	(0.004)	(0.01)
Above med. high shares $G-1$	0.07***	0.06***	0.08***	0.12***	0.05***	0.12***
	(0.02)	(0.01)	(0.01)	(0.02)	(0.004)	(0.01)
Above med. max lvl played $G-1$	0.05**	0.06***	0.03**	0.04*	0.02***	0.11***
	(0.02)	(0.01)	(0.01)	(0.02)	(0.004)	(0.01)
Adopted during weekend $G-1$	−0.02	−0.01	−0.002	0.0002	0.0003	−0.01**
	(0.01)	(0.01)	(0.01)	(0.02)	(0.003)	(0.01)
Adopted 1st 2 weeks $G-1$	−0.01	0.02**	−0.01	0.005	−0.001	−0.01
	(0.02)	(0.01)	(0.01)	(0.01)	(0.003)	(0.01)
Observations	6,950	5,162	4,949	3,430	6,224	26,715
R^2	0.02	0.17	0.06	0.13	0.14	0.22
Note:						*p<0.1; **p<0.05; ***p<0.01

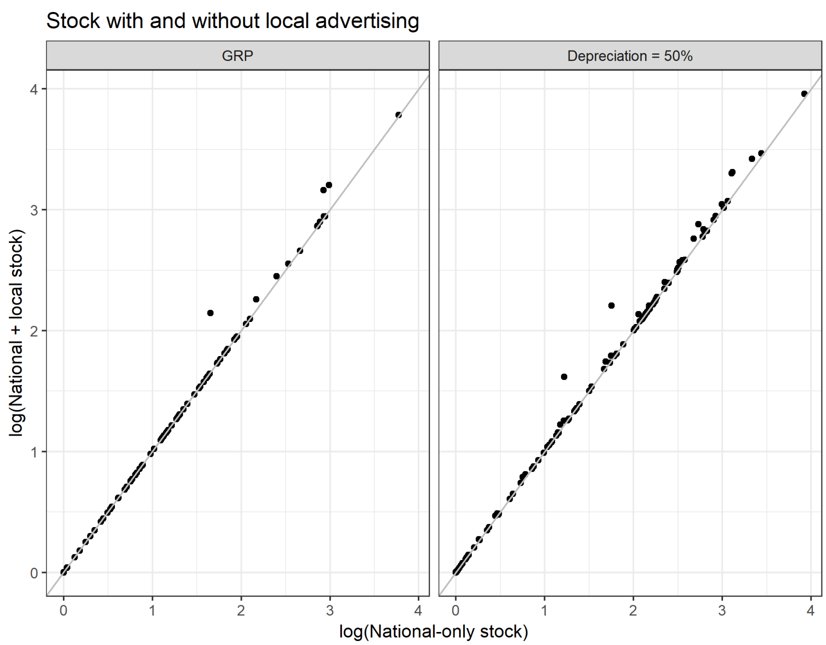
Table A.7: OLS of overall high shares G for each cluster and overall.

A.9 Advertising: evidence of geographic targeting

In Fig. A.10(a) I visualize cumulative gross rating points (GRP, which accounts for population size) over each DMA from 14 to 120 days post-release. For instance, the four most-targeted DMAs (New York, Los Angeles, Philadelphia, and Chicago) account for over a quarter of total local advertising exposure. In Fig. A.10(b) I show that the exclusion of local advertising does not qualitatively affect the measure of advertising (either flow or stock with 50% depreciation).



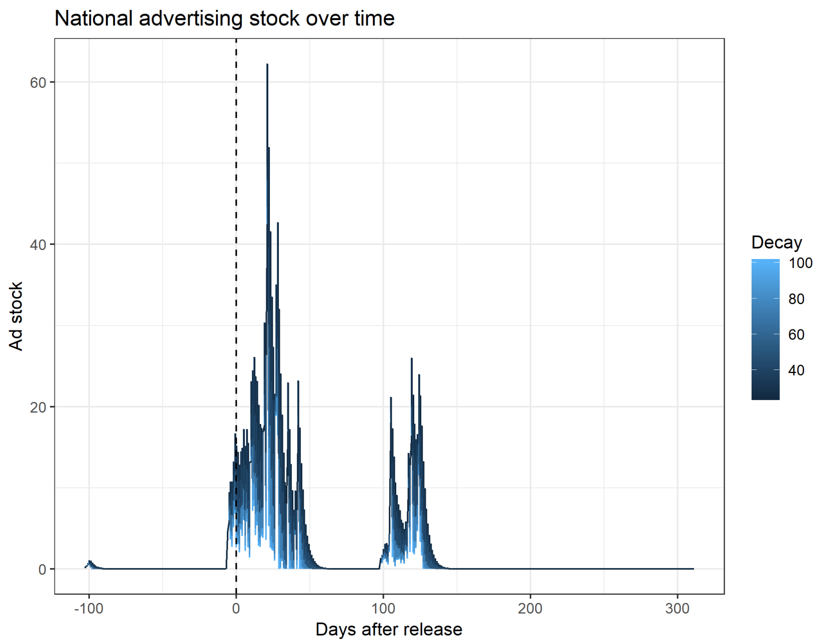
(a)



(b)

Figure A.10: Evidence of targeting in local ads. However, gross rating looks incredibly similar with and without inclusion of local ads.

A.10 Advertising: evolution of advertising stock, using different depreciation κ



(a)

Figure A.11: Advertising stock over time.

A.11 Regression of outcomes on console switching

Table A.8: Outcomes on console upgrade dummy and controls

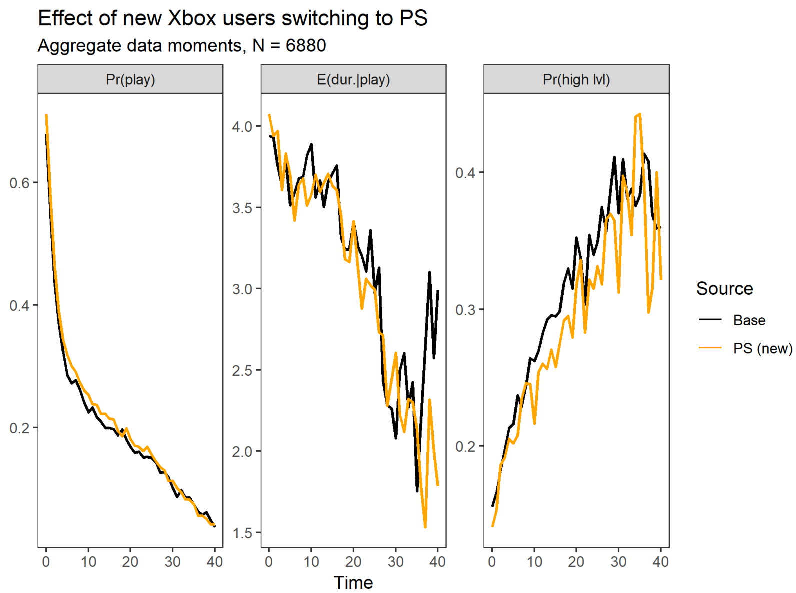
	Dependent variable:			
	log(Dur)	Pr(play)	Cond. dur	Shares (High)
	(1)	(2)	(3)	(4)
(Intercept)	0.880***	0.087***	1.715***	0.132***
	(0.193)	(0.024)	(0.280)	(0.030)
Upgraded console	0.497***	0.091***	0.222**	0.001
	(0.071)	(0.009)	(0.104)	(0.011)
log(Total play, $G-1$)	0.422***	0.026***	0.453***	−0.038***
	(0.055)	(0.007)	(0.080)	(0.009)
No. yrs in samp.	−0.037	−0.003	−0.014	−0.006
	(0.041)	(0.005)	(0.059)	(0.006)
No. games owned	0.129***	0.008	0.041	0.015**
	(0.046)	(0.006)	(0.067)	(0.007)
Adopt day, $G-1$	0.0004**	0.0001***	0.0004	0.0001**
	(0.0002)	(0.00002)	(0.0003)	(0.00003)
Adopted during weekend, $G-1$	−0.242***	−0.017*	−0.462***	−0.015
	(0.079)	(0.010)	(0.114)	(0.012)
Last wk. played, $G-1$	−0.004	−0.0001	−0.005	−0.001
	(0.004)	(0.0005)	(0.006)	(0.001)
Pct. wks. played, $G-1$	0.540**	0.235***	0.409	−0.015
	(0.259)	(0.032)	(0.376)	(0.040)
First 4 wks. play, $G-1$	0.007**	−0.0004	0.054***	0.001***
	(0.003)	(0.0004)	(0.005)	(0.001)
Overall high shares, $G-1$	0.238	0.042**	0.830***	0.501***
	(0.159)	(0.020)	(0.230)	(0.025)
Is above med. high lvls, $G-1$	−0.052	−0.008	−0.336*	0.117***
	(0.123)	(0.015)	(0.179)	(0.019)
Above med. last wk. played, $G-1$	0.142	0.018	−0.048	0.049**
	(0.125)	(0.016)	(0.181)	(0.019)
Above med. HHI, $G-1$	−0.277***	−0.029**	−0.634***	0.0003
	(0.104)	(0.013)	(0.150)	(0.016)
Above med. no. wks. played, $G-1$	0.014	0.057***	−0.547***	0.023
	(0.125)	(0.016)	(0.181)	(0.019)
Is high exp	0.171	0.013	0.366*	0.010
	(0.136)	(0.017)	(0.197)	(0.021)
R^2	0.258	0.260	0.249	0.338
Residual Std. Error (df = 2241)	1.623	0.203	2.357	0.252
F Statistic (df = 15; 2241)	52.071***	52.546***	49.629***	76.221***
Note:				*p<0.1; **p<0.05; ***p<0.01

Table A.10: Outcomes on PlayStation switch dummy and controls

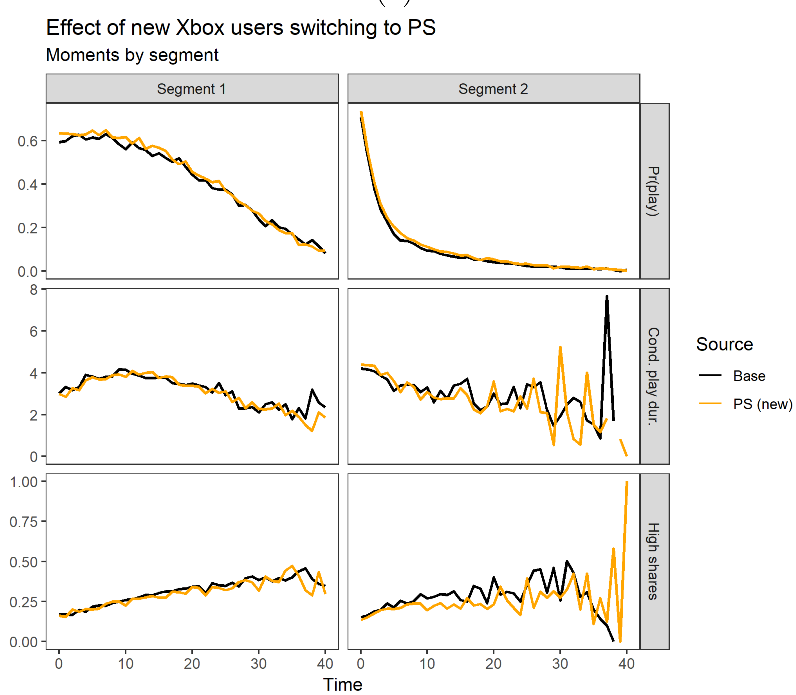
	Dependent variable:			
	log(Dur)	Pr(play)	Cond. dur	Shares (High)
	(1)	(2)	(3)	(4)
(Intercept)	0.852***	0.084***	1.314***	0.083***
	(0.106)	(0.015)	(0.176)	(0.017)
Xbox to PS	0.144	0.016	0.006	−0.013
	(0.101)	(0.014)	(0.168)	(0.016)
log(Total play, $G-1$)	0.450***	0.026***	0.616***	−0.012***
	(0.029)	(0.004)	(0.049)	(0.005)
No. yrs in samp.	−0.008	−0.011***	0.007	−0.011***
	(0.020)	(0.003)	(0.033)	(0.003)
No. games owned	0.105***	0.015***	0.075**	0.013***
	(0.022)	(0.003)	(0.037)	(0.004)
Adopt day, $G-1$	0.0005***	0.0001***	0.001***	0.0001***
	(0.0001)	(0.00002)	(0.0002)	(0.00002)
Adopted during weekend, $G-1$	−0.246***	−0.014**	−0.387***	−0.023***
	(0.043)	(0.006)	(0.072)	(0.007)
Last wk. played, $G-1$	0.003	0.0002	−0.002	−0.00003
	(0.002)	(0.0003)	(0.003)	(0.0003)
Pct. wks. played, $G-1$	0.779***	0.335***	0.590***	−0.014
	(0.124)	(0.017)	(0.207)	(0.020)
First 4 wks. play, $G-1$	0.001	−0.0003	0.021***	0.001***
	(0.001)	(0.0002)	(0.002)	(0.0002)
Overall high shares, $G-1$	−0.027	−0.008	0.157	0.550***
	(0.082)	(0.011)	(0.136)	(0.013)
Is above med. high lvls, $G-1$	0.036	0.011	0.219**	0.075***
	(0.059)	(0.008)	(0.099)	(0.010)
Above med. last wk. played, $G-1$	0.099*	0.006	−0.005	0.010
	(0.060)	(0.008)	(0.100)	(0.010)
Above med. HHI, $G-1$	−0.135**	−0.010	−0.543***	−0.0002
	(0.060)	(0.008)	(0.100)	(0.010)
Above med. no. wks. played, $G-1$	−0.006	0.018**	−0.694***	−0.006
	(0.061)	(0.008)	(0.102)	(0.010)
Is high exp	0.073	0.011	0.163	0.010
	(0.068)	(0.009)	(0.113)	(0.011)
R^2	0.344	0.304	0.228	0.375
Residual Std. Error (df = 7058)	1.467	0.202	2.446	0.236
F Statistic (df = 15; 7058)	246.681***	205.199***	139.141***	281.794***

Note: *p<0.1; **p<0.05; ***p<0.01

A.12 Counterfactual simulation visuals for New player console switching

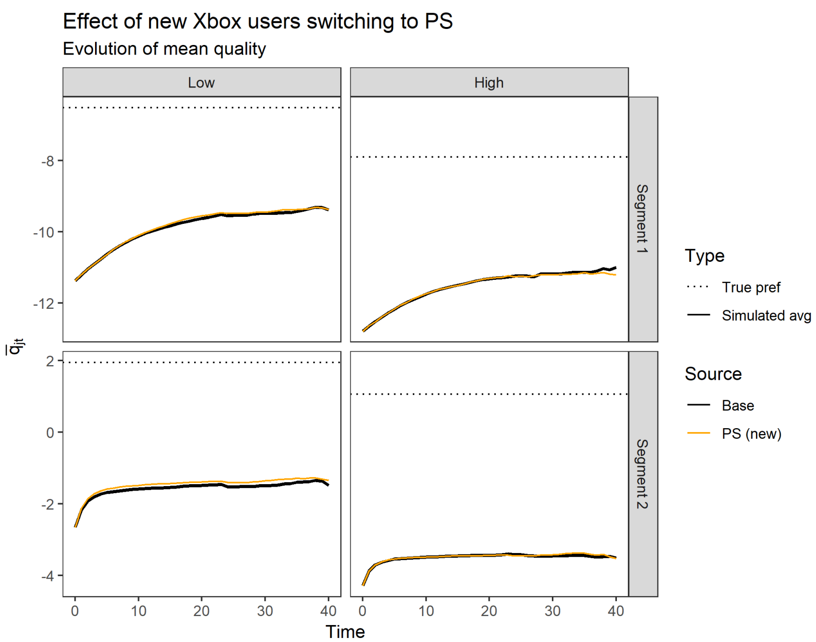


(a)

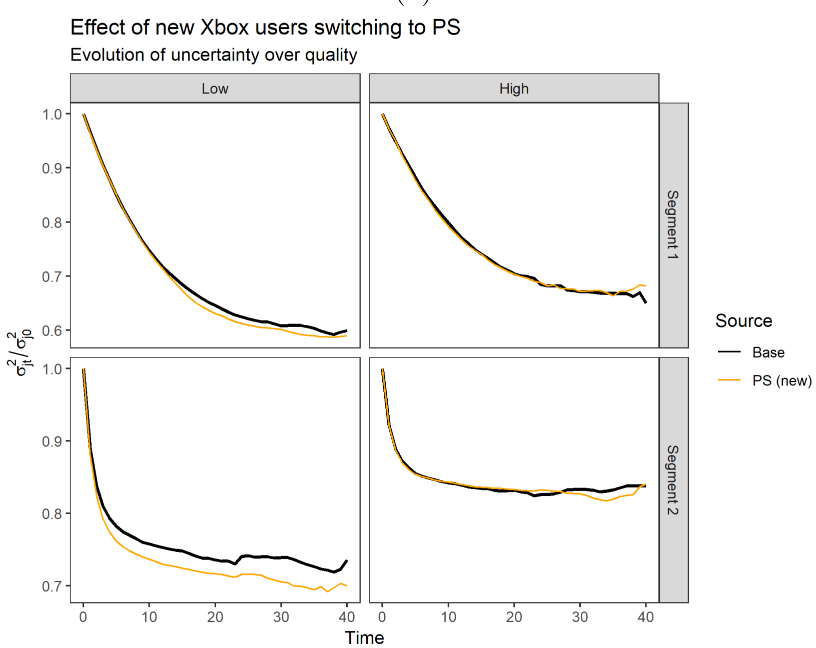


(b)

Figure A.12: The PlayStation counterfactual for new players is qualitatively similar as for returning players, with a net effect of 9.2%.

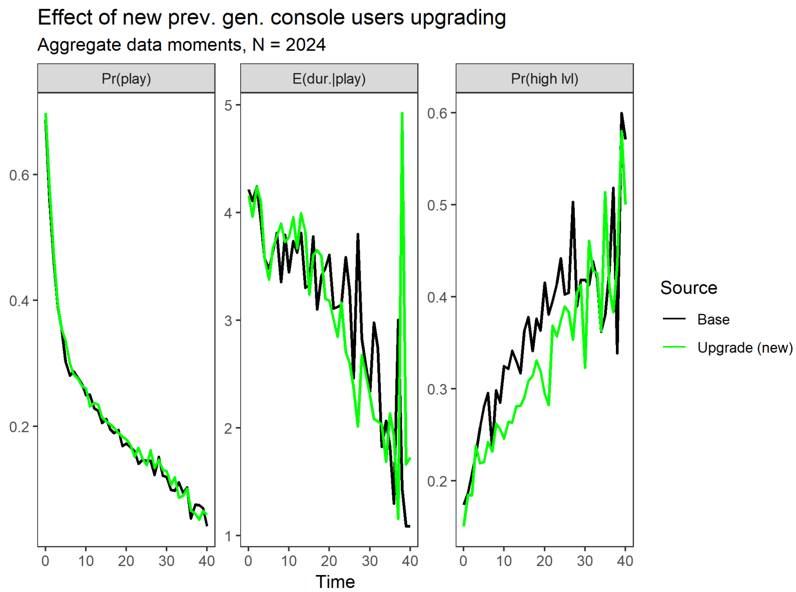


(a)

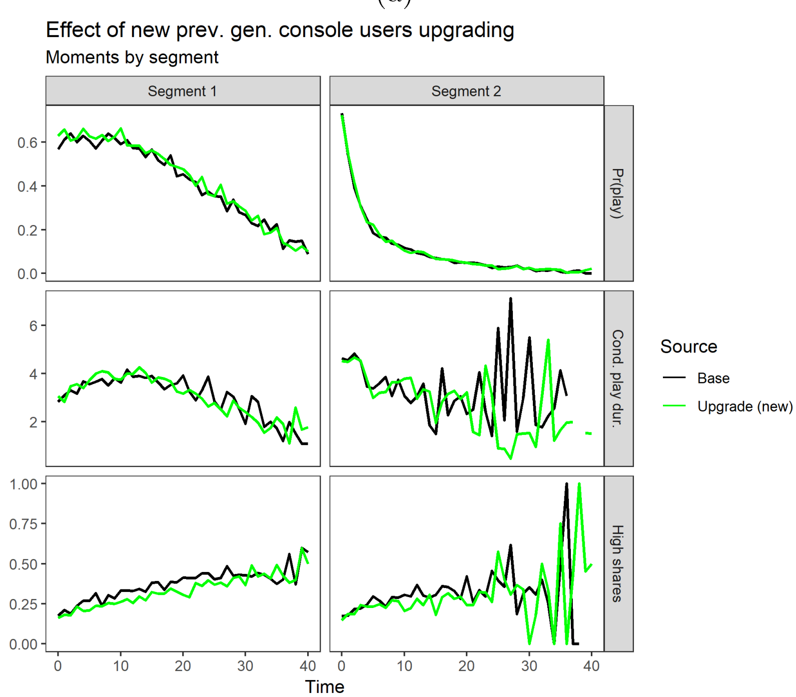


(b)

Figure A.13: Learning curves appear similar. Note that uncertainty resolution is broadly worse for new players (vs returning).

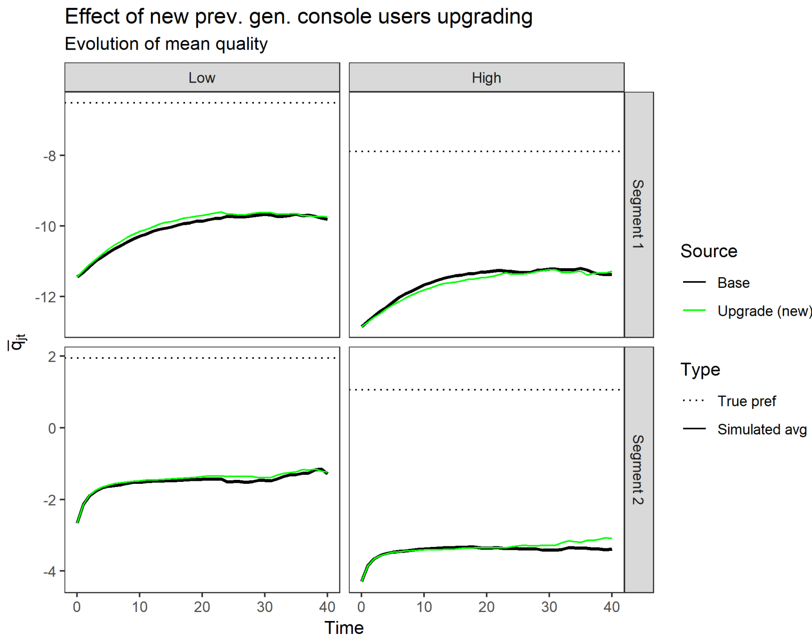


(a)

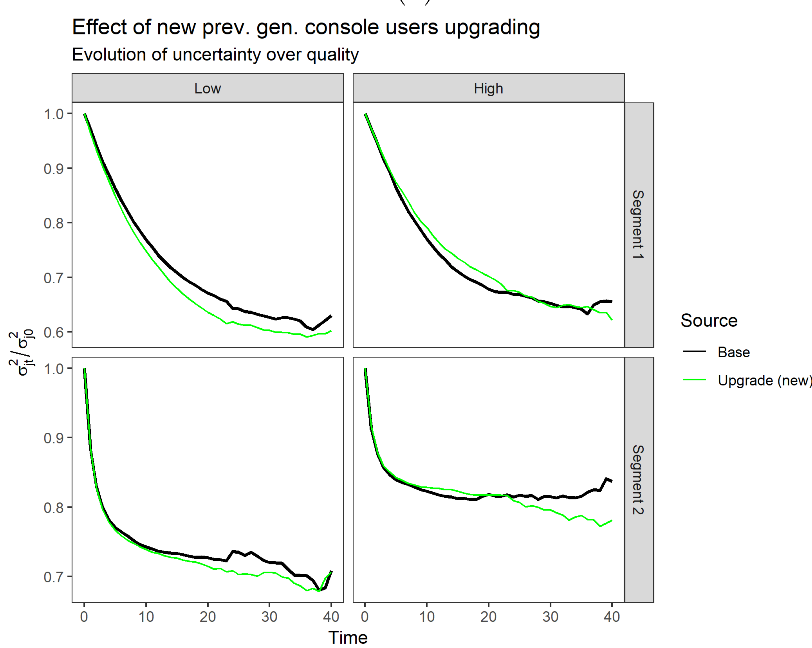


(b)

Figure A.14: With a relatively small sample, results are noisy. Note that Segment 2 players are actually predicted to play less! Net effect on play is around 1%.



(a)



(b)

Figure A.15: Observe the noisiness in learning curves, especially near the end of sample. Empirically retention of New players (especially Segment 2) is low.

A.13 Regressions of adoption/in-game purchases on play

The following tables contain only first-order effects. They may be more useful to get a sense of the size of effects other than the primary three discussed in the main article.

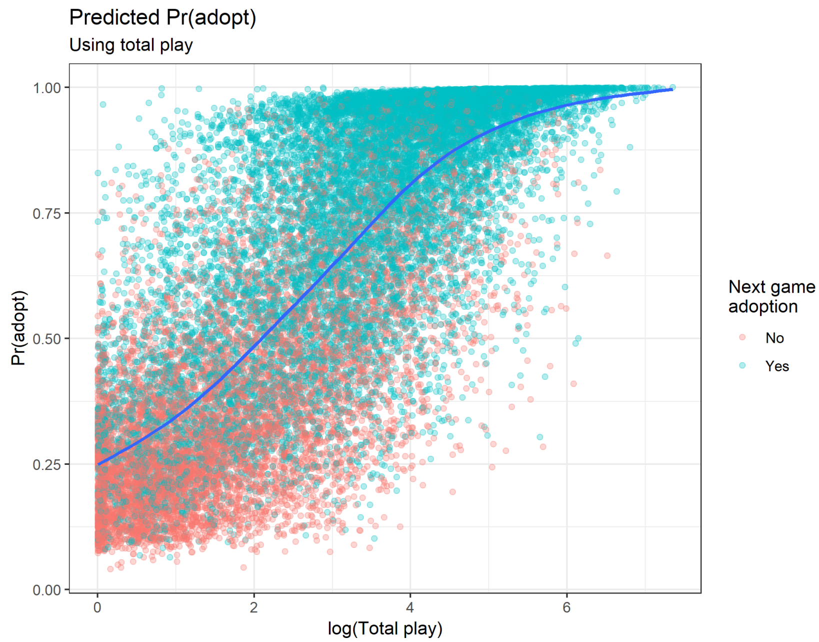
	Dependent variable:	
	Game G adoption	
	(1)	(2)
(Intercept)	-2.731^{***} (0.184)	-3.185^{***} (0.212)
log(Total play)	0.470^{***} (0.060)	
log(Low lvl. play)		0.311^{***} (0.049)
log(High lvl. play)		0.055 (0.036)
New player	0.008 (0.061)	0.013 (0.061)
Num. prev. games	0.638^{***} (0.030)	0.639^{***} (0.030)
Adoption time	0.0002 (0.0002)	0.0002 (0.0002)
Is weekend adoption	-0.130^{***} (0.040)	-0.125^{***} (0.040)
PlayStation	0.007 (0.035)	0.012 (0.035)
Curr. gen. console	1.175^{***} (0.039)	1.170^{***} (0.039)
Num. sessions	0.003^{**} (0.001)	0.003^{**} (0.001)
log(Weekend play)	-0.052 (0.048)	0.028 (0.046)
Avg. session length	-0.549^{***} (0.084)	-0.437^{***} (0.081)
Time from purch. to 1st play	0.004^{***} (0.001)	0.003^{***} (0.001)
Time from 1st to last play	0.003^{***} (0.0003)	0.003^{***} (0.0003)
High lvl. shares	-0.223^{***} (0.078)	0.360^{***} (0.135)
HHI shares	-0.196 (0.148)	0.297^{*} (0.178)
log(First month play)	-0.048 (0.040)	0.018 (0.037)
log(Last month play)	0.316^{***} (0.030)	0.319^{***} (0.030)
Avg. date played	0.00002 (0.001)	0.001 (0.001)
% of time played on wkend	0.140 (0.090)	0.027 (0.088)
Played any High lvl.	0.177 (0.171)	-0.109 (0.173)
Played any mult. lvls.	-0.083 (0.169)	0.212 (0.170)
Observations	22,181	22,181
Log Likelihood	$-10{,}367.340$	$-10{,}377.620$
Akaike Inf. Crit.	20,776.670	20,799.230
Note:		*p<0.1; **p<0.05; ***p<0.01

Table A.12: Adoption vs play, main effects only.

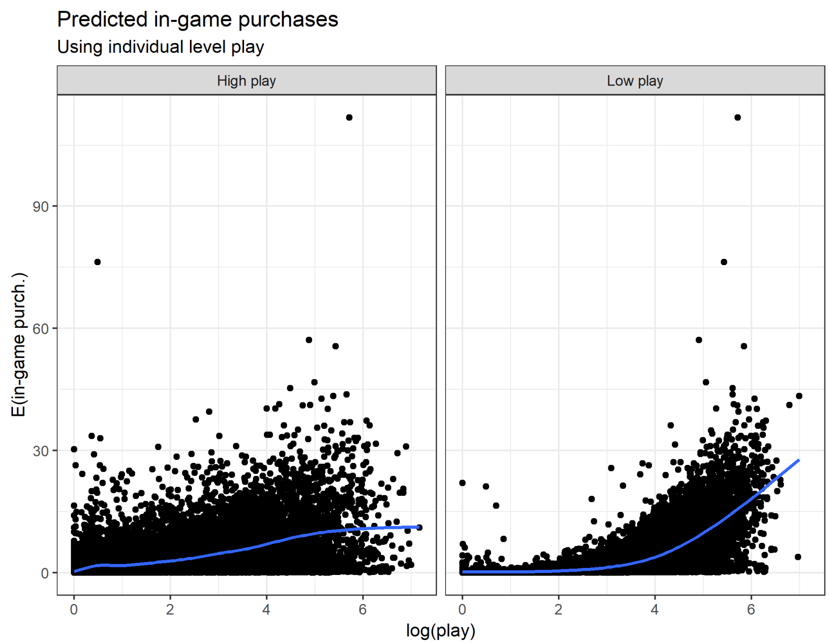
In Figs. A.13 I present predictions from the second-order interactions models. The adoption model has an AUC of 0.85.

	Dependent variable:	
	In-game purchases	
	(1)	(2)
(Intercept)	-0.886^{**} (0.373)	-2.704^{***} (0.481)
log(Total play)	0.846^{***} (0.155)	
log(Low lvl. play)		0.976^{***} (0.124)
log(High lvl. play)		0.093^{*} (0.050)
New player	-0.407^{***} (0.085)	-0.403^{***} (0.082)
Num. prev. games	-0.046 (0.028)	-0.045 (0.028)
Adoption time	-0.002^{***} (0.001)	-0.002^{***} (0.001)
Is weekend adoption	-0.051 (0.057)	-0.049 (0.055)
PlayStation	-0.255^{***} (0.044)	-0.251^{***} (0.043)
Curr. gen. console	0.315^{***} (0.066)	0.309^{***} (0.064)
Num. sessions	-0.001 (0.001)	-0.002^{**} (0.001)
log(Weekend play)	0.477^{***} (0.150)	0.352^{***} (0.126)
Avg. session length	-1.322^{***} (0.148)	-1.421^{***} (0.146)
Time from purch. to 1st play	0.010^{***} (0.001)	0.010^{***} (0.001)
Time from 1st to last play	0.00004 (0.0004)	-0.00003 (0.0004)
High lvl. shares	-1.371^{***} (0.110)	0.962^{***} (0.337)
HHI shares	-1.777^{***} (0.163)	-0.091 (0.304)
log(First month play)	-0.081^{**} (0.041)	-0.084^{**} (0.039)
log(Last month play)	-0.088^{***} (0.025)	-0.084^{***} (0.024)
Avg. date played	-0.006^{***} (0.001)	-0.006^{***} (0.001)
% of time played on wkend	-0.845^{*} (0.441)	-0.489 (0.378)
Played any High lvl.	1.716^{**} (0.758)	1.231 (0.749)
Played any mult. lvls.	-0.418 (0.753)	0.052 (0.745)
Observations	21,960	21,960
Note:		*p<0.1; **p<0.05; ***p<0.01

Table A.13: In-game purchases vs. play, main effects only (for purchases ≤ 90).



(a)



(b)

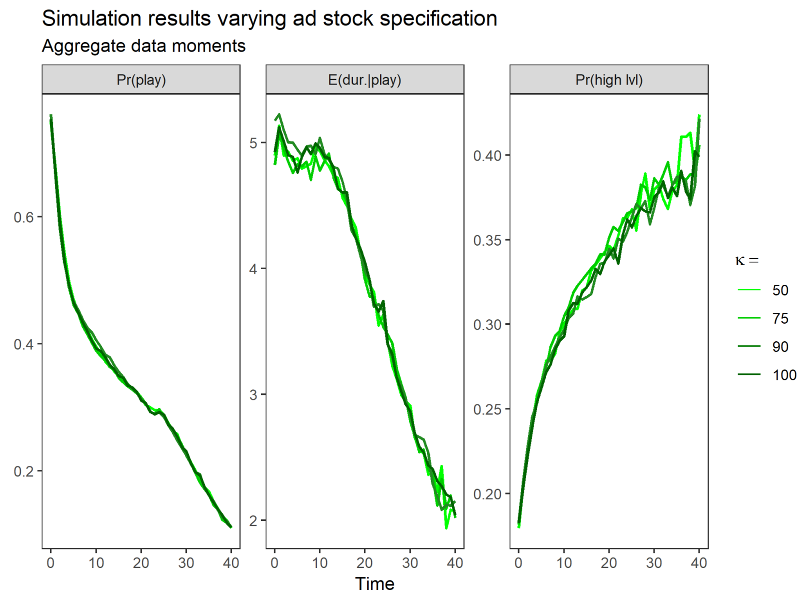
Figure A.16: Predicted purchasing behavior from the second-order interaction models. Note that the relationship between High play and in-game purchases is stronger than for Low play.

A.14 Varying ad stock specifications

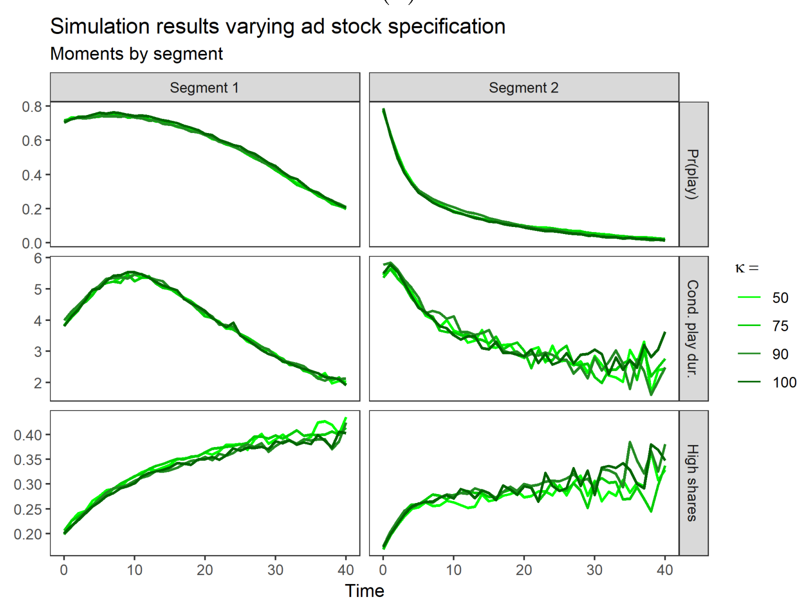
In Fig. A.17 I present simulated data moments using different capital depreciation rates κ (denoted as percentage). Recall the advertising stock is defined $S_t^A(\kappa) = \kappa A_t + (1-\kappa)S_{t-1}^A$. The implied moments are rather indistinguishable. I believe the specific advertising stock formulation matters far less than the fact that in general periods of high advertising and low advertising are identified. In Table A.14 I present the advertising coefficient estimates alongside the model likelihood. I find that model log-likelihood for $\kappa = 0.75$ is 0.016% better than complete decay. However, I do not believe that this finding warrants consideration of separate estimation of the advertising stock depreciation.

	$\kappa =$			
Ad coef.	100	90	75	50
Low level, Segment 1	2.18	2.00	3.01	3.37
High level, Segment 1	0.68	0.56	1.04	1.20
Low level, Segment 2	3.57	2.32	4.96	5.67
High level, Segment 2	0.58	0.21	0.84	1.12
-LL	1223833	1224267	1223620	1223804

Table A.14: Ad stock coefficients by κ.



(a)



(b)

Figure A.17: Predicted data moments are very similar under various advertising specifications.

APPENDIX B

THEORY

B.1 Derivation of closed-form posterior distribution for Bayesian learning specification

Here I derive the posterior given in Eq. (4.3). In particular, I consider the following:

Let the prior mean be distributed Gaussian, i.e. $p(\mu) \propto \exp\left[-\frac{1}{2}(\mu-\mu_0)'\Sigma^{-1}(\mu-\mu_0)\right]$ where μ is $J \times 1$ and Σ is assumed known. Assume we then observe $K = k_1 + \ldots k_J$ draws R with k_j coming from mean element μ_j. Again, we have the $J \times K$ collapsing matrix M defined as block diagonal composed of ones row vectors of sizes $k_1, \ldots, k_J$, such that $\mathbb{E}[R] = M'\mu$. Here we do not impose any distributional structure on R besides Gaussianity, with $K \times K$ covariance, say Σ_R. Then we have:

$$\begin{aligned} p(\mu)p(R|\mu) &\propto \exp\left\{-\frac{1}{2}\left[(\mu-\mu_0)'\Sigma^{-1}(\mu-\mu_0) + (R - M'\mu)'\Sigma_R^{-1}(R - M'\mu)\right]\right\} \\ &= \exp\left\{-\frac{1}{2}\left[\mu'(\Sigma^{-1} + M\Sigma_R^{-1}M')\mu - 2\mu'(\Sigma^{-1}\mu_0 + M\Sigma_R^{-1}R) + \text{const.}\right]\right\} \end{aligned}$$

Letting $\Sigma_{\text{post}} = (\Sigma^{-1} + M\Sigma_R^{-1}M')^{-1}$ we complete the square as usual:

$$p(\mu|R) \propto \exp\left\{-\frac{1}{2}\left[(\mu - \Sigma_{\text{post}}(\Sigma^{-1}\mu_0 + M\Sigma_R^{-1}R))'\Sigma_{\text{post}}^{-1}(\mu - \Sigma_{\text{post}}(\Sigma^{-1}\mu_0 + M\Sigma_R^{-1}R)) + \text{const.}\right]\right\}$$

Thus, we see that the posterior $\mu|R$ is distributed normal with:

$$\begin{aligned} \mathbb{V}[\mu|R] &= (\Sigma^{-1} + M\Sigma_R^{-1}M')^{-1} \\ \mathbb{E}[\mu|R] &= \mathbb{V}[\mu|R] \cdot (\Sigma^{-1}\mu_0 + M\Sigma_R^{-1}R) \end{aligned}$$

In the usual case when $k_1 = \ldots = k_J = 1$ (i.e. one signal associated with each μ element), M is simply $I_{J\times J}$ and the expressions above collapse into the standard Bayesian conjugacy for Gaussian distributions. Note that when $k_j > 0 \forall j$, there is at least one signal associated with each element of μ so rank$(M) = J$ and $M\Sigma_R^{-1}M'$ will be positive definite, meaning the posterior covariance will be PD (i.e. valid covariance). In the case that some $k_j = 0$ we have rank$(M) < J$ and $M\Sigma_R^{-1}M'$ will no longer be invertible. In particular, the rows and columns associated with the unobserved signals will all be zero, while the sub-matrix associated with the observed signals remains positive definite. Thus, the sum $\Sigma^{-1} + M\Sigma_R^{-1}M'$ remains invertible.

B.2 Alternate learning specifications

Here I discuss alternate learning specifications that incorporate correlated learning:

First, there is the model in Coscelli and Shum (2004), who study the spillover effects of omeprazole prescriptions for different patients and diagnoses. The authors use the same prior covariance I do (diagonal), and a one-factor variance components structure on signals. In particular, they define:

$$R_{jts} = \bar{q}_j + \rho_j \theta_t + \nu_{jts}$$

where $\theta_t \sim N(0, \sigma_\theta^2)$ are i.i.d. over t, $\nu_{jts} \sim N(0, \sigma_{\nu_j}^2)$ are i.i.d. over t, j, s. This induces the following correlation structure:

$$\begin{aligned}
&\text{Var}(R_{jts}) = \rho_j^2 \sigma_\theta^2 + \sigma_{\nu_j}^2 \\
&\text{Cov}(R_{jts}, R_{jts'}) = \rho_j^2 \sigma_\theta^2, \quad s \neq s' \\
&\text{Cov}(R_{jts}, R_{j'ts'}) = \rho_j \rho_j j' \sigma_\theta^2, \quad j \neq j'
\end{aligned}$$

This model allows for a somewhat flexible correlation structure while restricting the number of covariance parameters to $2j + 1$. Note that my current model uses a specialization of this model restricting $\rho_j = 0 \ \forall j$. The major drawback to this model is that updates on level j only show up in Σ_{Rt} if they are selected, i.e. $k_{jt} > 0$. This implies that no correlated learning occurs unless a level is selected.

Second, there is the model in Ching and Lim (2016), who study how doctors learn about the general efficacy of statins (category of drugs) from landmarks trials of individual statins.

In this paper, the authors instead induce correlated learning by assuming first-period priors are correlated while signals are uncorrelated. Recycling notation from above, they let Σ_0 be a general dense $J \times J$ matrix while $R_{jts} = \bar{q}_j + \nu_{jts}$ where $\nu_{jts} \sim N(0, \sigma^2_{\nu_j})$. In the context:

$$\begin{aligned}
&\text{Var}(R_{jts}) = \sigma^2_{\nu_j} \\
&\text{Cov}(R_{jts}, R_{j'ts'}) = 0, \quad s \neq s'
\end{aligned}$$

The major drawback of this model is that an dense covariance matrix consisting of $J(J+1)/2$ terms must be estimated for the priors. In their empirical application $J = 3$ so this was not a problem.

Note that updating for both the Coscelli and Shum (2004) and the Ching and Lim (2016) models falls under the general structure I describe in the Model structure. Next I consider models based on a missing data approach from Dominici and Parmigiani (2000). Here, consider the $J \times 1$ latent signal $R^{\dagger}_{ts} = \bar{q} + u_{ts}$, $u_{ts} \sim N(0, \Sigma_u)$ where Σ_u is a $J \times J$ potentially dense matrix. In this formulation, we can either think of a single play signal as a censored draw from this distribution (e.g. a play session from level 1 is a draw from $R^{\dagger}_{ts}$ with only the first element observed) or consider play signals jointly. The first approach is simply a generalization of the Ching and Lim (2016) model. I focus on the second below:

As a motivating example, consider the case $J = 3$ where a user plays the levels $3, 5, 6$ times (say starting from level 1 up). Here $K_t = 14$ but she is considered to have observed $\max(k_{jt}) = 6$ signals. In particular, the first 3 signals will consist of fully observed realizations of $R^{\dagger}_t$; the next 2 will be $R^{\dagger}_t$ with the first element censored; the last 1 will be $R^{\dagger}_t$ with both the first and second elements censored. We can then define the censoring matrices:

$$C_{1t} = I_3, \quad C_{2t} = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \quad C_{3t} = \begin{bmatrix} 0 & 0 & 1 \end{bmatrix}$$

Note that each matrix maps from the full J dimensions to the corresponding observed subset. In general there are $2^J - 1$ possible permutations of observed subsets but due to the ordered nature in each time period, there are only $J - 1$ possible censoring matrices. Fixing t, the marginal distribution of signals from each censoring group will be $N_{c_g}(C_g \bar{q}, C_g \Sigma_u C_g^T), g = 1, ..., G \leq J - 1$. Letting priors be $N(q_t, \Sigma_t)$, Dominici and Parmigiani (2000) shows that the posterior will be distributed $N(q_{t+1}, \Sigma_{t+1})$ with:

$$\Sigma_{t+1}^{-1} = \Sigma_t^{-1} + \sum_g n_g (C_g \Sigma_u C_g^T)^{-1}$$
$$q_{t+1} = \Sigma_{t+1}(\Sigma_t^{-1} q_t + \sum_g \sum_{s=1}^{n_g} (C_g \Sigma_u C_g^T)^{-1} R^{\dagger}_{ts,\ censored})$$

where n_g is the number of signals in each group (3,2,1 respectively in this example), and $R^{\dagger}_{ts,\text{censored}}$ is the censored signal. Note that unlike Coscelli and Shum (2004), this model will allow learning for levels that are unexplored and unlike Ching and Lim (2016) we do not have to assume correlated learning is completely driven by a priori beliefs. Further, Dominici and Parmigiani (2000) notes that $(\Sigma_u)_{j,j'}$ is identified by observations where levels j, j' are jointly observed. In my empirical application, all $2^J - 1$ permutations of censoring groups are observed, with the least common group containing at least 1,000 observations.

Finally, I consider the model in Chintagunta et al. (2009). In this paper, authors study whether media and network effects affect joint learning about doctor beliefs over Cox-2 inhibitors. To induce correlated learning, they additively decompose the signal into a common quality term and an idiosyncratic one. In other words, they let $R_{jts} = \bar{Q}_0 + [\bar{q}_j - \bar{Q}_0] + \nu_{jts}$. Consumers must then have priors over $(\bar{Q}_0, [\bar{q}_j - \bar{Q}_0])$. The authors show that the posterior precision has form:

$$\Sigma_{it+1}^{-1} = \begin{bmatrix} s & a_1 & \dots & \dots & a_J \\ a_1 & m_1 & & & \\ \vdots & & \ddots & & 0 \\ \vdots & 0 & & \ddots & \\ a_J & & & & m_J \end{bmatrix}$$

where:

$$s = \sum_j \frac{n_{jt}^R}{\sigma_{\nu_j}^2} + \frac{1}{\sigma_{Q_0}^2}$$

$$a_j = \frac{n_{jt}^R}{\sigma_{\nu_j}^2}$$

$$m_j = \frac{n_{jt}^R}{\sigma_{\nu_j}^2} + \frac{1}{\sigma_{q0}^2}$$

where n_{jt}^R is the number of total signals for level j up to and including time t, $(\sigma_{Q_0}^2, \sigma_{q0}^2)$ are prior variances for the common and idiosyncratic components, respectively, and $\sigma_{\nu_j}^2$ is the signal variance. The authors further provide an analytic form for the posterior variance and show it is fully dense, implying fully correlated learning between levels. Crucially, note that the authors observe both the common and idiosyncratic signals in their empirical application. Further, they apply correlated learning over patients, so there are J Σ_{jt}'s that grow in size over time. In contrast, I apply correlated learning over levels, so there would be N Σ_{jt}'s that remain static over time.

For a general review of learning models in marketing (including models of correlated learning), please see Ching et al. (2013).

B.3 Non-identification of the γ-model in the presence of learning

Here I present simulation-based evidence that the γ-model is not identified in the presence of Bayesian learning. In Fig. B.1 I plot the negative log-likelihood function using the same simulation parameters as my Main simulation except with $J = 1$. The true values are given by the point in red ($\gamma = 1, \sigma_\nu^2 = 4.5$). Note that in this specification, the likelihood is roughly flat for signal variance at any given γ value (except as σ_ν^2 tends towards 0, implying perfect knowledge). As a result, signal variance is not identified here. Empirically, I find that γ also tends to be difficult to pin down reliably.

(Negative) log-likelihood function with respect to γ and σ_ν^2, $J = 1$

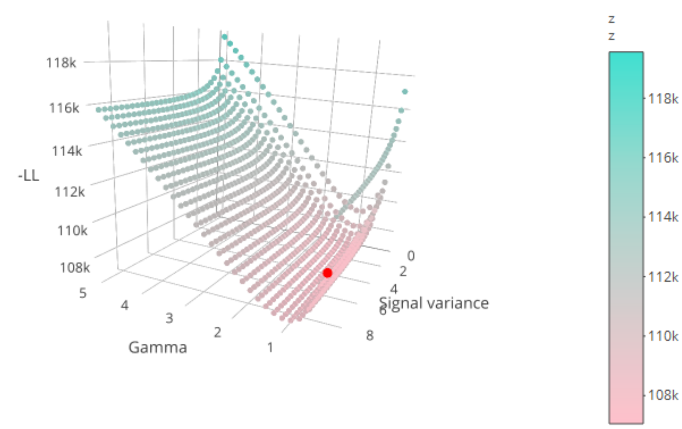


Figure B.1: Note that at any given γ value, the likelihood is flat with respect to signal variance except at very small values.

Next in Fig. B.2 I give the log-likelihood plane when $J = 3$, which corresponds to the number of inside goods in my simulation study. Again, note that the plane is almost flat with respect to signal variance at any given value of the translation parameter γ.

(Negative) log-likelihood function with respect to γ_1 and $\sigma^2_{\nu_1}$, $J = 3$

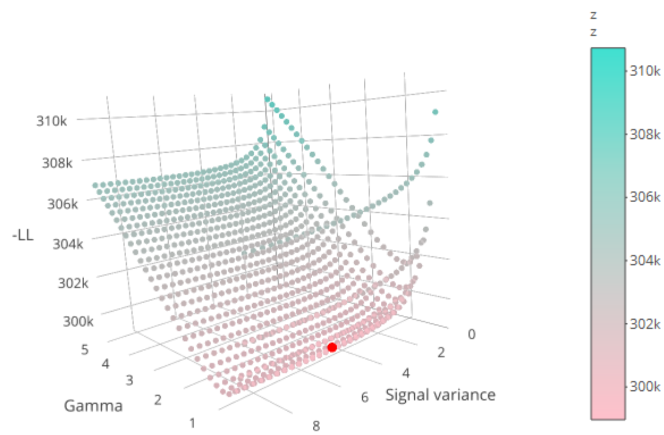


Figure B.2: Note that again at any given γ value, the likelihood is flat with respect to signal variance.

B.4 Additional properties of the SDCEV model

Here I consider some additional properties of the MDCEV model without learning, which in my application corresponds to the first observation from each user[1]. I focus on early adopters, and present first week statistics by group in Table B.1. Due to the heavy right skew in the data, note that the standard deviation of conditional play durations is about the same as the conditional mean. Here I investigate the mapping from MDCEV parameters to consumption distributions and conversely, what the empirical consumption patterns imply about the underlying parameters.

Adoption	Count	% played	E(duration|played)	SD(duration|played)	% playing high shares|played
0	10068	0.97	7.54	7.08	0.20
1	2125	0.95	4.43	4.90	0.16
2	664	0.94	3.88	4.76	0.16
3	512	0.89	3.12	3.93	0.17
4	396	0.93	3.54	4.33	0.16

Table B.1: First week statistics for early adopters, by week of adoption.

I reduce the problem to single discrete-continuous (SDC) by aggregating consumption of each level into overall 'inside' consumption (x_1) and 'outside consumption (x_0) (relative to a budget E). In that case, a consumer solves:

$$\max_{x_0+x_1\leq E} \frac{1}{\alpha_0} e^{\epsilon_0} x_0^{\alpha_0} + \frac{1}{\alpha_1} e^{Z_1\beta+\mu-r\sigma^2+\epsilon_1}[(x_1+1)^{\alpha_1}-1] \tag{B.1}$$

Recall $\epsilon_0, \epsilon_1 \overset{\text{iid}}{\sim}$ standard Gumbel and (α_0, α_1) are satiation parameters bounded from above by 1[2]. This form assumes some quantity of the outside good is consumed w.p. 1. I further impose homogeneity in consumers, so that the baseline log-MU is constant, say $\kappa = Z_1\beta + \mu - r\sigma^2$. Then the log FOC of this problem can be expressed:

1. Empirically, over 95% of users do record a session the week they adopt.

2. Satiation of 1 implies linear utility (no satiation), of 0 implies log utility (CRRA), and negative implies very high satiation.

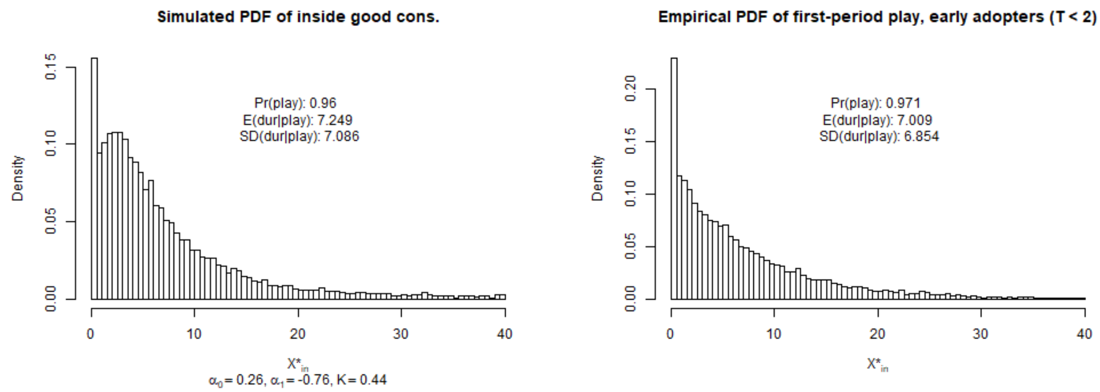
$$g(x_1^*, \kappa, \alpha_1, \alpha_0; E) \equiv \kappa + (\epsilon_1 - \epsilon_0) + (\alpha_0 - 1)\log(E - x_1^*) + (\alpha_1 - 1)\log(x_1^* + 1) \leq 0 \quad \text{(B.2)}$$

with equality iff $x_1^* > 0$ and inequality otherwise. Note then that the sum of the first two terms are distributed logistic with location κ and scale 1, call this random variable ν. Under homogeneity, the distribution of x_1^* is fully determined by ν. Further, it is simple to see that the distribution of x_1^* is monotonic increasing in κ, which can be interpreted as log ratio of MU's of inside vs outside good at the point of zero consumption, or a measure of the **relative preference for inside consumption**. The expansion path of x_1^* as a function of κ is governed by the satiation parameters (α_1, α_0).

To the first point, in Fig. B.3 I present a comparison of the simulated vs. empirical distribution of first-week play (early adopters). To construct the simulated distribution I note that it is parameterized by three values: $\alpha_0, \alpha_1, \kappa$. Then using $Pr(x_1^* > 0), \mathbb{E}[x_1^*|x_1^* > 0], \mathbb{V}[x_1^*|x_1^* > 0]$ as moments, I use SMM to fit the parameters to the observed data. Note that while I can capture the raw moments, the homogeneity assumption is overly restrictive in fitting the data's shape. In particular, while the empirical distribution resembles a zero-inflated exponential, the pure SDCEV has a "hump" in its strict positive domain. Further, the implied parameters here include $\alpha_1 < 0$, signifying very high inside satiation.

Besides this limitation, I note that the estimated results are also sensitive to the choice of budget scaling, E[3]. In Fig. B.4 I show how the estimated parameters vary as I scale E from 1 to 100 (baseline 40). The two primary trends are that $\hat{\alpha}_1$ increase with scale and that estimates become extreme, possibly unstable, at very small budget scales.

3. Lee et al. (2018) do recover the budget parameter in a Bayesian framework, but I am unable to estimate the budget parameter in my frequentist framework. Whether it is the Bayesian structure or the omission of learning that allows one to estimate budgets is a question for future work.



(a) (b)

Figure B.3: Simulated vs empirical distribution of first week play, X_1^*.

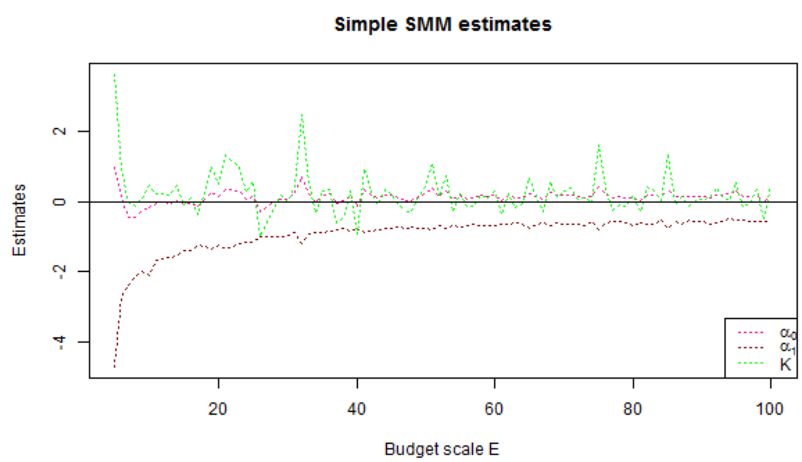


Figure B.4: SMM estimates of $(\alpha_0, \alpha_1, \kappa)$ varying the data scale

This relates to a second point: the expansion path of x_1^* wrt κ exhibits aberrant behavior when the budget is sufficiently small. In Fig. B.5 I plot the expansion path holding one $\alpha = 0.01$ (close to log-utility), and present curves varying the other α from -5 to 0.95. The LHS panels hold α_1 and vary α_0 while the RHS panels do the reverse. Note that for budget scales greater than 1, there exists a crossover point in the LHS curves. Below the crossover point, consumers behave rationally: increasing outside satiation decreases outside consumption. However, above this crossover, increasing outside satiation leads to **increasing** outside consumption–this is clearly an anomalous consumption pattern. This same pattern is not replicated in the RHS panels: increasing outside satiation always leads to decreasing outside satiation.

Analytically, the regime change is defined by $\mathrm{sgn}(\frac{\partial x_1^*}{\partial \alpha_0})$, i.e. $\frac{\partial x_1^*}{\partial \alpha_0} < 0$ represents "normal" behavior while $\frac{\partial x_1^*}{\partial \alpha_0} > 0$ is aberrant. Thus, given budget scale E the cross over point is $\{x_c : \frac{\partial x_1^*}{\partial \alpha_0}|_{x_1^*=x_c} = 0\}$. Assuming for now that the crossover point exists and is greater than 0, we can assume Eq. (B.2) holds with equality. Taking partials:

$$\begin{aligned}
\frac{\partial g}{\partial x_1^*} &= \frac{\alpha_1 - 1}{x_1^* + 1} + \frac{\alpha_0 - 1}{E - x_1^*} < 0 \\
\frac{\partial g}{\partial \alpha_0} &= -\log(E - x_1^*) \\
\frac{\partial g}{\partial \alpha_1} &= \log(x_1^* + 1) > 0 \\
\frac{\partial g}{\partial \kappa} &= 1
\end{aligned}$$

From this, we can see the following:

$$\begin{aligned}
\frac{\partial x_1^*}{\partial \kappa} &= -\left(\frac{\alpha_1 - 1}{x_1^* + 1} + \frac{\alpha_0 - 1}{E - x_1^*}\right)^{-1} > 0 \\
\frac{\partial x_1^*}{\partial \alpha_1} &= \frac{-\log(x_1^* + 1)}{\frac{\alpha_1 - 1}{x_1^* + 1} + \frac{\alpha_0 - 1}{E - x_1^*}} > 0 \\
\frac{\partial x_1^*}{\partial \alpha_0} &= \frac{\log(E - x_1^*)}{\frac{\alpha_1 - 1}{x_1^* + 1} + \frac{\alpha_0 - 1}{E - x_1^*}} \gtrless 0
\end{aligned} \tag{B.3}$$

The first two partials in Eq. (B.3) are intuitively correct but note the third changes sign at $\log(E - x_1^*) = 0 \Leftrightarrow x_0^* = 1$. In other words, while outside consumption is greater than one unit, consumers behave as we expect. Once outside consumption dips below that (i.e. inside consumption increases to "near" budget) that is when the econometric issues arise.

Thus, the so-called crossover point has an extremely simple formula:

$$x_{\text{crossover}} = E - 1 \tag{B.4}$$

I have shown that under this particular formulation of the SDCEV model, in which an outside good is assigned and consumed w.p. 1, scale invariance does not hold. Distortions are, however, small unless the budget scale is also small. In the special case where $E = 1$, i.e. the problem is rescaled in terms of proportion of total budget, the distortions are severe.

B.4.1 Closed-form characterization of inside consumption

Here I continue analysis using the setup above. Again, I emphasize that the SDCEV model is completely characterized by the complementarity condition:

$$g \cdot x_1^* = 0$$

From this we can derive $Pr(x_1^* > 0)$ as follows:

$$\begin{aligned} Pr(x_1^* = 0) &= Pr(g < 0 \forall x) \\ &= Pr(\max_{x \in (0,E)} g(x) < 0) \\ &= Pr(\lim_{x \downarrow 0} g(x) < 0) \\ &= Pr(\kappa + \nu - (\alpha_0 - 1)\log(E) < 0) \\ &= Pr(\nu < (\alpha_0 - 1)\log(E) - \kappa) \\ &= [1 + \exp(\kappa - (\alpha_0 - 1)\log(E))]^{-1} \end{aligned}$$

where the third line follows from the fact that $\frac{\partial g}{\partial x} < 0$, so $g < 0$ holding universally implies it must hold for the best case $x = \delta \to 0$. The final line follows from ν being distributed standard logistic. Importantly, note that this probability **is independent of** α_1! Intuitively, this makes sense as satiation should play no part in whether a good is consumed at the point of zero consumption. Further, this mirrors the single-discreteness case where utilities are defined:

$$u_0 = (\alpha_0 - 1)\log(E) + \epsilon_0$$
$$u_1 = \kappa + \epsilon_1$$

where $(\alpha_0 - 1)\log(E)$ serves as a (dis)utility baseline for allocating all consumption to the outside good[4]. Of course, the additional parameters (α_0, E) are not identified in this case and become absorbed into the baseline utility for the inside good, κ.

4. Note that in full single-discreteness, all satiation parameters are 1, which leads to linear utility and full allocation of consumption to whichever good has highest baseline utility. In that case, baseline $u_0 = \epsilon_0$.

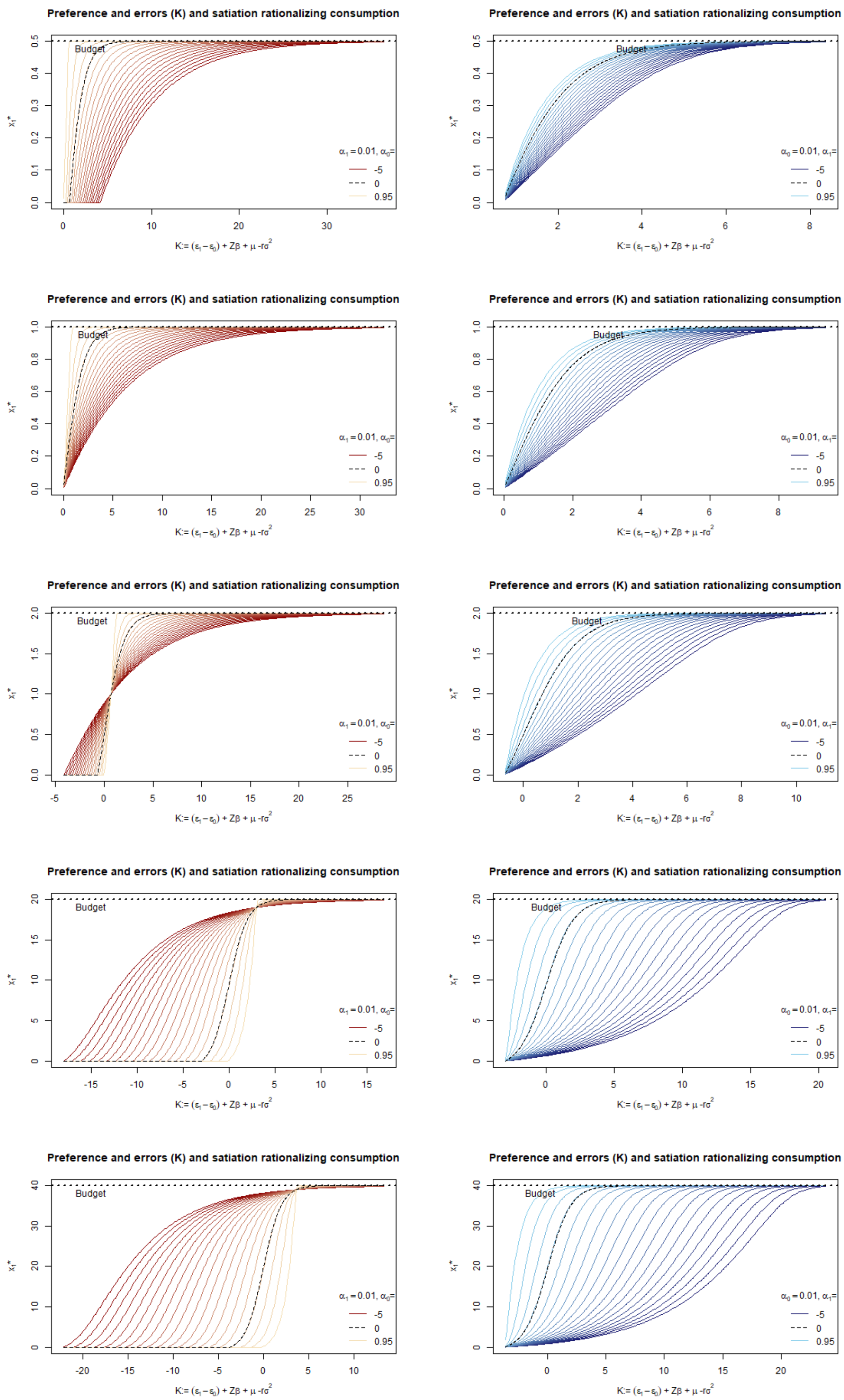


Figure B.5: Crossover point in budget scale, holding $\alpha_1 = 0.01$.

B.5 Likelihood of the nested extreme value (NEV) model

Bhat (2008) provides a constructive derivation of the likelihood under NEV errors and a toy example. Here I provide exact expressions using that "blueprint" under the two nest structure: $B_1 = \{0\}, B_2 = \{1,2,3,4\}$ which has a single parameter θ defining the correlation structure of the second nest. Here Nest 1 represents the outside good (consumed w.p. 1) and Nest 2 the inside goods (i.e. playing any level of the game).

Formally, let (x_1, x_2, x_3, x_4) be the vector of inside consumption, with outside consumption $x_0 = E - \sum_{j=1}^{4} x_j > 0$. Assume $(\epsilon_{1t}, \ldots \epsilon_{4t}) \sim NEV(\theta)$, and $\epsilon_{0t} \perp \epsilon_{jt} \forall j \neq 0$. Define the following:

$$\pi = \frac{1-\theta}{\theta}$$

$$A = e^{V_0}$$

$$B = \sum_{j>0} e^{\frac{V_j}{\theta}}$$

$$H = A + B^{\theta}$$

Then likelihoods for each pattern $\mathcal{L}_r$ where r inside goods are consumed are given:

$$
\begin{aligned}
\mathcal{L}_0(0,0,0,0) &= \frac{A}{H} \\
\mathcal{L}_1(x_1,0,0,0) &= \det(T_V) e^{V_1/\theta} \frac{AB^{\theta}}{H^2} \\
\mathcal{L}_2(x_1,x_2,0,0) &= \det(T_V) e^{\frac{V_1+V_2}{\theta}} \left[\frac{2AB^{2\theta-2}}{H^3} + \frac{AB^{\theta-2}}{H^2} \cdot \pi \right] \\
\mathcal{L}_3(x_1,x_2,x_3,0) &= \det(T_V) e^{\frac{V_1+V_2+V_3}{\theta}} \left[\frac{6AB^{3\theta-3}}{H^4} + \frac{6AB^{2\theta-3}}{H^3} \cdot \pi + \frac{AB^{\theta-3}}{H^2} \cdot \pi(1+2\pi) \right] \\
\mathcal{L}_4(x_1,x_2,x_3,x_4) &= \det(T_V) e^{\frac{V_1+V_2+V_3+V_4}{\theta}} \left[\frac{24AB^{4\theta-4}}{H^5} + \frac{36AB^{3\theta-4}}{H^4} \cdot \pi + \right. \\
&\quad \frac{2AB^{2\theta-4}}{H^3} \cdot (3\pi(3\pi+1) + \pi(2\pi+1)) + \\
&\quad \left. \frac{AB^{\theta-4}}{H^2} \cdot (3\pi+2)(2\pi+1)\pi \right]
\end{aligned}
$$

Note that for a model specification using less inside goods, the likelihoods correspondingly truncate. For example, in the case of only two inside goods, the possible observed consumption patterns can be described by $\mathcal{L}_0, \mathcal{L}_1, \mathcal{L}_2$.

B.6 Differentiation from existing projects

This paper is made possible through data granted by a joint partnership with the Wharton Consumer Analytics Initiative (WCAI) alongside a corporate data sponsor. Through attendance at the WCAI symposium, I have observed several other papers using the same data that either utilize a similar framework or attempt to address a similar question to me. Here I give a brief overview of those papers and explain how this particular paper differs from past work:

Mechanically, my proposed model is similar to the model employed in Sunada (2019), in which consumers optimize trial play and upgrade actions in a dynamic setting with Bayesian learning. However, my fundamental research question and specific model differ drastically. While Sunada (2019) aims to solve for the optimal trial design in a dynamic setting with a small discrete space, I am interested in characterizing the skill trajectories of players in a static setting in a continuous space where usage duration matters (e.g. beyond the binary use decision). Nevskaya and Albuquerque (2019), the main goal of my project is **not** to predict purchase.

Furthermore, Tao and Sweeting (2019) also considers skill growth among franchise players. The paper primarily aims to determine the impact that incompatible game modes across franchise games (e.g. across generations) has on port-back behavior (and quantifying the resultant revenue loss). Skill is defined in a capital accumulation sense and primarily refers to the physical construct of game save files, where compatibility means saved progress is ported forward to the next generation game. Operationally, this monotonically increasing construct directly enters into the consumer's utility.

Similar to Tao and Sweeting (2019), I show using reduced form evidence that there is a certain "switching cost" when upgrading game versions. I also observe a complementarity between (human) capital in the form of experience and increased utilization of "harder" (i.e. more competitive) game modes. Similar patterns of behavior are commonly observed in the literature for any substantial switch, e.g. when workers find a new job in a different field (Sanders, 2016) or when photographers upgrade to more advanced cameras (Huang, 2019).

However, while Tao explains this behavior in the context of forward-compatibility of certain game modes, I explain this behavior as a result of uncertainty and risk aversion, especially in the realm of competition and loss. For example, in a later section I show that even experienced players reduce their level of play right after upgrading, but eventually climb up to new highs. Uncertainty, for example, may arise from addition of new game modes as well as new core mechanics to be mastered (e.g. updated catch or throw mechanics). Online discussion boards are rife with questions from veteran players such as: "I just bought [game] and am wondering what the gameplay changes are that I need to learn to master"[5]. A core conceptual difference is that Tao entangles skill as both a physical construct (game progress) and mental one (player experience) and considers it as mode-specific. I instead consider skill as a purely psychic construct with potential spillover effects between modes. Intuitively, this means that all modes are compatible with respect to capital, but players may still have uncertainty about specific modes. Thus, when upgrading, there is a psychic switching cost but I show that in both short- and long-run, experience is not fully lost. Finally, the aims of my project differ from Tao's: rather that consider the gains from compatibility, I am concerned about possible gains from encouraging players to try more competitive game modes or reducing player uncertainty about game modes.

5. From: https://www.reddit.com/