The Consumption of Serial Media Products and the Optimal Release Strategy

Clarice Zhao[*], Nitin Mehta[†], and Mengze Shi[‡]

March 2023

Abstract

A digital publishing platform may choose to release serialized media content, such as chapters of a book or episodes of a drama, either *simultaneously* (all at once) or *sequentially* (over time). These alternative release strategies lead to different platform visiting and content consumption behavior: whereas a simultaneous release may induce binge consumption, a sequential release may generate increased platform visits and explorations of other books. To disentangle these opposing effects, we develop a structural model of consumers' serial purchases. We then estimate the model using data from an online book platform in China. Our estimation results confirm the hypothesized binge consumption effect of a faster release and the product exploration effect of a slower release. We also perform a counterfactual analysis to assess other forms of serial release strategies. We find that compared to the existing sequential release strategy, the platform is worse off with a simultaneous release strategy, but it could improve its sales revenue by implementing an optimized hybrid release strategy in which it releases some of a book's chapters simultaneously and the rest sequentially. Interestingly, a properly designed hybrid release can turn the binge consumption and product exploration effects into complementary forces.

[*] Desautels Faculty of Management, McGill University. E-mail: clarice.zhao@mcgill.ca
[†] Rotman School of Management, University of Toronto. E-mail: nmehta@rotman.utoronto.ca
[‡] Business School, Hong Kong University of Science and Technology. E-mail: mengzeshi@ust.hk

1. Introduction

A serial is a form of media in which a story is delivered in contiguous installments. Traditionally, the serial publishing model was used in media categories such as books, comics, television shows, and radio programs. More recently, serial publication has experienced explosive growth both in volume and in the variety of media products that use it, aided by emergence of digital platforms[1] and technological breakthroughs in creating/sharing digital content (Pittman and Sheehan 2015). In distributing serialized media content, digital technology provides platforms with more flexibility to manage the speed at which their content is released. Consider the book publishing industry as an example. Traditional publishing houses such as Harper Collins release physical books in their entirety. In contrast, digital book platforms follow a diverse set of release strategies: whereas platforms such as Barnes and Noble Press and Amazon Kindle Direct Publishing release entire books, platforms such as Webtoon and Wattpad release chapters sequentially over time. Despite the surge in digital serial content, there has been little research on how different release strategies may impact consumers' purchase behaviors and the ensuing platform's profits.

This paper is intended to bridge that gap with an empirical study of a digital book platform releasing serialized content (book chapters) over time. We investigate the platform's profitability for three types of release strategies[2]: *simultaneous release strategy* (releasing all the chapters of a book at once), *sequential release strategy* (releasing book chapters sequentially over time), and *hybrid release strategy* (releasing some chapters simultaneously and the rest sequentially). For sequential and hybrid strategies, we investigate the optimal release speed that maximizes profits of the platform. Our empirical investigation uses data from a large online book platform in China, which enjoys a 70% market share in the romance genre. It releases book chapters sequentially over time and uses a fixed pay-per-chapter pricing model.

The performance of a platform's serial release strategy depends on consumer responses to platform visits and book consumption. We propose two opposing forces that a book's release speed could have on consumer behavior: a *binge consumption effect* and a *product exploration effect*.

[1] Examples include Netflix, Amazon Prime Video, and Hulu for television series; Amazon Kindle, Wattpad, and China Literature for books; Webtoon and VIZ for comics; and iTunes and Spotify podcasts for radio shows.

[2] In this paper, we explore the release strategies from the platform's perspective, i.e., which chapter release strategy will maximize the platform's profits. The assumption here is that the platform has the power to mandate the writers to follow their release strategy. We will discuss later why this assumption is valid in the context of our platform.

Binge consumption refers to continuous and accelerated consumption of multiple chapters or episodes of a book/serial or multiple reading/viewing sessions in a short period (Schweidel and Moe 2016).[3] Faster release of new chapters, with simultaneous release being an extreme, makes more chapters of a book available to consumers at a given time and increases their propensity to binge-consume those chapters. And binge consumption in turn leads to a higher likelihood that a reader will finish reading a book as compared to staggered consumption (Schweidel and Moe 2016; Lu et al. 2019). The reason is a consumer typically needs to read enough chapters of a book before she is captivated, and likelihood of that happening is higher with binge consumption. With staggered consumption, a consumer can lose interest in a book before she is able to read enough chapters to get captivated.

Product exploration refers to consumers' increased likelihood of trying new books/serials on the platform. This can arise through two mechanisms. *First*, the slower the release speed of the focal book that the consumers are currently reading, the smaller will be the number of chapters of that book available to them at any point in time to read. To satisfy their reading needs, consumers may spread their consumption to books they otherwise may not have consumed. We call this phenomenon the *rationing effect*, following the convention in the economics literature, which studies the similar spillover effect to other substitutable goods when a focal good that a consumer would have otherwise purchased in large quantities is rationed (i.e., when there is a limited supply; Neary and Roberts 1980; Lee and Pitt 1987; Wang and Chen 1992). The magnitude of the rationing effect depends on availability of new chapters of the focal book and the extent of substitutability between the rationed book and other books. *Second*, platforms notify a book's current readers of the release of new chapters, and such notifications induce consumers to visit the platform more often than they would otherwise. Thus, the slower the release speed, the longer the period over which consumers receive notifications of an ongoing book they are currently reading and the more numerous their visits to the platform before finishing the book. Frequent visits give consumers more exposure to other books, especially those promoted by the platform on its webpage, prompting them to try them out. We call this phenomenon the *additional platform visits effect*.

[3] Binge consumption typically occurs when the content is engaging, chapters are sequentially connected (Lu et al. 2019), consumer has flexibility in terms of when to read, and the platform has enough chapters of the book available to consumers (Pittman and Sheehan 2015). These conditions were satisfied in the online book platform we studied.

Both rationing and additional platform visits effects can lead to increased consumption not only in the short run but also in the long run if consumers like the new books that they sample.

To disentangle these effects, we propose a structural model of consumer's platform visits and purchases in our online book platform. The platform lets the writers decide when to release their new chapters sequentially over time. We estimate our model on a panel data that consists of platform visits and purchases of randomly selected consumers. On a platform visit, consumers typically purchase multiple chapters across multiple books.

Our structural model considers a two-stage consumer decision process in which a consumer first decides whether to visit the platform and then, conditional on visiting the platform, decides which books to purchase and how many chapters of each of those books to purchase. We model purchases in the second stage as outcomes of consumer's utility maximization over the set of available books on the platform, subject to the budget, non-negativity, inventory, and availability constraints. As per the inventory constraint, a consumer will not purchase chapter(s) of a book if she has a positive inventory of chapters of it. The rationale for it follows from our institutional knowledge of how consumers purchase additional chapters of a book on the platform, which we will discuss it later in the paper. The availability constraint for a book refers to the maximum number of chapters of a book that a consumer can purchase on a platform visit. The release speed of a book, and thereby its chapter availability, impacts the consumer's utility maximization through the availability constraint. In the utility, we allow for complementarity/substitutability between books, and for consumers to purchase multiple chapters across multiple books on a platform visit.

In our data, we only observe purchases of consumers and not their consumption behavior. Thus, we cannot make direct inferences on their binge consumption behavior. However, as we will discuss later in the paper, we can infer a consumer's consumption rate of each book in each period by exploting the inventory constraint. We thus operationalize binge consumption of a book in terms of its consumption rate in that period, where a high consumption rate of a book implies a high propensity to binge consume it in that period. We model two types of binge consumption: (1) binge consumption of a book within a period, as captured by the impact of availability of chapters of a book on its consumption rate in that period. (2) Binge consumption of a book between periods, as captured by the impact of a consumer's consumption rate of that book in the previous period on its consumption rate in the present period. The rationing effect in the model is captured by the availability of chapters of the focal book through the availability constraint, and the extent of

substitutability between books. The more constrained the availability of the focal book, and the greater the substitutability between books, the more likely a consumer is to purchase other books. The additional platform visits effect is captured through the impact of notifications of new chapter releases on the probability of platform visits in the first stage, and the impact of a book being on the platform's homepage on purchases in the second stage. A strong impact of both variables will imply that notifications sent by the platform to consumers following an ongoing book will lead to additional platform visits, which may lead to purchases of new books promoted by the platform.

Our key empirical results are as follows. (i) The binge consumption effect within a period is significant - the consumption rate of a given book increases with availability of its chapters. Binge consumption behavior is also significant between periods - conditional on availability and controlling for book and consumer fixed effects, the consumption rate of a book in the last period has a significant positive impact on its consumption rate in the present period. (ii) With enough chapter availability, binge consumption results in consumers to read those available chapters, which then increases the probability of them being captivated by the book (and thereby finish reading it). On the other hand, low chapter availability increases the probability of consumers losing interest in the book. (iii) The substitutability between books is small, which implies that the rationing effect is small in magnitude. (iv) The additional platform visits effect is large and statistically significant. The number of notifications has a significant impact on consumer's platform visit probability—on average, a notification prompted by a new chapter release of an ongoing book that the consumer is currently following increases her platform visit probability by 21%. And conditional on platform visit, having a book listed on the platform's homepage increases the probability of the book being purchased by 101%.

Using estimates of the structural model, we conduct counterfactual analyses to compare the platform's performance under alternative release strategies. The platform's current strategy is to release chapters sequentially and let writers decide when they should release the new chapters. In the counterfactuals, we instead consider the case where the platform decides the release strategy and mandates the writers to follow it (the rationale for this assumption will be discussed later). The key results are as follows: (i) switching to simultaneous release strategy from the current strategy results in an 18% decrease in platform revenues. This is because of a significant (32%) reduction in the number of platform visits and thereby the number of books explored, which offsets the increase in binge consumption that comes with simultaneous release. (ii) Under the

counterfactual sequential release strategy, releasing 5 chapters/week for all books optimizes platform revenues, and generates a 7% increase in platform revenues over the existing release strategy. (iii) The hybrid release strategy, in which chapters in the first half are released simultaneously and the rest are released sequentially, yields 45% higher platform revenues over the current strategy, with increase in both the binge consumption and additional platform visits effects. The reason for increase in both effects is because simultaneous release of chapters in the first half enables consumers to binge consume many chapters, which increases their probability of being captivated by the book. Consequently, despite the slower release speed under sequential release in the second half (which may have otherwise resulted in consumers to lose interest in the book), they would continue to purchase new chapters of it. And while doing so, they would also explore other new books on the platform because of the additional platform visits effect. (iv) If chapters in the first half are released sequentially and chapters in the other half are released simultaneously, the platform revenues are lower than the current release strategy. This is because sequential release for first half results in relatively more consumers to lose interest in the book; and simultaneous release for the second half does not significantly increase binge consumption as many readers had quit reading the book in the first half. This shows that hybrid release, if designed properly, can turn binge consumption and platform visits effects into complementary forces.

Before we proceed, it is important to point out that we only focus on the impact of release strategy on consumer demand in this paper, while taking care of possible endogeneity from supply side (writer's) decisions. We do not focus on the impact of release strategy on writer's decisions such as what should be the book quality, what should be the number of chapters in a book, etc. As we will discuss later, the release strategies will not substantially impact the writer's decisions in our context. Thus, our study allows us to carefully identify the roles played by the demand side factors in the release strategy without having to worry about the supply side factors.

1.1 Related Literature

This paper is related to four streams of research. The first stream is on binge consumption (Schweidel and Moe 2016; Lu, Bradlow, and Hutchinson 2019; Lu, Karmarkar, and Venkatraman 2019; de Matos and Ferreira 2020; Zhang, Chan, and Luo 2021). Schweidel and Moe (2016) focus on the drivers of binge behavior and the impact of binge consumption on advertising exposure. Using consumption data from a streaming platform, they find that the number of episodes viewed

of a serial is positively associated with the continuation of viewing the same serial and returning quickly for a future viewing session, a phenomenon that they refer to as "viewing begets more viewing." Lu, Bradlow, and Hutchinson (2019) study binge behavior in the context of online educational platforms and refer to the across-day sequence of consumption as temporal binging and the within-day sequence of consumption as sessional binging. Some recent papers have examined the drivers of binge consumption for serialized media content, for example, the sequential connection of episodes through a storyline (Lu, Karmarkar, and Venkataraman 2019), and the digital media platforms providing flexibility in terms of when and how much to view (Pittman and Sheehan 2015). De Matos and Ferreira (2020) study the effect of binge-watching on platform subscription behavior and find that binge-watchers quickly deplete their content of interest and have lower willingness-to-pay for platform subscriptions. We extend this literature by introducing the release speed as a determinant of binge consumption. We also provide evidence for both within and across-periods binge behavior in a new empirical context.

The second stream is on rationing of goods. The literature has modeled rationing by adding an availability constraint for the quantity of the rationed good in the utility maximization (Neary and Roberts 1980; Lee and Pitt 1987). The empirical literature on rationing has found the spillover effects of rationing onto other goods. For instance, Wang and Chen (1991) find that rationing housing and staple foods can significantly increase the demand for unrationed non-staple foods. We extend the literature to a new empirical context in which the sequential release of digital content could lead to rationing of chapters at any given time. We apply the same methodology as the rationing literature and find a positive spillover effect on purchase of other books.

The third stream is on strategic mechanisms used to expand retail traffic. The main motivation for sequential release is to generate additional platform visits through notifications, with an expectation that the additional visits could lead to the exploration of new books. The idea of profiting from consumers' unplanned purchases and search costs resembles some conventional marketing practices in the retail industry. For instance, when a retailer engages in bait-and-switch practices, it attracts consumers by advertising a low-price product with limited availability. When consumers visit a store and realize the advertised product is out of stock, they purchase other products in the store (Wilkie et al. 1998; Li et al. 2013). Even if the low-price product is not out of stock, sacrificing the margin of one product can be profitable because of the profit margins from other products in the consumer's shopping basket. The focal product has often been referred to as

the loss leader (Lal and Matutes 1994). We extend these ideas to a new context. Unlike bait-and-switch and loss-leader pricing, which attract consumers using low prices, in this case, platforms generate visits through a sequential release followed by notifications of new chapter releases.

This paper is also related to studies of online platforms with serial digital content. These studies have examined the issues of intellectual property rights (Li et al. 2020), consumer herding (Ding and Li 2019) and motivating creators (Li et al. 2021). Zhang et al. (2022) use a similar dataset as ours but from a different platform that offers both pay per chapter and subscription-based pricing. They find evidence that many consumers overpay significantly by opting for pay-per-chapter, and that this behavior can be explained by time-inconsistent preferences and strategic self-control. We contribute to this line of literature by studying platforms' release strategies.

2. Institutional Details and Data

Our dataset comes from a leading Chinese online publishing platform for novels, whose identity we cannot disclose due to our non-disclosure agreement. China's online publishing industry is worth $4.3 billion, with 460 million consumers and more than 17 million writers by 2020. The focal platform is one of the largest in the industry, with more than 50 million registered readers and nearly 20 years in business. It enjoys a 70% market share in the romance genre, and it has exclusive contracts with its writers. Its books are published in a serialized format, chapter by chapter, rather than all at once. The platform follows a sequential release strategy in which it lets the writers decide when to release each chapter. The platform uses pay-per-chapter business model rather than subscriptions. The platform sets a price for each chapter according to length, with the unit price for all its books fixed at RMB0.04 per 1,000 Chinese characters. The platform did not offer any quantity discount or price promotions during our sample period.

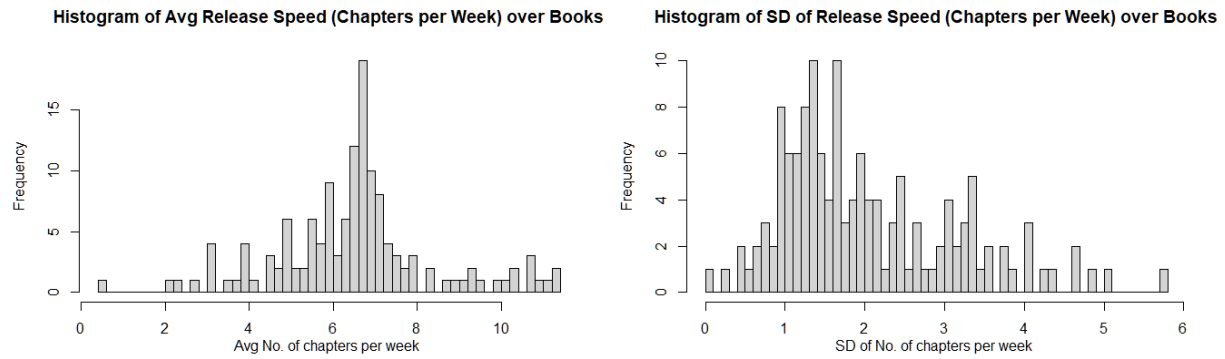
When a consumer visits the platform, the interface allows her to find both ongoing books (books that are still releasing chapters) and completed books (books that are fully published). There are multiple ways for a consumer to find a book in the platform's library, most significantly, through the homepage weekly sales ranking. The platform's home page lists the bestselling books categorized based on genres or predefined periods (e.g., last week, last month, etc.). A consumer can also find a book by searching the platform's library based on genres and subgenres, book age, completed or ongoing books, and writer reputation. The platform sends reminder notifications about new chapters to consumers who have purchased previous chapters. The purchased chapters

are saved either on the platform's cloud or on the consumer's local drive within the app. The book data are in text format and take up very little storage space; accordingly, the cost of storing the books is very low. The platform's app integrates user interfaces for reading already-purchased chapters and buying additional ones. After scrolling down the page to read all previously purchased chapters, the consumer reaches a page that displays the title of the next unpurchased chapter. The app does not display any of the content of the unpurchased chapter, only a message prompting the consumer to purchase the next chapters and a button leading to a menu of available chapters. After the purchase, the interface returns to the first unread chapter. This interface design eliminates interruptions and minimizes the effort required for a consumer to access the purchase page. According to our conversation with the platform's manager, this is the predominant method by which consumers buy additional chapters of a book. Given the integrated interface design for reading and purchasing, consumers are unlikely to buy new chapters before they have finished reading previously purchased chapters.

The data that we use for estimation includes 6 months in 2017 (from April to September). To keep our structural model estimation computationally manageable, we focus on the genre of "Pure Love" romance books to limit the size of the choice set. We include 610 consumers from a randomly selected consumer sample based on the following three criteria for each consumer: (i) at least 90% of the consumer's expenditures are on books in the chosen genre, and at least 75% of the consumer's expenditures are on books that are in the 138 top sellers in the genre; (ii) the consumer uses the platform before the sample period; and (iii) the consumer purchases chapters from at least three different books and spends more than 25 Chinese yuan during the sample period.

On average, a book consists of 170.7 chapters and takes 4 months to be fully released. The length of the weekly chapter releases of the books in our sample has a mean of 6.4 chapters and a standard deviation of 2.8 chapters. In Figure 1, we plot the histogram of the average (on the left) and standard deviation (on the right) of the speed of each book's weekly release.

Figure 1. Distributions of Release Speed Across and Within Books



The histogram on the left suggests a unimodal distribution with substantial variation in the average release speed across books. The histogram on the right shows that the standard deviations of release speed within each book are positive and around 2 chapters, indicating sufficient dispersion in release speed within books considering the average release speed of 6.4 chapters per week.

The unit of observation in our data is consumer/day. On any given day, a consumer makes two decisions: (1) she first decides whether to visit the platform and (2) if she decides to visit, she chooses which books to buy and how many chapters of each of those books to purchase. In our data, we observe both these decisions for each consumer. On average, a consumer purchases 1.7 books and 3.4 chapters per book, per visit. The purchase patterns for ongoing and completed books are different due to their availability. When reading an ongoing book, on average, consumers catch up with the book's release process within 2.9 days after their initial purchase; they then continue reading the newly released chapters as they are uploaded. When starting a completed book, on average, consumers complete their purchase of the book within 2.6 days.[4]

Within the genre under study, we choose 138 best-selling books in 2017, which make up 89% of the sales revenue within the genre. These books are written by experienced writers: on average, the writers have 3 years of writing experience, with 5 books written before 2017. Among these books, the average number of chapters is 170.7, and the average release period for an entire book is 223.6 days. Eighty of the 138 books begin publishing before April 2017, and 19 are completely published by April 2017; 58 begin publishing after April 2017. By the end of September 2017, 82 books are completed. In Table 1, we list the number of books based on an

[4] See section 1 of the Web Appendix for details on how consumers' purchases differ depending on the number of the chapters of a book already released when the consumer first started reading the book.

initial publishing time window and a final publishing time window, which suggests that book publishing windows are staggered during our data time window.

Table 1 Number of Books Based on Initial and Final Publishing Times

	Before 2017	2017 Q1	2017 Q2	2017 Q3	2017 Q4	After 2017
Initial Publishing Time	46	34	36	22	0	0
Final Publishing Time	7	12	25	38	31	25

The purchase data provided by the platform include platform visits at the daily level, individual purchase records at the order level, with the time of the order detailed to the second. We set each day as a single period. On average, for a given platform visit (defined by day and consumer), a consumer purchases 6.1 chapters total, or 3.6 chapters per book. The purchase of ongoing books is heavy in the initial weeks but light after the consumer catches up with the publishing process. On average, a consumer purchases 641.7 chapters from 9.4 different books, starts 7.7 new books, and finishes reading 3.5 books in the sample period.

We report the summary statistics of the key variables in Table 2. They are categorized into book-specific, book-week-specific, and consumer-week-specific variables. We index the 183 days in the sample period as days 1–183 based on their distance from April 1, 2017. We use this day index to construct the "First publishing day" variable.

Table 2 Summary Statistics

Statistic	N	Mean	S.D.	Min	1st Qu.	Median	3rd Qu.	Max
Book-specific variables:								
No. of paid chapters	138	170.7	137.3	28	86	135	203.8	1,060
No. of release days	138	223.6	187.6	46	109.2	156	266.8	1,350
First publishing day	138	-70.1	188.7	-816	-144.3	-31.5	62.0	158
Book-day specific variables: (conditional on at least one chapter released)								
Completed (vs. ongoing)	19,730	0.45	0.50	0	0	0	1	1
No. of chapters available	19,730	128.1	120.5	1	59	102	161	1,060
Consumer-day specific variables:								
Positive purchases	105,225	0.57	0.49	0	0	1	1	1
Chapters purchased (> 0)	58,152	5.4	18.8	1	1	2	3	1,060
Books purchased (> 0)	58,152	1.71	1.07	1	1	1	2	12

Chapters purchased/book (> 0)	58,152	3.4	13.6	1	1	1	1.5	1,060

3. Reduced-Form Evidence

In this subsection, we provide reduced-form evidence for the presence of the binge consumption, rationing, and additional platform visits effects of a book's release strategy.

3.1 Binge Consumption Effect

Binge consumption refers to the behavior of reading/viewing a large number of chapters in a short time (Schweidel and Moe 2016). In our data, we do not observe the reading behavior of consumers. As a result, we cannot make direct inferences on their binge consumption behavior. However, as we will discuss later in this section, we can infer a consumer's consumption rate of each book in each period from her purchase data. We thus operationalize binge consumption of a book in terms of its consumption rate in that period, where a high consumption rate implies a high propensity to binge consume in that period. The literature has examined two types of binge behavior (Schweidel and Moe 2016; Lu, Bradlow, and Hutchinson 2019): (a) binging within a period, and (b) binging across periods. In our context, we operationalize within period binge consumption by the impact of release speed of a book on its consumption rate in the same period, and the binge consumption across periods by the impact of consumption rate of a book in the last period on its consumption rate in the present period. In this section, we provide reduced-form evidence for both types.

3.11 Binge Behavior within a period

We demonstrate within period binge consumption in two steps. First, we show that as the release speed of a book increases, consumers purchase more chapters of that book in that period. Second, we show that when consumers purchase a larger number of chapters of a book, their consumption rate of that book increases in that period compared with when they purchase a smaller number of chapters. Together, these results imply that as a book's release speed increases, consumers consume more chapters of the book in that period, indicating binge consumption within that period.

Effect of Release Speed on Purchase Quantity. To examine the effect of release speed of a book on its purchase quantity (number of chapters purchased) in that period, we run a regression of a consumer (i)'s purchases of chapters of book (j) in period (t) against the book's release speed in

that period (measured in terms of number of chapters released per day), conditional on consumer i having caught up with book j's progress at least once in the past.

There is a potential endogeneity issue with a writer's new chapter release speed for three reasons. First, books' release speed may be correlated with their overall qualities, which are unobserved by the researcher. For instance, books released at a slower speed could be written more carefully and thus could be of higher quality than books released at a faster speed. Second, the release speed of a given book's chapters over time can be correlated with qualities of the chapters. That is, compared to chapters that the writer takes a shorter time to release, chapters that the writer takes a longer time to release could be of higher quality because the latter chapters may be written more carefully. Third, the decision on when to release a chapter is made by the writer, who may choose a release speed in response to market-level demand shocks unobserved by the researcher.

To deal with the first source of endogeneity, we include book-fixed effects in the regression to control for any time-invariant unobserved confounders, such as a book's overall quality. To deal with the second source of endogeneity, note that our dataset consists of best-selling books by experienced writers. In Section 2 of the web appendix, we have provided reduced form evidence of chapter quality not changing with the release speed for experienced writers. Another related source of endogeneity is that the release speed can impact the comments that readers post on webpages of chapters on the platform. This can impact the quality of subsequent chapters, since it is possible that writers read those comments and modify their writing based on them. This is not a source of concern for two reasons. First, we show in section 3 of the web appendix that the number of readers' comments on a chapter is not significantly correlated with the release speed. Second, it is mostly the inexperienced writers who modify their writing based on readers' comments.

Finally, regarding the third source (and all other potential sources) of endogeneity, we construct an instrument for the release speed of the focal book, as a function of the release speed of all other books written by the same author at least 6 months earlier. We divide the focal book and the aforementioned prior books into five parts, each with an equal number of chapters and calculate the average release speed for each quantile. To obtain the IV for the release speed of focal book j in week t, we first identify the quantile of the book to which the chapters released that week belong. For example, for book j in week t, if most recent chapters released are from the first quarter of book j, then the IV is the average release speed of the chapters in the first quarter

for all previous (at least 6 months prior) books written by the author of book j. We have discussed why this IV satisfies the relevance and exclusion conditions in section 4 of the Web Appendix.

Table 3 shows the results of the regression of purchase quantity of a book in each period on its release speed in that period, using the instrument. The results confirm that release speed has a significant positive effect on purchase quantity.

Table 3 The Effect of Release Speed on Purchase Quantity

	Dependent variable:
	Chapters purchased
Release speed: Chapters/week	0.63 * (0.26)
Consumer F.E.	✓
Book F.E.	✓
2SLS with instruments	✓

Note: Regressions are at the consumer-book-day level. Standard errors (in brackets) are three-way clustered at the consumer, book, and day levels. ***, **, and * indicate significance at the 0.1%, 1%, and 5% levels, respectively.

Impact of Purchase Quantity on Consumption Rate. We next provide evidence for the impact of purchase quantity of a book on its consumption rate. Since we do not observe consumption, we first discuss how we identify the consumption rate of each book in each period. We identify it by utilizing the inventory constraint that we discussed in section 1, as per which a consumer only purchases more chapters of a book after she finishes reading all the previously purchased chapters of that book. The rationale for it is as follows. First, the novels in our sample have interconnected storylines across chapters. To follow a story's development, a consumer should read the chapters in their natural order rather than jumping ahead to later chapters. Second, consumers have no incentive to engage in forward buying and stockpiling because the price of book chapters remains same over time, the storage cost is minimal, and the platform never runs a promotion. In retail settings in which forward buying is commonly observed, the most important driver of purchasing more goods before depleting inventory stock is the desire to take advantage of temporary promotions. Third, as discussed in Section 2, the platform's app integrates user interfaces for reading already-purchased chapters and buying additional ones. Since this is the predominant method by which consumers buy additional chapters of a book, consumers are unlikely to buy new chapters before they have finished reading previously purchased chapters of that book.

The inventory constraint implies that the number of chapters read by a consumer between any two consecutive purchase orders of a book is the number of chapters purchased in her first purchase order. If we denote the times of two consecutive purchases of a given book by *s* and *t*, with $s < t$, then the number of chapters of that book read during the period between *s* and *t* is the chapters purchased at time *s*. We then calculate the consumer's average consumption rate of the book during the period between two consecutive purchases (at times *s* and *t*) by dividing the number of chapters read by the number of days in the duration of that period (i.e., between periods *s* and *t*). We regress the consumer's average consumption rate of a book in each period on the number of purchase quantity of that book in the same period, controlling for book FE and consumer FE. If the regression coefficient of recent purchase order size is positive and significant, then it will imply that the average consumption rate of the book increases with the number of chapters last purchased. The regression in Table 4 results show a significant and positive coefficient (0.64). This result combined with the previous result that higher availability due to a faster release speed leads to more purchases suggests that that as availability increases, the average consumption rate also increases in that period, which is indicative of binge consumption in that period.

Table 4 The Effect of Order Size on Consumption Rate

	Dependent variable:
	Consumption rate in a week
Recent purchase order size: No. of chapters	0.64 *** (0.04)
Consumer F.E.	✓
Book F.E.	✓

Note: Regressions are at the consumer-book-day level. Standard errors (in brackets) are three-way clustered at the consumer, book, and day levels. ***, **, and * indicate significance at the 0.1%, 1%, and 5% levels, respectively.

3.12 Binge Behavior Across Periods

We demonstrate binge behavior across periods in two steps. First, we show that as the consumption rate of a book in the previous period increases, then conditional on availability in the present period, the purchase quantity of that book in the present period will also increase. Second, we show that as the purchase quantity of a book in the current period increases, its consumption rate in the current period also increases. Together, these results imply that high consumption rate of a book in the previous period leads to its high consumption rate in the present period. Since we have demonstrated the second result in section 3.11, we will only demonstrate the first result here.

We regress the number of chapters of a book purchased by a consumer in present period on her consumption rate for that book in the last period, controlling for both consumer- and book-fixed effects. Since this analysis is conditional on chapter availability, we consider only those observations in which a consumer has not caught up with the book's progress. The results are summarized in Table 5, which show that there is a positive and significant relationship between the current period purchase and the previous period binge. This result combined with the previous result that higher purchase quantity leads to higher consumption in that period purchases suggests binge consumption behavior across periods.

Table 5 Binge Consumption Across Periods

	Dependent variable:
	No. of chapters purchased in a day
No. of chapters consumed in the previous day	0.246 *** (0.03)
Consumer F.E.	✓
Book F.E.	✓

Note: Regressions are at the consumer-book-day level. Standard errors (in brackets) are three-way clustered at the consumer, book, and day levels. ***, **, and * indicate significance at the 0.1%, 1%, and 5% levels, respectively.

3.2 Rationing Effect

When a book is released slowly, only a limited number of book chapters are available for consumers to purchase. This limited availability affects the consumers following the book in a manner like a rationing policy, whereby the consumers may start purchasing substitute books. Thus, by affecting the level of chapter availability, a book's release speed can influence the purchases of new books on the platform. We test the rationing effect by regressing the release speed of a book (j) in a week (t) on the number of new books tried by the consumer (i) in the week (t). We include book and consumer FE and the IV discussed earlier to handle the endogeneity of release speed. The regression analysis uses observations conditional on consumer i visiting the platform that week, and conditional on consumer i purchasing chapters of book j before week t and catching up with the progress of book j's release. This is done to ensure that the consumer is facing an immediate constraint imposed by the release speed (instead of the scenario in which the consumer is not affected by the release speed as many of the available chapters are not yet read).

The results in Table 6 show that conditional on a consumer following a book and visiting the platform, release speed has a negative and significant effect on product exploration within the

same period. However, the coefficient is rather small, which implies a small magnitude of the rationing effect - all else being equal, if the release speed decreases by 50% on the platform, each consumer would buy only one additional book over the course of a year.

Table 6 Rationing Effect: The Effect of Release Speed on Purchase Variety

	Dependent variable:	
	Whether consumer tries a new book	Log (No. of new books tried + 1)
	(1)	(2)
Release speed: Chapters/week	-0.003 (0.001)	-0.003 (0.001)
Consumer F.E.	✓	✓
Book F.E.	✓	✓
2SLS with instruments	✓	✓

Note: Regressions are at the consumer-book-week level. Standard errors (in brackets) are three-way clustered at the consumer, book, and week levels.

3.3 Additional Platform Visits Effect

Release speed could also affect product exploration through changes in the number of platform visits. When consumers are following an ongoing book, they receive notifications from the platform when new chapters of that book are released. Since consumers look forward to reading the newly released chapters, these notifications would motivate them to visit the platform. The slower the release speed of a book, the greater will be the number of periods over which the consumers receive such notifications, and thus the greater will be the number of platform visits. Once they arrive at the platform, they are likely to buy other books, especially the ones that are promoted on the platform's home page. Thus, the slower the release speed, the greater will be the exploration of other books because of increased visits to the platform.

We present the reduced form evidence of notifications of new chapters bringing consumers to the platform, and that the slower the release speed, the greater their number of visits to the platform. We run a regression in which the DV is whether a consumer visited the platform on a given day. The regressors are (i) whether the day is a weekend, (ii) consumer's visit frequency in the last 7 days, (iii) consumer FE, (iv) number of notifications received by the consumer (about new chapters being released) for those books that she is up to date with their progress, and (v) number of notifications received by the consumer for those books that she has been following *but is not* up to date with their progress. Table 7 reports the results. We find that notifications released

by the platform for ongoing books that the consumer is currently reading *but is not up to* date with their progress has a much smaller impact on platform visits. This makes sense since notifications of new chapters are less likely to motivate those consumers to go to the platform who have yet purchased all its previously released chapters. However, the notifications released for books, for which the consumer is up to date with, has a significant, positive and greater effect on her platform visits. Moreover, we see that the impact of the first notification is much greater than any additional notification. This means that for a fixed number of chapters in a book, a decrease in the release speed from say 3 chapters/day to 1 chapter/day will result in the expected number of visits to increase over the release time of the book. This supports the additional platform visits effect.

Table 7 The Effect of Notifications on Platform Visits

	Dependent variable: Platform Visit
Weekend	-0.006 (0.005)
No. of visits in the last 7 days	0.092 *** (0.002)
No. of notifications from books that the consumer is up to date on ≥ 1	0.165 *** (0.009)
No. of notifications from books that the consumer is up to date on ≥ 2	0.027 *** (0.004)
No. of notifications from books that the consumer is up to date on ≥ 3	0.001 (0.004)
No. of notifications from books that the consumer is not up to date on ≥ 1	0.070 *** (0.007)
No. of notifications from books that the consumer is not up to date on ≥ 2	0.017 ** (0.006)
No. of notifications from books that the consumer is not up to date on ≥ 3	0.016 ** (0.006)
Consumer F.E.	✓

Note: Regressions are at the consumer–day level. Standard errors (in brackets) are clustered at the consumer level. ***, **, and * indicate significance at the 0.1%, 1%, and 5% levels, respectively.

To summarize, we find evidence for all three effects. First, for binge consumption effect, we find that consumers are more likely to binge read in a period when the release speed is faster than when it is slower, and single-period binge behavior has a positive effect on future binge behavior. Second, for rationing effect, we find that when a consumer is following an ongoing book, a slower release causes the consumer to explore other books on the platform than a faster release, but this effect is small. Third, for additional platform visits effect, we find that platform visits are positively affected by whether a new chapter of a previously purchased book is released, and that the slower the release speed, the greater will be the number of visits to the platform.

4. Model

We develop a two-stage model for a consumer's platform visit and book purchase behavior in each period. In the first stage, the consumer decides whether to visit the platform. In the second stage, conditional on visiting the platform, the consumer chooses books and chapters to purchase. We describe the model in a backward fashion, starting in Section 4.1 with our proposed model for the consumer's purchase decisions in the second stage. We then move back to the first stage and explain the platform visit decision in Section 4.2. In section 4.3, we discuss the updating of inventories and the specification of consumer's consumption rate. In Section 4.4, we discuss how our model captures the binge consumption, rationing and the additional platform visits effects.

4.1. Second Stage: Book and Chapter Purchase Decisions

In the second stage, a consumer visiting the platform selects which books and chapters to purchase. Recall from section 2 that in the data, consumers typically purchase multiple chapters across multiple books on a given platform visit. To model multiple discreteness in purchases, we adopt a direct utility maximization approach (Wales and Woodland 1983; Bhat 2008). A consumer, denoted by i, chooses from N_{it} books to purchase on the platform at time t. For consumer i, the choice set N_{it} can change over time for two reasons. First, a book (denoted by j) exits the choice set after consumer i purchases all its chapters. Second, a new book can enter the consumer's choice set when the writer releases the book to the market. The N_{it} books represent the 'inside goods' which consist of N_{it}-1=138 bestselling books in the romance genre, and a composite book which includes all other books that the consumers might purchase on the platform. Let x_{ijt} be the number of chapters of book j that consumer i purchases at time t, and $x_{it} = \{x_{ijt}\}_{j=1}^{N_{it}}$ be an $N_{it} \times 1$ vector of the purchase quantities of N_{it} books. Let good *0* be the outside good and x_{i0t} be the quantity of the outside good that consumer i purchases at time t. The outside good represents the purchases made outside the platform. We assume that the consumer always purchases the outside good (i.e., $x_{i0t} > 0$) and we normalize the price of the outside good to be 1.

Utility Function. We define the direct utility of consumer $i's$ purchases in time t as follows:

$$U_{it} = A_{it}^T x_{it} - \frac{1}{2} x_{it}^T B_i x_{it} + x_{i0t} \tag{4.1}$$

The utility is quadratic in the quantities of the N_{it} inside goods and linear in the outside good.[5] The quadratic utility has commonly been used in the prior literature (e.g., Wales and Woodland 1983; Economedies et al. 2008). The term $A_{it} = \{\alpha_{ijt}\}_{j=1}^{N_{it}}$ is an $N_{it} \times 1$ vector, where α_{ijt} is consumer i's baseline preference for book j in time t. The baseline preference for book *j* positively impacts both the purchase incidence and the purchase quantity decisions related to book *j*, and it is specified as follows:

$$\alpha_{ijt} = \alpha_{i0} + \alpha_j + \alpha_i^T Z_{ijt} + \varepsilon_{ijt} \quad (4.2)$$

The term α_{i0} is the category intercept, which is the same across all books but is assumed to be heterogenous across consumers. It captures the relative preference of inside goods to the outside good. The term α_j is book *j*'s specific intercept, which is normalized to zero for book *j* = 1.

The vector Z_{ijt} consists of the following explanatory variables: (1) an indicator of whether book *j* is fully published, which captures consumer *i*'s relative preference for purchasing ongoing vs. completed books; (2) an indicator of whether the platform's homepage promotes book *j* at time *t*; (3) an indicator of whether the consumer visited the platform after receiving a notification(s) from a book she is currently following and is up to date with the progress. This captures the extent to which the purchases of books will differ across organic platform visits and platform visits motivated by notifications; (4) an interaction of variables 2 and 3, which captures the extent to which the purchases of books promoted on the platform will differ across organic platform visits and platform visits motivated by notifications; (5) consumer *i*'s consumption rate of book *j* in the last 7 days, which captures whether the binge consumption of a book in the last period begets its binge consumption in the present period; (6) the time since the publication of the first chapter of book *j*, which captures consumer *i*'s preference for newer vs. older books; (7) an indicator for whether any chapter of the book is previously purchased by consumer *i*; (8) log of time elapsed since the last purchase of book *j* by consumer *i*, which captures the extent to which the preference for the book decreases with the time elapsed since its last purchase; (9) an indicator of whether book *j* is purchased by consumer *i* in the past 7 days; (10) log of total number of chapters of book

[5] It is important to note that economic theory does not require the utility to be monotonically increasing in the inside goods (Mas-Collel, Whinston and Greene, 1995). Instead, economic theory only requires local non-satiation of preferences, which is satisfied by the utility in (4.1) being strictly increasing in the quantity of the outside good.

j already purchased by consumer *i*, which captures the extent to which the consumer is captivated by the book, and (11) whether day *t* is a weekend or a weekday.

Finally, the econometrician's error term ε_{ijt} is assumed to have a standard normal distribution, which is IID across consumers and over time, but can be correlated across books. We assume the correlation to be the same across all pairs of books and equal to ρ_{ε}. This assumption is made for tractability, as we have 138 books in our dataset, which is reasonable in our context because all 138 books are best sellers in the same genre. The quadratic term $B_i = \{b_{i,jj'}\}$ in Equation (4.1) is an $N_{it} \times N_{it}$ dimensional symmetric and positive definite matrix.[6] The diagonal terms $b_{i,jj}$ capture the diminishing rate of returns of consumer *i*'s purchase of *j*. The off-diagonal terms $b_{i,jj'}$ $(j \neq j')$ capture the substitutability/ complementarity between book *j* and book *j'*. We assume that $B = \{b_{i,jj'}\}$ takes the following form:

$$\text{Diagonal: } b_{i,jj} = \beta_i$$

$$\text{Off diagonal: } b_{i,jj'} = \rho\beta_i \quad \forall j \neq j' \tag{4.3}$$

In Equation (4.3), we assume that consumer *i*'s diminishing rate of return from her purchases is the same across all books, and the degree of substitutability is also the same across all pairs of books. Once again, these assumptions are made for tractability and are reasonable in our context because all 138 books are best sellers in the same genre, and none of them are sequels to another. Parameter ρ captures the degree of complementarity/substitutability between books, which takes a value between -1 and 1. A value of ρ = 0 implies that the books are independent, and a positive (negative) value implies that the books are substitutes (complements).

The Constrained Utility Maximization Problem. Consumer *i*'s purchase decisions in the second stage of period *t* follow the constrained utility maximization problem formulated below:

$$\max_{x_{it}, x_{i0t}} A_{it}^T x_{it} - \frac{1}{2} x_{it}^T B_i x_{it} + x_{i0t} \tag{4.4}$$

$$\text{s.t.} \quad p^T x_{it} + x_{i0t} \leq M_{it} \quad \text{(Budget Constraint)}$$

$$x_{ijt} \geq 0 \;\; \forall j = 1..N_{it} \quad \text{(Non-negativity Constraints)}$$

[6] Symmetry and positive definiteness of B_t is required to satisfy the regularity properties required by economic theory, which ensure that (i) Slutsky matrix is symmetric and negative semidefinite, and (ii) utility maximization yields a unique solution.

$$x_{ijt} \leq c_{ijt} \;\; \forall\, j = 1..N_{it} \quad \text{(Availability Constraints)}$$

$$Inv_{ijt}.x_{ijt} = 0 \quad \forall\, j = 1..N_{it} \quad \text{(Inventory Constraints)}$$

The utility maximization in (4.4) is similar to the discrete-continuous choice utility maximization approaches used in the prior literature (Chintagunta 1993; Nair, Dube and Chintagunta 2005; Mehta and Ma 2012) to model consumers' purchase incidence and quantity decisions.[7] In the utility maximization in equation (4.4), the first constraint represents the budget constraint, in which M_{it} is the per-period budget, which as discussed earlier is larger than the amount spent on books in this platform. The term $p = \{p_j\}_{j=1}^{N_{it}}$ represents a vector of per-chapter prices of N_{it} books, which are constant over time and across chapters within a book. The second constraint is the non-negativity constraint, as per which the purchase quantity of any book cannot be negative. If the purchase quantity of a book *j* takes a value of 0, it implies that book *j* is not purchased. The third constraint is the availability constraint on the number of chapters of book *j* that are available for consumer *i* to purchase in period *t*. The release speed of a book, and thereby its chapter availability, impacts the consumer's utility maximization through the availability constraint. The term c_{ijt} represents the maximum number of chapters of book *j* that consumer *i* can purchase in period *t*. It is calculated as the difference between the total number of chapters of book *j* released by period t and the total number of chapters consumer i purchases from book *j* before period *t*.

The fourth constraint is the inventory constraint, in which the Inv_{ijt} represents consumer *i*'s inventory of purchased and unread chapters of book *j* just before her purchase decision in period *t.* As per the inventory constraint, the consumer does not purchase chapter(s) of book *j* if she has a positive inventory of the chapters of that book – i.e., if $Inv_{ijt} > 0$, then $x_{ijt} = 0$; and the consumer will only purchase chapters of a book if her inventory of that book is zero – i.e., if $x_{ijt} > 0$, then $Inv_{ijt} = 0$. The rationale for this constraint has been discussed earlier in section 3.1.

[7] There are five points to note regarding the discrete-continuous utility maximization in (4.4): (i) It assumes that purchase quantities of inside goods (i.e., number of chapters purchased), x_{it}, are continuous variables. (ii) It assumes that price per chapter of a book is the same across all chapters of that book. (iii) It does not have a time constraint. (iv) It assumes that consumers are myopic decision makers (v) It is unlike the dynamic inventory models (Erdem, Imai and Keane 2003; Ching and Osborne 2020) in which forward looking consumers purchase strategically in response to price promotions by stockpiling a good, and in which utility is maximized over quantities of goods consumed in each period (and not maximized over quantities purchased in each period, as in discrete continuous utility maximization models). In section 5 of the Web Appendix, we have explained our rationale in detail for each of these points.

The solution to the utility maximization gives the optimal decisions in both purchase incidence (whether to buy any chapter for each book) and purchase quantity (how many chapters to buy for each book) decisions. The Lagrangian is given as follows:

$$\mathcal{L} = A_{it}^T x_{it} - \frac{1}{2} x_{it}^T B_i x_{it} + M_{it} - \sum_{j=1}^{N_{it}} p_j x_{ijt} - \sum_{j=1}^{N_{it}} \mu_j x_{ijt} \quad - \sum_{j=1}^{N_{it}} \omega_j \left(x_{ijt} - c_{ijt}\right) - \sum_{j=1}^{N_{it}} \gamma_j x_{ijt} \quad (4.5)$$

In Equation (4.5), μ_j, ω_j, and γ_j are the Lagrange multipliers for non-negativity, availability, and inventory constraints, respectively. The complementary slackness conditions imply the following: (i) if $x_{ijt} > 0$, then $\mu_j = 0$; if $x_{ijt} = 0$, then $\mu_j \leq 0$. (ii) If $x_{ijt} = c_{ijt}$, then $\omega_j \geq 0$; if $x_{ijt} \leq c_{ijt}$, then $\omega_j = 0$. Given the constraints, the consumer's purchase amount of book j ($j = 1, \ldots, N_{it}$) at time t can belong to one of the following three cases.

Case 1: $x_{ijt} = 0$ (i.e., consumer i does not purchase any chapter of book j at time t). This is the case when either (a) the consumer's inventory of book j in period t is positive; (b) the consumer's inventory of book j in period t is zero, the number of non-purchased and available chapters of book j in period t, c_{ijt}, is positive, and the optimal purchase quantity of book j takes a boundary solution at zero; or (c) the number of non-purchased and available chapters of book j in period t, c_{ijt}, is zero, in which case book j is not in consumer i's choice set in period t.

Case 2: $x_{ijt} = c_{ijt} > 0$ (i.e., consumer i purchases book j at the availability constraint c_{ijt} at time t). In this case, consumer i has no inventory of book j in period t and she buys all the available chapters that she has not yet purchased.

Case 3: $x_{ijt} \in (0, c_{ijt})$ where $c_{ijt} > 0$, and thus consumer i purchases a subset of the available chapters of book j. This is the case when the consumer has no inventory of book j in period t and the optimal purchase quantity of book j takes an interior solution.

Following the above classification, we next categorize the entire choice set of books for consumer i in period t into three groups. Group 1 consists of books $j = 1 \ldots n_1$ corresponding to case 1, group 2 consists of books $j = n_1 + 1 \ldots n_1 + n_2$ corresponding to case 2, and group 3 consists of books $j = n_1 + n_2 + 1 \ldots n_1 + n_2 + n_3$ (where $n_1 + n_2 + n_3 = N_{it}$) corresponding to case 3. Based on the Kuhn Tucker conditions and complementary slackness conditions that follow from the Lagrangian, we derive the conditions below for each of the three groups.

Group 1: Conditions for non-purchase of books $j = 1..n_1$:

$Inv_{ijt} > 0$ or $\{Inv_{ijt} = 0,\ \alpha_{ijt} - p_j - \beta_i \rho s_{it} \leq 0\}$ (4.6)

Group 2: Conditions for purchase at the availability constraint for books $j = n_1 + 1..n_1 + n_2$:

$$\{Inv_{ijt} = 0, \alpha_{ijt} - p_j - \beta_i(1-\rho)c_{ijt} - \beta_i \rho s_{it} \geq 0\} \quad (4.7)$$

Group 3: Conditions for purchase at less than the availability constraint for books $j = n_1 + 1..n_1 + n_2$:

$$\{Inv_{ijt} = 0, \alpha_{ijt} - p_j - \beta_i(1-\rho)x_{ijt} - \beta_i \rho s_{it} = 0\} \quad (4.8)$$

Equations (4.6)–(4.8) represent the joint conditions that need to be satisfied for book $j = 1...n_1$ to not be purchased, books $j = n_1 + 1..n_1 + n_2$ to be purchased at the availability constraint, and books $j = n_1 + n_2 + 1..n_1 + n_2 + n_3$ to be purchased at a quantity less than the availability constraint in period t. These conditions will be used to construct the observational likelihood of consumers' purchases in the second stage. The term $s_{it} = \sum_{j=n_1+1}^{n_1+n_2} c_{ijt} + \sum_{j=n_1+n_2+1}^{n_1+n_2+n_3} x_{ijt}$, which represents the total number of chapters purchased by consumer i in period t. Note that the budget term M_{it} does not appear in the conditions above because in a quasi-linear utility function, M_i enters the Lagrangian as a constant and disappears in the Kuhn Tucker conditions. The randomness in these conditions stems from the randomness in the baseline preference a_{ijt} as given in equation (4.2).

4.2. First Stage: Platform Visit Decision

In the first stage of period t, a consumer decides whether to visit the platform. The decision to visit the platform depends on the consumer's expected utility of purchases at the platform and the cost of visiting the platform. If we were to model the platform visit decision in a structural fashion, we would need to calculate a consumer's expected utility of purchases over all possible combinations of purchases and non-purchases across the 138 books in the choice set, which would be of the order of 2^{138}. This is computationally prohibitive. As an alternative, in the same spirit as Ching, Erdem, and Keane (2009) and Mehta et al. (2017), we model the first-stage decision in a reduced-form manner. We specify consumer i's decision of visiting the platform in period t as:

$$\begin{aligned} Visit_{it} &= 1 \quad if\ \gamma_0 + Z_{it}\gamma_1 + \eta_{it} \geq 0 \\ &= 0 \qquad \text{otherwise} \end{aligned} \quad (4.9)$$

where η_{it} is a standard normal random variable. We allow η_{it} to be correlated with each of the errors ε_{ijt} in the first stage with a correlation of ρ_η. The term Z_{it} represents the following list of variables related to the consumer's recent visits and the state of the books that she has purchased: (1) whether day t is during the weekend; (2) consumer's visit frequency in the past 7 days; (3) log of the number of notifications received by the consumer about the release of new chapters of those

books that she is up to date with (log is used to capture the concavity of the notification effect); (4) log of the number of notifications received by the consumer about the release of new chapters of those books that she is *not* up to date with; (5) her total inventory over all books she is reading; (6) log of average of total number of chapters purchased by the consumer till time *t* across books the consumer is currently reading. This captures the average extent to which the consumer is captivated by the books she is currently reading; (7) log of average of the time elapsed since the last purchase of books the consumer is currently reading. This captures the average extent to which the consumer has lost interest in the books she is currently reading.

4.3. Consumption Rate and Inventory Updating

Recall that the consumer's decisions in both stages are functions of inventories of books. In this section, we specify the consumer's inventory of each book in a given period based on the inventory constraint as per which a consumer will purchase more chapters of a book only if she has read all the previously purchased chapters of that book. Define the period $l_j(t)$ as the last time before period t that consumer i makes a purchase of book j (we suppress subscript i for notational convenience). Let $x_{j,l_j(t)}$ be the amount purchased by the consumer in her most recent purchase in period $l_j(t)$. Based on the identification condition discussed above, $x_{j,l_j(t)}$ is also the consumer's inventory of book j immediately after her purchase of book j in period $l_j(t)$. We let CR_{jt} denote the consumption rate of book j in period t and specify the consumption rate as follows:

$$CR_{jt} = \min\left\{a + b \cdot x_{j,l_j(t)} + \sigma_c \varepsilon_{c,jt}, Inv_{jt}\right\} \tag{4.10}$$

where the first part of the expression, $a + b \cdot x_{j,l_j(t)} + \sigma_c \varepsilon_{c,jt}$,[8] indicates the unconstrained per-period consumption rate of book j between two consecutive purchases of chapters of the book. In the linear expression, a is the baseline consumption rate, b measures the impact of the most recent purchase quantity on the consumption rate, and $\sigma_c \varepsilon_{c,jt}$ is the random error in the consumption rate, where $\varepsilon_{c,jt}$ is a standard IID normal random error and σ_c is the standard deviation of that error. A positive b implies that the consumption rate increases with the quantity last purchased. We allow for the unconstrained consumption rate of book j to vary depending on the quantity of the last purchase. However, we assume it remains constant within two consecutive purchases of that book.

[8] We also tried a quadratic term for $x_{j,l_j(t)}$ in the consumption rate but found its coefficient to be insignificant.

The reason is because we cannot identify the variation in consumption rate of a book between two consecutive purchases of it from the data. The second part of the expression, Inv_{jt}, indicates that the consumer can consume up to Inv_{jt}. This ensures that inventory of book *j* is not negative. Given the consumption rate in Equation (4.10), the inventory of book j at period t is $Inv_{jt} = Inv_{j,t-1} - \min\left\{a + \sigma_c \varepsilon_{c,jt} + b \cdot x_{j,l_j(t)}, Inv_{j,t-1}\right\}$, where $Inv_{j,l_j(t)} = x_{ij,l_j(t)}$ immediately after the purchase of book j in $l_j(t)$ period. This yields the inventory of book j in period t as follows:

$$Inv_{jt} = \mathbb{I}_{\left(x_{j,l_j(t)} \cdot (1 - b \cdot \Delta t_{j,t}) - (a + \sigma_c \varepsilon_{c,jt}) \Delta t_{j,t} > 0\right)} \cdot \left[x_{j,l_j(t)} \cdot \left(1 - b \cdot \Delta t_{j,t}\right) - (a + \sigma_c \varepsilon_{c,jt}) \Delta t_{j,t}\right] \quad (4.11)$$

where $\Delta t_{j,t} = t - l_j(t)$ is the number of elapsed periods since the most recent purchase of book *j* and $\mathbb{I}$ is an indicator variable.

4.4 Discussion of the Three Effects

In this section, we discuss how our model captures the three effects. We model the additional platform visits effect via two parameters: (1) parameter γ_1 in Equation (4.9) which captures the impact of the notifications of new chapter releases of books (that the consumer is currently reading and is up to date with their progress) on the probability of visiting the platform in the first stage and (2) α_i^T in the baseline preference α_{ijt} in Equation (4.2) which captures the impact of a book being on the platform's homepage on consumer purchases in the second stage. Both variables having a strong impact implies that the release of new chapters of ongoing books leads to additional platform visits, which leads to the purchase of new books promoted on the platform's homepage. We model the rationing effect via two ways in the second stage: (1) availability constraint, which is captured by the magnitude of c_{ijt} in Equation (4.4) and (2) the degree of substitutability between books, which is captured by parameter ρ in Equation (4.3). The smaller the value of c_{ijt} and the greater the substitutability between the books, the greater is the rationing effect.

Our model captures two aspects of binge consumption effect, viz., the binge consumption effect in a period and whether the binge consumption in a period begets binge consumption in the next period. *First*, the binge consumption effect in a period is modeled using the following three terms: (1) the baseline preference for a book (α_{ijt}); (2) the diminishing marginal return parameter in the quadratic utility (β) in Equation (4.1); and (3) parameter b of the consumption rate in Equation (4.10), which measures the impact of the number of a book's chapters last purchased on its consumption rate. A lower value of β or a higher value of α_{ijt} in the utility implies that

conditional on purchasing a book and its availability on the platform, the consumer purchases large number of chapters of that book on a given visit. A positive value of *b* indicates that the consumer binge consumes the book if she buys large number of chapters. Putting these three factors together, a lower value of β, a higher value of α_{ijt}, and a larger value of *b* lead to an increased consumption rate in the period, thus indicating binge consumption. *Second*, regarding how much the binge consumption in one period could beget binge consumption in the next period, we specify the relation in the second-stage model using the impact of a consumer's consumption rate of the focal book in the last period on her baseline preference in the present period, as in Equation (4.2).

5. Estimation

We first discuss the parameters in our model. We estimate the following parameters in the first-stage model given in Equation (4.9): (i) intercept γ_0; (ii) coefficients of various variables in Z_{it}, γ_1; (iii) correlation between the errors in the first and the second stage, ρ_η; and (iv) coefficient of the control function. We estimate the following parameters in the inventory updating mechanism given in Equation (4.15): (i) baseline consumption rate, *a*; (ii) the impact of the quantity last purchased on the consumption rate, *b*; (iii) standard deviation of the error in inventory specification, σ_c.

For the second-stage model given in Equation (4.5), we estimate the following parameters: (i) the diminishing returns of the purchase, β; (ii) intercept α_{i0}, which is common to the baseline preference of all books; (iii) the coefficient of the indicator for "whether all of the chapters of the focal book are published" in the baseline preference; (iv) the coefficient of the indicator for "whether the focal book is currently on the platform's home page" in the baseline preference; (v) the coefficient of the variable "consumer's consumption rate of the focal book in the last period" in the baseline utility; (vi) the coefficient of the indicator for "whether the consumer purchased any chapters of the focal book before" in the baseline utility; (vii) the coefficient of the variable "the number of chapters of the focal book already purchased by the consumer" in the baseline preference; (viii) the book-specific dummy, α_j, in the baseline preference for each book j = 2...139; (ix) the degree of substitutability between books, ρ; (x) the correlation in the econometrician's errors across all books, ρ_ε; and (xi) the coefficient of the control function. an indicator of whether the platform's homepage promotes book *j* at time *t*; (xi) coefficient of whether the consumer visited the platform after receiving a notification(s) from a book she is currently following and is up to date with the progress; (xii) coefficient of the interaction variable which captures the extent to

which the purchases of books promoted on the platform will differ across organic platform visits and platform visits motivated by notifications.

We incorporate unobserved heterogeneity into the parameters by considering latent class segmentation. The candidate parameters for the basis of latent segmentation are those that vary the most across consumers and are related to the release strategy. We allow the following parameters to vary across the segments: parameters (i)–(ii) in the first stage and parameters (i)–(v) in the second stage. We do not allow for unobserved heterogeneity in the 138 book specific intercepts since that will lead to an explosion in the number of parameters. We estimate all the parameters simultaneously using maximum likelihood. Finally, since the release speed of books is endogenous, it follows that the number of available chapters in the availability constraint, c_{ijt}, in the second stage, as well as the number of notifications in the first stage will also be endogenous as both are functions of the release speed. To deal with this endogeneity, we use the control function approach (Petrin and Train 2010), in which we use the same instrument for release speed as the one described in Section 2.1. In section 6 of the Web Appendix, we have provided a detailed discussion of the formulation of the likelihood function.

Identification. We discuss the identification strategy for the key parameters in the model. First, the identification of the parameters of the baseline utility term, α_i, and the parameter capturing the diminishing returns of purchasing, β_i, comes from the change in the release speed of books over time. Typically, in the literature of discrete continuous models, α_i and β_i are identified by the impact of changes in a product's prices on its own demand. However, we do not observe price variation in our setting. Instead, we observe variations in the release speed of new chapters both across books and over time for a given book. Thus α_i and β_i are identified by the impact of a change in the release speed of an ongoing book on its own demand. If consumers purchase more chapters from the focal book when its release speed increases, high α_i and low β_i are implied. Specifically, α_i is identified by the impact of the release speed change on both the purchase quantities and the purchase incidence of the focal book. β_i is identified by the impact of the release speed change on the purchase quantities of the focal book, given the availability constraint.

Second, the identification of the substitution term ρ comes from two types of variations—change in the release speed of books and the change in the choice set over time. The first variation refers to the impact of the release speed change of a focal book on the consumer's decision to buy

other books. If consumers purchase an increased number of books when the release speed of the focal book decreases, higher substitutability is implied. The second variation refers to the following two scenarios: (a) a consumer's choice set expands when new books join the platform and (b) a consumer's choice set shrinks when she finishes buying all of the book's chapters and the book naturally drops out of her choice set. If a change in the choice set size leads to a significant change in the total number of books a consumer buys, high substitutability is implied.

Third, the identification of parameters in the consumption rate specifications, a and b, comes from the inventory constraint, which assumes that a consumer does not purchase more of a book's chapters unless she has read all the book's chapters that she purchased previously. This implies that a book's inventory level for a consumer (i.e., number of chapters that she has purchased but not yet read) immediately before a purchase order is zero. Similarly, a book's inventory level for a consumer immediately after a purchase order is the same as the order size for the book (i.e., the number of book chapters bought). This information gives us data points on inventory levels at different points in time, which then helps us quantify the consumption rate.

Finally, regarding the identification of the parameters in the first stage, recall that in our data, we observe whether the consumer makes a platform visit in each period. This allows us to identify the parameters in the first stage.

6. Estimation Results

The parameter estimates are reported in Table 8. We choose the model with four segments that maximizes the BIC. The top panel of Table 8 presents the weights of the four segments. The remaining parameter estimates are grouped into the following four parts: the baseline utility parameters from the second stage, other utility parameters in the second stage, parameters in the first stage, and consumption rate parameters.

Baseline Utility Parameters. We first look at the baseline utility parameters that are common across the four segments. Regarding the past purchase history, we include four variables: (1) whether the reader previously purchased the book, α_5; (2) the number of days since the most recent purchase of the book, α_6; (3) whether the reader purchased the book in the past 7 days, α_7; and (4) the log of the total number of chapters in a book purchased by the reader before the current period, α_8. All four of the parameter estimates are significant. The estimates $\hat{\alpha}_5 = 0.15$ and $\hat{\alpha}_7 = 0.71$ indicate that all else being equal, the consumer is more likely to purchase a book that she has

previously purchased, especially if the purchase is recent. The negative estimate of $\hat{\alpha}_6 = -1.17$ means that the preference for the book weakens as the time lapse between two consecutive purchases increases. This implies that the slower the release speed, the greater the likelihood that a consumer will lose interest in the book. The estimate $\hat{\alpha}_8 = 0.52$ is positive and significant, suggesting that the more chapters that a consumer reads, the greater is her propensity to continue purchasing the book. This suggests that the greater the release speed, the larger will be the number of chapters that the consumer can purchase on a trip, and consequently the greater will be the extent to which she will get captivated by the book and purchase more chapters in the future. We also include (1) the variable of whether the consumer visited the platform after receiving a notification(s) from a book she is currently following and is up to date with the progress (α_{10}), which captures the extent to which the purchase of books will differ across organic platform visits and platform visits motivated by notifications, and (2) an interaction variable $\hat{\alpha}_{11}$ which captures the extent to which the purchases of books promoted on the platform will differ across organic platform visits and platform visits motivated by notifications. The positive and significant estimates $\hat{\alpha}_{10} = 1.20$ and $\hat{\alpha}_{11} = 0.04$ suggest that visits motivated by notifications are more likely to induce purchase and specifically purchase of books promoted on the homepage.

Next, we discuss the heterogeneous parameters in the baseline utility: (1) the preference for completed books, α_1; (2) the effect of ranking (homepage promotion) on choosing a new book, α_2; (3) the consumption rate of the book in the prior week , α_3. The results show that, first, all segments have a positive preference for a completed book over an ongoing book, after controlling everything else, whereas segment 2 and 4 have the stronger tendency to choose completed books. Second, all segments react positively to a book that is promoted on the platform's homepage, Fourth, the sensitivity to a book's recent sales (popularity) varies cross segments - segment 2 and 4 are less likely than other segments to purchase the most popular book. This outcome is likely related to the fact that segment 2 and 4 has a stronger preference for completed books. Fifth, the estimate of the effect of the previous week's consumption rate is positive across all four segments. This suggests the binge effect across periods, an intertemporal effect in which reading a book in the past leads to reading more chapters of the same book in the future.

Other Parameters in the Utility in the Purchase Stage. The quadratic utility parameter (β) is positive and significant in all four segments. The substitution parameter ρ is assumed to be

homogenous across segments. The estimate $\hat{\rho} = 0.04$ is positive and small, suggesting that the books are relatively weak substitutes; thus, the rationing effect is small.

Parameters in the Platform Visit Stage and the Consumption Rate Equation. Recall that in the first stage, we model a consumer's decision to visit the platform using a probit model. First, the state dependence term for the platform visit decision is positive for all segments ($\hat{\gamma}_1 > 0$), suggesting that frequent visits in the past week lead to a higher chance of visiting the platform in the current period. Second, the relation is negative between the level of inventory stock of chapters and the likelihood of visiting the platform. Third, we have two notification/new updates variables associated with different sets of books—notifications received by the consumer about the release of new chapters of those books that she is up to date with(γ_2) and notifications received by the consumer about the release of new chapters of those books that she is not up to date with (γ_3). The results show a stronger and positive effect of the updates for recently purchased books ($\hat{\gamma}_2 > \hat{\gamma}_3$). These estimates imply that one notification for a recently purchased book can increase the platform visit probability in a day by an average of 19%. This result, along with the earlier result that the probability of purchasing a given book could increase if the book is promoted on the homepage, imply that the additional platform visits effect is large.

Finally, in the consumption rate equation, the estimate of $\hat{a} = 54.3$ and $\hat{b} = 0.69$ indicates that the consumption rate of a book is positive and significantly related to the recent purchase size. This suggests that the within period binge effect will be significant if the consumer purchases a large number of chapters of a book in a period.

Table 8 Parameters Estimates with Four Segments of Readers

Segments:	(1)	(2)	(3)	(4)
Segment weights	0.28 (0.02)	0.27 (0.02)	0.26 (0.01)	0.20 (0.02)
Baseline Utility Parameters				
Constant (α_0)	-2.08 (0.05)	-2.04 (0.04)	-1.89 (0.05)	-1.57 (0.04)
Completed book (α_1)	0.88 (0.04)	1.09 (0.03)	0.79 (0.05)	0.92 (0.04)
Promoted on homepage (α_2)	0.42 (0.03)	0.09 (0.04)	0.32 (0.03)	0.16 (0.04)
Consumption in the last 7 days (α_3)	0.04 (0.01)	0.05 (0.01)	0.15 (0.02)	0.08 (0.01)
Book age (α_4)	-0.65 (0.02)			
Whether purchased before (α_5)	0.15 (0.04)			

Time since last purchase of the book (α_6)	-1.17 (0.07)			
Whether purchased in last 7 days (α_7)	0.71 (0.03)			
Log (No. of chapters purchased before) (α_8)	0.52 (0.02)			
Weekend (α_9)	0.05 (0.06)			
Visited after notification from a recently bought and up-to-progress book (α_{10})	1.20 (0.03)			
Visited after notification from a recently bought and up-to-progress book * Promoted on homepage (α_{11})	0.04 (0.01)			
Second-stage control function (α_λ)	-0.85 (0.04)			
Other Purchase Stage Parameters				
Quadratic utility parameter (β)	0.06 (0.0003)	0.02 (0.0003)	0.10 (0.0004)	0.03 (0.0002)
Substitution parameter (ρ)	0.04 (0.002)			
Correlation in errors across books (ρ_ε)	0.06 (0.01)			
Correlation in errors between two stages (ρ_η)	0.22 (0.02)			
Platform Visit Stage Parameters				
Constant (γ_0)	-1.30 (0.03)	-1.39 (0.04)	-1.31 (0.04)	-1.41 (0.03)
Visit frequency in last 7 days (γ_1)	0.29 (0.02)	0.26 (0.02)	0.25 (0.03)	0.16 (0.03)
Notifications of up-to-date books (γ_2)	0.79 (0.01)	0.70 (0.01)	0.71 (0.01)	0.75 (0.01)
Notifications of not up-to-date books (γ_3)	0.20 (0.01)	0.05 (0.01)	0.24 (0.01)	0.09 (0.01)
Inventory (γ_4)	-0.10 (0.02)			
Weekend (γ_5)	-0.03 (0.002)			
Average No. of chapters purchased before of all recently bought books (γ_6)	-0.003 (0.001)			
Average time since last purchase of all recently bought books (γ_7)	-0.36 (0.02)			
First-stage control function (γ_λ)	0.06 (0.04)			
Consumption Rate Parameters				
Consumption rate constant (a)	54.3 (5.29)			
Consumption rate slope (b)	0.69 (0.08)			

Note: Book-fixed effects are included in the baseline utility. Standard errors are in parentheses.

7. Counterfactuals

We perform counterfactuals that evaluate the platform's revenues under alternative release strategies to address the following questions: (1) Should the platform release chapters of books simultaneously or sequentially? (2) If released sequentially, what should be the release speed that maximizes platform revenue? (3) Can the platform increase its revenue by using a hybrid release

strategy that combines simultaneous and sequential releases? In performing the counterfactuals, we assume that the platform can mandate writers on when to release their new chapters. Support for this assumption comes from the fact that the platform under study has more than 70% of the Chinese market share in romance genre, giving it enough power to enforce the release strategies. Further, the platform signs exclusive contracts with the writers which does not allow the writers to write for any other platform. We also assume that given the counterfactual changes in the platform's decisions, each writer continues writing for the platform and their creative production decisions, including the quality of the chapters and initial release times, remain unaffected. We have provided justifications for this assumption in section 7 of the Web Appendix.

Because the platform has a pay-per-chapter model, we measure the platform performance of each counterfactual based on the simulated purchases of consumers of our sample size across all books over our sample period using the estimates of the structural model. Table 9 reports all the results of our counterfactual analysis. We use (I) to index the existing release strategy in which the platform releases the chapters sequentially but lets the authors decide when to release new chapters, (II) to index the simultaneous release strategy, (III) to index the sequential release strategy with a constant release speed, (IV) to index the hybrid release strategy in which we have simultaneous release for the first half of the book and sequential release for the other half, and (V) to index the hybrid release strategy in which we have sequential release for the first half of the book and simultaneous release for the other half.

For each release strategy, we focus on four measures of outcomes: (a) total number of chapters purchased across all consumers during the entire year, (b) total number of platform visits across all consumers during the year, (c) average number of chapters bought during a visit with a positive purchase, and (d) the total number of books explored (i.e., the number of times the consumer purchases at least one chapter). Outcome (a) is the direct measure of the platform's overall sales during the entire year. Outcome (b) is the measure of platform visits, which is linked to the additional platform visits effect. Outcome (c), which is the average number of chapters purchased per visit, is related to the binge consumption effect. Outcome (d) is the measure of book purchases, which captures the general product exploration effects of additional platform visits. Since the estimate of the substitutability between books is small, we do not focus on the rationing effect. Next, we explain each of the counterfactuals and their corresponding results.

Simultaneous vs. Current Release Strategies. A simultaneous release strategy (II) releases all of a book's chapters in a single period. When simulating such a policy, the window of release for each book decreases from several months, as observed in the data, to a single day. We need to decide when the release date should be in the counterfactual. We consider two choices of release time, either the first day or the middle of a book's actual release window in our data to ensure that the counterfactual result is not driven by the choice of publication time. Under simultaneous release strategy, all books are completed when they are released, and the consumers do not receive any notifications of new chapter releases. Note that although all chapters are released simultaneously, we allow for consumers to purchase chapters at their own pace.

The results show that simultaneous release strategy (II) leads to 18% revenue decrease from the existing sequential strategy (I). Under the simultaneous strategy, on one hand, the number of chapters purchased in each platform visit is higher than under the existing strategy, specifically, 7.3 chapters vs. 5.4 chapters per visit with purchase, respectively. Since consumption rate increases with the purchase quantity, this suggests a higher binge effect under simultaneous strategy than under the existing sequential strategy (I). On the other hand, due to the additional platform visits effect, consumers are less likely to return to the platform, resulting in a much lower number of platform visits under the simultaneous strategy (43,761 vs. 64,812 visits under the existing strategy). Thus, the additional platform visits effect outweighs the binge consumption effect as we move from the existing sequential strategy to the simultaneous strategy.

Optimal Sequential Release Strategy. Here we examine the sequential release strategy (III) in which the release speed is the same over time and across all books. This release strategy is used by platforms such as Webtoon requires its creators to release a fixed number of chapters each week. In the counterfactual, we anchor a book's release window around the mid-release period[9] of the book's original release window. We find that the sequential release strategy (III) with a constant speed of 5 chapters per week leads to the highest demand (331,864 chapters), a 6.8% improvement in platform revenue over the existing release strategy (I). Comparing this to various measures under release speeds of four chapters per week and six chapters per week, we find five chapters per week to be the optimal point, as the two opposing effects (binge and platform visits)

[9] We tried different initial release periods in our simulation and found the counterfactual results to be robust.

offset each other. For the platform visits effect, under the optimal speed, there is an increase in platform visits (9.3%) and an increase in books purchased (6.3%) compared to the existing strategy. For the binge effect, the average binge consumption measured by chapters per visit with purchase is 5.2, compared to 5.4 under the existing strategy (I).

Hybrid Release Strategies (IV and V). We first consider a hybrid release strategy (IV) in which the platform releases the first half of a book's chapters at once and the second half of the chapters sequentially at a constant speed.[10] To the best of our knowledge, no platform has systematically implemented a hybrid strategy. However, this strategy is inspired by the recent practice of some TV shows. For example, consider the first two seasons of The Boys, released by Amazon Prime Video. In the first season, the platform released all eight episodes simultaneously. In the second season, the platform released the first three episodes in one day and the remaining five episodes once per week. While conducting the counterfactual, each book's initial release period is anchored at the middle of the publishing cycle of the book's original release window in the data so that it is close to the writing progress in the original data – since the split ratio is 50%-50% for the hybrid strategy, anchoring at middle of the cycle ensures we do not impose a very different schedule compared to the observed production progress in data.[11]

Table 9 Counterfactuals on Platform-Wide Release Strategy

	(a)	(b)	(c)	(d)	(e)	(f)
	Chapters	Visits	Visits with purchase	Chapters per visit with purchase	Books per visit with purchase	Books explored
(I) Existing release strategy						
	310,808	64,812	57,557	5.4	1.71	4,959
(II) Simultaneous release (releasing all chapters at once)						
All in initial period	**254,821**	43,761	34,907	7.3	1.82	4,896

[10] See section 8 of the Web Appendix for more counterfactuals on the hybrid release strategy with different split ratios for the simultaneous and sequential parts. Our analysis suggests that 50%-50% split (between simultaneous first and sequential later) is close to the optimal split that maximizes the platform revenue.

[11] A simple comparison is as follows. Under original release strategy, half of the book would have been written and released by the date of mid-release period. Under 50%-50% hybrid release strategy, we will also have half of the book being written and released by same date.

All in mid period	238,931	43,761	34,133	7.0	1.80	5,004
(III) Sequential release at a constant speed						
4 chapters/week	322,536	75,867	67,195	4.8	1.82	5,192
5 chapters/week	**331,864**	70,868	63,820	5.2	1.81	5,164
6 chapters/week	322,682	66,533	59,756	5.4	1.77	4,999
(IV) Hybrid release: Simultaneous for the first 50% of chapters + sequential for the rest at a constant speed						
1 chapter/week	358,524	85,182	79,672	4.5	2.06	4,544
2 chapters/week	**449,175**	77,843	77,444	5.8	1.96	5,320
3 chapters/week	393,545	67,965	63,475	6.2	1.82	4,966
(V) Hybrid release: Sequential for the first 50% of chapters at a constant speed + simultaneous for the rest						
4 chapters/week	251,767	60,796	51,381	4.9	1.73	4,885
5 chapters/week	**254,155**	57,793	48,876	5.2	1.72	4,716
6 chapters/week	252,779	55,547	46,811	5.4	1.71	4,680

The hybrid strategy (IV) yields a much higher revenue than other release strategies. The optimal release speed is 2 chapters per week for the second half of the book. For this release speed, the expected revenue is 45% better than the existing strategy. The optimal release speed for the second half of the hybrid strategy is nearly half of the optimal speed of the sequential release strategy (III), which indicates that the total duration of book release is similar between the two cases. Interestingly, in hybrid strategy (IV), there is an increase in both the binge consumption and additional platform visits effects over the existing release strategy. The reason for this increase is that releasing the first half of the book simultaneously enables consumers to binge consume, which then increases their probability of being captivated by it. Consequently, despite the slower release speed for the second half (which would have otherwise resulted in consumers to lose interest in the book), the consumers would continue reading the book, and would thereby also explore new books on the platform because of the additional platform visits effect. In contrast, in the hybrid strategy (V) in which the first half of chapters are released sequentially and the rest are released simultaneously, the platform revenues are even lower than the current release strategy. This is because sequential release for the first half results in consumers to lose interest in the book; and simultaneous release for the second half does not increase binge consumption much as many readers had quit reading the book in the first half. This shows that hybrid release, if designed properly, can turn binge consumption and platform visits effects into complementary forces.

There are two caveats regarding our counterfactual analysis. First, our analysis suggests that the reason why sequential release outperforms simultaneous release is because the additional platform visits effect outweighs the binge consumption effect. Recall that the additional platform visits effect stems from the notifications sent by the platform whenever new chapters are released, which then motivates its current readers to visit the platform. A natural question that arises here is whether the platform could instead follow a simultaneous release, but it could send other types of notifications to consumers to attract them to the platform, such as sending a notification when a new book being released? Our conversations with the platform managers suggest that notifications about new books being released are lot less effective as compared to notifications of new chapters being released from an ongoing book that the consumer is currently reading. This is because if a consumer is reading an ongoing book, she is invested in it and is motivated to purchase newly released chapters of that book as compared to the case when a new book is released.

The second issue is whether moving from the existing strategy (I) to sequential release strategy (III) or hybrid release strategy (IV) can create operational difficulties for the platform. Recall that in the existing release strategy, the chapters are released sequentially, with an average of 6.4 chapters per week, while in the sequential release strategy (III), the optimal release speed is 5 chapters per week. This means that on average, the writer's release speed in the current strategy (I) is not far from the optimal release speed in the sequential release strategy (III). Thus, if the platform were to adopt the optimal sequential release strategy (III), the cost of change to the writers would be small. As for the hybrid release strategy, no platform has systematically implemented it. Platforms may anticipate resistance or costs associated with implementing this practice. Because a hybrid release strategy requires saving up content before the initial release, the creator must delay the product release. In return, the platform may need to compensate its creators for revenue losses due to discounting. However, our counterfactuals show that the revenue upside of the hybrid strategy is sufficiently large that its benefits are likely to outweigh its implementation costs.

8. Conclusions

The digitalization of media products permits more flexible release strategies for serial products, creating both new opportunities and new challenges for platforms in selecting the proper release strategy. To make the right decision, it is important to have a good understanding of how the market will respond to the alternative release strategies (simultaneous vs. sequential) and

different release speeds. We develop a two-stage structural model to estimate three effects of release speed on consumer behavior: the binge consumption, rationing, and additional platform visits effects. Our empirical analysis of an online book publishing platform shows the evidence for each of these three effects. All else being equal, with a faster release speed, the binge consumption effect suggests a greater consumption of books, and the other two suggest a smaller variety of books purchased by consumers. In our results, the product exploration effect is mainly driven by additional platform visits rather than by rationing and substitution between books.

Our counterfactual analysis highlights the interaction between the product exploration effect and the binge consumption effect under alternative release policies. Our results show that the simultaneous release strategy performs worse than the sequential release strategy. This is because the additional platform visits effect plays a stronger role than the binge consumption effect as we move from sequential to simultaneous strategy. This suggests that platform managers should primarily aim to build long-term consumption habits among platform users, which is more important than providing users with a few short bursts of consumption experiences and ensuring that they finish their chosen series on hand. Furthermore, our results show that the platform does not have to follow a pure strategy, since both are outperformed by a hybrid release strategy. This is because the hybrid strategy turns the otherwise opposing effects into complementary effects.

Future research could extend our framework to not only different digital media platforms but also those with different business models. For instance, consider TV series on Netflix and podcasts on I-Tunes. Even though it is possible that the optimal release strategies for TV series and podcasts are different from that for digital book platforms, however, the framework that we have introduced in the paper would still apply to TV series and podcasts. This is because the decision to go for sequential vs. simultaneous vs. hybrid release strategies for TV series and podcasts would depend on the relative magnitudes the three effects that we have identified in our paper. Similarly, whereas the online book platform under study has a pay-per-chapter model, many digital platforms such as Netflix follow a subscription-based pricing model. Extending our paper to this context requires adding a consumer decision on whether to renew a subscription to our model. We expect both binge consumption and platform visits effects of the release speed to remain significant under the subscription model. However, unlike the pay per chapter model in which the annual business performance is measured by the overall demand over the year, the business performance in the annual subscription model will be assessed by the number of consumers whose willingness to pay

is greater than the value they anticipate from consumption of books during the year. Lastly, this paper is silent on the impact of release strategy on the entry of new customers or market expansion, as our dataset only contains the existing customers. It is likely that simultaneous release could attract new customers to the platform as the consumers have higher taste for binge consumption.

References

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

Ching, A., Erdem, T., & Keane, M. (2009). The price consideration model of brand choice. *Journal of Applied Econometrics, 24*(3), 393–420.

Ching A., and M. Osborne (2020), "Identification and Estimation of Forward Looking Behavior: The case of Stockpiling", *Marketing Science,* 39(4), pp. 707-726

Chintagunta, P. K. (1993), "Investigating Purchase Incidence, Brand Choice and Purchase Quantity Decisions of Households," *Marketing Science*, 12(2), 184-208.

CNNIC. (2019). *The 44th report on the Internet development in China.* Chinese Internet Information Center.

Ding, A.W, and S. Li (2019), "Herding and the Consumption of Digital goods and Moderators of the Herding Bias," Journal of Academy of Marketing Science, 47, pp. 460-478

Godinho de Matos, M., & Ferreira, P. (2020). The effect of binge-watching on the subscription of video on demand: Results from randomized experiments. *Information Systems Research, 31*(4), 1337–1360.

Lal, R., & Matutes, C. (1994). Retail pricing and advertising strategies. *Journal of Business*, 345–370.

Lazear, E. P. (1995). Bait and switch. *Journal of Political Economy, 103*(4), 813–830.

Lee, L.-F., & Pitt, M. M. (1987). Microeconometric models of rationing, imperfect markets, and non-negativity constraints. *Journal of Econometrics, 36*(1–2), 89–110.

Li, X., Gu, B., & Liu, H. (2013). Price dispersion and loss-leader pricing: Evidence from the online book industry. *Management Science, 59*(6), 1290–1308.

Li, X., C. Liao, Y. Xie. (2020), "Digital piracy, creative productivity, and customer care eort: Evidence from the digital publishing industry," *Marketing Science* forthcoming

Li, X., Shi, M., & Zhao, C. Y. (2021). *Incentivizing mass creativity: An empirical study of the online publishing market*. SSRN working paper: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3842153

Lu, J., Bradlow, E., & Hutchinson, J. (2019). *Multiple dimensions of binging: The hidden costs and benefits*. SSRN working paper: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3493759

Lu, J., Karmarkar, U., & Venkatraman, V. (2019). *Planning to binge: How consumers choose to allocate time to view sequential versus independent media content.* Wharton working paper: https://faculty.wharton.upenn.edu/wpcontent/uploads/2017/06/Lu_Karmarkar_Venkatraman_June2017.pdf

Mehta, N. and Y. Ma (2012), , "A Multi Category Model of Consumers' Purchase Incidence,

Brand Choice and Quantity Decisions: Methodological Issues and Implications on Promotional Decisions," *Journal of Marketing Research*, 49 (5), 435 -451

Mehta, N., Ni, J., Srinivasan, K., & Sun, B. (2017). A dynamic model of health insurance choices and healthcare consumption decisions. *Marketing Science, 36*(3), 338–360.

Merikivi, J., Bragge, J., Scornavacca, E., & Verhagen, T. (2020). Binge-watching serialized video content: A transdisciplinary review. *Television and New Media, 21*(7), 697–711.

Nair, H., Dubé, J.-P., & Chintagunta, P. (2005). Accounting for primary and secondary demand effects with aggregate data. *Marketing Science, 24*(3), 444–460.

Neary, J. P., & Roberts, K. W. (1980). The theory of household behaviour under rationing. *European Economic Review, 13*(1), 25–42.

Petrin, A., & Train, K. (2010). A control function approach to endogeneity in consumer choice models. *Journal of Marketing Research, 47*(1), 3–13.

Pittman, M., & Sheehan, K. (2015). Sprinting a media marathon: Uses and gratifications of binge-watching television through Netflix. *First Monday*.

Schweidel, D. A., & Moe, W. W. (2016). Binge watching and advertising. *Journal of Marketing, 80*(5), 1–19.

Villas-Boas, J. M., & Winer, R. S. (1999). Endogeneity in brand choice models. *Management Science, 45*(10), 1324–1338.

Wales, T. J., & Woodland, A. D. (1983). Estimation of consumer demand systems with binding non-negativity constraints. *Journal of Econometrics, 21*(3), 263–285.

Wang, Z., & Chern, W. S. (1992). Effects of rationing on the consumption behavior of Chinese urban households during 1981–1987. *Journal of Comparative Economics, 16*(1), 1–26.

Wilkie, W. L., Mela, C. F., & Gundlach, G. T. (1998). Does "bait and switch" really benefit consumers? *Marketing Science, 17*(3), 273–282.

Zhang, S., Chan, T., & Luo, X. (2022). Time-inconsistent preferences and strategic self-control in digital-content consumption. *Marketing Science, 41*(3), 616–636.

Web Appendix for

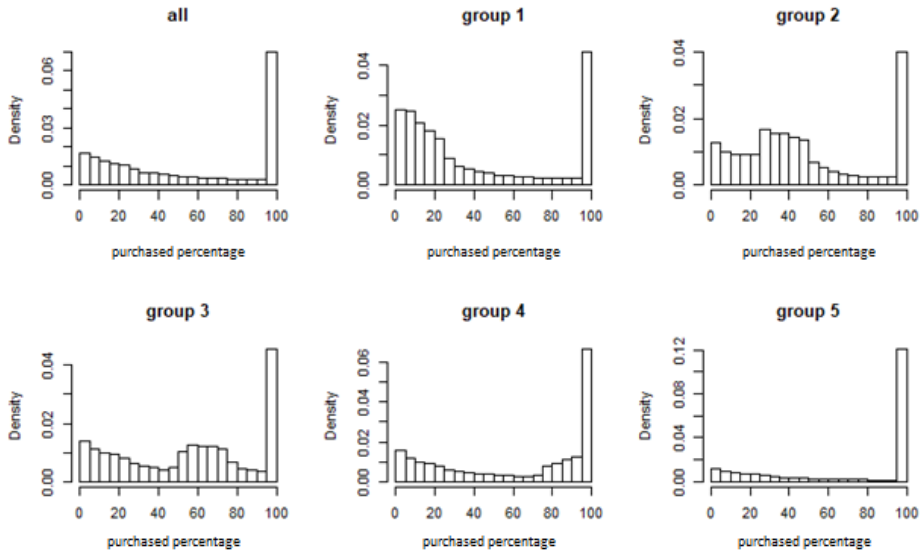
The Consumption of Serial Media Products and the Optimal Release Strategy

Table of Contents

Section 1. Chapter Availability at the time of First Purchase on Subsequent Purchases

We explore the variation in the time when a consumer starts reading a book during the publishing process. We separate the consumer–book pairs into five groups based on the proportion of chapters available when the consumer starts purchasing the book: group 1 starts to purchase before 25% of the chapters are released, group 2 starts to purchase when 25% to 50% of the chapters are released, group 3 starts to purchase when 50% to 75% of the chapters are released, group 4 starts to purchase when 75% to 100% of the chapters are released, and group 5 starts to purchase after all of the chapters are released. The latter groups have more available chapters at the time that purchases begin.

Figure A1 Distributions of the Percentage of Chapters Purchased



We check whether higher availability (groups 1–5) leads to a higher purchase rate. In Figure A.1, we plot the histogram of the percentage rate of a book both for the entire sample and for each group. First, the figure shows that the 100% bar is the highest for group 5, the second highest for group 4, and much lower for groups 1–3. This means that the latter groups (with higher availability) tend to read a higher percentage of a book than the earlier groups. Second, once the consumers catch up with a book's publication, they are more likely to stop purchasing future chapters than before they catch up. For example, group 2 is more likely to stop reading between 25% and 50% (where they are likely to have caught up with the chapter releases) than between 0% and 25% (where these chapters are already released at the beginning of their purchase), and a similar pattern holds for groups 3 and 4. Thus, consumers are less likely to stop reading when many chapters are available. Third, the distributions tend to dip after finishing the initial percentage of available chapters. For example, the purchase rate of group 1 is mainly distributed either before 25% or 100% and rarely between 25% or 100%. This seems to suggest that consumers who decide to follow a book after catching up with its publication are likely to keep reading to the end.

Section 2. Impact of Release Speed Affect Chapter Quality (Experienced vs. Inexperienced Writers)

For a given book, the release speed of chapters over time can be correlated with the qualities of those chapters. This is because the chapters that the writer takes a longer time to release could be higher quality because they are written more carefully. Regarding this source of endogeneity, note that our dataset consists of best-selling books whose writers are experienced (averaging 3 years of writing experience, with at least five books written before 2017). For such writers, we do not expect a significant correlation between a chapter's quality and the speed of its release.

To validate this expectation, we run a regression in which the dependent variable is a chapter's *post-completion retention rate*, which is a measure of its quality. The retention rate of a chapter (k) is defined as amongst consumers who purchased the chapter (k), the percentage who purchase the next chapter $(k+1)$. Importantly, we calculate the retention rate using only the purchase data after the full book is released (i.e., the *post-completion* retention rate), where consumers no longer face any availability constraint due to release progress. Therefore, each chapter's release speed should not affect the *post-completion* retention rate via the availability constraint on the demand side. The regressors are (1) the release speed associated with the day that the chapter is uploaded and (2) whether the book is from an experienced writer. The regression is at chapter level and restricted to the books in the romance genre released in 2017. We control for the chapter index using second-degree polynomials, whether the chapter is the last chapter of a book section, book-fixed effects, and month-fixed effects. The regression results in Table A1 show that chapter retention rate is not significantly related to the release speed of the books in our sample. However, for other books from the same genre that are written by inexperienced writers, release speed has a negative and significant impact on chapter quality. This implies that we can rule out any concern related to the second source of endogeneity for our chosen sample of books.

Table A1 The Effect of Release Speed on Chapter Quality

	Dependent variable:
	Chapter retention rate (%)
Log (No. of characters released) * Experienced	-0.04
	(0.09)
Log (No. of characters released) * Inexperienced	-0.14 *
	(0.05)
Chapter Index/100	0.86 ***
	(0.16)
(Chapter Index/100)^2	-0.10 *
	(0.05)
Last chapter in a book section	-0.94 *
	(0.34)
Book F.E.	✓
Month F.E.	✓

Note: Regressions are at the book-chapter level. Standard errors (in brackets) are two-way clustered at the book and week levels. ***, **, and * indicate significance at the 0.1%, 1%, and 5% levels, respectively.

Section 3. The Effect of Release Speed on the Number of Comments Written by Readers

We investigate how a book's release speed can affect consumers' comment behaviors. We run regressions at the consumer-book-week level and test whether a consumer is more likely to leave comments (or an increased number of comments) when the book that she is following is being updated faster than when it is not, while controlling for consumer-fixed effects and book specific month-fixed effects and using the instrument to deal with the potential endogeneity of release speed. The regression results in Table A2 show that neither consumers' propensity to comment nor the number of comments is significantly affected by the book's release speed.

Table A2 The Effect of Release Speed on Comments

Dependent Variable:	(1)	(2)
	Whether to leave comment	Log (No. of comments + 1)
No. of chapters released per week	0.00043 (0.0008)	0.0016 (0.0013)
Consumer F.E.	Y	Y
Book specific month F.E.	Y	Y
2SLS with instruments	Y	Y

Note: Regressions are at the consumer-book-week level. Standard errors (in brackets) are three-way clustered at the consumer, book, and week levels. ***, **, and * indicate significance at the 0.1%, 1%, and 5% levels, respectively.

Section 4. Validity of the Instrumental Variable

In this section, we discuss the validity of the instrument discussed in section 3 of the paper, in terms of its relevance and whether it satisfies the exclusion conditions. We construct an instrument for the release speed, using the release speed of other books written by the same author at least 6 months earlier. We divide each book into five parts, each with an equal number of chapters, and calculate the average release speed for each quantile. To obtain the instrument variable of the release speed of focal book j in week t, we first identify the quantile of the book to which the chapters released that week belong. For example, for book j in week t, if the chapters released are from the first quantile of book j, then the instrument value is the average release speed of the chapters in the first quantile for all of the previous (at least 6 months prior) books written by the author of book j.

The instrument is relevant because a writer tends to have similar writing habits and a similar writing pace across time, and hence a similar cost structure for updating specific parts of their book chapters. The first-stage regression reports the incremental F-statistic for the instrument as 41.9, which suggests that we do not have a weak instrument problem.

The underlying condition for the instrument to satisfy the exclusion restriction is that the time-varying demand shocks[1] are uncorrelated across books written at different times (at least 6 months apart) by the same author. This is likely to hold because the writers in our data never release multiple books on the platform at the same time. Instead, they appear to focus on writing one book at a time and take long breaks—at least several months—between books. Given a

[1] The reason that we are only focusing on time varying demand shocks is because we control for the book-fixed effects in our analysis.

sufficient time gap between books, the demand shocks faced by one book should be different from those faced by another book by the same author.

Lastly it is possible that there could be an unobserved demand shock pattern across chapters that is common to all books written by the same author. This can potentially violate the exclusion condition of our instrument. The common demand shock pattern could arise because a writer may have a similar unobserved writing style/story formula across all her books. This is not a source of concern because this common demand shock pattern across chapters (that stems from a common writing style/story formula across books) would only impact the demand of a book through the (unobserved) qualities of the chapters. As discussed in our reduced-form analysis in section 2 of the web appendix, the release speed does not vary with chapter quality for experienced writers. This implies that the common demand shock pattern will not be correlated with release speed, which is the variable of interest that triggers an endogeneity concern.

Section 5. Justification of Assumptions in the Utility Maximization Model in Equation (4.1) of the Paper

In this section, we provide justification for the five assumptions in the utility maximization model. These five assumptions are: (i) It assumes that purchase quantities of inside goods (i.e., number of chapters purchased), x_{it}, are continuous variables. (ii) It assumes that price per chapter of a book is the same across all chapters of that book. (iii) It does not have a time constraint. (iv) It assumes that consumers are myopic decision makers. (v) It is unlike the dynamic inventory models (Erdem, Imai and Keane 2003; Ching and Osborne 2020) in which forward looking consumers purchase strategically in response to price promotions by stockpiling a good, and in which utility is maximized over quantities of goods consumed in each period (and not maximized over quantities purchased in each period, as in discrete continuous utility maximization models). We discuss our justification for these assumptions as follows.

Section 5.1: Assumption of purchase quantities of inside goods, x_{it}, being continuous variables.

Justification: This assumption is in keeping with prior literature that derives purchase conditions from a utility-maximizing framework (e.g., Chiang 1991; Nair, Dube, and Chintagunta 2005; Song and Chintagunta 2007). There are two reasons for this assumption:

i. Nair, Dube, and Chintagunta (2005) empirically compared the predictions of the continuous quantity assumption in these models with those of the discrete quantity. They find that the results are similar across the two, implying that the continuous quantity assumption is robust.
ii. The assumption of continuous quantities allows us to use the Kuhn Tucker conditions to solve the utility maximization problem. If the quantity is not continuous, then we can no longer use the first order conditions and we would need to resort to comparing the utilities across all possible combinations of purchases and non-purchases of books and their chapters in the entire set. This is an NP Hard problem since it will require comparing utilities the utility across all combinations of purchase/non purchase of 138 books/alternatives and across all their possible quantity choices for each alternative that can range from 1 to 100 chapters/units per book/alternative on a given trip. Note that the prior literature has relaxed continuous quantity assumption for the simpler cases when the

number of brands/alternatives is less than 5 and when the quantity choice per alternative is limited to less than 3 units has only been relaxed.

Section 5.2 Assumption of Price per chapter of a book is the same across all chapters of the book.

Justification: There are two reasons for this assumption:

i. The first reason for this assumption follows from the fact that the platform charges a fixed price per 1000 characters (which does not vary over time), and in our data, the average within book coefficient of variation in the number of characters per chapter is 0.15, which suggests that the chapter-level price variation for a given book is limited. Thus, we assume that the price per chapter is the same across all chapters in a book.

ii. The assumption of 'price per chapter not changing across chapters in a given book' allows us to use the Kuhn Tucker conditions to solve the utility maximization problem. If the price per chapter changes for each chapter, then we can no longer use the first order conditions and we would need to resort to comparing the utilities across all possible combinations of purchases and non-purchases of books and their chapters in the entire set. This is an NP Hard problem since it will require comparing utilities the utility across all combinations of purchase/non purchase of 138 books and across their quantity choices that can range from 1 to 100 chapters per book on a given trip.

Section 5.3 Assumption of Consumers as Myopic Decision Makers

Justification: We model readers as myopic decision makers for three reasons.

First, purchasing online books is a low-involvement purchase decision and there is minimal storage cost, no price variation, and no quantity discount on the platform. Strategically altering purchase timing cannot benefit consumers, as it does as in many other purchase settings.

Second, given that we use a discrete–continuous model with multiple discreteness with 138 alternatives at the purchase stage, it is very difficult to incorporate consumer's forward looking behavior. To the best of our knowledge, no prior paper has modeled consumer's forward looking behavior with even one alternative in a discrete continuous choice model.

Third, we investigate the possibility of a potential dynamic behavior in our setting using the model-free analysis. The dynamic behavior that we examine is whether consumers strategically wait for a certain minimum number of chapters to be released before purchasing them so that they can read those multiple chapters together. This is discussed as follows.

We investigate the strategic waiting behavior of consumers, which is a result of forward-looking behavior. Strategic waiting means that consumers would wait for a certain number of chapters to be published before making a purchase. This can be of three types: (i) Consumers waiting for a minimum number of chapters of an ongoing book to be released before they make their first purchase of chapters of that book. These consumers would typically start reading an ongoing book when the chapters released so far have received a good reception from the existing readers. (ii) Even after catching up with the progress of an ongoing book, consumers preferring to wait for a minimum number of chapters to be released so that they can purchase those chapters in bulk, which would them allow them to binge read them later. (iii) consumers preferring to purchase fully published books. We next discuss each of these three types.

We capture the first type of strategic waiting behavior in a reduced form fashion by including the following variables in the baseline utilities of the books: age of the book (which is the time elapsed since the release of the first chapter of the book), sales of a book and whether the book

was promoted on the platform's webpage (the platform typically promotes the best sellers on its homepage, where the best sellers could be ongoing as well as completed books. Consumers who engage in the first type of strategic waiting can get information on how good it is from these sources). In our results, we find that two of the four segments prefer purchasing older books and they prefer purchasing the books on the platform's homepage. On the other hand, we find that the other two segments prefer purchasing books whose first chapters have recently been released and their choices of books are not influenced much by whether the books are promoted on the platform's homepage.

Moving on to the second type of strategic waiting behavior, consider the consumer–book combinations in which the consumer begins to purchase book chapters before publication is complete. For these consumer-book combinations, we find that after a consumer catches up with the book's progress (that is, when the consumer purchases the most recently released chapter of the book), 98% of subsequent purchases happen within a week. The fact that the average release speed is 6.4 chapters in a week and the average length of a book is 150 chapters indicate that these consumers do not engage in strategic waiting after catching up.

Finally moving on to the third type of strategic waiting behavior, we looked at the consumer–book combinations in which the consumer begins to purchase book chapters after publication is complete. Most of these consumers start purchasing within 2 weeks of the book's finale, and they constitute approximately 16% of the consumer–book combinations. There are multiple reasons that a consumer may start reading a book after publication is complete, including strategic waiting, reduction in uncertainty, and the consumer's lack of previous awareness of the book. To distinguish the effect of strategic waiting from the effects of the other reasons, we run a regression in which we regress the number of new readers who start reading a book 2 weeks before and after its finale. The regressors are: a dummy for after-finale release, book-fixed effects, book's cumulative sales, and whether the book was on the platform's best-seller list. The extent of uncertainty about the book is similar across the four periods because, as discussed earlier, most learning happens early on. The variables "whether the book was on the best-seller list" and "cumulative sales" control for any potential decrease in uncertainty and consumer awareness about the book. The post-finale time dummy captures the effect of strategic waiting on demand.

The regression results are in Table A3. Based on the estimates of the dummy for after-finale release, we find the number of readers who start reading because of strategic waiting to be insignificant. Dividing this number by the average number of readers per book, we find that the fraction of consumer–book combinations in our data that may implicate strategic waiting is 1.5%. Because this is a small fraction, our structural model assumes the consumers to be myopic. And we capture the forward-looking behavior of these 1.5% consumers in our model in a reduced form manner by including the variable 'an indicator of whether book *j* is fully published', which captures consumer *i*'s relative preference for purchasing ongoing vs. completed books.

Table A3 The Effect of Release Speed on Comments

Dependent Variable:	No. of new readers who pick up book *j* in week *t*
After finale release of the book	0.529 (0.858)
Being on homepage ranking	2.762 * (1.195)
Cumulative sales of the book	9.569 *** (2.450)
Book F.E.	Y

Note: Regressions are at the book-week observation level, which is restricted to 4 weeks for each book, including the 2 weeks before and after each book's final release. Standard errors (in brackets) are two-way clustered at the book-week level. ***, **, and * indicate significance at the 0.1%, 1%, and 5% levels, respectively.

Section 5.4 Not Having a Time Constraint in the Utility Maximization

Justification: In utility maximization in equation (4.4) of the paper, we have four constraints: budget constraint, non-negativity constraint, availability constraint and inventory constraint. Note that the time constraint is not needed in this utility maximization. This is because we do not observe the reading the reading speeds of the consumers (which are the weights of the quantities of books in the time constraint). In the Kuhn Tucker conditions, these unobserved reading speeds will be subsumed into the intercept in the term α_{ijt} in equations (4.6) – (4.8) in the main paper. Thus, including the time constraint will not make any difference to the estimation.

Section 5.5. Not using a Dynamic Inventory Model:

The utility maximization in equation (4.4) of the paper is similar to the discrete-continuous choice utility maximization approaches used in the prior literature (Chintagunta 1993, Nair, Dube and Chintagunta 2005, and Mehta and Ma 2012) to model the consumers' purchase incidence and quantity decisions. In these models, the utility in each period is maximized over the quantities of the goods that are purchased in that period, and not the quantities of goods that are consumed in that period. This approach is in direct contrast to the dynamic inventory models (Ching and Osborne 2020; Erdem, Imai and Keane 2003) in which the per period utility is a function of the quantity of a book that will be consumed in the present period. This consumed quantity of a book in the present period will include the unconsumed chapters of that book from the past (which will be consumed in the present period) as well as a fraction of chapters purchased in the present period which will be consumed in the present period itself. Thus in such models, the past inventory would enter the quantity arguments in the utility in period t since the inventory from period t-1 would be consumed in the present period t. There are two reasons why we do not use the approach used in the dynamic inventory models:

i. In the dynamic inventory models used in the prior literature, the number of alternative/brands is small (typically around 4-5) and the brands are perfect substitutes (i.e., only one brand will be purchased if a purchase is made in the category). These assumptions are made for tractability. In our case, we have 138 alternatives that are imperfect substitutes. Keeping track of the inventories of all 138 alternatives that are not perfect substitutes is computationally infeasible. Moreover, the prior dynamic inventory models have assumed constant consumption rate over time for a given consumer-brand, which once again is made for computational convenience. In our model, we do not make this assumption since it would defeat the purpose of modeling binge consumption.
ii. Consumption utility vs. Purchase Utility: As discussed above, in the discrete continuous utility maximization model such as ours, the utility is conceptualized as a purchase utility - that is, in each period it is maximized over the quantities of goods that are purchased in that period, and not the quantities of goods that are consumed in that period. Now it is true that in the dynamic inventory models, it been conceptualized in terms of consumption utility, and not purchase utility. Now note that the consumption utility (that is, the utility which is a function of the quantities of goods that are consumed) will be identical to the purchase utility (that is, the utility which is a function of the quantities of goods that are consumed) if there is no leftover inventory from last period – that is, if the

quantity purchased in a given period is entirely consumed in that period. Since in reality that may not be so (that is, since in reality, there would be positive inventory from last period), we approximate the consumption utility as a purchase utility, but with the following two fixes: (i) the inventory of a book will impact the utility maximization through the inventory constraint (which as discussed in the paper implies that that a consumer will not purchase more chapters from a book unless they have consumed all the inventory of this book), (ii) the overall inventory level which will impact the visit decision. With these two fixes, the utility maximization approach that we use in discrete continuous models (in which we use purchase utilities) can be seen as an approximation of the utility maximization approach in which we use consumption utilities.

Section 6. Maximum Likelihood Estimation

We first derive the observational likelihood in the absence of the endogeneity issue. We start with deriving the second-stage likelihood and then incorporate the first-stage likelihood. Following that, we incorporate the control function into the observational likelihood to deal with endogeneity of release speed.

6.1 Specification of the Observational Likelihood without Endogeneity Correction.

Second-stage Observational Likelihood. Recall from Section 4 of the main paper that any observation can be classified into three groups. The conditions for the three groups follow from equations (4.6)-(4.8) of the paper as:

Group 1: Conditions for non-purchase of books $j = 1..n_1$:

$$Inv_{ijt}>0 \text{ or } \{Inv_{ijt}=0,\ \alpha_{ijt} - p_j - \beta_i \rho s_{it} \leq 0\} \quad \text{(TA.1)}$$

Group 2: Conditions for purchase at the availability constraint for books $j = n_1 + 1..n_1 + n_2$:

$$\{Inv_{ijt}=0,\ \alpha_{ijt} - p_j - \beta_i(1-\rho)c_{ijt} - \beta_i \rho s_{it} \geq 0\} \quad \text{(TA.2)}$$

Group 3: Conditions for purchase at less than the availability constraint for books $j = n_1 + 1..n_1 + n_2$:

$$\{Inv_{ijt}=0,\ \alpha_{ijt} - p_j - \beta_i(1-\rho)x_{ijt} - \beta_i \rho s_{it} = 0\} \quad \text{(TA.3)}$$

In the above equations, the expression for α_{ijt} follows from equation (4.2) of the paper as

$$\alpha_{ijt} = \alpha_{i0} + \alpha_j + \alpha_i^T Z_{ijt} + \varepsilon_{ijt} \quad (TA.4)$$

Recall that the econometrician's error term ε_{ijt} is assumed to have a standard normal distribution, which is IID across consumers and over time, but can be correlated across books. We assume the correlation to be the same across all pairs of books and equal to ρ_ε. We represent these equicorrelated errors as

$$\varepsilon_{ijt} = v_{ijt} + v_{i0t} \quad \text{(TA.5)}$$

where v_{ijt} is IID across all i, j, and t, v_{i0t} is IID across all i and t but the same across the books for a given consumer-period combination, and $v_{ijt} \perp v_{i0t}$. Their distributions are given by $v_{ijt} \sim N(0, 1-\rho_\varepsilon)$ and $v_{i0t} \sim N(0, \rho_\varepsilon)$. The expressions for $Inv_{ijt}>0$ and $Inv_{ijt}=0$ follow from equation (4.11) of the paper as:

Inv_{ijt}>0 iff $x_{j,l_j(t)} \cdot (1 - b \cdot \Delta t_{j,t}) - (a + \sigma_c \varepsilon_{cjt}) \Delta t_{j,t} > 0$ (TA.6)

Inv_{ijt}=0 iff $x_{j,l_j(t)} \cdot (1 - b \cdot \Delta t_{j,t}) - (a + \sigma_c \varepsilon_{cjt}) \Delta t_{j,t} \leq 0$ (TA.7)

Substituting the above expressions for the expressions for α_{ijt}, Inv_{ijt}>0 and Inv_{ijt}=0 into equations (TA.1)-(TA.3) of the paper, we obtain the following conditions for groups 1–3.

Group 1: Conditions for non-purchase of books $j = 1..n_1$:

$$\left\{ \begin{matrix} x_{j,l_j(t)} \cdot (1 - b \cdot \Delta t_{j,t}) - (a + \sigma_c \varepsilon_{cjt}) \Delta t_{j,t} \leq 0, \\ v_{ijt} \leq \beta_i \rho_i s_{it} + p_j - \alpha_{0i} - \alpha_j - \alpha_i^T Z_{ijt} - v_{i0t} \end{matrix} \right\} \cup \left\{ x_{j,l_j(t)} \cdot (1 - b \cdot \Delta t_{j,t}) - (a + \sigma_c \varepsilon_{cjt}) \Delta t_{j,t} > 0 \right\} \quad \text{(TA.8)}$$

Group 2: Conditions for purchase at the availability constraint for books $j = n_1 + 1..n_1 + n_2$:

$$\left\{ \begin{matrix} x_{j,l_j(t)} \cdot (1 - b \cdot \Delta t_{j,t}) - (a + \sigma_c \varepsilon_{cjt}) \Delta t_{j,t} \leq 0, \\ v_{ijt} \geq \beta_i (1 - \rho_i) c_{ijt} + \beta_i \rho_i s_{it} + p_j - \alpha_{0i} - \alpha_j - \alpha_i^T Z_{ijt} - v_{i0t} \end{matrix} \right\} \quad \text{(TA.9)}$$

Group 3: Conditions for purchase at less than the availability constraint for books $j = n_1 + 1..n_1 + n_2$:

$$\left\{ \begin{matrix} x_{j,l_j(t)} \cdot (1 - b \cdot \Delta t_{j,t}) - (a + \sigma_c \varepsilon_{cjt}) \Delta t_{j,t} \leq 0, \\ v_{ijt} = \beta_i (1 - \rho_i) x_{ijt} + \beta_i \rho_i s_{it} + p_j - \alpha_{0i} - \alpha_j - \alpha_i^T Z_{ijt} - v_{i0t} \end{matrix} \right\} \quad \text{(TA.10)}$$

where ε_c is a standard IID normal random measurement error in the inventories. The error v_{ijt} is IID across all i, j, and t and is distributed as $v_{ijt} \sim N(0, 1 - \rho_\varepsilon)$, and v_{i0t} is IID across all i and t but the same across the books for a given consumer–period combination and is distributed as $v_{i0t} \sim N(0, \rho_\varepsilon)$. The contribution of groups 1 and 2 to the second-stage observational likelihood follows from Equations (TA.8) and (TA.9) as

$$L_{it,g1}^{stage\ 2}(x_{ijt} | \theta_i, v_{i0t}) = \prod_{j=1}^{n} \left[\begin{matrix} \Phi\left(\frac{x_{j,l_j(t)} \cdot (1 - b \cdot \Delta t_{j,t}) - a \Delta t_{j,t}}{\sigma_c \Delta t_{j,t}} \right) + \\ \Phi\left(\frac{-x_{j,l_j(t)} \cdot (1 - b \cdot \Delta t_{j,t}) + a \Delta t_{j,t}}{\sigma_c \Delta t_{j,t}} \right) \Phi\left(\frac{\beta_i \rho_i s_{it} - \alpha_{0i} - \alpha_j - \alpha_i^T Z_{ijt} - v_{i0t}}{\sqrt{1 - \rho_\varepsilon}} \right) \end{matrix} \right] \quad \text{(TA.11)}$$

$$L_{it,g2}^{stage\ 2}(x_{ijt} | \theta_i, v_{i0t}) = \prod_{j=n_1+1}^{n_1+n_2} \left[\begin{matrix} \Phi\left(\frac{-x_{j,l_j(t)} \cdot (1 - b \cdot \Delta t_{j,t}) + a \Delta t_{j,t}}{\sigma_c \Delta t_{j,t}} \right) \\ \left(1 - \Phi\left(\frac{\beta_i (1 - \rho_i) c_{ijt} + \beta_i \rho_i s_{it} - \alpha_j - \alpha_{0i} - \alpha_i^T Z_{ijt} - v_{i0t}}{\sqrt{1 - \rho_\varepsilon}} \right) \right) \end{matrix} \right] \quad (TA.12)$$

In Equations (TA.11) and (TA.12), the likelihoods are conditional on the normal random variable v_{i0t} and the parameters θ_i that vary across consumers $i = 1..I$. Also note that we remove the per-chapter price term p_j in the likelihoods in Equations (TA.11) and (TA.12) because there is no change in the price over time. Thus, the price term merely rescales the brand intercept α_j in the baseline preference. For group 3, the joint density of the observed quantities is obtained through the transformation of variables from v_{ijt} to x_{ijt} from equation (TA.10) as

$$L_{it,g3}^{stage\ 2}\left(x_{ijt}\middle|\theta_i, v_{i0t}\right) = \frac{\beta_i^{n_3}\left(1+\rho_i(n_3-1)\right)(1-\rho_i)^{n_3-1}}{\left(\sqrt{1-\rho_\varepsilon}\right)^{n_3}} * \prod_{j=n_1+n_2+1}^{n_1+n_2+n_3}\left[\phi\left(\frac{\Phi\left(\frac{-x_{j,l_j(t)}\cdot\left(1-b\cdot\Delta t_{j,t}\right)+a\Delta t_{j,t}}{\sigma_c\Delta t_{j,t}}\right)}{\frac{\beta_i(1-\rho_i)x_{ijt}+\beta_i\rho_i s_{it}-\alpha_{0i}-\alpha_j-\alpha_i^T Z_{ijt}-v_{i0t}}{\sqrt{1-\rho_\varepsilon}}}\right)\right] \quad (TA.13)$$

where ϕ is the standard normal probability density function, and the first term is the Jacobian of the transformation from v_{ijt} to x_{ijt}.

The observational likelihood for the second stage of consumer i at period t conditional on v_{i0t} and θ_i (which is $L_{it}^{stage\ 2}\left(x_{ijt}\middle|\theta_i, v_{i0t}\right)$) is a product of the RHS of Equations (TA.11), (TA.12), and (TA.13) as

$$L_{it}^{stage\ 2}\left(x_{ijt}\middle|\theta_i, v_{i0t}\right) = \left(L_{it,g1}^{stage\ 2}\left(x_{ijt}\middle|\theta_i, v_{i0t}\right)\right)^{d_{1it}}\left(L_{it,g2}^{stage\ 2}\left(x_{ijt}\middle|\theta_i, v_{i0t}\right)\right)^{d_{2it}}\left(L_{it,g3}^{stage\ 2}\left(x_{ijt}\middle|\theta_i, v_{i0t}\right)\right)^{d_{3it}} \text{(TA.14)}$$

Where d_{1it} take a value of 1 if the observation lies in group 1 and 0 otherwise, d_{2it} take a value of 1 if the observation lies in group 2 and 0 otherwise, and d_{3it} take a value of 1 if the observation lies in group 3 and 0 otherwise. Now note that this second-stage likelihood, $L_{it}^{stage\ 2}\left(x_{ijt}\middle|\theta_i, v_{i0t}\right)$, in equation (TA.14) is conditional on the normal random error v_{i0t} that is distributed as $v_{i0t}\sim N(0,\rho_\varepsilon)$.

Incorporating the First-stage Model into the Observational Likelihood. We first specify the probability of visiting the platform. Following that we will discuss we incorporate the platform visit probability into the likelihood of the second stage to get the observational likelihood. now incorporate the first stage, the consumer's platform visit decision, into the observational likelihood.

$$Pr_{it}(Visit) = Pr(\gamma_0 + Z_{it}\gamma_1 + \eta_{it} \geq 0) \quad \text{(TA.15)}$$

where η_{it} is a standard normal random variable. We next allow for the errors in the two stages to be correlated. Specifically, we allow for the error η_{it} in the first stage to be correlated with the error ε_{ijt} in the second stage with a correlation of ρ_η. We allow for this correlation by representing η_{it} as

$$\eta_{ijt} = \omega_{it} + \frac{\rho_\eta}{\rho_\varepsilon} v_{i0t} \quad \text{(TA.16)}$$

where ω_{it} is IID across all i and t, v_{i0t} is the error from the second stage that is IID that is across all i and t but the same across the books for a given consumer-period combination, with $\omega_{it} \perp v_{i0t}$ and $\omega_{it} \perp v_{ijt}$. Note that the correlation between η_{it} tin the first stage and the error ε_{ijt} in the second stage stems from the common error v_{i0t} in the expressions of both η_{it} and ε_{ijt} in equations (TA.16) and (TA.5) respectively. Recall that the distribution of v_{i0t} is $v_{i0t}\sim N(0,\rho_\varepsilon)$. Since η_{it} has a standard normal distribution, it follows from equation (TA.16) that ω_{it} will be distributed as $\omega_{it}\sim N(0,1-\frac{\rho_\eta^2}{\rho_\varepsilon})$. Given this, we get the probability of platform visit by consumer i at time t conditional on the error v_{i0t} as

$$Pr_{it}(Visit|\theta_i, v_{i0t}) = \Phi\left(\frac{\gamma_0 + Z_{it}\gamma_1 + \frac{\rho_\eta}{\rho_\varepsilon} v_{i0t},}{\sqrt{1-\frac{\rho_\eta^2}{\rho_\varepsilon}}}\right) \quad \text{(TA.17)}$$

Given the visit probability in (TA.17) and the likelihood for the second stage in (TA.14), we can write the observational likelihood conditional on the error v_{i0t} as

$$L_{it}(x_{ijt}|\theta_i, v_{i0t}) = \begin{cases} 1 - P_{it}(Visit|\theta_i, v_{i0t}) & \text{if } visit_{it} = 0 \\ P_{it}(Visit|\theta_i, v_{i0t}) \times L_{it}^{stage\ 2}(x_{ijt}|\theta_i, v_{i0t}) & \text{if } visit_{it} = 1 \end{cases}$$

which can be written as

$$L_{it}(x_{ijt}|\theta_i, v_{i0t}) = \left(1 - P_{it}(Visit|\theta_i, v_{i0t})\right)^{1-Visit_{it}} \left(P_{it}(Visit|\theta_i, v_{i0t}) \times L_{it}^{stage\ 2}(x_{ijt}|\theta_i, v_{i0t})\right)^{Visit_{it}} \quad \text{(TA.18)}$$

Where $Visit_{it}$=1 if consumer i makes visit to platform in period t and 0 otherwise. We integrate out v_{i0t} in the observational likelihood in equation (TA.17) by using the Gauss-Hermite quadrature. To do so, we define $\tau = \frac{v_{i0t}}{\sqrt{2\rho_\varepsilon}}$ and rewrite the likelihood above as

$$\begin{aligned} L_{it}(x_{ijt}|\theta_i) &= \int L_{it}(x_{ijt}|\theta_i, v_{i0t}) \frac{1}{\sqrt{\rho_\varepsilon}} \phi\left(\frac{v_{i0t}}{\sqrt{\rho_\varepsilon}}\right) dv_{i0t} \\ &= \int L_{it}(x_{ijt}|\theta_i, v_{i0t} = \tau\sqrt{2\rho_\varepsilon}) \frac{1}{\sqrt{\pi}} e^{-\tau^2} d\tau\,. \end{aligned} \quad (TA.19)$$

Next, we use the Gauss-Hermite quadrature to approximate the integral above to get the observational likelihood conditional on the consumer-specific parameters θ_i.

$$L_{it}(x_{ijt}|\theta_i) \approx \sum_{m=1}^{12} L_{it}(x_{ijt}|\theta_i, v_{i0t} = \tau_m\sqrt{2\rho_\varepsilon}) \frac{w_m}{\sqrt{\pi}}\,. \quad (TA.20)$$

where the expression for the visit probability, *P*(*Visit*), is given as where we use 12 pairs of (τ_m, w_m) of the quadrature.

6.2 Modification of the Observational Likelihood with the Control Function Approach.

Recall that the control function enters both models of the first and second stages. Let μ_{jt} be the errors in the control function which is assumed to be normally distributed with mean zero and standard deviation of σ_μ. Let $\hat{\mu}_{jt}$ be the residuals retained from the first step of the control function approach and let the estimate of the standard deviation be $\hat{\sigma}_\mu$.

We start with the discussion of how we include the control function in the second stage and how that changes the observational likelihood in that stage. As discussed above, we add the control function part into the baseline utility in the second stage if and only if the book is still ongoing and consumer i is purchasing at the availability constraint in period t, i.e., x_{ijt} belongs to group 2.

Recall that in the second stage in equation (TA.5), we specified the error term ε_{ijt} as the sum of two zero-mean IID normal random variables: (1) v_{ijt} with a variance of $1 - \rho_\varepsilon$; and (2) v_{i0t} with a variance of ρ_ε. When adding the control function to the model, we need to readjust the error

term variances for the observations where the control function matters. We derive the modification below.

Suppose μ_{jt} and v_{ijt} are correlated with ρ_{c2}. Given that they are bivariate normally distributed, the expected value and variance of v_{ijt} conditional on $\mu_{jt} = \hat{\mu}_{jt}$ and $\sigma_\mu = \hat{\sigma}_\mu$ is

$$E\left(v_{ijt}\middle|\hat{\mu}_{jt}\right) = \frac{\rho_{c2}\hat{\mu}_{jt}\sqrt{1-\rho_\varepsilon}}{\hat{\sigma}_\mu}, Var\left(v_{ijt}\middle|\hat{\mu}_{jt}\right) = (1-\rho_\varepsilon)(1-\rho_{c2}^2). \quad (TA.21)$$

Therefore, we can rewrite v_{ijt} conditional on $\hat{\mu}_{jt}$ as

$$v_{ijt} = \frac{\rho_{c2}\hat{\mu}_{jt}\sqrt{1-\rho_\varepsilon}}{\hat{\sigma}_\mu} + \hat{v}_{ijt}, \quad (TA.22)$$

where $\hat{v}_{ijt}$ is a normally distributed random variable and $\hat{v}_{ijt} \sim N\left(0, (1-\rho_\varepsilon)(1-\rho_{c2}^2)\right)$. Substituting Equation (TA.22) into the condition for group 2 in Equation (TA.9), the modified condition for books $j = n_1 + 1, \ldots, n_1 + n_2$ is

Group 2: Conditions for purchase at the availability constraint for books $j = n_1 + 1..n_1 + n_2$:

$$\left\{x_{j,l_j(t)} \cdot \left(1 - b \cdot \Delta t_{j,t}\right) - a\Delta t_{j,t} + \sigma_m \varepsilon_m \le 0,\ \hat{v}_{ijt} \ge\ \beta_i(1-\rho_i)c_{ijt} + \beta_i\rho_i s_{it} - \alpha_{0i} - \alpha_j - \alpha_i^T Z_{ijt} - \frac{\rho_{c2}\hat{\mu}_{jt}\sqrt{1-\rho_\varepsilon}}{\hat{\sigma}_\mu} - v_{i0t}\right\} \quad (TA.23)$$

Accordingly, we obtain the modified contribution to the second-stage observational likelihood for group 2 as follows:

$$L_{it,g2}^{stage\ 2}\left(x_{ijt}\middle|\theta_i, v_{i0t}\right) = \prod_{j=n_1+1}^{n_1+n_2}\left[\begin{array}{c}\Phi\left(\frac{-x_{j,l_j(t)}\cdot\left(1-b\cdot\Delta t_{j,t}\right)+a\Delta t_{j,t}}{\sigma_c\Delta t_{j,t}}\right) \\ \left(1-\Phi\left(\frac{\beta(1-\rho)c_{ijt}+\beta\rho s_{it}-\alpha_0-\alpha^T Z_{ijt}-\frac{\rho_{c2}\hat{\mu}_{jt}\sqrt{1-\rho_\varepsilon}}{\hat{\sigma}_\mu}--v_{i0t}}{\sqrt{1-\rho_\varepsilon}\sqrt{1-\rho_{c2}^2}}\right)\right)\end{array}\right] \quad (TA.24)$$

Following this, we simply replace the second-stage likelihood contribution for group 2 in Equation (TA.12) with the second-stage likelihood contribution for group 2 in the above Equation (TA.24). This completes the control function modification of the observational likelihood in the second stage.

We next discuss the modification of the visiting probability in the first stage (equations TA.15 and TA.17). Suppose the residual errors in the control function, μ_{jt}, and the normal random error, ω_{ijt}, in the probit model in the first stage (as given in equation TA.16) are correlated with ρ_{c1}. Recall that ω_{it} is distributed as $\omega_{it} \sim N(0, 1-\frac{\rho_\eta^2}{\rho_\varepsilon})$. Given that ω_{ijt} and μ_{jt} are bivariate normally distributed, the expected value and variance of ω_{ijt} conditional on $\mu_{jt} = \hat{\mu}_{jt}$ and $\sigma_\mu = \hat{\sigma}_\mu$ are

$$E\left(\omega_{ijt}\middle|\hat{\mu}_{jt}\right) = \frac{\rho_{c1}\hat{\mu}_{jt}}{\hat{\sigma}_\mu}\sqrt{1-\frac{\rho_\eta^2}{\rho_\varepsilon}},$$

$$Var\left(\omega_{ijt}\middle|\hat{\mu}_{jt}\right) = (1-\rho_{c1}^2)\left(1-\frac{\rho_\eta^2}{\rho_\varepsilon}\right) \quad (TA.25)$$

Therefore, we can rewrite ω_{ijt} conditional on $\hat{\mu}_{jt}$ as

$$\omega_{ijt} = \frac{\rho_{c1}\hat{\mu}_{jt}}{\hat{\sigma}_{\mu}}\sqrt{1-\frac{\rho_{\eta}^2}{\rho_{\varepsilon}}} + \widehat{\omega}_{ijt}, \qquad (TA.26)$$

where $\widehat{\omega}_{ijt}$ is a normally distributed random variable and $\widehat{\omega}_{ijt} \sim N(0, 1-\rho_{c1}^2)\left(1-\frac{\rho_{\eta}^2}{\rho_{\varepsilon}}\right)$. Substituting Equation (TA.26) into the visit probability into equations (TA.15) and (TA.16), we get the modified visit probability conditional on the error v_{i0t} as

$$Pr_{it}(Visit|\theta_i, v_{i0t}) = \Phi\left(\frac{\gamma_0 + Z_{it}\gamma_1 + \frac{\rho_{\eta}}{\rho_{\varepsilon}}v_{i0t} + \frac{\rho_{c1}\hat{\mu}_{jt}}{\hat{\sigma}_{\mu}}\sqrt{1-\frac{\rho_{\eta}^2}{\rho_{\varepsilon}}}}{\sqrt{1-\rho_{c1}^2}\sqrt{1-\frac{\rho_{\eta}^2}{\rho_{\varepsilon}}}}\right) \qquad (TA.27)$$

Following this, we simply replace the visiting probability in Equation (TA.17) with the visiting probability in Equation (TA.27). This completes the control function modification of the observational likelihood in the first stage.

The Overall Likelihood. Previously, we specified the observational likelihood in Equation (TA.20) conditional on the consumer-specific parameters θ_i, $L_{it}\left(x_{ijt}|\theta_i\right)$. Given this, we get the likelihood for consumer i across all weeks $t = 1, \ldots, T$ in the data conditional on θ_i as

$$L_i(x_i \equiv \{x_{it}\}_{t=1}^{T}|\theta_i) = \prod_{t=1}^{T} L_{it}\left(x_{ijt}|\theta_i\right) \qquad (TA.28)$$

We model heterogeneity using the latent class segmentation method, in which θ_i takes different values across different latent segments. Let there be G segments with weights as $\{w_g\}_{g=1}^{G}$. For each segment g, the segment-specific parameters are θ_g.The unconditional joint likelihood of consumer i over all periods is

$$L_i(x_i) = \sum_{g=1}^{G} w_g \cdot L_i\left(x_i|\,\theta_g\right). \qquad (TA.29)$$

The log-likelihood for the entire sample is

$$l = \sum_{i=1}^{I} \ln[L_i(x_i)] \qquad (TA.30)$$

We maximize the log-likelihood with respect to all of the parameters to obtain the model estimates.

Section 7. Justification of Assumptions in the Counterfactuals Regarding Writer Behavior

In the counterfactuals, we assume that the following aspects of writer behavior do not change with a change in release strategy: (1) whether the writer continues writing books on the platform/switches to another platform, (2) qualities of the chapters in the book, (3) the initial release time of the book, (4) the number of chapters in the book, and (5) the number of books written in a fixed time frame.

Regarding the first aspect, it is unlikely that the change in release strategy will stop writers from writing books on the platform or drive them to another platform. The platform is a virtual

monopoly with a 70% market share in the romance genre, and all the writers have signed a long-term exclusive contract that prevents them from switching to other platforms or multihoming.

Regarding the second aspect, the analysis in section 2 of Web Appendix suggests that the qualities of the book chapters in our sample are not significantly affected by different release strategies. This is likely because the books in our sample are written by *experienced* writers, who tend to be more skilled and have high standards for producing consistently high-quality content. Therefore, we do not expect their chapter qualities to change much with release speed under our proposed counterfactuals. Furthermore, compared with inexperienced writers, experienced writers are likely to be more confident in their own skills and taste, value their creative independence more, and thus care less about and react less to readers' comments. Therefore, even if a release strategy change (such as adopting a hybrid release strategy) leads to much less reader feedback during an early stage, our sample writers are less likely to respond by changing the content of their books. However, given that the same reduced-form analysis shows that chapter quality from inexperienced writers. Thus if anything, the platform can implement our proposed release strategy for *experienced* writers.

Regarding the third aspect, note that our model and counterfactuals focus on release speed while taking the initial release timing as exogeneous. There are two reasons why we would not expect the change in release strategy to impact the initial release time.

i. In the counterfactual, we keep the initial release time the same as they were observed in data. We also tried varying the initial release time in the counterfactual and it doesn't affect the main results – the reason for that is because the extent of substitutability between books is small and the extent of seasonality effect is also low.
ii. Regarding the seasonality impact on the release of the first chapter, we plot the histogram of the books' initial release months in Figure A2. The distribution is rather even across all 12 months, unlike in the movie industry, in which many more movies are released during peak moviegoing seasons such as December. This outcome suggests that writers do not time their initial releases very strategically throughout the year. The only minor exception is that a slightly lower number of books begin their release in February, which is usually when the Chinese New Year falls. This is most likely because traditionally the Chinese New Year is a vacation period, and online creators are likely to take time off (see Figure A2). This finding implies that there is little seasonality effect when it comes to a book's initial release month, suggesting that the release strategy may not impact the book's initial release time.

Regarding the fourth aspect, like the argument around the second aspect, experienced writers are less likely to change their book content based on release strategy than inexperienced writers. To demonstrate this finding, we run regression testing on how the number of chapters in a book is affected by the book's average release speed, controlling for book initial release year-fixed effect, writer-fixed effect, and book sales. The results in Table A4 show that release speed does not significantly affect the length of the books in our sample. Therefore, it is reasonable to assume that counterfactuals on different release strategy for these books do not change their length.

Figure A2 Distributions of Books' Initial Release Month

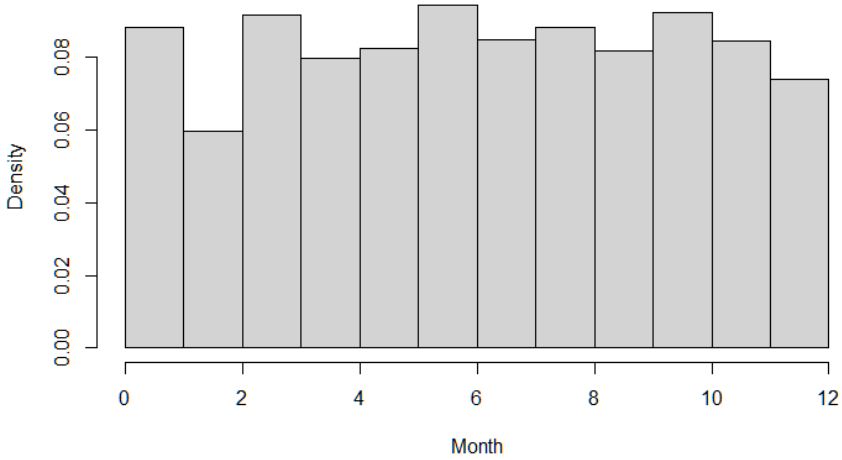


Table J The Effect of Release Speed on Book Length

Dependent variable:	Number of chapters in a book			
	(1)	(2)	(3)	(4)
Release speed: No. of chapters per day	2.24	13.07	3.20	7.46
	(6.17)	(11.32)	(6.35)	(8.765)
Log (daily sales during release period)			-9.673	34.77
			(10.63)	(18.37)
Book Initial Release Year F.E.	✓	✓	✓	✓
Writer F.E.	×	✓	×	✓

Note: Regressions are at the book level. Standard errors (in brackets) are clustered at the writer level.

Regarding the fifth aspect, our counterfactual requires the assumption that the total number of books produced by these writers does not change when they are told to release the books differently. We would not expect a different release strategy to significantly change an experienced writer's overall creative production process during the entire year. Thus if the book content and length do not change under the counterfactual (as in the second and fourth aspects), it is unlikely that the writer will write more or fewer books under different release strategies.

Section 8. Counterfactuals on Different Splits of Sequential and Simultaneous Release for Hybrid Release Strategy IV

Table A5 Optimal Hybrid Release Strategy

	(a)	(b)	(c)	(d)	(e)	(f)
	Chapters	Visits	Visits with purchase	Chapters per visit with purchase	Books per visit with purchase	Books explored

(1) Simultaneous for the first 40% of chapters + sequential for the rest at a constant speed						
2 chapters/week	386355	78688	74299	5.2	1.93	5021
3 chapters/week	**421638**	72557	70273	6.0	1.89	5188
4 chapters/week	367970	65269	58408	6.3	1.81	4907
(2) Simultaneous for the first 45% of chapters + sequential for the rest at a constant speed						
1 chapter/week	323020	84619	75121	4.3	1.98	4236
2 chapters/week	**429099**	79539	78018	5.5	1.96	5264
3 chapters/week	409851	69279	66105	6.2	1.86	4971
(3) Simultaneous for the first 50% of chapters + sequential for the rest at a constant speed						
1 chapter/week	358524	85182	79672	4.5	2.06	4544
2 chapters/week	**449175**	77843	77444	5.8	1.96	5320
3 chapters/week	393545	67965	63475	6.2	1.82	4966
(4) Simultaneous for the first 55% of chapters + sequential for the rest at a constant speed						
1 chapter/week	360979	82892	76804	4.7	2.00	4550
2 chapters/week	**436276**	74966	73945	5.9	1.88	5176
3 chapters/week	373626	64919	58379	6.4	1.79	4824
(5) Simultaneous for the first 60% of chapters + sequential for the rest at a constant speed						
1 chapter/week	383788	82285	78324	4.9	1.98	4708
2 chapters/week	**403056**	71293	67176	6.0	1.84	4968
3 chapters/week	367348	63292	56515	6.5	1.77	4802