Intertemporal Price Discrimination with Complementary Products: E-Books and E-Readers

Hui Li*

December 2017

Abstract

This paper studies intertemporal price discrimination (IPD) with complementary products in the context of e-readers and e-books. Using individual-level data (2008-2012), I estimate a dynamic demand model for e-reader adoption and subsequent book quantity, reading format, and retailer choices in several book genres. I use the estimates to simulate a monopolist's optimal dynamic pricing strategies when facing forward-looking consumers. The results illustrate how skimming/penetration pricing incentives for e-readers and harvesting/investing incentives for e-books interact in this novel setting. The optimal joint IPD strategy is skimming for e-readers and investing for e-books. Counterfactual results suggest that combining IPD with complementary product pricing improves firm profitability because it attenuates the limitations of each pricing approach. In a single-product IPD setting, firms' pricing power is limited when consumers anticipate future price changes and delay purchases. Adding complementary products offers firms two pricing instruments; opposite price trajectories provide conflicting incentives for consumers, limiting intertemporal arbitrage. In a static complementary product setting, firms' pricing power is limited when the relative elasticity between the two products is heterogeneous and conflicting among consumers. Adding IPD sorts heterogeneous consumers into different periods and reduces the need to balance across consumer types.

Keywords: intertemporal price discrimination; complementary products; e-book market; skimming and penetration pricing; harvesting and investing

*Hui Li, Assistant Professor of Marketing, Tepper School of Business, Carnegie Mellon University. Email: huil1@andrew.cmu.edu.

1 Introduction

In many industries, especially digital and online businesses, a single firm jointly sells durable primary hardware and complementary software. When pricing these two goods, firms usually set high prices to "skim" high-valuation consumers and later cut prices to appeal to low-valuation consumers. However, software pricing is either flat (e.g., Amazon Kindle and e-books) or considered separately from hardware (e.g., consoles and video games). The possibility of jointly conducting intertemporal price discrimination (IPD) for both hardware and software remains underexplored.

The fact that the same firm sells both hardware and software presents a novel and challenging setting for IPD. First, firms have mixed pricing incentives for both hardware and software. For hardware, firms can set high prices to "skim" high-valuation consumers or set low prices to "penetrate" the market and earn from subsequent software sales. For software, firms can set high prices to "harvest" existing consumers or set low prices to "invest" in new consumers. Because the mix of consumers evolves over time, firms have incentives to dynamically price both products. Although skimming/penetration and harvesting/investing strategies have been widely used and studied, it remains unclear how they interact in this setting.

Second, firms can potentially benefit from coordinating the two price instruments because the demand for the two products is closely linked. Consumers must buy the hardware to consume the software, and their hardware adoption is driven by their software usage. It is thus necessary to build a demand model that explicitly models software usage and consumers' self-selection into hardware adoption based on their heterogeneous software preferences. However, in the traditional durable product demand models, software usage is either ignored or modeled in a reduced-form manner.

This paper empirically investigates optimal IPD strategies with complementary products in the context of e-readers and e-books. I demonstrate the advantages of combining IPD and complementary product pricing and how the skimming vs. penetration strategies for e-readers interact with the harvesting vs. investing strategies for e-books in this novel setting. Here, skimming or harvesting (penetration or investing) is defined as decreasing (increasing) mark-ups over time so that the pricing strategies represent strategic interactions rather than simply cost changes. I start with estimating a discrete-continuous choice model of e-readers and books using individual transaction data from 2008 to 2012. Consumers first choose to buy or upgrade their e-readers and then decide the book quantity, reading format (e-books or print books), and retailers for print books (Amazon.com, other online retailers, or offline bookstores) in a number of book genres. Their e-reader adoption decision is driven by their book usage, which is further endogenized to be a function of prices and heterogeneous reading tastes. Given the estimates, I numerically solve for the optimal joint IPD strategy of e-readers and e-books for a monopolist who maximizes total profits from e-readers, print books, and e-books.[1] I focus on the pricing problem and take cost, product availability, and quality from the data. In the pure-strategy Markov-perfect Nash equilibrium (MPNE), both consumers and Amazon are forward-looking. Forward-looking consumers may anticipate future price changes and intertemporally arbitrage. I compare the optimal joint IPD strategy with the static complementary product pricing strategy and the hardware-IPD-only strategy. The results illustrate how combining IPD and complementary product pricing can help attenuate the limitations of each approach and improve firm profitability.

I find that the optimal joint IPD strategy – or the interaction between skimming vs. penetration and harvesting vs. investing – depends on a new dimension of consumer heterogeneity in this setting. The demand estimates reveal two major unobserved consumer types: avid readers who have high reading tastes and general readers who have low reading tastes. Traditional single-product IPD exploits the heterogeneity *across* consumer types: Avid readers are less price elastic to both e-readers and e-books than general readers in absolute terms. The joint IPD policy further exploits the heterogeneity *within* each consumer type. The demand estimates suggest that avid readers are relatively more price elastic to e-books than to e-readers, while general readers are relatively more price elastic to e-readers than to e-books.[2] The optimal joint IPD policy exploits this new dimension of consumer heterogeneity; for any given general (avid) reader penetration rate, as the avid (general) reader penetration rate increases, the firm should reduce (raise) e-reader prices and raise (reduce) e-book prices.[3] The overall price path balances the two consumer types. The firm should use a skimming strategy for e-readers and an investing strategy for e-books.

[1] Note that the objective of this analysis is not to explain or fit Amazon's observed strategy. I take a normative view and estimate the demand system without assuming that the observed prices are optimal. I use the demand estimates to solve for what the firm *should* do in the monopolist scenario.

[2] Intuitively, avid readers buy more books and spend more on books than on e-readers relative to general readers. They care more about subsequent book prices when buying e-readers. Note that this difference is not an imposed assumption; it comes from consumers' endogenous choices and heterogeneous preferences in the model and is a result of the estimation.

[3] Consumer types are still unobserved to the firm. This result is a characteristic of policy functions.

To illustrate the advantages of this joint IPD policy, I compare the market outcomes under three scenarios: 1) the static scenario in which the firm solves a static complementary product pricing problem as in the traditional "razor-and-blade" setting period by period; 2) the single IPD scenario in which the firm dynamically prices the e-readers only and statically prices the e-books; and 3) the joint IPD scenario. The results show that firm profitability in the static case is smaller than in the single IPD case, while both are smaller than in the joint IPD case because combining IPD and complementary product pricing can reduce the limitations of each approach and enhance firms' overall ability to price discriminate.

On one hand, in a traditional static razor-and-blade setting, firms set lower (higher) prices on products with more (less) elastic demand. Their pricing power is limited when the relative elasticity is heterogeneous and conflicting among consumers; firms have to weight and balance across consumer types, so prices tend to be driven closer to marginal costs (Rosen and Rosenfield 1997). In my new setting, adding IPD helps reduce the need to balance heterogeneous consumers, as illustrated by Scenarios 1 and 2. As consumers of different types are sorted to different periods (i.e., avid readers buy earlier), firms can set different price combinations in different periods that are less distorted by consumer heterogeneity and earn higher profits. I refer to this mechanism as *"sorting"*.

On the other hand, in a traditional single-product IPD setting, firms' pricing power is limited when consumers anticipate future price changes and delay purchases. In an extreme case, firms are better off committing to a fixed price rather than conducting IPD, which is commonly known as the Coase conjecture. In my new setting, adding a complementary product helps limit consumers' ability to intertemporally arbitrage, as illustrated by Scenarios 2 and 3. Given two pricing instruments, firms can use opposite price trajectories to provide conflicting incentives for consumers to buy earlier and delay purchase, which serves as a coordinated solution to the Coase conjecture and improves profitability. I refer to this mechanism as *"incentive-mixing"*.

I provide evidence of the "*sorting*" and "*incentive-mixing*" mechanisms by comparing e-reader sales over time and the fraction of avid readers across the three scenarios. I further analyze how the joint IPD strategy induces changes in micro-level consumer book and Kindle purchases and how these changes contribute to the profit gains of the joint IPD strategy. The joint IPD strategy induces five types of changes in book and Kindle purchases, which can be mapped to "sorting" and "incentive-mixing" mechanisms. Overall, the "sorting" related changes account for 63% of the total profit increase, and the "incentive-mixing" related changes account for 34% of the total profit increase.[4] I also find that the Kindle-side changes contribute more to the profit gains than the book-side changes, yet the magnitude of the Kindle-side benefits decreases over time while that of the book-side increases over time; the book-side benefits can predominate in the long run.

The contributions of this paper are three-fold. First, I contribute to the pricing literature by studying IPD in the complementary product setting, which is increasingly relevant as more firms attempt to build complex "ecosystems" of products to lock in consumers. I propose a novel joint IPD policy that exploits the link between hardware and software heterogeneity. I discuss how combining IPD and complementary product pricing can enhance firms' pricing ability and improve profitability. The results have managerial implications for other industries, such as consoles and video games, Apple TV and digital content in iTunes, and printers and cartridges. Second, I study how widely used pricing strategies – skimming/penetration and harvesting/investing – can be combined and interact with each other. The key logic of the optimal strategy combination is that different software usage intensity leads to different relative demand elasticities between the two products, which, in turn, drives the pricing policy. Third, I develop a demand framework for modeling dynamic durable hardware adoption with software usage. Previous studies largely ignore software usage or model it in a reduced-form way. Individual data on book consumption allow me to explicitly model software usage as a function of software prices and unobserved heterogeneous software tastes, which is critical to identifying the novel price discrimination opportunity in the complementary product setting. It also helps produce more accurate estimates, as hardware adoption contains information about software tastes. This framework is applicable to other scenarios in which both intensive margins and extensive margins are of interest.

2 Literature Review

The paper builds upon the literature on price discrimination. There have been theoretical and empirical studies on dynamic pricing of a single product (e.g., Stokey 1979, Besanko and Winston 1990, Nair 2007, Hendel and Nevo 2013) and static pricing of complementary products (e.g., Oi 1971,

[4] One of the changes, the "retailer competition effect", cannot be mapped to either "sorting" or "incentive-mixing" and account for 3% of the profit gains. More details are in Section 7.3.

Rosen and Rosenfield 1997, Gil and Hartmann 2009). Little is known, however, about dynamic pricing of complementary products. Nair (2007) and Liu (2010) empirically study IPD in the video game and console industry. They focus on the IPD of one product and abstract from the pricing of the other one. I contribute to the literature by modeling the joint IPD decision of both products.

This novel setting provides a unique opportunity to demonstrate how two types of pricing strategies, skimming vs. penetration pricing and harvesting vs. investing strategies, interact in the context of complementary products. Previous literature has studied each of the two types of strategies separately: when conducting IPD on durable products, firms have an incentive to skim high-valuation consumers (e.g., Stokey 1979, Besanko and Winston 1990, Nair 2007, Hendel and Nevo 2013) or penetrate the market if there are indirect network effects and/or complementary products (e.g., Dubé, Hitsch and Chintagunta 2010, Liu 2010); when there are switching costs, firms have incentives to harvest locked-in consumers or invest in acquiring new consumers to earn from subsequent sales (e.g., Klemperer 1995, Dubé, Hitsch and Rossi 2009). In my setting, firms have skimming/penetration incentives for hardware and harvesting/investing incentives for software. By modeling the dynamic pricing decisions of both products, I illustrate the optimal combination of the two types of pricing strategies and how firms can benefit from coordinating the two price instruments.

The demand model builds on the literature on dynamic durable product demand (e.g., Gowrisankaran and Rysman 2012, Melnikov 2013). It also joins the literature on complementary product demand, including tying and bundling (e.g., Gil and Hartmann 2009) and dynamic demand of hardware-software products (e.g., Gandal, Kende and Rob 2000, Nair, Chintagunta and Dubé 2004, Clements and Ohashi 2005, Hartmann and Nair 2010, Lee 2013). This paper contributes to the literature in two ways. First, it provides a framework to model durable hardware demand with software usage. Most of the previous studies focus on hardware demand. Software usage is either abstracted away from or modeled in a reduced-form way (e.g., as a function of software variety). Data on book consumption allow me to explicitly model software usage as a continuous choice, which is endogenized to be a function of software prices and unobserved heterogeneous software tastes. This framework is applicable to other scenarios where both intensive margin and extensive margin are of interest. Second, besides modeling the demand side, this paper further jointly solves the supply pricing problem. The supply-side model contributes to the nascent empirical literature on dynamic pricing in which both firms and consumers are forward-looking and form expectations about future states (e.g., Nair 2007, Goettler and Gordon 2011, Lee 2013).

3 Data and Industry Background

3.1 The U.S. E-Book Industry

The e-book market did not experience rapid growth until Amazon released its first e-reader, the Kindle, in November 2007. E-book sales enjoyed triple-digit annual growth rates from 2008 to 2011 and accounted for 23.8% of the total trade-book unit sales by 2013 (Book Industry Study Group, BookStats 2011, 2014). Amazon substantially promoted the diffusion of e-reading. Its existing relationship with publishers enabled it to offer a wide variety of e-books. By providing high-quality, affordable e-readers over time and pricing e-books of new releases and *New York Times* best sellers at $9.99, Amazon's market share reached nearly 90% by the end of 2009. Because the Kindle enjoyed a monopoly position from 2007 to 2009 and was the dominant device from 2010 to 2012 (Bowker Market Research, 2012), I focus on the optimal IPD strategies of Amazon, which had a monopoly on e-readers and e-books.[5] Consumers can still buy print books from all major retailers: Amazon.com, other online retailers, and offline bookstores.

There are three vertical players (publishers, retailers, and consumers) and three relevant products (print books, e-books, and e-readers) in the market. Publishers sell books to retailers at a wholesale price, which remained stable during the sample period and is taken as given in the supply-side simulation.[6] Retailers then set book retail prices for consumers.[7] They also launch their own e-readers and set e-reader prices. Discussions with industry practitioners suggest that publishing and pricing decisions for print books are unaffected by e-book pricing. Thus, I take print book prices as exogenously given from the data and focus on Kindle and e-book pricing in the simulation.

[5] Barnes & Noble entered the e-book market in October 2009, accounted for approximately 20% of all e-book sales by 2011, but has struggled to remain profitable. Apple began to sell e-books in iBookstore in January 2010 and eventually accounted for only approximately 10% of total e-book sales (Gilbert, 2015).

[6] As a robustness check, I allow publishers to optimally change wholesale prices. The predictions on dynamic pricing for retailers remain qualitatively unchanged. The details are available from the author upon request.

[7] E-book pricing followed the wholesale contract from 2008 to 2010 in which Amazon set book retail prices and paid wholesale prices to the publishers. The contract was switched to the agency contract from 2010 to 2012, in which publishers set book retail prices and Amazon obtained 30% of the book revenues. Print book pricing always follows the wholesale contract. I build the IPD problem based on the wholesale contract in which retailers set both e-reader and e-book prices.

3.2 Data Description

I combine three individual-level online transaction datasets and supplement them with data on aggregate offline book sales, costs, Kindle penetration rates, and e-book availability. The first dataset is the individual-level online book transaction records from 2008, when there was no Kindle, to 2012, gathered by comScore.[8] Each purchase record contains the retail website, purchase time, book title, format (print book or e-book), price, and quantity information.[9] It also includes demographics such as household income and age (each categorized into three groups), family size, and zip code, etc. Consumers were resampled every year, and 41% of them bought at least one book a year. There are 20,637 book buyers and 72,619 book purchases over the five-year sample period. Hardcovers account for only 5% of the transactions. I thus group them into "paperbacks" and use "paperbacks" to refer to all print books.

The second dataset contains book genre information that I collected from Amazon using web scrapers. For each book title in the first dataset, I collect its genre information and prices for both paperback and e-book formats. There are 122,068 pieces of title-format information. I group Amazon's subgenres into three genres, "lifestyle," "casual," and "practical," which account for 30%, 47%, and 23% of the sales, respectively. Subgenres within the same genre have similar prices, reading purposes, and consumer consumption patterns.[10] In particular, "lifestyle" books usually contain more pictures, "casual" books usually serve entertainment purposes, and "practical" books usually require in-depth reading and note-taking. These features will affect how consumers perceive e-books as substitutes for paperbacks and thus are relevant for categorization purposes.

[8] The comScore Web Behavior Database panel captures the detailed browsing and buying behaviors of a random sample from a cross-section of more than two million Internet users. It is weighted so that the distribution of the demographics matches that of the U.S. Internet user population.

[9] Book format information is available only for 2011 and 2012. For observations in 2008, 2009, and 2010, I integrate the format choice when calculating the likelihood. For instance, the probability of buying q books equals the sum of the following three probabilities: the probability of buying q paperbacks for a Kindle nonowner, the probability of buying q paperbacks for a Kindle owner, and the probability of buying q e-books for a Kindle owner. See the likelihood function section for the details on constructing the latter three probabilities. In Figure 2, the observed aggregate e-book sales for 2008, 2009, and 2010 are linearly interpolated between zero and the year 2011 value for plotting purposes only.

[10] The "lifestyle" genre includes "lifestyle & home," "cooking," "travel," "fitness & dieting," "crafts, hobbies & home," "arts & photography," and "children's book," etc. The "casual" genre includes "fiction," "science fiction," "humor," "nonfiction," and "biographies & memoirs," etc. The "practical" genre includes "computers & technology," "business & investing," "medical books," and "education & reference," etc. A typical consumer in the sample buys only one or two books a year. As the number of genres G increases, both the number of zero-consumption choices and consumers' ex ante book utility (integrated over G error terms) increase by construction. G is chosen to remain representative of the book heterogeneity while avoiding too many zero choices.

Given that households were resampled yearly and that Kindle prices and qualities changed annually in the data, I choose one year as the time period. For each consumer in every period, I aggregate their book purchase records to obtain their genre-format-retailer-level book quantity choices in the demand model. For instance, consumer i bought two "casual" e-books and one "practical" paperback in 2010 from Amazon. I also calculate the average book prices at the genre-format-retailer level every year. The sales-weighted and unweighted prices differ by less than 2%. I use the unweighted prices in the estimation.

The third dataset contains individual-level Kindle purchase records for 2008 through 2012 from comScore. I observe purchase time, Kindle version, price, quantity, and household demographics.[11] Due to yearly resampling, I cannot fully distinguish between first-time purchase and upgrading or link individuals' book transactions in the first dataset to their device transactions in the third dataset unless they happened during the same year. I use the model-predicted Kindle ownership probabilities to construct the likelihood in Section 5.1. I further obtain data on e-reader penetration rates in the U.S. from 2009 to 2012, gathered by Pew Research Center.[12] These aggregate data are combined with the individual-level Kindle sales data to further identify first-time purchase and upgrading. I discuss details on the identification and estimation strategy in Section 5.1 and 5.2.

I supplement the individual-level data with other relevant information. First, I impute aggregate offline book sales by genre from online and offline retailer market shares (Bowker's Books & Consumers report, 2012) and genre market shares (Nielsen BookScan). I obtain the offline population size from the fraction of consumers who have purchased books online (Nielsen Online Shopping Trend report, 2012).[13] Second, I obtain the number of e-books available in the Kindle Store, which increased from 126,630 in 2008 to 1,428,500 in 2012 as shown in the first column of Table 1, from a widely cited blog that takes monthly snapshots of Amazon.[14] Finally, I impute Kindle costs and book wholesale prices from industry reports.[15] The cost of the most popular Kindle version dropped

[11] Consumers bought more than one Kindle in 2.8% of the transactions over the five years, half of which happened during the holiday season indicating that a small fraction of Kindles might be purchased as gifts.

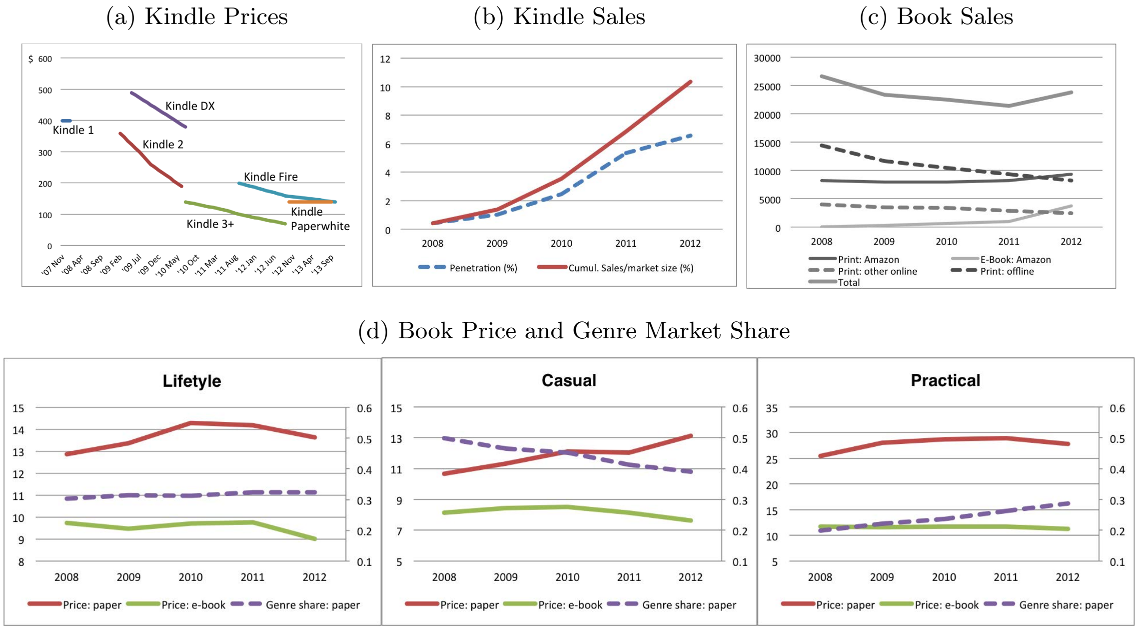
[12] See http://www.statista.com/statistics/249642/penetration-rate-of-e-reading-devices-in-the-us-by-type/.

[13] E-commerce constituted 25.1%, 35.1%, and 43.8% of the U.S. trade book sales from 2010 to 2012, respectively. Among book buyers, 44% have purchased books online.

[14] See http://ilmk.wordpress.com/category/analysis/snapshots/.

[15] In the publishing industry, the list price of e-books is 80% of the list price of paperbacks. Amazon sells books at 60% of the list price on average. The wholesale price for both paperbacks and e-books is 50% of the list price. I use these rules and the observed Amazon paperback prices to calculate the wholesale prices, which are \$15 for paperbacks and \$12 for e-books. Kindle costs are imputed from firms that release teardown reports almost every year (http://www.isuppli.com/Teardowns/News/Pages/Amazon-Kindle-Fire-Costs-\$201-

Figure 1: Observed Kindle and Book Prices and Sales



(a) Kindle Prices (b) Kindle Sales (c) Book Sales

(d) Book Price and Genre Market Share

Notes: Sub-graph (b) shows the cumulative Kindle sales from the comScore data and the penetration rates from the Pew Research survey. Sub-figure (c) shows the genre-specific paperback and e-book prices and the market shares of each genre among paperbacks. The left y-axis represents prices ($), and the right y-axis represents market share.

from $236 in 2008 to $89 in 2012.

3.3 Observed Pricing and Consumption Patterns

As shown in Figure 1, Amazon annually launched new Kindle generations and cut the prices of existing ones. The prices of paperbacks and e-books vary across genres and over time. The market shares of each genre among paperbacks vary as the prices change.

The data reveal considerable heterogeneity in consumers' book quantity choices; 13.8% of the consumers accounted for 46.8% of the total book purchases. Regarding genre, the correlation between consumers' book genre consumption and observed household characteristics is low, suggesting that the genre choice may be better explained by unobserved heterogeneous genre-specific reading tastes. As for format, consumers' choices seem to differ by genre. "Casual" books constitute a disproportionally larger share of the e-book format (71%) than the paperback format (44%). There is strong substitution between e-books and paperbacks in the same genre; 98.66% of households

70-to-Manufacture.aspx, https://www.engadget.com/2009/04/22/isuppli-359-kindle-2-costs-185-to-build-whispernet-says-shhh/). For years without these reports, I extrapolate data by assuming that the cost drops at the same rate as that of computer parts.

Table 1: E-Book Availability, Book Consumption and E-Reader Ownership

	#E-Books Available (unit: thousand)	Book Consumption by Kindle Ownership (unit: #books)			
		Kindle nonowner	Std. Err.	Kindle owner	Std. Err.
2008	126.6	3.91	(0.044)	7.63	-1.68
2009	300.6	3.47	(0.043)	9.43	-1.91
2010	586.6	3.28	(0.046)	4.49	-0.35
2011	957.3	3.13	(0.039)	4.72	-0.19
2012	1428.5	2.90	(0.038)	5.18	-0.13

purchased within a particular genre in at most one reading format. Regarding retailers, Amazon's market share increased from 32% in 2008 to 55% in 2012 at the expense of the sales of other online retailers and offline bookstores; Amazon's print book and e-book sales steadily increased while other retailers' print book sales decreased, as shown in Figure 1. The industry-wide book sales increased because of e-reading.

The data also suggest that book usage is correlated with e-reader adoption. Table 1 tabulates the yearly average number of books bought online per person among Kindle nonowners and owners. Kindle owners bought substantially more books than nonowners, suggesting that heavier book readers are more likely to buy Kindles. The difference between the two groups is larger in earlier years, suggesting that heavier book readers buy Kindles earlier. I also run a probit regression using Kindle ownership as the dependent variable. Controlling for household demographics and year fixed effects, the coefficient on the number of books bought is positive and significant. These results suggest that book consumption might drive Kindle adoption, which is the key demand-side feature that would drive the supply-side pricing results.

4 Model Setup

4.1 Consumer Problem

In each period, consumers first make ex ante dynamic device decisions based on the current Kindle price and quality, book price and e-book availability, their beliefs on the future values of these variables, and idiosyncratic device-side shocks. In particular, Kindle nonowners choose whether to buy a Kindle or wait for a better-quality and lower-priced Kindle in the future. Buying Kindles enables them to buy e-books that are potentially cheaper, more convenient to read, and better available over time. These consumers need to trade-off between the gain in discounted book flow utilities from the time of purchase onward and the one-time payment of the Kindle price. Kindle owners choose whether to upgrade to the latest Kindle generation or wait. Given their device-adoption statuses, the idiosyncratic book-side shocks are realized, and consumers make decisions in each genre about book purchases (buy or not), format (paperback or e-book), and paperback retailer (Amazon, other online retailers, or offline bookstores). They never drop out of the market.

Consumers' book consumption is modeled at the genre-format-retailer level instead of book title level because aggregate book sales are more relevant in the pricing problem than single-title sales.[16] I assume that consumers have persistent heterogeneous book tastes. Their book consumption changes in response to time-varying book prices, availability, and idiosyncratic shocks. They have perfect foresight on prices, Kindle quality, and book availability.[17] Kindle launches and book availability are taken from the data. Kindle qualities are taken as given and estimated in the model. For years beyond the sample period from 2008 to 2012, I assume that these variables stop evolving and remain at year 2012 levels.[18] I also make the following assumptions for tractability and data limitation reasons. First, I assume that consumers read e-books only on e-readers and not on other screens such as PCs and tablets.[19] Second, I assume that consumers use only one Kindle at a time and that Kindles have no resale value. I also assume that Amazon offers only the most popular Kindle version per period, which accounts for at least 70% of the sales in the data.[20]

Book quantity and format choices. Consumers make quantity-format choices in each genre. Index the three genres "lifestyle," "casual," and "practical" by $g = 1, 2, 3$, respectively. Let superscript E denote e-books and P denote paperbacks. Let superscript 0 denote Kindle nonowners and q_{igt}^{P0}

[16]Moreover, modeling at the title level would require strong assumptions about the books that enter consumers' choice set. We should not assume that consumers must decide from the millions of books that are available or from best sellers only, as 99.94% of the titles were purchased fewer than 10 times in the data. Modeling at the title level also requires estimating title fixed effects to account for price endogeneity issues. I do not have title-level aggregate book sales data and cannot estimate such fixed effects. Section 5.2 provides more discussion on how the price endogeneity issue is treated at the genre level.

[17]The results are robust to another rational expectation assumption for which consumer exceptions follow an AR(1) process and the coefficients in the AR(1) model are empirically estimated. I assume perfect foresight because Amazon changed prices annually over the five-year period, leading to a short panel and making the AR(1) model less appealing.

[18]As a validation of this assumption, Kindle prices experienced a significant decrease from 2007 to 2011 and have remained in the \$139-\$199 range since 2011.

[19]As a robustness check, I allow consumers to buy other reading devices after 2010 in the demand estimation. The estimated Kindle qualities are smaller, while the key demand-side results remain unchanged. In another robustness check, I account for e-book reading on other devices by adding book profits generated on other devices to Amazon's profit function. In this case, Amazon has weaker incentives to set low e-book prices to induce Kindle adoption because some consumers already own other devices. However, the joint IPD strategy does not qualitatively change.

[20]Goettler and Gordon (2011) also make this single-product assumption because multiproduct firm pricing is computationally prohibitive.

their quantity choices. Let superscript 1 denote Kindle owners and $\left\{q_{igt}^{P1}, q_{igt}^{E}\right\}$ their quantity choices. A Kindle owner (nonowner) i maximizes his period utility from both paperbacks and e-books (only paperbacks):

$$\max_{\left\{q_{igt}^{P1}, q_{igt}^{E}\right\}_g} \quad u_{it}^{book,1} = \sum_g \frac{1}{b_i}\left(a_{igt}^{P} q_{igt}^{P1} + a_{igt}^{E} q_{igt}^{E} - \frac{\left(q_{igt}^{P1} + q_{igt}^{E}\right)^2}{2}\right) - \sum_g \left(p_{gt}^{P} q_{igt}^{P1} + p_{gt}^{E} q_{igt}^{E}\right) \tag{1}$$

$$\max_{\left\{q_{igt}^{P0}\right\}_g} \quad u_{it}^{book,0} = \sum_g \frac{1}{b_i}\left(a_{igt}^{P} q_{igt}^{P0} - \frac{\left(q_{igt}^{P0}\right)^2}{2}\right) - \sum_g p_{gt}^{P} q_{igt}^{P0}$$

where a_{igt}^{P} and a_{igt}^{E} are heterogeneous book tastes to be parameterized later. b_i can be interpreted as a heterogeneous price coefficient because it enters the optimal quantity choice (below) linearly before the price. The optimal quantity choices in each genre for owners and nonowners are

$$\left\{q_{igt}^{P1*}, q_{igt}^{E*}\right\} = \begin{cases} \{0, 0\} & \text{if } p_{gt}^{P} > \frac{a_{igt}^{P}}{b_i}, p_{gt}^{E} > \frac{a_{igt}^{E}}{b_i} \\ \left\{a_{igt}^{P} - b_i p_{gt}^{P}, 0\right\} & \text{if } p_{gt}^{P} < \frac{a_{igt}^{P}}{b_i}, p_{gt}^{E} > \frac{a_{igt}^{E}}{b_i}, \text{ or} \\ & p_{gt}^{P} < \frac{a_{igt}^{P}}{b_i}, p_{gt}^{E} < \frac{a_{igt}^{E}}{b_i}, a_{igt}^{P} - b_i p_{gt}^{P} > a_{igt}^{E} - b_i p_{gt}^{E} \\ \left\{0, a_{igt}^{E} - b_i p_{gt}^{E}\right\} & \text{if } p_{gt}^{P} > \frac{a_{igt}^{P}}{b_i}, p_{gt}^{E} < \frac{a_{igt}^{E}}{b_i}, \text{ or} \\ & p_{gt}^{P} < \frac{a_{igt}^{P}}{b_i}, p_{gt}^{E} < \frac{a_{igt}^{E}}{b_i}, a_{igt}^{P} - b_i p_{gt}^{P} < a_{igt}^{E} - b_i p_{gt}^{E} \end{cases} \tag{2}$$

$$q_{igt}^{P0*} = \begin{cases} a_{igt}^{P} - b_i p_{gt}^{P} & \text{if } p_{gt}^{P} < \frac{a_{igt}^{P}}{b_i} \\ 0 & \text{otherwise} \end{cases}$$

This solution suggests that book usage is an endogenous function of heterogeneous book reading tastes $\left\{a_{igt}^{P}, a_{igt}^{E}\right\}$ and book prices $\left\{p_{gt}^{P}, p_{gt}^{E}\right\}$. The solution conditions represent, for instance, that Kindle owners buy paperbacks if only paperbacks are worth buying or if paperbacks are more attractive when both formats are worth buying. I choose the quadratic utility functional form instead of the discrete choice logit utility or constant elasticity of substitution (CES) utility, because its optimal quantity solution contains multiple-unit and zero consumption cases and is linear in prices.[21] The utility form implies that utilities from different book genres do not interact and that there is perfect substitution between paperbacks and e-books of the same genre.[22] Note that perfect

[21] More flexible quadratic utility specifications yield qualitatively similar demand-side predictions even though the optimal quantity solutions are more complex. Economides, Seim, and Viard (2008) adopt a similar quadratic functional form without allowing for substitution. For a good survey on direct utility models of consumer choice in marketing, see Chandukala, Kim, Otter, Rossi, and Allenby (2008).

[22] 1) Although the utilities across genres do not interact here, consumers' genre-specific preferences θ_{ig} are allowed to be correlated across genres. The data reveal four heterogeneous preference segments. One prefers both lifestyle and practical books, while another prefers all three genres. This correlation of genre-specific preferences cannot be

substitution does not imply that a single consumer cannot buy both paperbacks and e-books; he or she can choose paperbacks in one genre and e-books in another genre.

Paperback retailer choices. Once a consumer chooses to buy paperbacks, he decides whether to buy from Amazon.com, other online retailers, or offline bookstores. Sales of online retailers had been growing almost linearly as a share of the total retail sales, even before the introduction of e-books.[23] I use a discrete choice logit structure to parsimoniously capture this trend while allowing e-reading to influence it. Denote the three retailers by A, B, and O. The retailer utilities contain genre-specific retailer fixed effects, time trends, and idiosyncratic shocks. They are $u_{igt}^{rA} = A_{0g} + A_{1g} \cdot t + A_2 \cdot 1\{owner\} + \zeta_{igt}^{A}$ for Amazon, $u_{igt}^{rB} = B_{0g} + B_{1g} \cdot t + \zeta_{igt}^{B}$ for other online retailers, and $u_{igt}^{rO} = \zeta_{igt}^{O}$ for offline bookstores.[24] Offline bookstores serve as the baseline choice, and its fixed effect and time trend are normalized to zero. I allow Kindle ownership to affect the probability of buying paperbacks from Amazon, which is captured by A_2. The probability of choosing retailer j at time t for consumer i is $r_{ijgt} = \exp\left(u_{igt}^{rj} - \zeta_{igt}^{j}\right) / \left[\sum_{j'} \exp\left(u_{igt}^{rj'} - \zeta_{igt}^{j'}\right)\right]$. It is r_{ijgt}^{1} for Kindle owners and r_{ijgt}^{0} for Kindle nonowners. They differ in that r_{ijgt}^{1} contains A_2, while r_{ijgt}^{0} does not. Consumers' retailer choices will respond to alternative IPD strategies in the supply-side simulation because the Kindle ownership status is endogenous and affects retailer choices through A_2.

Consumer heterogeneity. The taste parameters contain a baseline taste (shared by the two

separately identified from the positive interaction of genre-specific utilities. I thus allow for correlation among genre preferences and do not allow for interaction among genre utilities. 2) By construction, this model cannot generate $\left\{q_{igt}^{P1*} > 0, q_{igt}^{E*} > 0\right\}$, which represents only 1.34% of the data. I treat these observations as two independent observations $\left\{q_{igt}^{P1}, 0\right\}$ and $\left\{0, q_{igt}^{E}\right\}$ from two shopping occasions in a period, which mildly overestimates the substitution between paperbacks and e-books.

[23] See http://www.census.gov/retail/mrts/www/data/pdf/ec_current.pdf.

[24] I abstract from paperback pricing responses and assume that paperback prices remain unchanged in the simulation. I use Amazon's paperback prices as print book prices in the demand estimation and supply simulation. One can add retailer-specific paperback prices in the retailer choice model. Yet the price gaps across retailers are stable so that these terms would be absorbed by the retailer fixed effects. Print books are a "now-mature" product (Reimers and Waldfogel 2014) with stable retailer-specific pricing rules. Typically, a traditional bookstore takes a 30% – 50% discount off the retail price. Online bookstores can take up to a 55% discount (http://www.smithpublicity.com/2014/03/determining-retail-price-printed-book/). Although average prices vary over time due to changes in the set of available books, the price gaps across retailers remain stable due to the stable retailer-specific pricing rules on individual titles. Therefore, if I specify the retailer utilities as $u_{igt}^{rA} = \tilde{A}_{0g} + A_{1g} \cdot t + A_2 \cdot 1\{owner\} + b^r p_{gt}^{PA} + \zeta_{igt}^{A}$ for Amazon, $u_{igt}^{rB} = \tilde{B}_{0g} + B_{1g} \cdot t + b^r p_{gt}^{PB} + \zeta_{igt}^{B}$ for other online retailers, and $u_{igt}^{rO} = b^r p_{gt}^{PO} + \zeta_{igt}^{O}$ for offline bookstores, then $b^r p_{gt}^{PA} - b^r p_{gt}^{PO} (= b^r \triangle p_g^{AO})$ and $b^r p_{gt}^{PB} - b^r p_{gt}^{PO} (= b^r \Delta p_g^{BO})$ would be absorbed by the genre-specific retailer fixed effects $\tilde{A}_{0g}$ and $\tilde{B}_{0g}$ when calculating the retailer probabilities. The retailer utilities can thus be re-written as those in the main specification.

formats), an e-format-specific taste, and idiosyncratic shocks:

$$
\begin{aligned}
a_{igt}^{P} &= \theta_{ig} + \beta_1 D_i^{age} + \xi_t + \eta_{igt}^{P} \\
a_{igt}^{E} &= \underbrace{\theta_{ig} + \beta_1 D_i^{age} + \xi_t}_{\text{Baseline taste}} + \underbrace{\left(\theta_g^{E} + \beta_2 D_i^{age} + \beta_3 \log\, n_t^{E}\right)}_{\text{E-format taste}} + \eta_{igt}^{E}
\end{aligned} \tag{3}
$$

Here, θ_{ig} are the genre fixed effects, and θ_g^E are the e-format genre fixed effects. n_t^E is the number of e-books available which is time-varying and can affect the e-format taste.[25] ξ_t are the time fixed effects that capture any other time-varying unobserved book characteristics.

First, consumers differ in their unobserved book reading tastes θ_{ig}; some might enjoy reading "casual" books, while others might enjoy reading "practical" books. I use a finite mixture specification and assume that these genre fixed effects $\{\theta_{ig}\}_{g=1,2,3}$ belong to one of K segments. Each segment k is characterized by its own parameter vector $\{\theta_g^k\}_{g=1,2,3}$, and the segment size is λ^k. I also allow for segment-specific unobserved heterogeneous preferences on the device side, which I describe in the device decision below. Second, consumers differ in their observed demographics $\{D_i^{income}, D_i^{age}\}$, each categorized into three groups in the data; $D_i^{income} = 1, 2, 3$ represent low, medium, and high income groups and $D_i^{age} = 1, 2, 3$ represent young, middle, and senior age groups. Both the baseline taste and the e-format taste can vary by age D_i^{age}, as senior consumers generally read more books and are less tech-savvy. Their price coefficients can vary by income, as $b_i = b_0 + b_1 D_i^{income}$. The unobserved types and the observed demographics are independent. Finally, consumers receive idiosyncratic shocks $\left\{\eta_{igt}^P, \eta_{igt}^E\right\}$ that are assumed to be i.i.d. normally distributed with mean zero and standard deviation σ[26].

The taste parameters in $\left\{a_{igt}^P, a_{igt}^E\right\}$ affect consumers' book consumption, as in Equation 2. Consumers' book consumption responds to time-varying prices, e-book availability, and idiosyncratic shocks. As later shown in the estimation results, the unobserved heterogeneous genre fixed effect θ_{ig} is the major difference between avid and general readers. This book-side heterogeneity drives the device-side behavioral differences; avid readers buy Kindles earlier and have different relative demand elasticities between Kindles and e-books than general readers do.

[25] The number of e-books available is not directly correlated with the e-reader userbase because e-book introduction is often not retailer exclusive; Amazon.com and Barnesandnoble.com both had approximately three million e-books available by 2012, but their e-reader userbases differed by a factor of five.

[26] Robustness checks show that allowing η_{igt}^P and η_{igt}^E to be correlated within the same genre does not change the implied substitution patterns and price elasticities.

Indirect flow utility from books. The ex ante indirect flow utilities from books for Kindle nonowners and owners $v_{it}^{book,0}$ and $v_{it}^{book,1}$ are obtained by substituting Equation 2 into Equation 1 and taking expectations over the error terms η_{igt} in a_{igt}:

$$\begin{aligned}
v_{it}^{book,0} &= \sum_g E[\frac{\left(a_{igt}^P - b_i p_{gt}^P\right)^2}{2b_i} \mid q_{igt}^{P0*} > 0] \cdot \Pr\left(q_{igt}^{P0*} > 0\right) \\
v_{it}^{book,1} &= \sum_g \Bigg\{ E[\frac{\left(a_{igt}^P - b_i p_{gt}^P\right)^2}{2b_i} \mid q_{igt}^{P1*} > 0,\, q_{igt}^{E*} = 0] \cdot \Pr\left(q_{igt}^{P1*} > 0,\, q_{igt}^{E*} = 0\right) \\
&\qquad + E[\frac{\left(a_{igt}^E - b_i p_{gt}^E\right)^2}{2b_i} \mid q_{igt}^{E*} > 0,\, q_{igt}^{P1*} = 0] \cdot \Pr\left(q_{igt}^{E*} > 0,\, q_{igt}^{P1*} = 0\right) \Bigg\}
\end{aligned} \tag{4}$$

Device adoption decision. Dynamically, given the utilities from books, consumers decide ex ante whether to buy or upgrade their Kindles. Book utilities enter device utilities so that book usage can drive device adoption.

$$\begin{aligned}
\bar{u}_{it}^0 &= \Gamma v_{it}^{book,0} + \bar{\varepsilon}_{it}^0 \\
\bar{u}_{it}^1 &= \Gamma v_{it}^{book,1}\left(p_t^E\right) + \bar{Q}_{it} + \bar{\varepsilon}_{it}^1 \\
u_{it} &= \Gamma v_{it}^{book,1}\left(p_t^E\right) + Q_{it} - \alpha_i P_t + \varepsilon_{it}
\end{aligned} \tag{5}$$

Here, $\bar{u}_{it}^0$ represents the device flow utility of a Kindle nonowner who chooses to wait and receives book utility only from paperbacks. $\bar{u}_{it}^1$ represents the utility of a Kindle owner who chooses not to upgrade and receives book utility from both paperbacks and e-books and the quality of his old Kindle $\bar{Q}_{it}$. u_{it} represents the utility of a consumer who chooses to buy/upgrade and receives book utility from both paperbacks and e-books plus the new Kindle quality Q_{it} at a cost of the Kindle price P_t. I allow consumers to have heterogeneous perceived quality of Kindles (see further details below). I also allow the price coefficient $\alpha_i = \alpha_0 + \alpha_1 D_i^{income}$ to vary across income groups. The idiosyncratic shocks $\left\{\bar{\varepsilon}_{it}^0, \bar{\varepsilon}_{it}^1, \varepsilon_{it}\right\}$ are identically and independently distributed extreme value type I errors, which are also independent of the book-side error terms. The mean is the negative of the Euler constant, and the variance is normalized to be 1. Γ represents how much consumers care about books when adopting Kindles. It is mainly identified from how Kindle sales respond to cross-sectional book usage variation and time-varying e-book availability.[27]

[27] As the variance of the shock is normalized to 1, Γ also captures the variance in consumer device choices. Equivalently, one can drop the coefficient Γ and estimate the variance of the error term. Lee (2013) adopts a similar setup in which software utility $\Gamma_{jt}\left(\alpha_i^{p,sw}, \alpha_i^{\gamma}; \iota\right)$ enters hardware utility $u_{ijt\iota}$ with coefficient α^{Γ} as

The Kindle qualities are estimated as dummies. To avoid overfitting, I capture the values of the qualities using $Q_{it} = Q_0^k + Q_1^k \log t$, where $t = 1$ to 5 represent years 2008 to 2012, and $\left\{Q_0^k, Q_1^k\right\}$ are coefficients to be estimated for consumer segment k.[28] The Kindle qualities do not enter book utilities, meaning that Kindles of higher qualities do not offer higher book utilities and that the upgrading decision is driven by higher qualities of the device itself. I do not allow Kindle qualities to affect book utilities because the data cannot identify such a relationship. In a robustness check, I add Kindle quality dummies to the book utilities such that consumers with better Kindles can have higher book utilities, buy more books, and have different upgrading probabilities. The estimated dummies are insignificant. I present the details of the robustness check in the Appendix.

As book utilities enter device utilities, lower e-book prices make Kindles more attractive. E-book and Kindle prices jointly affect the Kindle purchase decision. As shown later in the demand-side estimation results, consumers differ in how sensitive they are to e-book prices when adopting Kindles; the cross-elasticity of Kindles with respect to e-book prices differs across consumer types. Firms can leverage this new dimension of consumer heterogeneity and use different Kindle and e-book price combinations to induce different consumers to purchase. This forms the basis for the joint IPD strategy.

The device flow utility enters the dynamic programming problem. To make the notation more general, I use $\bar{u}_{it}$ to jointly denote the flow utility of waiting for nonowner and owner $\left\{\bar{u}_{it}^0, \bar{u}_{it}^1\right\}$. Kindle quality is $\bar{Q}_{it} = 0$ for a nonowner. The state space contains (1) the current Kindle ownership status $\bar{Q}_{it}$, which evolves based on the device adoption choice; (2) the e-book and paperback prices $\left\{p_t^E, p_t^P\right\}$, which enter the book flow utility; (3) the offered Kindle price P_t, quality Q_{it}, and cost c_t;[29] and (4) the idiosyncratic shocks on the device side $\vec{\varepsilon}_{it} \equiv \{\bar{\varepsilon}_{it}, \varepsilon_{it}\}$. Define $\Omega_t = \left\{p_t^E, p_t^P, P_t, Q_{it}, c_t\right\}$ and let $V\left(\bar{Q}_{it}, \Omega_t, \vec{\varepsilon}_{it}\right)$ denote the value function of a consumer with device $\bar{Q}_{it}$ at the beginning of the period. $d_{it} = 1$ indicates buying/upgrading and $d_{it} = 0$ indicates waiting. The Bellman

$u_{ijt\iota} = \alpha^x x_{jt} + \alpha_i^{p,hw} p_{jt} + \alpha^\Gamma \Gamma_{jt}\left(\alpha_i^{p,sw}, \alpha_i^\gamma; \iota\right) + D(\iota) + \xi_{jt} + \epsilon_{ijt\iota}$.

[28] I estimated another specification with a full set of $\{Q_t\}_{t=2008}^{t=2012}$ for each consumer segment. The estimates exhibit a similar pattern as the estimates from this log form specification.

[29] Kindle cost does not enter the state space in the demand estimation because the Kindle price is sufficient to solve the consumer-side Bellman equation. Kindle cost enters the state space in the supply-side simulation because cost enters the firm problem and affects the firm's Kindle pricing decision. Consumers need cost information to form expectations over future prices.

equation is

$$V\left(\bar{Q}_{it},\Omega_t,\vec{\varepsilon}_{it}\right)=\max\Big\{\bar{u}_{it}+\delta E\left[V\left(\bar{Q}_{it},\Omega_{t+1},\vec{\varepsilon}_{it+1}\right)\mid\Omega_t,d_{it}=0\right],$$
$$u_{it}+\delta E\left[V\left(Q_{it},\Omega_{t+1},\vec{\varepsilon}_{it+1}\right)\mid\Omega_t,d_{it}=1\right]\Big\} \tag{6}$$

The first and second elements of the max operator are the choice-specific value functions of waiting and buying/upgrading. Conditional on waiting, the device adoption status remains at $\bar{Q}_{it+1}=\bar{Q}_{it}$. Conditional on buying/upgrading, the device adoption status evolves deterministically as $\bar{Q}_{it+1}=Q_{it}$. The rest of the state space Ω_t evolves to Ω_{t+1}, according to consumers' expectation about next-period values $h\left(\Omega_{t+1}|\Omega_t\right)$. In particular, the Kindle quality evolves deterministically according to a Markov process with a degenerate transition matrix. The transition probability is $p\left(Q_{i,t+1}|Q_{it}\right)=1$ for $t=2008$ to 2011 and $p\left(Q_{i,2012}|Q_{i,2012}\right)=1$ from 2012 onwards.

Intuitively, consumers are motivated to buy Kindles for three reasons: the gain from current-period book utility, the attractiveness of device prices and qualities, and the option value of device adoption and waiting. They form expectations about the price trajectories of e-books and Kindles over time and decide whether and when to purchase. Specifically, lower e-book prices (lower Kindle prices) in the current period raise current-period book utility (attractiveness of Kindles) so that more consumers would like to buy Kindles now. Higher expected e-book prices (higher expected Kindle prices) in the future reduce the continuation value of waiting so that consumers would like to buy Kindles earlier. Therefore, firms can use the price trajectories of e-books and Kindles to influence both the volume and timing of consumers' Kindle purchases.

Let $EV(\cdot)=\int_{\varepsilon}V\left(\cdot,\vec{\varepsilon}\right)dg_{\vec{\varepsilon}}$ denote the expectation of the value function integrated over $\vec{\varepsilon}_{it}$. The expected value function is

$$EV\left(\bar{Q}_{it},\Omega_t\right)=\ln\Big[\exp\left(\bar{u}_{it}-\bar{\varepsilon}_{it}+\delta E\left[V\left(\bar{Q}_{it},\Omega_{t+1},\vec{\varepsilon}_{it+1}\right)\mid\Omega_t,d_{it}=0\right]\right)$$
$$+\exp\left(u_{it}-\varepsilon_{it}+\delta E\left[V\left(Q_{it},\Omega_{t+1},\vec{\varepsilon}_{it+1}\right)\mid\Omega_t,d_{it}=1\right]\right)\Big] \tag{7}$$

Note that there is a unique expected value function for consumers in each observed demographic group and unobserved segment k. The taste preference shocks η_{igt}'s are not relevant here because they are integrated out when calculating the ex ante book flow utility in Equation 4. For consumer i with demographics $\left\{D_i^{age},D_i^{income}\right\}$ and unobserved taste type k, denote his/her device utilities and shocks as $\left\{u_{it}^k, \bar{u}_{it}^k, \varepsilon_{it}^k, \bar{\varepsilon}_{it}^k\right\}$. His/her probability of buying/upgrading conditional on having Kindle $\bar{Q}_{it}$ is (A and B are scalars):

$$\begin{aligned}
\phi\left(d_{it}=1 \mid \bar{Q}_{it}, \Omega_t, k\right) &= \frac{B}{A+B} \\
A &= \exp\left(\bar{u}_{it}^k - \bar{\varepsilon}_{it}^k + \delta E\left[V\left(\bar{Q}_{it}, \Omega_{t+1}, \vec{\varepsilon}_{it+1}\right) \mid \Omega_t, k, d_{it}=0\right]\right) \\
B &= \exp\left(u_{it}^k - \varepsilon_{it}^k + \delta E\left[V\left(Q_{it}, \Omega_{t+1}, \vec{\varepsilon}_{it+1}\right) \mid \Omega_t, k, d_{it}=1\right]\right)
\end{aligned} \tag{8}$$

The key feature of the demand system is that Kindle adoption is driven by book usage intensity, which is further endogenized to be a function of consumers' book tastes and book prices. In this sense, e-book prices affect Kindle attractiveness. The book-side and device-side decisions are linked because (1) the ex ante flow utilities from books affect the Kindle adoption decisions, and (2) Kindle adoption statuses influence the book formats from which consumers can choose. Consumers who like reading will benefit more from having Kindles and adopt earlier.

4.2 Firm Problem

I present the full firm problem in this subsection. A simple two-period model is presented in the Appendix to illustrate the basic trade-offs and the new features of complementary product IPD.

I take a normative stance and estimate the demand system without assuming the optimality of the observed prices. I use the demand estimates to compute Amazon's optimal Kindle and e-book pricing strategies. The consumers' and the firm's dynamic problems are jointly solved in the simulation. I take Kindle costs and book wholesale prices from industry reports.[30] I abstract from paperback price responses because it is computationally prohibitive to solve and requires supply-side data on other retailers. Consumers' paperback retailer choices still respond to alternative pricing strategies as Kindle ownership affects these probabilities and is endogenous to Kindle and e-book pricing. To keep the model tractable, I make several simplifications from the demand model. First,

[30] The wholesale prices are \$15 for paperbacks and \$12 for e-books. However, the observed Amazon e-book price is \$9.72, and the simulation cannot generate such a low price level because unobserved factors – such as spillover effects into Amazon's other product business, negotiated quantity discounts that are not publicly observed, and competition pressure – change Amazon's actual marginal cost. To obtain a more realistic marginal cost value, I allow for a spillover effect per book transaction in the simulation. The predictions on dynamic pricing are very robust to different magnitudes of this spillover effect. I choose one magnitude so that the simulated e-book price *level* is comparable to the first-period observed price. Note that this does not match the entire price path because I still take a normative view on the dynamic pricing policy. Gentzkow (2007) adopts a similar approach when rationalizing the zero prices of online newspapers.

I abstract from the quality improvements and use the average estimated quality in the simulation. Quality improvements are often intertwined with price changes. The lack of R&D data and the computational burden also prevent me from jointly solving for qualities and prices. I assume that the cost still declines exogenously over time at the rate of computer parts. Robustness checks show that the pricing results are robust if the cost is constant or decreases at different rates. The details are in the Appendix. Consumers have only two device ownership statuses: Kindle owner and nonowner.[31] Second, I restrict the pricing policy to be functions of only the two unobserved types and average over the observed demographic types, which helps to considerably reduce the state space, from 36 dimensions to two dimensions, while maintaining the major heterogeneity. Third, I solve for one e-book price and let the prices in the demand model change with it uniformly across genres.

Amazon sets the Kindle price P and e-book price p^E to maximize its total discounted profits from Kindles, paperbacks, and e-books. The firm's state space Δ is a vector that contains the number of Kindle nonowners for each type at the beginning of the current period. The demand system provides two key inputs to the firm's pricing problem: (a) the vector of Kindle adoption probabilities ϕ_t and (b) the book profits that Amazon earns from each Kindle owner $R_t^1 = \left(p_t^E - w^E\right) \cdot q_t^E \left(p_t^E\right) + \left(p_t^P - w^P\right) \cdot q_t^{P1} \left(p_t^E\right) r_{At}^1$ and nonowner $R_t^0 = \left(p_t^P - w^P\right) \cdot q_t^{P0} r_{At}^0$, where $\left\{w^P, w^E\right\}$ are the wholesale prices paid to publishers and $\left\{r_{At}^1, r_{At}^0\right\}$ are the paperback retailer probabilities. As shown later in the estimation results, $r_{At}^1 > r_{At}^0$ holds so that consumers prefer Amazon as a paperback retailer after they adopt Kindles; $q_t^E + q_t^{P1} > q_t^{P0}$ holds so that consumers buy more books after they adopt Kindles. Both results suggest that $R_t^1 > R_t^0$ holds so that Amazon benefits from converting a nonowner to an owner and has incentives to "invest" in new Kindle adopters. The number of nonowners in the next period Δ_{t+1} equals the probability of not buying/upgrading times the number of nonowners in this period Δ_t, which indicates that the state space evolves deterministically as $\Delta_{t+1} = \left[I - \phi_t\right] \Delta_t$. The Bellman equation of the firm is

$$\begin{aligned} EW_t\left(\Delta_t\right) &= \max_{P_t, p_t^E} \pi_t\left(P_t, p_t^E, \Delta_t\right) + \delta E\left[W_{t+1}\left(\Delta_{t+1}\right) \mid P_t, p_t^E, \Delta_t\right] \\ \pi_t\left(P_t, p_t^E, \Delta_t\right) &= \underbrace{\left(\phi_t \cdot \Delta_t\right)\left[P_t - c_t\right]}_{\text{Kindle profit}} + \underbrace{\left(I - \phi_t\right) \cdot \Delta_t \cdot R_t^0}_{\text{book profit: nonowner}} + \underbrace{\left(\Delta_0 - \Delta_t + \phi_t \cdot \Delta_t\right) \cdot R_t^1\left(p_t^E\right)}_{\text{book profit: owner}} \end{aligned} \tag{9}$$

where Δ_0 is the initial market size at period 0. Note that I need to compute the value function

[31] The upgraders are modeled in a simplified way. They have proportionally higher device flow utilities than first-time adopters. The proportion is calculated from the average value in the demand system.

separately for each period because the Kindle cost c_t, paperback prices p_t^P, and retailer probabilities $\left\{r_{At}^1, r_{At}^0\right\}$ differ across periods. I assume that these variables stop evolving and remain at year 2012 levels to keep the problem stationary. Taking the F.O.C. with respect to the Kindle price yields

$$\underbrace{\Delta_t \cdot \frac{\partial \phi_t}{\partial P_t}\left[P_t - c_t\right] + \phi_t \cdot \Delta_t}_{\text{static Kindle profit change}} + \underbrace{\Delta_t \cdot \frac{\partial \phi_t}{\partial P_t} \cdot \left(R_t^1 - R_t^0\right)}_{\text{static book profit change}} - \underbrace{\Delta_t \cdot \frac{\partial \phi_t}{\partial P_t} \cdot \delta \frac{\partial W_{t+1}\left(\Delta_{t+1}\right)}{\partial \Delta_{t+1}}}_{\text{dynamic future state change}} = 0 \tag{10}$$

Statically, the firm needs to manage the size and the mix of Kindle owners (affected by p_t^E and P_t) and earn higher profits from each owner (affected by p_t^E). A higher Kindle price increases the marginal gain on the existing Kindle sales (the first term) at the expense of gains from new adopters (the second term) and their associated book profits (the third term). The demand elasticities dictate the magnitudes of these effects. Dynamically, two effects are captured in the fourth term: (1) a higher current Kindle price reduces the future market size and changes the future mix of the two consumer types; (2) the current prices affect consumers' expectation of future prices, which, in turn, affects current adoption. Taking the F.O.C. with respect to the e-book prices yields the following equation with similar trade-offs:

$$\begin{aligned} &\underbrace{\Delta_t \cdot \frac{\partial \phi_t}{\partial p_t^E} \cdot \left(R_t^1 - R_t^0\right) + \left[\Delta_0 - \Delta_t + \phi_t \cdot \Delta_t\right] \cdot \frac{\partial R_t^1}{\partial p_t^E}}_{\text{static book profit change}} + \\ &\quad \underbrace{\Delta_t \cdot \frac{\partial \phi_t}{\partial p_t^E}\left[P_t - c_t\right]}_{\text{static Kindle profit change}} - \underbrace{\Delta \cdot \frac{\partial \phi_t}{\partial p_t^E} \cdot \delta \frac{\partial W_{t+1}\left(\Delta_{t+1}\right)}{\partial \Delta_{t+1}}}_{\text{dynamic future state change}} = 0 \end{aligned} \tag{11}$$

I consider the pure-strategy Markov-perfect Nash equilibrium (MPNE) in which both consumers and the firm are forward-looking. The setup is similar to the frameworks in Nair (2007) and Goettler and Gordon (2011). The equilibrium requires that the consumer's expectation over the future state is consistent with the firm's optimal strategy. The equilibrium is defined as the set $\left\{V^*, W^*, P^*, p^{E*}, h^*\right\}$, which contains the equilibrium value functions for the consumers and the firm, the optimal pricing policy functions for Kindles and e-books, and the beliefs about the next period's state space.

5 Demand Estimation Methods

5.1 Likelihood Function

To combine individual-level book and Kindle purchase data with aggregate offline book sales data, I use maximum likelihood methods in which the log-likelihood of the micro data serves as the objective function and the aggregate data serve as a set of over-identifying restrictions.[32] The constrained optimization problem is

$$\begin{aligned} \max \quad & \mathcal{L}\left(\Theta | X, \left\{q^{P1}, q^{E}\right\}, q^{P0}, j, d\right) \\ & s.t.\ \hat{H}_t\left(\Theta\right) = H_t \end{aligned}$$

where Θ represents the model parameter set, X contains the observed household demographics, Kindle and book prices, and e-book availability. $\left\{q^{P1}, q^{E}, q^{P0}, j, d\right\}$ represent the vectors of book quantity choices for Kindle owners and nonowners, retailer choices, and device choices. $\mathcal{L}\left(\Theta | X, \left\{q^{P1}, q^{E}\right\}, q^{P0}, j, d\right)$ is the log-likelihood of the micro data, $\hat{H}_t\left(\Theta\right)$ are the model-predicted aggregate offline book sales and device penetration rates, and H_t are the observed ones.[33] I discuss how $\mathcal{L}\left(\Theta | X, \left\{q^{P1}, q^{E}\right\}, q^{P0}, j, d\right)$ and $\hat{H}_t\left(\Theta\right)$ are constructed below.

The log-likelihood function for the individual observations is the log of the *joint* probabilities of the individual's device adoption choices and book format-quantity-retailer choices. These probabilities are conditional probabilities, all of which condition on the unobserved heterogeneous consumer segment. These decisions are jointly modeled and linked through the unobserved consumer heterogeneity. In the device adoption data, individuals' probabilities of purchasing a Kindle or waiting contribute to the likelihood function as follows:

[32] This approach adopts the maximum-likelihood estimation (MLE) framework and uses macro data as constraints of the maximization problem. The idea is similar to the MPEC (mathematical program with equilibrium constraints) approach, which chooses the structural parameters to maximize the likelihood of the data subject to the equilibrium constraints that contain the structural parameters (Su and Judd 2012). The idea is also similar to the integrated estimation procedure in Chintagunta and Dubé (2005), where they maximize the likelihood of the micro data in the first step and match the model-predicted aggregate shares to the observed shares in the second step. Another approach of combining micro and macro data is to adopt a generalized method of moments (GMM) framework and convert likelihood into moments as in Imbens and Lancaster (1994). The idea is to treat the score functions of the micro likelihood as moments and combine them with the aggregate moments. In my case, calculating the score function requires numerical approximation of derivatives for 40 variables and is computationally prohibitive. I adopt the MLE framework for convenience and tractability reasons.

[33] Here I assume that the observed sales are the population values without sampling error, as the offline sales data are from an industry-wide report. The observed sales value is divided by population and multiplied by sample size to account for the size difference between the industry report and the data set.

$$
\begin{aligned}
\ell_i\left(\Theta|d_{it}=1\right) &= \sum_k \lambda^k \left\{ \Psi^{k,0}_{i,t-1} \cdot \phi\left(d_{it}=1 \mid 0, \Omega_t, k\right) + \sum_{\tau=2008}^{t-1} \Pr\left(\bar{Q}_{i,t-1} = Q_{i\tau} \mid k\right) \cdot \phi\left(d_{it}=1 \mid Q_{i\tau}, \Omega_t, k\right) \right\} \\
\ell_i\left(\Theta|d_{it}=0\right) &= \sum_k \lambda^k \left\{ \Psi^{k,0}_{i,t-1} \cdot \left[1-\phi\left(d_{it}=1 \mid 0, \Omega_t, k\right)\right] \right. \\
&\quad \left. + \sum_{\tau=2008}^{t-1} \Pr\left(\bar{Q}_{i,t-1} = Q_{i\tau} \mid k\right) \cdot \left[1-\phi\left(d_{it}=1 \mid Q_{i\tau}, \Omega_t, k\right)\right] \right\} \qquad (12)
\end{aligned}
$$

where $\phi\left(d_{it}=1 \mid Q_{i\tau}, \Omega_t, k\right)$ is the conditional probability of buying/upgrading in Equation 8. $\Pr(\bar{Q}_{it} = Q_{i\tau} \mid k)$ is the probability of having a particular Kindle version $Q_{i\tau}$ at the end of period t and $\left\{\Psi^{k,1}_{it}, \Psi^{k,0}_{it}\right\}$ are the probabilities of having and not having *any* Kindles at time t. The latter two probabilities can be recursively derived from the conditional probabilities $\phi\left(d_{it}=1 \mid Q_{\tau}, \Omega_t, k\right)$ (details are provided below). Note that I do not need to observe consumers' book purchases (or the realized preference shocks η_{igt}s) when calculating these device-side probabilities because they contain only the ex ante indirect book utilities (η_{igt}s are integrated out when calculating the indirect book utilities in Equation 4).

In the book purchase data, individuals' device adoption status is not observed due to yearly resampling, meaning that I need to integrate out the device ownership probabilities when constructing the likelihood. Gowrisakanran and Rysman (2012) adopt a similar approach and use the model-predicted ownership distribution to construct the market share. I assume that consumers who bought e-books in a year are Kindle owners in that year. For consumers who have not bought any e-books in a year, I assume that they are Kindle owners with probability $\Psi^{k,1}_{it}$ and are Kindle nonowners with probability $\Psi^{k,0}_{it}$. Consumers' contribution to the likelihood function equals the joint probability of their format-quantity choices $\left\{q^{P1}_{igt}, q^{E}_{igt}, q^{P0}_{igt}\right\}$, their retailer choices j, and their device ownership statuses (1 represents having Kindles or e-book purchases and 0 represents no e-book purchases):

$$
\begin{aligned}
\ell_i\left(\Theta|\left\{q^{P1}_{igt}, q^{E}_{igt}\right\}, j, 1\right) &= \sum_k \lambda^k \cdot \Pr\left(\left\{q^{P1}_{igt}, q^{E}_{igt}\right\} \mid k\right) \cdot r^1_{ijgt} \cdot \Psi^{k,1}_{it} \qquad (13) \\
\ell_i\left(\Theta|\left\{q^{P0}_{igt}\right\}, j, 0\right) &= \sum_k \lambda^k \cdot \left[\Pr\left(\left\{q^{P0}_{igt}\right\} \mid k\right) \cdot r^0_{ijgt} \cdot \Psi^{k,0}_{it} + \Pr\left(\left\{q^{P0}_{igt}, 0\right\} \mid k\right) \cdot r^1_{ijgt} \cdot \Psi^{k,1}_{it}\right]
\end{aligned}
$$

where $\Pr\left(\left\{q^{P1}_{igt}, q^{E}_{igt}\right\} \mid k\right)$ and $\Pr\left(\left\{q^{P0}_{igt}\right\} \mid k\right)$ are the format-quantity probabilities (derived below).[34] The total log-likelihood of the book purchase data and the device purchase data is

$$\begin{aligned}
&\mathcal{L}\left(\Theta|X,\left\{q^{P1},q^{E}\right\},q^{P0},j,d\right)= \\
&\sum_t\sum_i\left[1\left\{q^{P1}_{igt},q^{E}_{igt},j,1\right\}\log\ell_i\left(\Theta|q^{P1}_{igt},q^{E}_{igt},j,1\right)+1\left\{q^{P0}_{igt},j,0\right\}\log\ell_i\left(\Theta|q^{P0}_{igt},j,0\right)\right] \\
&+\sum_t\sum_i\left[1\left\{d_{it}=1\right\}\log\ell_i\left(\Theta|d_{it}=1\right)+1\left\{d_{it}=0\right\}\log\ell_i\left(\Theta|d_{it}=0\right)\right]
\end{aligned}$$

Finally, the model-predicted aggregate variables contain offline paperback sales of genre g at time t and device penetration rates at time t, $\hat{H}_t(\Theta)\equiv\left[\left\{\hat{H}^1_{gt}(\Theta)\right\}_{g=1,2,3},\hat{H}^2_t(\Theta)\right]$. The predicted device penetration rate equals the cumulative probabilities of new adoption up to time t, $\hat{H}^2_t(\Theta)=\sum_{\tau=1}^{\tau=t}\sum_k\lambda^k\Psi^{k,0}_{i,\tau-1}\cdot\phi\left(d_{i\tau}=1\mid 0,\Omega_\tau,k\right)$. To obtain $\left\{\hat{H}^1_{gt}(\Theta)\right\}_{g=1,2,3}$, I first simulate the error terms in the taste parameter $\left\{\eta^P_{igt},\eta^E_{igt}\right\}$ 10,000 times for each individual. Given these error terms and the parameter values Θ, for each consumer segment k (i.e., $\{\theta_{ig}\}=\left\{\theta^k_g\right\}$), I use Equation 2 and 3 to calculate the predicted paperback purchase quantities for a Kindle nonowner $\hat{q}^{P0,k}_{igt}$ and an owner $\hat{q}^{P1,k}_{igt}$. The individual offline paperback consumption equals the probability of being a Kindle owner/nonowner $\left\{\Psi^{k,1}_{it},\Psi^{k,0}_{it}\right\}$ times the corresponding paperback quantity $\left\{\hat{q}^{P1,k}_{igt},\hat{q}^{P0,k}_{igt}\right\}$ times the probability of choosing offline retailers $\left\{r^1_{iOgt},r^0_{iOgt}\right\}$. Summing over the individuals, we can obtain the aggregate predicted sales $\hat{H}^1_{gt}(\Theta)=\sum_i\sum_k\lambda^k\left[\hat{q}^{P1,k}_{igt}\cdot r^1_{iOgt}\cdot\Psi^{k,1}_{it}+\hat{q}^{P0,k}_{igt}\cdot r^0_{iOgt}\cdot\Psi^{k,0}_{it}\right]$. The set of simulated error terms is fixed throughout the estimation to keep the problem stationary.

Derivation of Kindle ownership probabilities. Start with 2008, when no Kindles were owned $\Psi^{k,0}_{i,2008}=1$. The probability of having a Kindle $Q_{i,2008}$ at the end of 2008 equals the probability of buying, so $\Pr\left(\bar{Q}_{i,2008}=Q_{i,2008}\mid k\right)=\phi\left(d_{i,2008}=1\mid 0,\Omega_{i,2008},k\right)$. For the following periods, the probability of having the latest Kindle version Q_t at time t equals the probability of buying for Kindle non-owners plus the probability of upgrading for previous Kindle owners. The probability of holding a previous Kindle version $Q_{i\tau}$ equals the probability of owning $Q_{i\tau}$ times the probability of not upgrading:

[34] The device-side and book-side probabilities are only linked through the observed demographics $\left\{D^{age}_i,D^{income}_i\right\}$ and the unobserved segment k; the idiosyncratic shocks η_{igt}'s affect only the book-side probabilities and do not affect the device-side probabilities. Therefore, the joint device- and book-side probability, given $\left\{D^{age}_i,D^{income}_i\right\}$ and k, can be written as the product of the two marginal probabilities. The retailer choice probabilities $\left\{r^0_{ijgt},r^1_{ijgt}\right\}$ are not indexed by k because k affects the the retailer choice only through the Kindle ownership status $\{1,0\}$; conditional on $\{1,0\}$, k does not affect $\left\{r^0_{ijgt},r^1_{ijgt}\right\}$.

$$\Pr\left(\bar{Q}_{it}=Q_{i\tau}\mid k\right)=\begin{cases}\Psi^{k,0}_{i,t-1}\cdot\phi\left(d_{it}=1\mid 0,\Omega_t,k\right)\\ \quad+\sum_{\tau'=2008}^{t-1}\Pr\left(\bar{Q}_{i,t-1}=Q_{i\tau'}\mid k\right)\cdot\phi\left(d_{it}=1\mid Q_{i\tau'},\Omega_t,k\right), & \text{if } \tau=t\\ \Pr\left(\bar{Q}_{i,t-1}=Q_{i\tau}\mid k\right)\left[1-\phi\left(d_{it}=1\mid Q_{i\tau},\Omega_t,k\right)\right], & \text{if } \tau<t\end{cases}\tag{14}$$

Given these probabilities of having a particular Kindle $Q_{i\tau}$ at time t, the probability of having and not having *any* Kindles are $\Psi^{k,1}_{it}=\sum_{\tau\leqslant t}\Pr\left(\bar{Q}_{it}=Q_{i\tau}\mid k\right)$ and $\Psi^{k,0}_{it}=1-\Psi^{k,1}_{it}$.

Derivation of format-quantity probabilities. $\Pr\left(\left\{q^{P1}_{igt},q^{E}_{igt}\right\}\mid k\right)$ and $\Pr\left(\left\{q^{P0}_{igt}\right\}\mid k\right)$ are constructed based on the feasible range of the normally distributed error terms $\left\{\eta^{P}_{igt},\eta^{E}_{igt}\right\}$ implied by the optimal quantity choices in Equation 2. To simplify the notation, I drop the i, g, and t subscripts and the type notation k for now. Define the realized error terms given the quantity choice q^P as $\eta\left(q^P\right)\equiv q^P+bp^P-\bar{a}^P$ and $\eta\left(q^E\right)\equiv q^E+bp^E-\bar{a}^E$. Define the thresholds of worth buying as $\bar{\eta}^P\equiv bp^P-\bar{a}^P$ and $\bar{\eta}^E\equiv bp^E-\bar{a}^E$. The format-quantity probabilities for Kindle nonowners and owners are

$$\begin{aligned}
\Pr\left(\left\{q^{P0}=0\right\}\right) &= \Pr\left(\eta^{P0}\leqslant\bar{\eta}^P\right)=\Phi\left(\bar{\eta}^P/\sigma\right)\\
\Pr\left(\left\{q^{P0}=q^{P0*}>0\right\}\right) &= f\left(\eta^{P0}=\eta\left(q^{P0*}\right)\mid\eta^{P0}>\bar{\eta}^P\right)\Pr\left(\eta^{P0}>\bar{\eta}^P\right)=\frac{1}{\sigma}\phi\left(\eta\left(q^{P0*}\right)/\sigma\right)\\
\Pr\left(\left\{q^{P1}=0,q^E=0\right\}\right) &= \Pr\left(\eta^{P1}\leqslant\bar{\eta}^P,\eta^E\leqslant\bar{\eta}^E\right)=\Phi\left(\bar{\eta}^P/\sigma\right)\Phi\left(\bar{\eta}^E/\sigma\right)\\
\Pr\left(\left\{q^{P1}=q^{P1*}>0,q^E=0\right\}\right) &= f\left(\eta^{P1}=\eta\left(q^{P1*}\right)\mid\eta^{P1}>\max\left\{\bar{\eta}^P,\,\eta^E+\left(\bar{\eta}^P-\bar{\eta}^E\right)\right\}\right)\\
&\quad\cdot\Pr\left(\eta^{P1}>\max\left\{\bar{\eta}^P,\,\eta^E+\left(\bar{\eta}^P-\bar{\eta}^E\right)\right\}\right)\\
\Pr\left(\left\{q^E=q^{E*}>0,q^{P1}=0\right\}\right) &= f\left(\eta^E=\eta\left(q^{E*}\right)\mid\eta^E>\max\left\{\bar{\eta}^E,\,\eta^{P1}-\left(\bar{\eta}^P-\bar{\eta}^E\right)\right\}\right)\\
&\quad\cdot\Pr\left(\eta^E>\max\left\{\bar{\eta}^E,\,\eta^{P1}-\left(\bar{\eta}^P-\bar{\eta}^E\right)\right\}\right)
\end{aligned}\tag{15}$$

where $\phi\left(\cdot\right)$ and $\Phi\left(\cdot\right)$ are the PDF and CDF of the standard normal distributions.

5.2 Identification

The book-side parameters include taste parameters $\left\{\theta_{ig},\,\theta^E_g,\,\beta_1,\,\beta_2\,,\beta_3,\,\sigma\right\}$, retailer fixed effects and time trends, and price coefficient b_i. The genre fixed effects in the baseline taste θ_{ig} and the e-format taste θ^E_g are identified from genre- and format-specific book sales. The price coefficient b_i is identified mainly from the price gap between paperbacks and e-books of the same genre (it differs across genres and varies over time) and the corresponding relative shares of paperbacks and

Table 2: Identification: Simulation Results of Varying Book Price Coefficient

Price coefficient b_0 (change)	0.1541 (- 0.010)	0.1591 (- 0.005)	0.1641 0	0.1691 (+ 0.005)	0.1741 (+ 0.010)
Kindle cumulative sales / market size by 2012 (%)	13.07	12.56	12.09	11.67	11.28
# paperbacks that a typical Kindle owner buys	2.38	2.30	2.23	2.16	2.10
# e-books that a typical Kindle owner buys	5.80	5.75	5.71	5.64	5.59

e-books. It is also identified from the link between Kindle and book consumption.[35] In particular, if consumers are more price sensitive, they will care more about the price advantage of e-books over paperbacks and thus be more attracted to Kindles. To further illustrate the identification of the price coefficient, I conduct a simulation by varying the book price coefficient b_0 around its estimated value and simulating consumers' Kindle and book consumption. As shown in Table 2, the change in the price coefficient causes significant changes in both the aggregate cumulative Kindle sales (as a percentage of total market size) and the number of paperbacks and e-books that a typical Kindle owner buys in a year. Finally, consumers' consumption patterns across age groups identify the coefficients on age β_1 and β_2. The time-varying substitution patterns between e-books and paperbacks identify the coefficient on time-varying e-book availability β_3. The retailer market shares and their variations over time identify the retailer fixed effects and time trends.

The device-side parameters include $\left\{\Gamma,\, \alpha_i,\, \{Q_{it}\}_{t=2008}^{t=2012}\right\}$. These parameters enter the conditional purchase probabilities $\phi\left(d_{it}=1 \mid Q_{i\mathcal{T}}, \Omega_t, k\right)$ which further determines the model-predicted yearly sales, as shown in Equations 12 and 14. The observed demographics enter book preferences and affect the book and Kindle price coefficients. The Kindle sales across the nine demographic groups over the five years identify the seven unknown parameters in $\phi\left(d_{it}=1 \mid Q_{i\mathcal{T}}, \Omega_t, k\right)$. Intuitively, the coefficient on book utility Γ is cross-sectionally identified from, given the same Kindle price and quality, the different device adoption probabilities of consumers based on the model-predicted distribution of device ownership statuses. The price coefficient α_i and Kindle quality dummies $\{Q_{it}\}_{t=2008}^{t=2012}$ are jointly identified from two sources: (1) cross-sectionally, the different adoption/upgrade probabilities of consumers given their model-predicted device ownership statuses; and

[35] To observe this, substitute Equation 3 into Equation 2 and drop the demographic terms and error terms for now. We have $q_{igt}^{P0} = \theta_{ig} + \xi_t - b_i p_{gt}^{E}$ and $q_{gt}^{P1} - q_{igt}^{P0} = \theta_g^E - b_i\left(p_{gt}^P - p_{gt}^E\right)$. In each period, there are seven unknowns (six θ_{ig}, θ_g^E and one b_i) and six book sales observations/conditions (three genres and two formats). The device side imposes another condition; the optimal device choice in Equation 8 contains the indirect book utilities in Equation 4, which are also functions of these unknowns. These conditions jointly identify the price coefficient.

(2) intertemporally, the adoption/upgrade probabilities for each consumer type. They are separately identified because price is incurred only once and quality enters utility every period. The link between book sales and Kindle sales imposes over-identifying restrictions on these parameters.

An identification challenge is that the individual-level data are yearly re-sampled, meaning that I do not fully observe consumers' device ownership status and cannot link consumers' device choices to book choices in the data. I identify first-time buyers and upgraders by combining the Kindle sales data and the supplementary penetration rate data. In general, new buying and upgrading can be identified even with aggregate sales data (Gowrisankaran and Rysman 2012). The penetration rate data further facilitate identification, as the difference between sales and the incremental number of households owning a device represents the extent of repeat purchasing. Gowrisankaran and Rysman (2012) and Gordon (2009) adopt similar identification strategies when studying the dynamic durable product demand. They combine aggregate sales data and penetration data to identify new purchases versus upgrades.

My data set contains two additional information sources that help identification of consumer heterogeneity and new buying versus upgrading compared to Gowrisankaran and Rysman (2012) and Gordon (2009). First, my model contains the book purchase decision in addition to the device decision. The degree to which e-book sales change with respect to Kindle sales further helps identification. Specifically, e-book sales are directly linked to the number of Kindle owners. In the extreme case, if additional Kindle sales in a year do not lead to additional e-book sales in that year, these Kindle sales do not generate additional Kindle owners and would have come from upgraders. Second, the book-side data contain additional device ownership information: I can partially observe consumers' device ownership status in that consumers who bought e-books should have a Kindle. These consumers' book purchases can help identify the utility of owning Kindles relative to the value of the outside option of no purchase.

I further conduct a Monte Carlo study to access 1) the model's ability to use yearly re-sampled data to identify new buying versus upgrading and unobserved consumer heterogeneity and 2) the role of the supplementary penetration rate data.[36] I first use the true parameter values to simulate

[36] Gordon (2009) uses a similar Monte Carlo simulation to illustrate how combining aggregate sales data and penetration data can identify new purchase versus upgrading. Gowrisankaran and Rysman (2012) compare the estimates with and without penetration data to illustrate the role of the supplementary penetration data and find similar biases in the estimates without penetration data.

book and Kindle purchase decisions at the individual level. I then construct a yearly re-sampled data set, which has the same features as the main data set, and construct the supplementary penetration rate data. I estimate the model using the yearly re-sampled data, with and without the supplementary penetration rate data, and compare the estimates with the true parameters. I also compare the implied statistics on consumer device adoption decisions such as cumulative Kindle sales, penetration rates, and the percentage of sales that comes from upgrading every period. The results suggest that the yearly re-sampled data alone can produce close estimates of the true parameters. The supplementary penetration data further facilitate identification, and estimates will be biased if they are excluded from the estimation. Combining the sales data with penetration data is able to recover both new buying versus upgrading and unobserved consumer heterogeneity. The details are in Section 8.5 in the Appendix.

Price endogeneity. The demand estimation is conducted without imposing pricing optimality conditions. In general, the observed prices might be endogenized to unobserved qualities and demand shocks. In my setting, neither standard instruments nor standard bias-correction approaches, such as BLP (Berry, Levinsohn and Pakes 1995) and the control function (Petrin and Train 2010), are readily usable.[37] I thus use fixed effects to control for the unobservables. I explicitly model and estimate the qualities of different Kindle generations. I use genre fixed effects to capture time-invariant genre quality and time fixed effects to capture time-varying unobservables.[38] Market- and individual-specific unobservables seem to be less troubling in my setup. First, book prices do not seem to be endogenous to individual-level or market-level unobservables. Amazon does not price discriminate based on location or demographics.[39] I also regress the genre-specific individual book prices on household demographics (a proxy for observed consumer types), a dummy for e-book purchase (a proxy for unobserved consumer taste types), the number of bookstores in his/her zip

[37] First, standard instruments such as marginal cost shifters and geographical price variations are not suitable in the book industry because book wholesale prices are very stable over time and book prices and characteristics do not vary across markets. Second, the BLP approach requires market shares with relatively little sampling error, which cannot be produced using individual-level transaction data in this paper. The demand model is also too complicated to estimate BLP controls. Third, the control function approach requires recovering controls using a regression such as OLS. This type of regression is not appealing, given that the observed Kindle prices changed annually and that the sample contains only five years of data.

[38] The time fixed effects ξ_t are the same across genres and formats. I cannot include genre- and format-specific time fixed effects because they would absorb all the variation in prices.

[39] Amazon's CEO Jeff Bezos promised that Amazon never will price based on customer demographics after a failed "test" of DVDs in 2000 (White House Report "Big Data and Differential Pricing" 2015). Several test searches for a sample of books support this claim.

code (from the Esri Business and Demographics Database, as a proxy for local tastes), and time fixed effects. The coefficients on household age and number of bookstores are insignificant, which seems to suggest no systematic price differences across markets and individuals.[40] Second, there is no systematic change in the observed book prices over time. Price variations come mainly from the difference in the mix of available books over time (similar to choice set changes in logit choice models, affecting sales-unweighted prices) and the switch of contract scheme between publishers and retailers in 2010. Note that the price variations do not come from consumer preferences or book qualities. First, if the price variations indeed come from consumer preferences for individual book titles, then one would see a difference between the sales-weighted and the sales-unweighted average prices. However, the two types of average prices differ by only 2% on average, and one is not systematically higher or lower than the other. Using either one produces robust estimation results. Second, bestsellers (reflecting book quality and consumer preference) do not seem to drive the average genre prices either, as 92.82% of the book titles had only one purchase record per year, 5.53% had two purchases, and 99.94% had fewer than 10 purchases.[41]

5.3 Computational Methods

To estimate the demand model, I use the Nested Fixed Point algorithm (NFXP) proposed by Rust (1987). For each iteration, I solve the dynamic programming problem for each consumer observed demographics group and unobserved taste segment holding different Kindle versions in the inner loop and use MLE in the outer loop.

Function approximations are used in the demand estimation. The demand side involves calculating conditional expectations and probabilities of a truncated normal error, of which the truncation point is a result of a maximization operator. Gauss-Hermite quadrature is used to calculate the conditional expectations inside $v_{it}^{book,1}$ and $v_{it}^{book,0}$ in Equation 4. The Gauss-Chebychev quadrature and Gauss-Laguerre quadrature are used to calculate the format-quantity choice probabilities in Equation 15. Details are presented in the Appendix.

[40] The coefficients on household income and the e-book purchase dummy are significant, as captured by the heterogeneous price coefficient b_i and unobserved reading tastes θ_{ig} in the full empirical model.

[41] The sample bestseller was bought 67 times, representing 0.46% of the total yearly sales, which is comparable to real books in the industry. The consumption is highly dispersed for all genres and reading formats.

6 Demand-Side Estimation Results

6.1 Model Fit

Table 3 displays the observed and predicted Kindle cumulative sales (as a percentage of total market size) and penetration rates over time. Figure 2 compares the observed and predicted book sales by format, genre, and retailer over time. Data from 2008 to 2012 are used in the estimation, and 2013 data are used as an out-of-sample fit test. The model fits the aggregate-level Kindle and book sales well.

At the individual level, model predictions can also be validated using survey data (The Pew Research Center, February 2012). According to the survey, an e-reader owner bought 1.7 times more books than a nonowner did in the past 12 months. The model predicts that a typical Kindle owner buys 1.64 times more books than a Kindle nonowner does.[42] Both the aggregate-level and individual-level model fits indicate that the model can recover the values of different book formats, genres, retailers, and device adopting and waiting.

Table 3: Model Fit: Kindle Sales and Penetration Rates

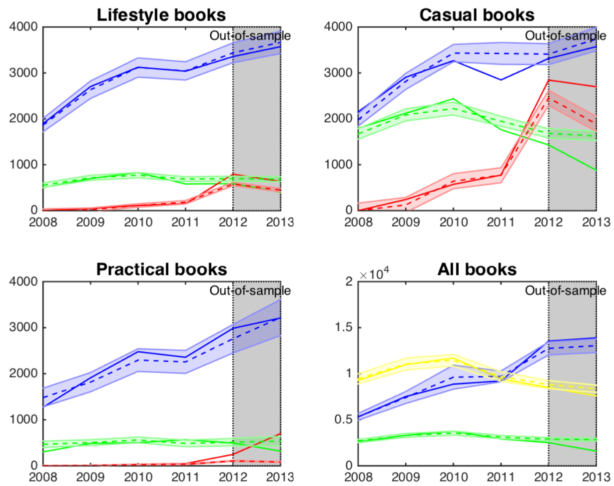
(%)	Cumulative Sales / Market Size			Penetration Rates		
Year	Observed	Predicted	Std.	Observed	Predicted	Std.
2008	0.41	0.29	(0.16)	0.41	0.28	(0.15)
2009	1.36	1.30	(0.21)	1.03	1.25	(0.21)
2010	3.54	4.13	(0.32)	2.46	3.69	(0.31)
2011	6.84	7.89	(0.44)	5.33	6.42	(0.45)
2012	10.34	12.09	(0.93)	6.56	7.09	(0.91)
2013	14.54	16.73	(1.34)			

6.2 Parameter Estimates

Parameter interpretations. Table 4 reports the parameter estimates. The estimates on the baseline book tastes show that consumers are highly heterogeneous in their unobserved genre-specific reading

[42] According to the model predictions in Table 5, avid readers' book consumption increases from 11.8 books to 15.8 books once they become Kindle owners, and general readers' book consumption increases from 1.1 books to 3.5 books. Of the total book buyer population, 6.8% are avid readers and 93.2% are general readers. By 2012, 26.0% of avid readers and 3.69% of general readers owned a Kindle. These findings imply that (6.8%*26.0%)/(6.8%*26.0%+93.2%*3.69%)=34.0% of Kindle owners are avid readers and that 100%-34.0%=66.0% are general readers. Overall, the model predicts that a typical Kindle owner buys 11.8*34.0%+1.1*66.0%=4.7 books before buying a Kindle and 15.8*34.0%+3.5*66.0%=7.7 books after buying one.

Figure 2: Model Fit: Books



Notes: Solid lines represent observed values, and dashed lines represent predicted values. The shaded areas represent 95% confidence intervals. Graphs 1-3 show the sales of Amazon paperbacks, other online retailer paperbacks, and Amazon e-books from top to bottom. Graph 4 show the sales of offline retailers, Amazon, and other online retailers from top to bottom.

tastes. The data identify two taste levels for each genre fixed effect θ_g^k , high and low, denoted by $\left\{\theta_g^H, \theta_g^L\right\}$ and four segments.[43] Segments 1, 2, and 3 have taste vector $\left\{\theta_1^L, \theta_2^H, \theta_3^L\right\}, \left\{\theta_1^H, \theta_2^L, \theta_3^H\right\}$, $\left\{\theta_1^H, \theta_2^H, \theta_3^H\right\}$, which represent consumers who have high reading tastes for "casual" books, for "lifestyle" and "practical" books, and for all books, respectively. They constitute 3.4% of the total population or 6.8% of the book buyers and share similar demand elasticities and consumption patterns. They also have perceived Kindle quality coefficients $\left\{Q_0^H, Q_1^H\right\}$. Segment 4 represents consumers who have low tastes for all genres and have perceived Kindle quality coefficients $\left\{Q_0^L, Q_1^L\right\}$. For the remainder of the discussion, I refer to the first three segments as "avid readers" and the fourth segment as "general readers." The estimates imply that an avid reader buys 9.62 more books than a general reader on average every year. I also find that older consumers enjoy reading more and that consumers in higher-income groups have lower price elasticities.

The estimates on consumers' heterogeneous preferences for Kindle qualities suggest that the avid readers have higher perceived qualities of Kindles than general readers do for earlier versions of Kindles, yet general readers have higher perceived quality *improvements* for newer versions of

[43] A complete combination of three genres and two levels leads to 2^3 segments. The estimated segment sizes are significantly different from zero for four out of the eight types.

Table 4: Parameter Estimates

Book	Lifestyle: $g=1$	Casual: $g=2$	Practical: $g=3$
		Baseline FE θ_{ig}	
θ_g^H	10.81*** (0.0705)	11.52*** (0.0457)	10.15*** (0.0301)
θ_g^L	0.6659*** (0.0252)	0.7058*** (0.0115)	2.236*** (0.0381)
		E-format FE	
θ_g^E	0.0544*** (0.0117)	2.714*** (0.0820)	-3.552*** (0.0291)
		Retailer FE & time trends	
A_{0g}	-0.3773*** (0.0180)	-0.7757*** (0.0227)	-0.4527*** (0.0391)
A_{1g} (time)	0.1375*** (0.0067)	0.1431*** (0.0092)	0.2267*** (0.0210)
B_{0g}	-1.578*** (0.0212)	-0.9207*** (0.0256)	-1.564*** (0.0391)
B_{1g} (time)	0.0581*** (0.0108)	0.0012*** (2.320e-4)	0.1046*** (0.0129)

	Device		**Book**		
α_0	0.0070***(0.0005)	m_1	0.0244*** (0.0030)	β_1	0.0865***(0.0059)
α_1	-1.543e-4***(4.352e-6)	m_2	0.0095*** (0.0040)	β_2	-3.702e-4***(8.120e-5)
Γ	12.01***(0.0256)	m_3	6.73e-5 (0.0019)	β_3	0.0004**(0.0002)
Q_0^L	-0.9737***(0.0520)	ξ_{2009}	0.0083 (0.0230)	b_0	0.1641***(0.0159)
Q_1^L	0.3981***(0.0352)	ξ_{2010}	0.0327 (0.0410)	b_1	-0.0120***(0.0021)
Q_0^H	-0.6067***(0.0490)	ξ_{2011}	0.0141 (0.0371)	σ	2.343***(0.0238)
Q_1^H	0.0027***(0.0315)	ξ_{2012}	0.0085 (0.0263)	A_2	0.5899***(0.0248)
	MLE Obj.: 197,115		# Obs: 89,382		

Notes: ***, **, * represent significance at the 1, 5, and 10 percent levels, respectively. Price coefficients enter utility negatively and vary by income group as $\alpha_i = \alpha_0 + \alpha_1 D_i^{income}$ and $b_i = b_0 + b_1 D_i^{income}$, where $D_i^{income} = 1,2,3$ and $D_i^{age} = 1,2,3$. The four consumer segments $k = 1,2,3,4$ have population mass $\{m_1, m_2, m_3, 1-m_1-m_2-m_3\}$ and genre baseline fixed effects $\left\{\theta_g^k\right\}_{g=1,2,3} = \left\{\theta_1^L, \theta_2^H, \theta_3^L\right\}, \left\{\theta_1^H, \theta_2^L, \theta_3^H\right\}, \left\{\theta_1^H, \theta_2^H, \theta_3^H\right\}, \left\{\theta_1^L, \theta_2^L, \theta_3^L\right\}$, respectively. Device qualities $\{Q_{it}\}_{t=2008}^{t=2012}$ are captured by $Q_{it} = Q_0^k + Q_1^k \log t$, where $t = 1$ to 5 represent years 2008 to 2012 and $\left\{Q_0^k, Q_1^k\right\}$ take the values $\left\{Q_0^H, Q_1^H\right\}$ for $k = 1,2,3$ and take the values $\left\{Q_0^L, Q_1^L\right\}$ for $k = 4$. ξ_{2008} is normalized to 0 for identification purposes.

Kindles. In other words, general readers value more the quality of Kindles than avid readers do; avid readers' Kindle adoption decisions are driven more by the book utility gain than the Kindle quality. This is consistent with the key finding on the relative demand elasticities (as discussed below): avid readers are relatively more price sensitive to books, while general readers are relatively more price sensitive to Kindles.

The estimates on e-format tastes show that consumers enjoy extra utilities from reading "lifestyle" and "casual" e-books and face disutilities from reading "practical" e-books; if there were no price differences between the two formats, the same consumer would buy 0.054 more "lifestyle" books, 2.71 more "casual" books, and 3.55 fewer "practical" books in the e-format than in the paperback format.[44] I also find that older consumers dislike the e-format and that e-book variety positively

[44] Besides utility-related genre differences, θ_g^E might also capture e-book availability differences across genres, if

affects e-format attractiveness.

Finally, the paperback retailer choice estimates show that consumers are migrating from offline to online and from other online retailers to Amazon.com. An interesting finding is that there is a positive correlation between Kindle ownership and Amazon retailer choice, as captured by a positive and significant estimate of A_2; a Kindle owner has a 53% probability of buying paperbacks from Amazon, while this probability is 39% for a Kindle nonowner.

Consumer heterogeneity. The estimates show that the unobserved heterogeneity, compared with the observed heterogeneity in income and age, leads to a much larger difference in demand elasticities and consumption patterns. I thus focus on the distinction between avid readers and general readers for the rest of the analysis.

Table 5 compares the demand elasticities and consumption behaviors of a typical avid reader and a typical general reader with medium levels of age and income ($D_i^{age} = 2, D_i^{income} = 2$), given average observed prices. I find that both avid readers and general readers buy more books after adopting Kindles. E-Books seem to be priced on the inelastic region of demand, which is consistent with the findings in previous literature (e.g., Reimers and Waldfogel 2014) and may suggest that Amazon invests in e-book pricing to stimulate Kindle sales in practice.

The key demand-side finding that drives the supply-side pricing strategy is the heterogeneous relative demand elasticities between Kindles and e-books. I find that avid readers have lower price elasticities for both Kindles and e-books than general readers in absolute terms, while they are relatively more price elastic to e-books than to Kindles: the ratio of the own-elasticity of Kindles to the own-elasticity of books is smaller for avid readers; the cross-elasticity of Kindles with respect to book prices is higher for avid readers. Both results suggest that avid readers are relatively more price sensitive to books than to Kindles. General readers are more price elastic to Kindles than to e-books. Intuitively, avid readers buy more books and spend relatively more on books than on devices. They care more about subsequent book purchases when considering buying Kindles. The different relative elasticities stem mainly from the unobserved heterogeneous baseline reading tastes θ_{ig}, which is the major difference between avid and general readers in the model.[45] Note

any, as the number of e-books available in the data is aggregate and not genre-specific.

[45] Although avid readers and general readers differ in how much they value earlier versus newer Kindle versions, their average perceived Kindle qualities across versions are very similar. The device-side heterogeneity is thus not the major driver of the results on relative demand elasticities.

Table 5: Consumer Heterogeneity in Kindle and Book Purchases

	Avid reader	General reader
Segment size	3.40%	96.60%
Demand elasticity		
Kindle	-2.93 (0.82)	-6.66 (1.65)
E-book	-0.30 (0.003)	-0.58 (0.005)
Ratio: Kindle/E-book	9.8	11.5
Cross-elasticity: Kindle w.r.t. E-book price	-1.24	-0.68
Book consumption per person per year		
Kindle nonowner: # paperbacks q^{P0}	11.8 (0.015)	1.08 (0.013)
Kindle owner: # paperbacks q^{P1}	4.71 (0.060)	0.86 (0.008)
Kindle owner: # e-books q^{E}	11.1 (0.140)	2.60 (0.076)
Kindle penetration rate by 2012 (%)	26.0(2.8)	3.69(0.3)

Notes: Standard errors are in parentheses.

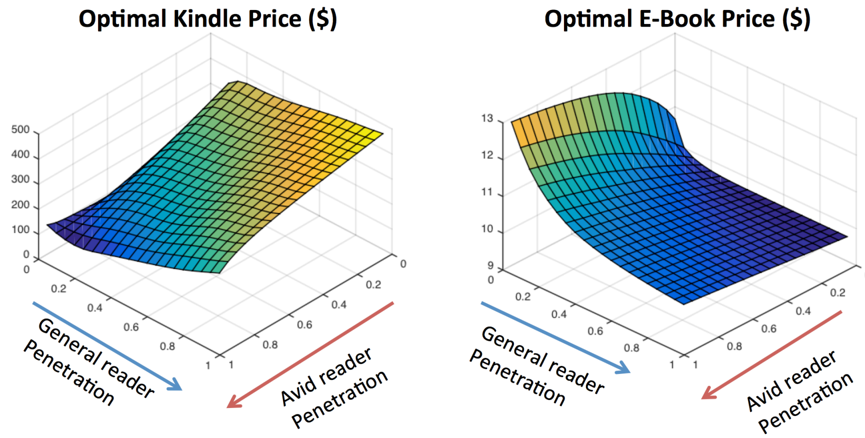
that the findings on relative demand elasticity come from estimation results and are not assumed by the model; if the estimated price coefficient doubles, the results would flip - avid readers would be relatively more price elastic to Kindles than to e-books.

7 Supply-side Simulation Results

Given the estimated demand system, I numerically solve for Amazon's optimal IPD policies. To solve for the supply-side equilibrium, I solve the firm's and the consumers' maximization problems in an inner loop and calculate the next-period states. The inner loop iterates until the consumers' beliefs and the updated next-period state space converge to a fixed point. The outer loop updates the value function guesses and iterates until convergence. The computation algorithm can be found in the Appendix. Function approximations are used in the supply simulation. First, the demand-side indirect book flow utilities $v_{it}^{book,1}$ and $v_{it}^{book,0}$ (which affect Kindle adoption probabilities ϕ_t) and book profits $\left\{R_t^0, R_t^1\right\}$ enter the supply side and are functions of the Kindle price P and the e-book price p^E in Equation 9. I evaluate them on a set of grid points for P and p^E and approximate them using splines.[46] Second, I discretize the state space into 20 grid points along each dimension. The value functions are approximated using cubic splines by interpolating between the grid points so that the functions are differentiable when computing the firm's first-order conditions.

[46]The functions are highly linear and level off as P and p^E increase. Linear splines with 11 breakpoints provide better approximation than cubic splines do, as cubic splines produce small fluctuations around the steady value and make the derivatives inaccurate when solving for the firm's first-order conditions.

Figure 3: Policy Functions



7.1 Optimal Pricing Policy

I find that the shapes of the optimal pricing functions differ for avid readers and general readers, as shown in Figure 3. Optimal policies are functions of the penetration rates of consumer types. Define skimming or harvesting (penetration or investing) as decreasing (increasing) mark-ups over time so that the pricing strategy represents strategic interactions rather than simply cost changes. I find that as the penetration rate of avid readers increases, the optimal strategy is to skim consumers with Kindles and invest in e-books. The opposite is true for general readers. The joint IPD policy exploits a new dimension of consumer heterogeneity, namely, the different relative demand elasticities between Kindles and e-books across consumers. Avid readers are more price elastic to e-books than to Kindles. The opposite is true for general readers. The overall price path balances the incentives for both consumer types and depends on the mix of consumers in the market.

7.2 The Benefit of Combining IPD and Complementary Products

What is the advantage of jointly conducting IPD on complementary products? I discuss the intuitions of combining IPD and complementary product pricing in this subsection and present counterfactual evidence in the next subsection.

In a traditional single-product IPD case, firms set high initial prices to skim high-valuation consumers and cut prices over time to appeal to low-valuation consumers. Their pricing power is limited by consumers' forward-looking behavior. Consumers anticipate future price changes and wait until the price drops. Firms need to account for this intertemporal substitution behavior when setting the price path. In the extreme case, firms are better off committing to a fixed price rather than conducting IPD, which is commonly known as the Coase conjecture.

In a traditional static complementary product case, firms set lower prices on products with more elastic demand and higher prices on products with less elastic demand. Their pricing power is limited when the relative elasticity is heterogeneous and conflicting among consumers; firms have to weight and balance across consumer types, meaning that prices tend to be driven closer to marginal costs (Rosen and Rosenfield 1997).

Combining IPD and complementary product pricing can reduce the limitations of each approach and enhance firms' overall ability to price discriminate. First, IPD helps complementary product pricing by reducing the need to balance heterogeneous consumers. Because consumers of different types are sorted to different periods (i.e., avid readers buy earlier), firms can set different complementary product prices in different periods that are less distorted by consumer heterogeneity. I refer to this mechanism as *"sorting"*. Second, complementary product pricing helps IPD by limiting consumers' forward-looking behavior. Given two pricing instruments, firms can use opposite price trajectories to provide conflicting incentives for consumers to buy earlier and delay purchase. Consumers' ability to intertemporally arbitrage is limited. I refer to this mechanism as *"incentive-mixing"*.

7.3 Counterfactual Analysis

To illustrate the above advantages of jointly conducting IPD on complementary products, I consider three scenarios. Consumers are forward-looking in all scenarios, while the firm solves different pricing problems. In Scenario 1, the firm solves a "static" pricing problem for complementary products period by period without considering the impact of current pricing on future states of the market (i.e., the dynamic nature is missing). It represents the traditional "razor-and-blade" setting. Given the state of the market every period, the firm maximizes the Kindle profits in a period and the sequence of book profits (i.e., the durable product nature is captured here):

$$\begin{aligned}
EW_t\left(\Delta_t\right) &= \max_{P_t, p_t^E} \pi_t\left(P_t, p_t^E, \Delta_t\right) \\
\pi_t\left(P_t, p_t^E, \Delta_t\right) &= \phi_t \cdot \Delta_t\left[P_t - c_t\right] + \left[I - \phi_t\right] \cdot \Delta_t \cdot \frac{R_t^0}{(1-\delta)} + \left(\Delta_0 - \Delta_t + \phi_t \cdot \Delta_t\right) \cdot \frac{R_t^1\left(p_t^E\right)}{(1-\delta)}
\end{aligned}$$

The F.O.C.'s can be derived accordingly. In Scenario 2, the firm dynamically prices Kindles and statically coordinates e-book pricing. By adding dynamics to Kindle pricing, the firm can benefit from the "sorting" mechanism. The maximization problem and the F.O.C. for Kindle pricing are the same as in Equations 9 and 10, while the F.O.C. for e-book pricing becomes

$$\Delta_t \cdot \frac{\partial \phi_t}{\partial p_t^E} \cdot \left(R_t^1 - R_t^0\right) + \left[\Delta_0 - \Delta_t + \phi_t \cdot \Delta_t\right] \cdot \frac{\partial R_t^1}{\partial p_t^E} + \Delta_t \cdot \frac{\partial \phi_t}{\partial p_t^E} \left[P_t - c_t\right] = 0$$

In Scenario 3, the firm solves a dynamic pricing problem of both Kindles and e-books, which is the full firm problem in the model setup section. The firm dynamically coordinates the two pricing instruments and can further benefit from the "incentive-mixing" mechanism.

I simulate the price trajectories and market outcomes under each scenario and present the results in Table 6. In the static case (Scenario 1), given the relative demand elasticities of the two products, the firm should set high Kindle prices and low e-book prices for avid readers and low Kindle prices and high e-book prices for general readers. Given that general readers represent a much larger size of the population, the firm's strategy leans heavily towards general readers and sets much lower Kindle prices than in other scenarios, as shown in the first column of Table 6.

When adding Kindle IPD to the static case (Scenario 2), the firm realizes that consumers are sorted into different periods (i.e., the "sorting" mechanism). As avid readers buy earlier, the fraction of avid readers in the remaining market is the highest initially and declines over time. The firm can set Kindle prices that are higher today than tomorrow to skim avid readers. As shown in the second column of Table 6, the firm charges much higher Kindle prices than in the static case and earns higher profits both on Kindles and overall. Note that Kindle and e-book prices are coordinated only in a static sense.

When conducting IPD on both Kindles and e-books (Scenario 3), the firm fully dynamically coordinates the pricing of the two products by skimming for Kindles and investing in e-books, as shown in the third column of Table 6. Compared with IPD only on Kindles (Scenario 2), the firm sets higher Kindle prices and lower e-book prices. Both Kindle and book profits increase because consumers have two conflicting incentives: they should delay purchase, given the declining Kindle prices, but they should adopt earlier, given the increasing e-book prices. Their ability to delay purchase and damage profitability is thus limited (i.e., the "incentive-mixing" mechanism).

Figure 4: Kindle Sales Over Time

Notes: To focus on the sales distribution over time, I normalize the total sales of each consumer type to 1 under each scenario.

I provide evidence for the "incentive-mixing" and "sorting" mechanisms by examining the results on the Kindle sales over time and the percentage of avid readers. First, Figure 4 plots the Kindle sales by consumer type over time in the single IPD case and the joint IPD case. The total Kindle sales are normalized to 1 for each type in each case to focus on the consumer arbitrage behavior. I find that avid readers adopt Kindles earlier when firms conduct IPD on both products, even though the Kindle price is higher and drops more quickly, which illustrates the "incentive-mixing" mechanism. General readers are much less responsive. Second, as shown in Figure 5, IPD on both products induces the highest fraction of avid readers, followed by IPD on Kindles only, and static pricing induces the lowest fraction of avid readers. This suggests that IPD on both products offers the firm a better screening device and induces a higher fraction of avid readers, who are more profitable, to adopt Kindles. To intuitively illustrate how the screening device works, consider two scenarios. With only one product, raising the price will discourage both avid readers and general readers from buying; they respond in the same direction. With two products, raising the Kindle (e-book) price and reducing the e-book (Kindle) price properly will attract avid (general) readers and discourage general (avid) readers; they respond in the opposite directions. The firm can use different price path combinations to induce different consumer types to purchase.

Mechanism Decomposition. To further illustrate the mechanism through which the joint IPD strategy benefits the firm, I compare the micro-level consumer purchase behaviors under the joint IPD strategy and the single IPD strategy. The firm's profits have two components: Kindle profits and book profits. Kindle profits are results of Kindle sales and prices over time (plotted

Table 6: Counterfactual Market Outcomes

	(i) Static	(ii) IPD on Kindles	(iii) IPD on both
Kindle price path ($, markups in parentheses)			
2008	83.9 (-152.1)	215.4 (-20.6)	314.1 (78.1)
2009	72.0 (-113.0)	164.3 (-20.7)	241.2 (56.2)
2010	61.9 (-82.1)	125.9 (-18.1)	176.6 (32.6)
2011	56.7 (-56.3)	97.5(-15.5)	130.1 (17.1)
2012	10.0 (-79.0)	75.4 (-13.9)	98.3 (9.3)
E-book price path ($)			
2008	10.86	11.83	10.30
2009	10.97	12.22	11.23
2010	10.93	12.44	11.72
2011	10.91	12.62	12.12
2012	11.11	12.77	12.44
Penetration rate by 2012			
avid readers	71.7%	48.6%	56.5%
general readers	16.7%	0.37%	0.30%
Discounted profits (in sample 2008-2012, $)			
total	2.36e5	2.40e5	2.64e5
from Kindles	-4.29e4	-4.94e3	1.51e4
from books	2.79e5	2.45e5	2.49e5

in the left graph of Figure 6) while book profits are results of the cumulative number of Kindle owners and e-book prices over time (plotted in the right graph of Figure 6). In both plots, the bars represent Kindle sales and the number of owners (the left y-axis) and the lines represent prices (the right y-axis). Darker bars and lines represent the joint IPD strategy and lighter ones represent the single IPD strategy. As shown in the graphs, the joint IPD strategy has higher Kindle prices and lower e-book prices. It induces consumers to adopt Kindles earlier at higher Kindle prices due to lower e-book prices. It also induces more consumers to adopt Kindles overall and a larger cumulative number of Kindle owners in all periods.

These pricing differences and consumer behavior changes contribute to higher firm profits. The joint IPD strategy raises Kindle profits in two ways: as lower e-book prices make Kindles more attractive, 1) the firm can charge higher Kindle prices ("Kindle price effect"), and 2) there are more consumers who adopt Kindles; these additional sales come earlier at higher prices ("Kindle quantity effect"). The joint IPD strategy affects book profits in three ways: 3) there are more Kindle owners due to earlier and increased Kindle adoption; they purchase e-books and generate more book profits

Figure 5: Fractions of Avid Readers in Kindle Sales

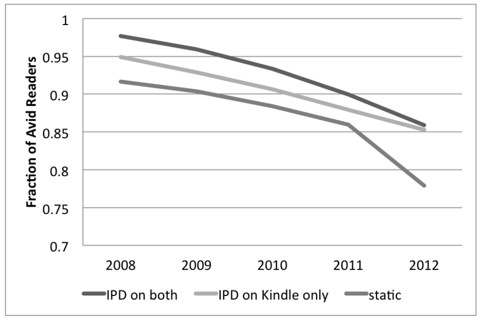
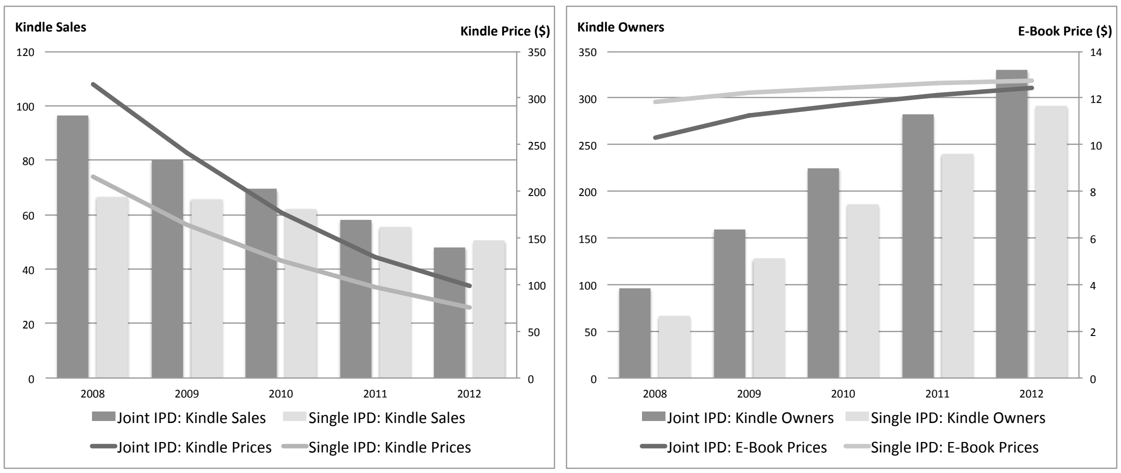


Figure 6: Consumer Purchases



than nonowners do ("owner quantity effect"); 4) the firm charges lower e-book prices and induces more e-book consumption per owner (not plotted); the quantity effect outweighs the pricing effect so that each owner generates higher book profits ("book price effect"); and 5) Kindle owners also prefer buying paperbacks from Amazon (not plotted), making Amazon a stronger paperback retailer. The probability of choosing Amazon as the paperback retailer is 39% for non-owners and 53% for owners, which also raises Amazon's paperback profits ("retailer competition effect").

I quantify the five effects and evaluate their contribution to the profit gains delivered by the joint IPD strategy. As shown in Figure 7, Effect 1 ("Kindle price effect", 58%) and Effect 3 ("owner quantity effect", 22%) contribute the most to the total profit increase, followed by Effect 2 ("Kindle quantity effect", 12%), Effect 4 ("book price effect", 5%), and Effect 5 ("retailer competition effect", 3%).

The results suggest that the firm benefits more from the Kindle side than from the book side.

Figure 7: Profit Effects Decomposition

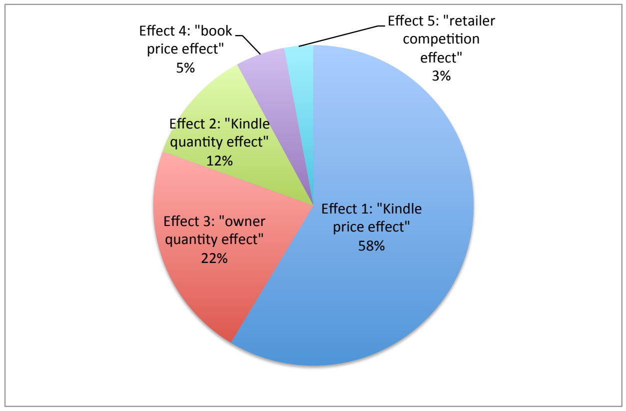
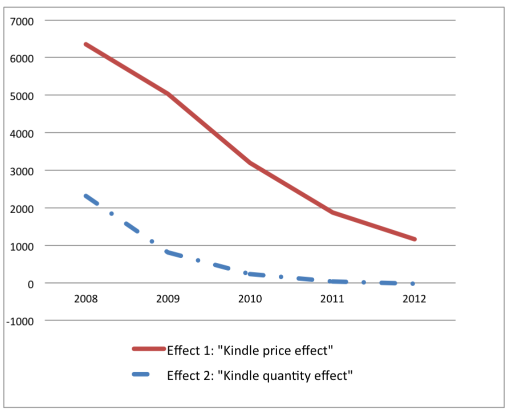
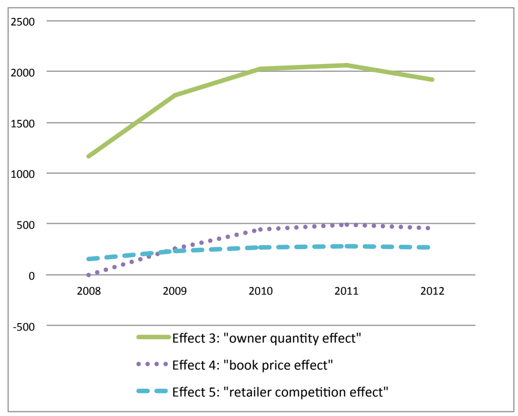


Figure 8: Profit Effects Over Time

(a) Kindle-side Effects



(b) Book-side Effects



However, the book-side benefits can predominate in the long run: I find that the Kindle-side effects decline over time while the book-side effects increase over time as shown in Figure 8. This is because the Kindle-side benefits of the joint IPD strategy mainly come from skimming consumers in the early periods, while the book-side benefits mainly come from there being more Kindle owners in later periods. Overall, the joint IPD strategy can be more beneficial as one considers a longer time horizon.

The five effects can be mapped to the "sorting" and "incentive-mixing" mechanisms. Effects 1 ("Kindle price effect") and 4 ("book price effect") correspond to the "sorting" mechanism in the static sense. As heterogeneous consumers are sorted into different periods due to higher Kindle prices and lower e-book prices, the firm's complementary product pricing is less distorted by conflicting consumer elasticities. Effects 2 ("Kindle quantity effect") and 3 ("owner quantity effect") correspond to the “incentive-mixing” mechanism in the dynamic sense. Given conflicting incentives for consumers to buy earlier and delay purchase, more consumers adopt Kindles earlier and start buying e-books earlier. Overall, the effects related to the “sorting” mechanism account for 63% of the total profit increase and the effects related to the “incentive-mixing” mechanism account for 34% of the total profit increase. Effect 5 (“retailer competition effect”) is an additional way in which, beyond the “sorting” and “incentive-mixing” mechanisms, Amazon benefits from promoting e-reading; the joint IPD strategy offers Amazon a competitive edge on paperback sales among Kindle owners.

It is important to note that the objective of this paper is not to explain or fit Amazon’s observed strategy. I abstract from product innovation, competition, and the e-book contract switch after 2010 in the supply-side simulation. I also abstract from paperback pricing responses. Incorporating all these factors would require R&D and competitor data. It is also computationally prohibitive and beyond the scope of this paper. I take a normative view and focus on a monopolist’s dynamic pricing problem in the novel setting of IPD with complementary products, which is itself both important and interesting to address. It also serves as a foundation to address more complex problems.

8 Conclusion

Software usage intensity drives hardware adoption in many industries, such as consoles and video games, Apple TV and digital content on iTunes, razors and blades, printers and cartridges, and K-cups and espresso machines. This paper empirically examines IPD in this novel complementary product setting. The demand-side estimation reveals a new dimension of consumer heterogeneity (i.e., relative demand elasticities between hardware and software) that firms can exploit. The supply-side simulation proposes a new joint IPD strategy, which attenuates the limitations of single-product IPD and static complementary product pricing. The results illustrate how skimming and penetration incentives for hardware interact with harvesting and investing incentives for software. The results can be generalized to other firms and industries where firms jointly sell a combination of multiple products and the usage of one product drives the purchase of another. The key mechanism is that the difference in usage intensity across consumer types drives the difference in their relative demand elasticities between hardware and software, which, in turn, drives the different combinations of hardware and software pricing strategies. The monopoly case studied in this paper demonstrates the fundamental trade-offs of complementary product pricing within a single firm. The mechanism would apply to each single firm in a competitive case as well, except that firms need to further consider competition-specific trade-offs.

The demand-side analysis develops a framework to understand the dynamic demand for durable products with continuous usage choices. It allows forward-looking consumers to self-select into buying hardware based on their heterogeneous tastes for software. Software usage intensity is explicitly modeled as an endogenous and continuous choice, given prices and consumer tastes.

The supply-side analysis helps illustrate retailers' trade-offs behind the policy debate of e-book pricing. The e-book pricing contract between publishers and retailers switched from a wholesale contract to an agency contract in 2010, which drew widespread public attention and close scrutiny from the Department of Justice.[47] The central difference between the two contract schemes is whether retailers have the right to price e-books. Addressing which contract is better is beyond the scope of this paper. However, the results for the joint IPD policy demonstrate how a monopolist retailer can benefit from having e-books as an additional pricing instrument.

A possible avenue for future research would be to model innovation and quality choices in addition to pricing. These factors may become more important as the e-reader market matures and more sales come from upgrading. Another avenue would be to study competition among retail platforms. In practice, Amazon and Barnes & Noble sell their own e-readers and compete for device and book buyers. It is challenging to solve for a dynamic competition model with both e-readers and e-books, but the multi-product setting can lead to potentially interesting competition patterns. The monopoly model in this paper illustrates the fundamental trade-offs and can serve as a starting point for even more complex dynamic strategy analysis. Finally, I do not address a number of issues that might be important in diffusion contexts, such as network effects, uncertainty, or switching costs, as one would need a long panel of individual-level data to address these issues. These factors can affect the magnitudes of the price increase or decrease in the optimal pricing strategy.[48] However, the key driving force of the joint IPD strategy (i.e., the heterogeneous relative demand elasticities between

[47]See http://www.nysd.uscourts.gov/cases/show.php?db=special&id=306

[48]If there were network effects or social influence, Amazon would have stronger incentives to "penetrate" the market and "invest" in new Kindle adopters to induce the diffusion of e-reading by setting lower Kindle prices and e-book prices in early periods. Similarly, if there was uncertainty about Kindle or e-book format quality, Amazon would have stronger incentives to "invest" by setting lower prices. Finally, if there were switching costs, Amazon would have stronger incentives to "penetrate" the market in early periods using lower Kindle prices and "harvest" existing consumers in later periods (as they are partially "locked-in") using higher e-book prices.

Kindles and e-books across heterogeneous consumers) is unlikely to be qualitatively affected.

References

[1] Besanko, D., Winston, W. L. 1990. "Optimal Price Skimming by a Monopolist Facing Rational Consumers," *Management Science*, 36(5), 555-567.

[2] Chandukala, S. R., Kim, J., Otter, T., Rossi, P. E., Allenby, G. M. 2008. "Choice Models in Marketing: Economic Assumptions, Challenges and Trends," *Foundations and Trends (R) in Marketing*, 2(2), 97-184.

[3] Chintagunta, P.K., Dubé, J.P., 2005. "Estimating a Stockkeeping-unit-level Brand Choice Model that Combines Household Panel Data and Store Data," *Journal of Marketing Research*, 42(3), 368-379.

[4] Clements, M. T., and Ohashi, H., 2005. "Indirect Network Effects and the Product Cycle: U.S. Video Games, 1994 - 2002," *Journal of Industrial Economics*, 53(4): 515-542.

[5] Dubé, J.P., Hitsch, G.J. and Rossi, P.E., 2009. "Do Switching Costs Make Markets Less Competitive?" *Journal of Marketing Research*, 46(4), pp.435-445.

[6] Dubé, J.P., Hitsch, G.J. and Chintagunta, P. 2010. "Tipping and Concentration in Markets with Indirect Network Effects," *Marketing Science*, 29(2): 216- 249.

[7] Economides, Seim, Viard. 2008. "Quantifying the Benefits of Entry into Local Phone Service," *RAND Journal of Economics*, 39(3), 699-730.

[8] Gandal, N., Kende, M., Rob, R. 2000. "The Dynamics of Technological Adoption in Hardware/Software Systems: the Case of Compact Disc Players." *RAND Journal of Economics*, 31: 43- 61.

[9] Gentzkow, M. 2007. "Valuing New Goods in a Model with Complementarity: Online Newspapers," *American Economic Review*, 97(3), 713 - 44.

[10] Gil, R., Hartmann, W. R. 2009. "Empirical Analysis of Metering Price Discrimination: Evidence from Concession Sales at Movie Theaters," *Marketing Science*, 28(6), 1046-1062.

[11] Gilbert, R. 2015. "E-books: A Tale of Digital Disruption," *Journal of Economic Perspectives*, 29(3), 165-184.

[12] Goettler, R., Gordon, B. 2011. "Does AMD Spur Intel to Innovate More?" *Journal of Political Economy*, 119(6), 1141-1200.

[13] Gordon, B. 2009. "A Dynamic Model of Consumer Replacement Cycles in the PC Processor Industry," *Marketing Science*, 28(5), 846-867.

[14] Gowrisankaran, G., Rysman, M. 2012. "Dynamics of Consumer Demand for New Durable Goods," *Journal of Political Economy*, 120(6), 1173-1219.

[15] Hartmann, W. R., Nair, H. S. 2010. "Retail Competition and the Dynamics of Demand for Tied Goods," *Marketing Science*, 29(2), 366-386.

[16] Hendel, I., Nevo, A. 2013. "Intertemporal Price Discrimination in Storable Goods Markets," *American Economic Review,* 103(7), 2722-2751.

[17] Hu, Y. J., Smith, M. D. 2013. "The Impact of Ebook Distribution on Print Sales: Analysis of a Natural Experiment." Available at SSRN: http://ssrn.com/abstract=1966115.

[18] Klemperer, P., 1995. "Competition When Consumers Have Switching Costs: An Overview with Applications to Industrial Organization, Macroeconomics, and International Trade," *The Review of Economic Studies*, 62(4), pp.515-539.

[19] Lee, R. 2013. "Vertical Integration and Exclusivity in Platform and Two-Sided Markets," *American Economic Review*, 103(7), 2960-3000.

[20] Liu, H. 2010. "Dynamics of Pricing in the Video Game Console Market: Skimming or Penetration?" *Journal of Marketing Research*, 47(3), 428-443.

[21] Melnikov. 2013. "Demand for Differentiated Durable Products: The Case of the US Computer Printer Market," *Economic Inquiry*, 51(2), 1277-1298.

[22] Nair, H., Chintagunta, P. and Dubé, J.P. 2004. "Empirical Analysis of Indirect Network Effects in the Market for Personal Digital Assistants," *Quantitative Marketing and Economics*, 2(1): 23-58.

[23] Nair, H. 2007. "Intertemporal Price Discrimination with Forward-looking Consumers: Application to the US Market for Console Video-games," *Quantitative Marketing and Economics*, 5(3), 239-292.

[24] Oi, W. 1971. "A Disneyland Dilemma: Two-part Tariffs for a Mickey Mouse Monopoly," *Quarterly Journal of Economics*, Vol. 85, No. 1, pp. 77-96.

[25] Reimers, I. and Waldfogel, J. 2014. "Throwing the Books at Them: Amazon's Puzzling Long Run Pricing Strategy," Mimeo. Available at SSRN: http://ssrn.com/abstract=2442747.

[26] Rosen, S. and Rosenfield, A. 1997. "Ticket Pricing," *Journal of Law and Economics*, 40(2), 351-76.

[27] Rust, J. 1987. "Optimal Replacement of GMC Bus Engines: An Empirical Model of Harold Zurcher," *Econometrica*, 55(5), 999-1033.

[28] Stokey, N. L. 1979. "Intertemporal Price Discrimination," *The Quarterly Journal of Economics*, 93(3), 355-371.

[29] Su, C.L., Judd, K.L., 2012. "Constrained Optimization Approaches to Estimation of Structural Models," *Econometrica*, 80(5), 2213-2230.

Appendix:

8.1 Robustness Checks

8.1.1 Allowing Kindle Quality to Affect Book Utility

Kindle quality does not affect book utility in the consumer problem. I conduct robustness checks and estimate another model specification in which Kindle quality affects book consumption. In particular, I add Kindle quality dummies to the e-format taste in Equation 3. This allows the Kindle version to affect consumer book purchase and device upgrading decisions. The likelihood function derivation is also affected.

The taste parameter specification in Equation 3 becomes

$$
\begin{aligned}
a^P_{igt} &= \theta_{ig} + \beta_1 D^{age}_i + \xi_t + \eta^P_{igt} \\
a^E_{igt} &= \underbrace{\theta_{ig} + \beta_1 D^{age}_i + \xi_t}_{\text{Baseline taste}} + \underbrace{\left(\theta^E_g + \beta_2 D^{age}_i + \beta_3 \log\, n^E_t + \bar{Q}^{kindle}_{it}\right)}_{\text{E-format taste}} + \eta^E_{igt}
\end{aligned}
$$

where $\bar{Q}^{kindle}_{it}$ is a dummy for the Kindle version that consumer i holds at time t. It can take values $\left\{Q^{kindle}_t\right\}^{t=2012}_{t=2008}$. Intuitively, given a Kindle version with better quality, consumers would enjoy e-reading more and thus more likely to prefer the e-format (as $a^E_{igt} - a^P_{igt}$ affects the substitution between paperbacks and e-books) and buy more e-books (as a^E_{igt} affects the e-book consumption through $q^{E*}_{igt} = a^E_{igt} - b_i p^E_{gt}$).

The device flow utility becomes

$$
\begin{aligned}
\bar{u}^0_{it} &= \Gamma v^{book,0}_{it} + \bar{\varepsilon}^0_{it} \\
\bar{u}^1_{it} &= \Gamma v^{book,1}_{it}\left(p^E_t, \bar{Q}^{kindle}_{it}\right) + \bar{Q}_{it} + \bar{\varepsilon}^1_{it} \\
u_{it} &= \Gamma v^{book,1}_{it}\left(p^E_t, Q^{kindle}_t\right) + Q_{it} - \alpha_i P_t + \varepsilon_{it}
\end{aligned}
$$

where Q^{kindle}_t is a dummy for the new Kindle offered at time t. The Kindle owner's book flow utility $v^{book,1}_{it}$ now becomes a function of $\left\{\bar{Q}^{kindle}_{it}, Q^{kindle}_t\right\}$; a higher $\bar{Q}^{kindle}_{it}$ leads to a higher book flow utility $\bar{u}^1_{it}$. Intuitively, given a Kindle version with better quality, consumers are less likely to upgrade their Kindles.

The likelihood function for the individuals in the book purchase data in Equation 13 has also changed. For Kindle owners, their format-quantity probability now further depends on the Kindle version $\bar{Q}^{kindle}_{it}$. Their device-side probability becomes the probability of having a particular Kindle version $\Pr\left(\bar{Q}_{it} = Q_{i\tau} \mid k\right)$:

$$
\begin{aligned}
\ell_i\left(\Theta \middle| \left\{q_{igt}^{P1}, q_{igt}^{E}\right\}, j, 1\right) &= \sum_k \lambda^k \cdot \left[\sum_{\tau=2008}^{t} \underbrace{\Pr\left(\left\{q_{igt}^{P1}, q_{igt}^{E}\right\} \mid \bar{Q}_{it}^{kindle} = Q_{\tau}^{kindle}, k\right) \cdot r_{ijgt}^{1}}_{book\,purchase} \cdot \underbrace{\Pr\left(\bar{Q}_{it} = Q_{i\tau} \mid k\right)}_{Kindle}\right] \\
\ell_i\left(\Theta \middle| \left\{q_{igt}^{P0}\right\}, j, 0\right) &= \sum_k \lambda^k \cdot \underbrace{\Pr\left(\left\{q_{igt}^{P0}\right\} \mid k\right) \cdot r_{ijgt}^{0}}_{book\,purchase} \cdot \underbrace{\Psi_{it}^{k,0}}_{Kindle}
\end{aligned}
$$

Similarly, the model-predicted aggregate offline paperback sales become

$$
\hat{H}_{gt}(\Theta) = \sum_i \sum_k \lambda^k \left\{\sum_{\tau=2008}^{t} \left[\hat{q}_{igt}^{P1,k}\left(\bar{Q}_{it}^{kindle} = Q_{\tau}^{kindle}\right) \cdot r_{iOgt}^{1} \cdot \Pr\left(\bar{Q}_{it} = Q_{i\tau} \mid k\right)\right] + \hat{q}_{igt}^{P0,k} \cdot r_{iOgt}^{0} \cdot \Psi_{it}^{k,0}\right\}
$$

The identification of the Kindle dummies in the e-format taste $\left\{Q_t^{kindle}\right\}_{t=2008}^{t=2012}$ mainly comes from the difference in the book format-quantity choices across consumers holding different Kindle versions. Given Kindle ownership status, a model in which the Kindle version does not affect book consumption (i.e., the estimates of $\left\{Q_t^{kindle}\right\}$ are insignificant) would predict lower e-book sales compared to a model in which higher Kindle versions lead to more e-book consumption (i.e., the estimates of $\left\{Q_t^{kindle}\right\}$ are positive and significant). Although I do not observe the Kindle version that a particular consumer owns, I can recursively calculate the probability that consumer i holds a particular Kindle version from the model and use it to construct the likelihood. $\left\{Q_t^{kindle}\right\}$ are separately identified from the book-side time fixed effect ξ_t because ξ_t is the same across consumers while $\bar{Q}_{it}^{kindle}$ differs across consumers in the same period. $\left\{Q_t^{kindle}\right\}$ are separately identified from the device-side Kindle qualities $\{Q_{it}\}$ because $\{Q_{it}\}$ are identified from the device adoption choices and $\left\{Q_t^{kindle}\right\}$ are mainly identified from the book consumption choices.

The estimation results are presented in Table 7. Q_{2008}^{kindle} is normalized to zero as it cannot be separately identified from the fixed effects. Similar to the device-side Kindle quality specification $Q_t = Q_0 + Q_1 \log t$, I let $Q_t^{kindle} = Q^{kindle} \log t$, where $t = 1$ to 5 represent years 2008 to 2012, and estimate Q^{kindle} to avoid overfitting problem. The estimates of other parameters are robust, and the estimate of Q^{kindle} is insignificant, which suggests that the data do not identify an interaction between Kindle quality and book utility. Intuitively, a Kindle owner who had a new Kindle in 2012 did not buy significantly different numbers of e-books compared with a Kindle owner in 2008. I thus keep the main model specification.

Table 7: Robustness Check: Allow Kindle Quality to Affect Book Utility

Book	Lifestyle: $g=1$	Casual: $g=2$	Practical: $g=3$
		Baseline FE	
θ_{ig}: High	12.01*** (0.0116)	9.182*** (0.0162)	8.525*** (0.0002)
θ_{ig}: Low	0.5322*** (0.0005)	0.7005*** (0.0173)	2.564*** (0.0148)
		E-format FE	
θ_g^E	0.0391*** (0.015)	2.718*** (0.0141)	-3.359*** (0.0130)
		Retailer FE & time trends	
A_{0g}	-0.3583*** (0.0103)	-0.7682*** (0.0131)	-0.4505*** (0.0145)
A_{1g} (time)	0.1308*** (0.0122)	0.1322*** (0.0091)	0.2214*** (0.0101)
B_{0g}	-1.609*** (0.0091)	-0.9368*** (0.0212)	-1.753*** (0.0222)
B_{1g} (time)	0.0653*** (0.0012)	0.0013*** (3.418e-5)	0.0733*** (0.0025)

	Device		**Book**		
α_0	0.0079***(0.0002)	m_1	0.0206*** (0.0020)	β_1	0.1502***(0.0025)
α_1	-1.873e-4***(8.402e-6)	m_2	0.0004 (0.0012)	β_2	-3.799e-4***(4.733e-5)
Γ	11.19***(0.0110)	m_3	0.0131*** (0.0015)	β_3	3.176e-4***(0.0001)
Q_0^L	-1.010***(0.0052)	ξ_{2009}	0.0092 (0.0060)	b_0	0.1718***(0.0028)
Q_1^L	0.1489***(0.0011)	ξ_{2010}	0.0425 (0.0390)	b_1	-0.0077***(0.0029)
Q_0^H	-0.4572*** (0.0048)	ξ_{2011}	0.0143 (0.0169)	σ	2.336***(0.0112)
Q_1^H	0.0534*** (0.0123)	ξ_{2012}	0.0109 (0.0139)	A_2	0.8081***(0.0341)
Q^{kindle}	**0.0066 (0.0132)**				
	MLE Obj.: 197,113		# Obs: 89,382		

Notes: ***, **, and * represent significance at the 1, 5, and 10 percent levels, respectively. This model specification allows Kindle quality to affect book utility. Similar to the device-side Kindle quality specification $Q_t = Q_0 + Q_1 \log t$, I let $Q_t^{kindle} = Q^{kindle} \log t$ and estimate Q^{kindle} to avoid overfitting problem. Q_{2008}^{kindle} is normalized to zero, as it cannot be separately identified from the fixed effects. Similar to the main specification, price coefficients enter utility negatively and vary by income groups as $\alpha_i = \alpha_0 + \alpha_1 D_i^{income}$ and $b_i = b_0 + b_1 D_i^{income}$. ξ_{2008} is normalized to 0 for identification purposes. The four consumer segments $k = 1,2,3,4$ have population mass $\{m_1, m_2, m_3, 1-m_1-m_2-m_3\}$ and genre baseline FEs $\{\theta_g^k\}_{g=1,2,3} = \{\theta_1^L, \theta_2^H, \theta_3^L\}, \{\theta_1^H, \theta_2^L, \theta_3^H\}, \{\theta_1^H, \theta_2^H, \theta_3^H\}, \{\theta_1^L, \theta_2^L, \theta_3^L\}$, respectively. Device qualities $\{Q_{it}\}_{t=2008}^{t=2012}$ are captured by $Q_{it} = Q_0^k + Q_1^k \log t$, where $t = 1$ to 5 represent years 2008 to 2012 and $\{Q_0^k, Q_1^k\}$ take the values $\{Q_0^H, Q_1^H\}$ for $k = 1,2,3$ and take the values $\{Q_0^L, Q_1^L\}$ for $k = 4$.

8.1.2 Varying Kindle Cost Decline Rate

To ensure that the pricing strategies are due to strategic actions and not simply due to falling costs, I conduct robustness checks and simulate the price paths when the cost declines at different rates or remains constant at the average value over time. In Table 8, the default scenario is that presented in the main results of the paper. The second and the third columns represent the scenarios in which the cost declines faster (+20%) or slower (-20%) than the default case. The fourth column represents constant cost over time. The optimal price paths are comparable as the firm faces the same demand curve; the prices are relatively higher when the cost is higher. In all scenarios, the optimal joint IPD strategy (skimming consumers with Kindles and investing in e-books) remains

Table 8: Robustness Check: Varying Evolution of Kindle Costs

	(i) Default	(ii) Faster drop	(iii) Slower drop	(iv) Constant
Kindle costs				
2008	236	265	207	144
2009	185	221	177	144
2010	144	117	148	144
2011	113	133	118	144
2012	89	89	89	144
Kindle price path ($, markups in parentheses)				
2008	314.1 (78.1)	321.9 (56.9)	278.1 (71.1)	239.3 (95.3)
2009	241.2 (56.2)	272.1 (51.1)	229.4 (51.9)	173.5 (29.5)
2010	176.6 (32.6)	210.2 (33.2)	173.8 (25.8)	150.4 (6.4)
2011	130.1 (17.1)	154.9 (21.9)	131.7 (13.2)	146.4 (2.4)
2012	98.3 (9.3)	102.3 (13.3)	97.4 (8.4)	94.7 (-49.3)
E-book price path ($)				
2008	10.30	10.24	10.73	10.11
2009	11.23	11.27	11.36	11.45
2010	11.72	11.72	11.82	12.04
2011	12.12	12.09	12.19	12.40
2012	12.44	12.35	12.45	12.55
Penetration rate by 2012				
avid readers	56.5%	52.4%	57.8%	63.3%
general readers	0.30%	0.26%	0.31%	0.34%

qualitatively the same across all scenarios.

8.2 Format-Quantity Probability and Indirect Utility Calculation

Format-quantity probability calculation. This subsection derives the format-quantity choice probability for a Kindle owner in Equation 15

$$\begin{aligned}\Pr\left(\left\{q^{P1}=q^{P1*}>0,q^{E}=0\right\}\right) = f\ &\left(\eta^{P1}=\eta\left(q^{P1*}\right)\mid \eta^{P1}>\max\left\{\bar{\eta}^{P},\,\eta^{E}+\left(\bar{\eta}^{P}-\bar{\eta}^{E}\right)\right\}\right)\\ &\cdot\Pr\left(\eta^{P1}>\max\left\{\bar{\eta}^{P},\,\eta^{E}+\left(\bar{\eta}^{P}-\bar{\eta}^{E}\right)\right\}\right)\end{aligned}$$

where η^{P1} and η^{E} are i.i.d. normally distributed error terms with mean 0 and variance σ^2. $\bar{\eta}^{P}$ and $\bar{\eta}^{E}$ are the thresholds of worth buying, as defined in Section 5.1. Define $\Lambda \equiv \left(\bar{\eta}^{P}-\bar{\eta}^{E}\right)$ and drop the ownership superscript 1 to simplify the discussions below. It is easier to start with calculating the following CDF instead of the PDF:

$$\Pr\left(\eta^P \leqslant \eta\left(q^P\right) \mid \eta^P > \max\left\{\bar{\eta}^P,\, \eta^E + \Lambda\right\}\right) \cdot \Pr\left(\eta^P > \max\left\{\bar{\eta}^P,\, \eta^E + \Lambda\right\}\right)$$

Define $\eta^P \leqslant \eta\left(q^P\right)$ as event A, $\eta^P > \bar{\eta}^P$ as event B, $\eta^P > \eta^E + \Lambda$ as event C, and $\bar{\eta}^P > \eta^E + \Lambda$ as event D. The above CDF can be written as $\Pr\left(A \mid B \cap C\right)\Pr\left(B \cap C\right)$. Notice that event $B \cap D$ implies event C, and event $C \cap \neg D$ implies event B. Event B and event D are independent. For the first component,

$$\begin{aligned}
\Pr\left(A \mid B \cap C\right) &= \Pr\left(A \mid B \cap C \cap D\right)\Pr\left(D\right) + \Pr\left(A \mid B \cap C \cap \neg D\right)\Pr\left(\neg D\right) \\
&= \Pr\left(A \mid B \cap D\right)\Pr\left(D\right) + \Pr\left(A \mid C \cap \neg D\right)\Pr\left(\neg D\right) \\
&= \frac{\Pr\left(A \cap B \cap D\right)}{\Pr\left(B \cap D\right)}\Pr\left(D\right) + \frac{\Pr\left(A \cap C \cap \neg D\right)}{\Pr\left(C \cap \neg D\right)}\Pr\left(\neg D\right)
\end{aligned}$$

where $\Pr\left(B\right) = \left[1 - \Phi\left(\frac{\bar{\eta}^P}{\sigma}\right)\right]$, $\Pr\left(D\right) = \Phi\left(\frac{\bar{\eta}^P - \Lambda}{\sigma}\right)$, $\Pr\left(\neg D\right) = 1 - \Phi\left(\frac{\bar{\eta}^P - \Lambda}{\sigma}\right)$, $\Pr\left(B \cap D\right) = \Pr\left(B\right)\Pr\left(D\right)$, $\Pr\left(A \cap B \cap D\right) = \left[\Phi\left(\frac{\eta\left(q^P\right)}{\sigma}\right) - \Phi\left(\frac{\bar{\eta}^P}{\sigma}\right)\right]\Phi\left(\frac{\bar{\eta}^P - \Lambda}{\sigma}\right)$ and

$$\begin{aligned}
\Pr\left(A \cap C \cap \neg D\right) &= \int_{\bar{\eta}^P + 1}^{\eta\left(q^P\right)} \left[\Phi\left(\frac{x - \Lambda}{\sigma}\right) - \Phi\left(\frac{\bar{\eta}^P - \Lambda}{\sigma}\right)\right] dF_x \\
&= \int_{\bar{\eta}^P + 1}^{\eta\left(q^P\right)} \Phi\left(\frac{x - \Lambda}{\sigma}\right) dF_x - \Phi\left(\frac{\bar{\eta}^P - \Lambda}{\sigma}\right)\Phi\left(\frac{\eta\left(q^P\right)}{\sigma}\right) \\
\Pr\left(C \cap \neg D\right) &= \int_{\bar{\eta}^P}^{+\infty} \left[\Phi\left(\frac{x - \Lambda}{\sigma}\right) - \Phi\left(\frac{\bar{\eta}^P - \Lambda}{\sigma}\right)\right] dF_x \\
&= \int_{\bar{\eta}^P}^{+\infty} \Phi\left(\frac{x - \Lambda}{\sigma}\right) dF_x - \Phi\left(\frac{\bar{\eta}^P - \Lambda}{\sigma}\right)\left[1 - \Phi\left(\frac{\bar{\eta}^P}{\sigma}\right)\right]
\end{aligned}$$

For the second component

$$\begin{aligned}
\Pr\left(B \cap C\right) &= \Pr\left(B \cap C \mid D\right)\Pr\left(D\right) + \Pr\left(B \cap C \mid \neg D\right)\Pr\left(\neg D\right) \\
&= \Pr\left(B \cap C \cap D\right) + \Pr\left(B \cap C \cap \neg D\right) \\
&= \Pr\left(B \cap D\right) + \Pr\left(C \cap \neg D\right) \\
&= \Pr\left(B\right)\Pr\left(D\right) + \Pr\left(C \cap \neg D\right) \\
&= \int_{\bar{\eta}^P}^{+\infty} \Phi\left(\frac{x - \Lambda}{\sigma}\right) dF_x
\end{aligned}$$

In all,

$$CDF\left(\eta\left(q^P\right)\right) = \Pr\left(A \mid B \cap C\right)\Pr\left(B \cap C\right) = \left[\frac{(a-b)\,c}{1-b} + \frac{I_1 - ac}{I_2 - c(1-b)}\,(1-c)\right] I_2$$

where $a = \Phi\left(\frac{\eta\left(q^P\right)}{\sigma}\right)$, $b = \Phi\left(\frac{\bar{\eta}^P}{\sigma}\right)$, $c = \Phi\left(\frac{\bar{\eta}^P - \Lambda}{\sigma}\right)$, $I_1 = \int_{\bar{\eta}^P + 1}^{\eta\left(q^P\right)} \Phi\left(\frac{x - \Lambda}{\sigma}\right) dF_x$ and $I_2 = \int_{\bar{\eta}^P}^{+\infty} \Phi\left(\frac{x - \Lambda}{\sigma}\right) dF_x$.

Now we are ready to calculate the PDF by taking a derivative of the CDF:

$$f\left(\eta^T = x \mid \eta^T > \max\left\{\bar{\eta}^T, \, \eta^{-T} + \left(\bar{\eta}^T - \bar{\eta}^{-T}\right)\right\}\right)$$
$$= \begin{cases} \frac{a'c}{1-b} + \frac{I_1' - a'c}{I_2 - c(1-b)}(1-c) & \text{if } x > \bar{\eta}^T \\ 0 & \text{otherwise} \end{cases}$$

where $a' = \frac{1}{\sigma}\phi\left(\frac{x}{\sigma}\right)$, $b = \Phi\left(\frac{\bar{\eta}^T}{\sigma}\right)$, $c = \Phi\left(\frac{\bar{\eta}^T - \Lambda}{\sigma}\right)$, $I_1' = \Phi\left(\frac{x-\Lambda}{\sigma}\right) f_x(x)$ and $I_2 = \int_{\bar{\eta}^T}^{+\infty} \Phi\left(\frac{x-\Lambda}{\sigma}\right) dF_x$. f_x and F_x are the PDF and CDF of $N\left(0, \sigma^2\right)$.

Given that book quantities are integers, the ultimate probability to calculate is

$$\begin{aligned} & \Pr\left(\eta\left(q^P\right) \leqslant \eta^P < \eta\left(q^P+1\right) \mid \eta^P > \max\left\{\bar{\eta}^P + 1, \, \eta^E + \Lambda\right\}\right) \\ & \cdot \Pr\left(\eta^P > \max\left\{\bar{\eta}^P + 1, \, \eta^E + \Lambda\right\}\right) \\ = \; & CDF\left(\eta\left(q^P+1\right)\right) - CDF\left(\eta\left(q^P\right)\right) \\ = \; & \left[\frac{\tilde{a}c}{1-b} + \frac{\tilde{I}_1 - \tilde{a}c}{I_2 - c(1-b)}(1-c)\right] I_2 \end{aligned}$$

where $\tilde{a} = \Phi\left(\frac{\eta\left(q^P+1\right)}{\sigma}\right) - \Phi\left(\frac{\eta\left(q^P\right)}{\sigma}\right)$ and $\tilde{I}_1 = \int_{\eta\left(q^P\right)}^{\eta\left(q^P+1\right)} \Phi\left(\frac{x-\Lambda}{\sigma}\right) dF_x$. There are two integrals to calculate: $\tilde{I}_1 = \int_{\eta\left(q^P\right)}^{\eta\left(q^P+1\right)} \Phi\left(\frac{x-\Lambda}{\sigma}\right) dF_x$ and $I_2 = \int_{\bar{\eta}^P}^{+\infty} \Phi\left(\frac{x-\Lambda}{\sigma}\right) dF_x$. I use Gauss-Chebychev quadrature with 10 nodes to calculate the first one and Gauss-Laguerre quadrature with 10 nodes to calculate the second one.

Indirect Flow Utility. This subsection derives the ex-ante indirect flow utilities from books in Equation 4. To simplify the notations, I drop i, t subscripts for now and define $\Lambda_g \equiv \left(\bar{\eta}_g^P - \bar{\eta}_g^E\right)$. The indirect flow utility for a Kindle nonowner is

$$\begin{aligned} v^{book,0} &= y + \sum_g E\left(\frac{\left(a_g^P - bp_g^P\right)^2}{2b} \mid q_g^P > 0\right) \Pr\left(q_g^P > 0\right) \\ &= y + \frac{1}{2b}\sum_g E\left(\left(\eta_g^P - \bar{\eta}_g^P\right)^2 \mid \eta_g^P - \bar{\eta}_g^P > 0\right) \Pr\left(\eta_g^P - \bar{\eta}_g^P > 0\right) \end{aligned}$$

where $X \equiv \eta_g^P - \bar{\eta}_g^P \sim N\left(-\bar{\eta}_g^P, \sigma^2\right)$ and $\Pr\left(\eta_g^P - \bar{\eta}_g^P > 0\right) = 1 - \Phi\left(\frac{\bar{\eta}_g^P}{\sigma}\right)$. From the truncated normal distribution properties, we know that

$$\begin{aligned} E\left(X^2 \mid X > 0\right) &= Var\left(X \mid X > 0\right) + \left[E\left(X \mid X > 0\right)\right]^2 \\ &= \sigma^2\left[1 - \lambda(\alpha)\left(\lambda(\alpha) - \alpha\right)\right] + \left[-\bar{\eta}_g^P + \sigma\lambda(\alpha)\right]^2 \end{aligned}$$

where $\alpha = \frac{\bar{\eta}_g^P}{\sigma}$ and $\lambda(\alpha) = \frac{\phi(\alpha)}{1-\Phi(\alpha)}$. This is a closed-form solution. The indirect flow utility for a Kindle owner is

$$
\begin{aligned}
v^{book,1} &= y + \sum_g \Big\{ E[\frac{\left(a_g^P - bp_g^P\right)^2}{2b} \mid q_g^{P1*} > 0,\, q_g^{E*} = 0] \cdot \Pr\left(q_g^{P1*} > 0,\, q_g^{E*} = 0\right) \\
&\quad + E[\frac{\left(a_g^E - bp_g^E\right)^2}{2b} \mid q_g^{E*} > 0,\, q_g^{P1*} = 0] \cdot \Pr\left(q_g^{E*} > 0,\, q_g^{P1*} = 0\right) \Big\} \\
&= y + \frac{1}{2b} \cdot \\
&\quad \sum_g E\left(\left(\eta_g^P - \bar{\eta}_g^P\right)^2 \mid \eta_g^P > \max\left\{\bar{\eta}_g^P, \eta_g^E + \Lambda_g\right\}\right) \Pr\left(\eta_g^P > \max\left\{\bar{\eta}_g^P, \eta_g^E + \Lambda_g\right\}\right) \\
&\quad + E\left(\left(\eta_g^E - \bar{\eta}_g^E\right)^2 \mid \eta_{ig}^E > \max\left\{\bar{\eta}_g^E, \eta_g^P - \Lambda_g\right\}\right) \Pr\left(\eta_g^E > \max\left\{\bar{\eta}_g^E, \eta_g^P - \Lambda_g\right\}\right)
\end{aligned}
$$

where the probability $\Pr\left(\eta_g^P > \max\left\{\bar{\eta}_g^P, \eta_g^E + \Lambda_g\right\}\right)$ was already calculated in the last subsection. To calculate the two conditional expectations, I use the conditional expectation definitions $E[X \mid \mathcal{H}] = \int_{-\infty}^{+\infty} x \cdot f(x \mid \mathcal{H})\, dx$ and $E\left[X^2 \mid \mathcal{H}\right] = \int_{-\infty}^{+\infty} x^2 \cdot f(x \mid \mathcal{H})\, dx$. The conditional density $f(x \mid \mathcal{H})$ was calculated in the last section. Given the conditional density, I calculate the conditional expectations using Gauss-Hermite quadrature with 10 nodes. Again, I account for the fact that book quantities are integers when deriving the equations in the final calculation.

8.3 A Simple Two-Period Model

To illustrate the pricing incentives, I present a two-period model in which a firm sells durable primary hardware at price P and complementary software at price p^E to a unit mass of consumers. The hardware serves as a gateway product to the software and does not bear any stand-alone value.[49] Consumers are heterogeneous in their tastes for the software. The value of a unit of the software v is uniformly distributed on $[0, 1]$. The utility of the hardware comes from the utility generated by subsequent software consumption $u = \lambda\left(v - p^E\right) - P$. The coefficient λ is the quantity of the software.[50] Consumers and the firm share the same discount factor δ and live for two periods. The marginal costs are assumed to be zero. The firm chooses software and hardware prices $\vec{p}_1 = \left\{p_1^E, P_1\right\}$

[49] I allow hardware to bear some qualities in the full empirical model.

[50] In the full empirical model, the usage intensity λ is endogenized to be a function of the taste parameter v and the book price p^E. I assume that it is a constant here to keep the analytical solution simple while illustrating the same qualitative results.

and $\vec{p}_2 = \left\{p_2^E, P_2\right\}$ in periods 1 and 2. Consumers have rational expectations about the firm's pricing policies; their beliefs about prices are consistent with the firm's strategy in equilibrium. The marginal consumer in period 1, v_1^*, is indifferent between buying and waiting:

$$\lambda \left[v_1^* - p_1^E + \delta \left(v_1^* - p_2^E\right)\right] - P_1 = \delta \left[\lambda \left(v_1^* - p_2^E\right) - P_2\right] \geqslant 0$$

where $v_1^* = \frac{P_1 - \delta P_2}{\lambda} + p_1^E$. Consumers who buy hardware and software in period 1 are in the range $[v_1^*, 1]$. Similarly, the marginal consumer in period 2, v_2^*, satisfies $\lambda \left(v_2^* - p_2^E\right) - P_2 = 0$, where $v_2^* = \frac{P_2}{\lambda} + p_2^E$. Consumers who buy hardware and consume software in period 2 are in the ranges $[v_2^*, v_1^*]$ and $[v_2^*, 1]$, respectively.

This simple setup captures the main features of the traditional single-product IPD and new features of IPD with complementary products. In particular, the firm's target is to first extract the most from high-valuation consumers through the hardware and then appeal to low-valuation consumers while earning the most from the software sales. As in the traditional IPD case, the firm faces a shrinking market and lower average willingness-to-pay for the product over time; both the market size and the consumer mix change. A decrease in P_1 reduces the hardware demand in period 2, changes the optimal P_2, and in turn changes consumer expectations of P_2 as consumers' beliefs are consistent with the optimal strategy. v_1^* summarizes the mass of consumers remaining in the market at the beginning of period 2 and is the relevant state variable for the pricing problem.

Three features are novel in the complementary product setup. First, consumers self-select into buying the hardware based on their heterogeneous tastes for the software. Second, the demands of the two products are interrelated. Consumers trade off between the utility from a current hardware purchase and the value of waiting, both of which further depend on the current and future software prices. Third, the firm needs to coordinate the pricing of the two products. p^E affects the profits from a hardware owner, while p^E and P jointly affect the number of hardware owners. The full model captures all the features of the simple model while allowing for richer heterogeneity and nonlinear demand elasticities.

Using backward induction to solve for period 2 and period 1 prices, I obtain

$$\begin{aligned}
\vec{p}_2\left(\vec{p}_1\right) &= \arg\max_{\vec{p}_2} \pi_2 = \left(1 - v_2^*\right)\lambda p_2^E + \left(v_1^* - v_2^*\right) P_2 \\
\vec{p}_1 &= \arg\max_{\vec{p}_1} \pi_1 + \delta\pi_2 = \left(1 - v_1^*\right)\left(\lambda p_1^E + P_1\right) + \frac{\lambda\delta}{4}\left(1 + \frac{1}{\delta}\left(p_1^E + \frac{P_1}{\lambda} - 1\right)^2\right)
\end{aligned}$$

The optimal prices in period 2 are $P_2^* = \frac{\lambda}{2}$ and $p_2^{E*} = 0$. The optimal prices in period 1 satisfy $p_1^{E*} + \frac{P_1^*}{\lambda} = 1 + \delta$. In particular, $P_1^* = \lambda\left(1 + \delta\right)$ and $p_1^{E*} = 0$ if $\lambda \geqslant 1$, and $P_1^* = 0$ and $p_1^{E*} = 1 + \delta$ if $\lambda < 1$.

The optimal strategies with complementary products differ from the harvesting strategy in the traditional single-product IPD case in two ways. First, both harvesting and investing can be optimal. If $\lambda \geqslant 1$, it is optimal to harvest on the hardware and invest in the software. The opposite is true for $\lambda < 1$. Similarly in the full model, I find that the firm should harvest on Kindles and invest in e-books for the avid readers with high λ and should do the opposite for the general readers with low λ. Second, the firm needs to coordinate p^E and P. The optimal pricing condition $p_1^{E*} + \frac{P_1^*}{\lambda} = 1 + \delta$ indicates that the optimal P increases as p^E decreases within the same period. The results from the full model echo the results from this simple model.

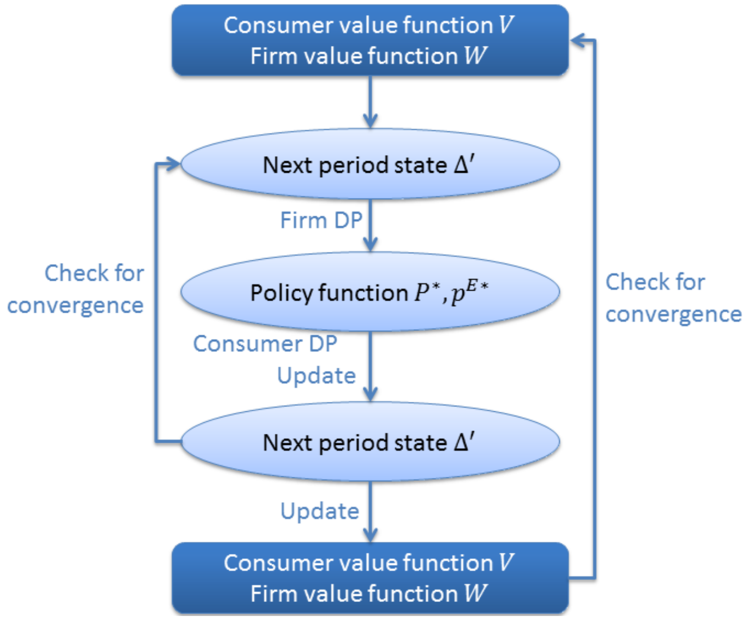
8.4 Computation Algorithm for the Dynamic Pricing Problem

The numerical algorithm is similar to that in Goettler and Gordon (2011). I summarize the algorithm in Figure 9. It contains an inner loop and an outer loop. The inner loop solves the firm's and the consumers' maximization problems along with the next-period state space given the value function guess. The outer loop updates the value function guesses and iterates until convergence. For each iteration $k = 1, 2, ...,$

1) Guess the value functions for the consumers and the firm $\left\{V^{k-1}, W^{k-1}\right\}$.

2) Given the value function guess, simultaneously solve the firm's first-order conditions at each state. Since the first-order conditions depend on consumers' current choices and next-period Δ', which in turn depend on consumers' rational expectations of Δ', I solve for a fixed point in Δ' such that consumers' expectations for Δ' are realized according to the state space evolution equation. In

Figure 9: Computation Algorithm



particular, to solve for the fixed point, I first guess the next-period state space Δ'^{m-1} and the firm's optimal pricing policy $\{P^{m-1}, p^{E,m-1}\}$, where m is the iteration number for the fixed point in the inner loop. Given the guess, I solve the consumers' device adoption problem to obtain the updated next-period state space Δ'^{m}. Given the updated Δ'^{m}, I solve the firm's first-order conditions at each state and obtain the updated optimal pricing policies $\{P^{m}, p^{E,m}\}$. Check convergence of $\mid \Delta'^{m} - \Delta'^{m-1} \mid$, $\mid P^{m} - P^{m-1} \mid$, and $\mid p^{E,m} - p^{E,m-1} \mid$. If converged, let Δ'^{k} and $\{P^{k}, p^{E,k}\}$ denote this fixed point. This is the solution to the inner loop given the value function guess $\{V^{k-1}, W^{k-1}\}$.

3) Update the value functions given the firm's policy and the next-period state space. Denote them $\{V^{k}, W^{k}\}$.

4) Check for convergence of the outer loop $\mid V^{k} - V^{k-1} \mid$ and $\mid W^{k} - W^{k-1} \mid$ at the state space grid points Δ. If convergence is not achieved, return to step 2).

Throughout the computation, I evenly discretize the state space into 20 grid points on both dimensions. The range of the state space is between 0 and the initial market size of each type. I use a cubic spline to interpolate between the grid points for the value functions and the policy functions. This is because solving the firm's first-order condition requires differentiable continuation values. The convergence is checked at the grid points.

8.5 Monte Carlo Study: Identification Using Yearly Re-sampled Data

An identification challenge is that the main data set is re-sampled yearly. Following Gordon (2009), I conduct a simulation analysis to illustrate 1) how the yearly re-sampled data identify consumer heterogeneity and new buying versus upgrading and 2) how the supplementary aggregate penetration data further help identification. I conduct the simulation for 4,500 consumers in two consumer segments. Given the observed Kindle and e-book prices, I simulate consumers' book and Kindle purchase decisions at the individual level. Based on it, I construct the supplementary aggregate penetration rate data in each period.

I estimate the model using the yearly re-sampled data as in the main model. Specifically, I integrate over the Kindle ownership statuses using the model-predicted ownership probabilities rather than using the actual Kindle ownership in the simulated data, as if I did not observe whether the sales are new purchases or upgrades. I also treat the device choices and the book choices as if they were from two samples and construct the likelihood similarly as in the main model. I conduct the estimation with and without the supplementary penetration rate data and compare the parameter estimates. I generate 50 sets of simulation data sets for the same set of true parameter values and obtain the estimates for each simulated data set to calculate the standard errors.

In Table 9, I compare the true parameter values (Column 1) with the estimates using the yearly re-sampled data, with (Column 2) and without (Column 3) the supplementary penetration rate data. I also compare the implied statistics on consumers; device adoption decisions such as cumulative Kindle sales, penetration rates, and the percentage of sales that comes from upgrading every period. The standard errors are based on the estimates from the 50 simulations.

The results in Columns 1 and 2 demonstrate that the model is able to recover consumer heterogeneity and the upgrading behavior from yearly re-sampled data. The true parameter values are within the 95% confidence intervals of the estimates. The predicted consumer device adoption statistics are also very close to the true values. The results in Columns 2 and 3 suggest that, although the yearly re-sampled data can produce estimates very close estimates to the true parameters, the supplementary penetration data further help identification. The models estimated with and without penetration data can both produce close estimates for the fraction of avid readers, while the model estimated without penetration data produces biased estimates of consumer preferences and price coefficients and generates biased results for the upgrading behavior. Specifically, the model estimated without the penetration data is not able to generate enough sales from upgrading; the percentage of sales from upgrading is substantially lower (32.7%) than that in the true model (45.2%) and that estimated with the supplementary data (44.4%). The reason is that it under-estimates the coefficients on Kindle quality improvement (Q_1^k, especially for avid readers because they are the major contributors to upgrading) and over-estimates the device-side price coefficient, both of which lead to weaker incentives to upgrade. These biases also potentially lead to lower incentives for new purchases and generate further biases on other coefficients. Because the model under-produces sales from upgrading, it over-produces sales from new purchases to rationalize the observed Kindles sales. To provide more incentives to new purchase, the model further under-estimates the book-side price coefficient and over-estimates the importance of the book utility in the Kindle utility (Γ); consumers are over-motivated by book-side incentives and under-motivated by Kindle-side incentives in their adoption decisions. Gowrisankaran and Rysman (2012) compare the estimates with and without penetration data and find similar biases. [51]

[51] Gowrisankaran and Rysman (2012) find that consumer price sensitivity is larger and quality matters less without penetration data. The baseline model predicts a very low upgrading rate because the coefficients on product characteristics are small. Adding penetration rate data corrects for the bias. To rationalize the percentage of upgrading in the penetration rate data, the full model produces a smaller price sensitivity and a larger coefficient on the quality that improves over time.

Table 9: Monte Carlo Simulations

Parameter Estimates		True	All data		No penetration data	
Book Side			Est.	Std err.	Est.	Std err.
	$\theta_{ig=1}$: High	10.00	9.81	(0.28)	9.26	(0.30)
	$\theta_{ig=2}$: High	10.00	9.90	(0.20)	9.39	(0.22)
Baseline FE	$\theta_{ig=3}$: High	10.00	9.76	(0.19)	8.76	(0.20)
	$\theta_{ig=1}$: Low	1.00	0.90	(0.09)	0.34	(0.10)
	$\theta_{ig=2}$: Low	1.00	0.92	(0.08)	0.40	(0.11)
	$\theta_{ig=3}$: Low	2.00	1.74	(0.15)	0.58	(0.18)
	$\theta^E_{g=1}$	0.050	0.049	(0.005)	0.033	(0.006)
E-format FE	$\theta^E_{g=2}$	2.75	2.74	(0.11)	2.96	(0.15)
	$\theta^E_{g=3}$	-3.50	-3.87	(0.32)	-3.87	(0.33)
	β_1	0.100	0.071	(0.015)	0.090	(0.019)
	β_2	-0.0003	-0.0006	(0.0002)	-0.0009	(0.0002)
	β_3	0.0005	0.0005	(0.0001)	0.0008	(0.0001)
Price	b_0: Baseline	0.1650	0.1529	(0.0124)	0.1202	(0.0130)
Coefficient	b_1: Income	-0.0100	-0.0091	(0.0012)	-0.0111	(0.0015)
	σ	2.350	2.361	(0.051)	2.411	(0.082)
Device Side						
Price	α_0: Baseline	0.0070	0.0074	(0.0003)	0.0106	(0.0005)
Coefficient	α_1: Income	-0.0002	-0.0002	(0.0001)	-0.0002	(0.0001)
	Γ	12.00	12.09	(0.12)	13.10	(0.26)
	Q^L_0: General	-1.000	-1.019	(0.015)	-0.863	(0.032)
Device	Q^L_1: General	0.400	0.431	(0.027)	0.406	(0.030)
Quality	Q^H_0: Avid	-0.600	-0.551	(0.049)	-0.504	(0.052)
	Q^H_1: Avid	0.050	0.052	(0.0017)	0.004	(0.0028)
Fraction of avid	m_1	0.100	0.088	(0.013)	0.092	(0.016)

Device adoption behavior	True	All data		No penetration data	
		Est.	Std err.	Est.	Std err.
Cumulative Sales / Market Size by 2012 (%)	11.94	12.43	(0.83)	13.95	(1.01)
Penetration Rate by 2012 (%)	8.61	9.04	(0.50)	11.48	(0.62)
% Sales from Upgrading					
2009	7.20	6.84	(1.32)	2.90	(1.92)
2010	18.5	17.9	(1.67)	10.0	(1.89)
2011	34.6	33.9	(2.32)	23.4	(2.57)
2012	45.2	44.4	(3.11)	32.7	(3.34)

Notes: "All data" refers to the yearly re-sampled data with penetration data. "No penetration data" refers to the yearly re-sampled data without penetration data. The percentage of sales from upgrading is 0 for 2008. The price coefficients enter utility negatively and vary by income groups as $\alpha_i = \alpha_0 + \alpha_1 D_i^{income}$ and $b_i = b_0 + b_1 D_i^{income}$. Device qualities $\{Q_t\}_{t=2008}^{t=2012}$ are captured by $Q_t = Q_0 + Q_1 \log t$. The two consumer segments $k = 1, 2$ have population mass $\{m_1, 1 - m_1\}$ and genre baseline FEs $\left\{\theta_g^k\right\}_{g=1,2,3} = \left\{\theta_1^H, \theta_2^H, \theta_3^H\right\}, \left\{\theta_1^L, \theta_2^L, \theta_3^L\right\}$, respectively.