Abstract

Firms jointly sell durable primary hardware and complementary software in many industries, but it remains unclear how they can coordinate the two products to conduct intertemporal price discrimination (IPD). In addition to the harvesting (i.e., price-cutting over time) incentive in the traditional single-product IPD literature, firms selling complementary products have another investing (i.e., price-raising over time) incentive to penetrate the market early and earn from subsequent complementary product sales. We study the optimal pricing strategy and profitability of joint IPD in the context of e-readers and e-books. A forward-looking monopolist firm makes dynamic pricing decisions for both e-readers and e-books, and forward-looking consumers make dynamic e-reader and book purchase decisions anticipating future price changes. We structurally estimate the demand system using individual-level transaction data of e-readers and books from 2008 to 2012. The estimation reveals two consumer types, “avid readers” and “general readers,” who self-select into buying e-readers based on their unobserved heterogeneous tastes for books. Given the demand estimates, we numerically solve for the optimal joint IPD policy and find that harvesting on e-readers and investing in e-books is optimal. In particular, the optimal pricing policies are functions of the composition of consumer types. The monopolist should harvest on e-readers and invest in e-books for avid readers, while it should invest in e-readers and harvest on e-books for general readers. We explain this difference by comparing the relative demand elasticities between e-readers and e-books for avid and general readers. The joint IPD policy provides a better screening device for the monopolist because different price path combinations can induce different consumer types to purchase. It also limits consumers’ ability to arbitrage on price changes over time. The profitability of the joint IPD policy depends on the composition of heterogeneous consumers in the initial market. We find that conducting IPD on both products may hurt firm profits if the fraction of avid readers is too low.
1 Introduction

In many industries, especially digital and online businesses, a single firm jointly sells durable primary hardware and complementary software. A typical example is e-readers and e-books. E-reading has been growing rapidly since 2007; by November 2012, 28 percent of all U.S. book purchases were in e-book format (Bowker Market Research, 2012). In this market, consumers first buy an e-reading device, e-reader, and then buy e-books from the same retailer.\(^1\) Other complementary product examples include consoles and video games, Apple TV and digital content in iTunes, razors and blades, printers and cartridges.

Given the dynamic nature of consumption, firms usually open with high prices for new products to exploit high-valuation consumers and later cut the prices to appeal to low-valuation consumers. They either conduct this intertemporal price discrimination (IPD for short) separately for hardware and software (e.g., consoles and video games) or only for hardware, keeping the software price stable (e.g., e-books). However, the two products are complementary: (1) consumers need to buy the primary product to consume the complementary product, and (2) the usage intensity of the complementary product drives adoption of the primary product. Is it possible to coordinate the two pricing instruments and generate higher profits?

To address this question, we develop a dynamic structural model of consumer demand and firm pricing to empirically investigate optimal joint IPD strategies in the e-reader and e-book market. There have been studies on dynamic pricing of a single product (e.g., Stokey 1979, 1981; Besanko and Winston 1990; Nair 2007) and static pricing of complementary products

\(^1\)There are three players (publishers, platform retailers, and consumers) and three products (print books, e-books, and e-readers) in the industry. Platform retailers launch their own e-readers and set e-reader prices. Two pricing contracts exist for books. Under the wholesale contract, retailers set book prices and pay wholesale prices to the publishers. Under the agency contract, publishers set book prices and share a fraction of book revenue with the retailers. E-book pricing followed the wholesale contract from 2007 to 2010 and the agency contract from 2010 to 2012. Print book pricing has always followed the traditional wholesale contract. This paper focuses on the pricing of e-readers and e-books under the wholesale contract where retailers set both prices. Discussion with industry practitioners suggests that print book launching and pricing are unaffected by e-book pricing. Thus, we take print book prices as exogenously given from the data.
(e.g., loss-leader strategy, or “metering” strategy, as in Gil and Hartmann 2009). Little is known, however, about dynamic pricing of complementary products. In a single-product setting, harvesting (price-cutting over time, skimming) is optimal. When accounting for the complementary product in the e-book and e-reader market, firms have another investing incentive (i.e., price-raising over time, penetration pricing). They can penetrate the market with low initial e-reader prices and earn from subsequent e-book sales. Both harvesting and investing incentives exist for both e-reader and e-book pricing, which leads to various possible price trajectory combinations. On the demand side, forward-looking consumers may anticipate future price changes and intertemporally arbitrage. Firms also need to account for this strategic behavior when setting prices.

We provide a framework to incorporate all these incentives. Using individual-level transaction data from 2008 to 2012, we first estimate a dynamic demand system of consumers buying and upgrading e-readers and buying books. Consumers maximize their direct utility from books by choosing book quantity, reading format (e-books or print books), and retail channel for print books (Amazon.com, other online websites, local bookstores) in a number of book genres. They self-select into buying e-readers based on their unobserved heterogeneous tastes for books. The demand system is estimated without imposing pricing optimality conditions. Using the demand estimates, we numerically solve for the optimal joint IPD strategies given cost information from industry reports. A monopolist retailer maximizes total profits from e-readers, print books, and e-books. Both consumers and the retailer are forward-looking in the pure-strategy Markov-perfect Nash equilibrium (MPNE).

We find that the optimal joint IPD strategy is to harvest on e-readers and invest in e-books. Conducting IPD on both products benefits the retailer in two ways. First, it offers the retailer a better screening device to induce a higher fraction of avid readers to buy. Second, it limits consumers’ ability to intertemporally arbitrage by providing incentives both to delay purchase and to buy earlier.² We discuss the profitability of the joint IPD policy

²Policy simulations indicate that instead of delaying their purchase, as shown in the traditional IPD literature, consumers buy e-readers earlier, even under higher and faster-declining e-reader prices, due to increasing e-book prices over time.
by comparing it to a single IPD policy in which, the firm only conducts IPD on e-readers and commits to a fixed e-book price. We find that profitability depends on the composition of heterogeneous consumers in the initial market. The joint IPD policy increases e-reader profits by 24.5% and total profits by $396 million if 30% of consumers are avid readers. If the fraction of avid readers is too low, it is better to commit to a fixed e-book price.

The novelty of the complementary setting is that the retailer can exploit a new dimension of consumer heterogeneity, namely, the relative demand elasticities between the two products. Our demand estimates reveal two consumer types, “avid readers” and “general readers,” with unobserved heterogeneous tastes for books. Traditional single-product IPD exploits the heterogeneity in the demand elasticities across consumer types: avid readers are less price elastic than general readers to both e-readers and e-books. The joint IPD policy further exploits the heterogeneity in the relative demand elasticities between the two products within each consumer type. The key demand-side finding that drives the supply-side pricing policy is that avid readers are more price elastic to e-books than to e-readers and that general readers are more price elastic to e-readers than to e-books. As the e-reading penetration rate of avid readers increases, the optimal pricing policy is to harvest on e-readers and invest in e-books. The opposite is true for general readers. In general, the retailer can use different joint price trajectories (i.e., harvesting on e-readers and investing in e-books, or investing in e-readers and harvesting on e-books) to induce different consumer types to purchase. The joint IPD policy serves as a better screening device for more profitable consumers.

Why do consumers exhibit different relative demand elasticities between the two products? The key intuition is that the usage intensity of the complementary product drives the adoption of the primary product. In general, our results are applicable to other industries that exhibit the same feature. We endogenize the usage intensity (i.e., book purchase) to be a function of heterogeneous consumer tastes and complementary product prices. Avid readers enjoy reading, spend more on books than on e-readers, and care more about subsequent book

\textsuperscript{3}Notice that consumer types are still unobserved to the firm. This result is a characteristic of the policy functions, which are functions of the penetration rates in each consumer type.
prices. Meanwhile, they benefit more from adopting e-readers and tend to buy e-readers earlier than others. This self-selection process drives the relative demand elasticities, which in turn drive the pricing policy. To see the importance of modeling this self-selection process, consider two pricing arrangements in a static setting: 1) high e-reader prices and low e-book prices, and 2) low e-reader prices and high e-book prices. A model without self-selection ignores the fact that consumers differ in their consumption of the two products and react to the two pricing arrangements differently. The model would thus predict that the two pricing arrangements (at some price levels) induce the same number of homogeneous adopters and make no difference to the retailer. However, our model with self-selection would predict that the former induces more avid-reader adopters while the latter induces more general-reader adopters. The compositions of adopters under the two scenarios have different pricing implications for the retailer. Similar logic applies to the intertemporal case, once we replace “high price” with harvesting and “low price” with investing. On the demand side, accounting for self-selection also produces more efficient estimates, as the adoption behavior contains information on consumers’ heterogeneous tastes.

To our knowledge, this is the first paper to empirically study IPD on complementary products. A number of empirical papers have studied optimal IPD strategy for a single durable good (e.g., Hendel and Nevo 2013, Lazarev 2013). In a complementary product setting, Leung (1997) and Koh (2006) theoretically study durable product IPD in the existence of a flat-rate complementary product. Nair (2007) and Liu (2010) empirically study IPD in the video game and console industry. They focus on single-product IPD and abstract from either software pricing or hardware pricing. We diverge from the extant literature by modeling the dynamic pricing decisions of both products. In particular, we also allow for IPD on the complementary product and for self-selection based on heterogeneous tastes. This enables us to discover new joint IPD strategies that take advantage of the complementarity.

Our paper also relates to the literature on complementary products, including tying and bundling. Most studies in this area are developed in a static setting. Turning to empirical
work in a dynamic setting, Hartmann and Nair (2010) estimate a dynamic demand system for tied goods incorporating competition across tied good systems and downstream retailers. Sriram, Chintagunta, and Agarwal (2010) examine the dynamic demand of durable goods in related technology product categories. Derdenger and Kumar (2013) study the dynamic, long-term impacts of bundling on demand. Our paper extends the literature by modeling both the dynamic demand and supply pricing problem. It also contributes to the nascent empirical literature on dynamic pricing problems where both firms and consumers are forward-looking (e.g., Nair 2007; Goettler and Gordon 2011). State-of-the-art numerical methods allow us to solve for a model where consumers’ dynamic e-reader adoption is endogenous to the retailer’s pricing strategy rather than an exogenously evolving diffusion process. The demand model shares features with dynamic models of technology adoption (e.g., Gowrisankaran, Rysman 2012; Lee 2013; Melnikov 2013).

In the next section, we describe the data and some stylized facts that motivate the demand model specification. Section 3 presents the dynamic demand model of e-readers and books, and the firm’s dynamic pricing problem. Section 4 discusses identification strategies and estimation methods, and presents the estimation results and model fit. The optimal IPD policy and profitability results are in Section 5. Section 6 offers a conclusion.

2 Data and Industry Background

2.1 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 2007. Since then, the market size of e-books has grown from $20 million to $969.9 million in 2011 (Association of American Publishers). Amazon, also a major bookseller of hardcovers and paperbacks, is the dominant player in this market. 4 Its

4We do not distinguish between hardcovers and paperbacks in this paper. From hereon, we use “paperbacks” to refer to all print books.
existing relationship with publishers enables it to offer a wide variety of e-books, which it sells at a substantial discount in the Kindle Store. Newly released *New York Times* bestsellers were sold at $9.99 in the early years of e-reading, a much lower price than the average retail print book price $17.66. Amazon also offered the gateway product to e-reading, Kindle, at affordable prices and increasing qualities over time. By providing various low-priced e-books and attractive e-readers, Amazon’s market share reached nearly 90% by the end of 2009. The Kindle enjoyed a monopoly position in 2007-2009 and was the dominant e-reader during the sample period of 2008-2012.

This paper focuses on the optimal IPD strategies for a monopolist retailer, Amazon. Consumers buy Kindles and e-books from Amazon, and buy paperbacks from all major retail channels: Amazon.com, other online websites, and offline bookstores. The data used contain prices and sales information for e-readers and books at the individual level, as discussed in the next subsection.

### 2.2 Data Description

We combine three individual-level online transaction data sets. The first data set is the transaction records of consumer online book purchases in years 2008-2012, gathered by comScore.\(^5\) Each purchase record contains the retail website, purchase time, book title, format (paperback or e-book), price, and quantity information. It also includes demographics such as household income, age, family size, zip code, etc.\(^6\) Consumers were re-sampled every year. There are 20,637 households and 72,619 book purchases. Among the online shoppers sampled by comScore, 41% bought at least one book every year. Amazon’s market share was 60% on average and increased over time as shown in Figure 2.

\(^5\)The comScore Web Behavior Database captures detailed browsing and buying behaviors of 100,000 Internet users across the United States. The panel is based on a random sample from a cross-section of more than two million global Internet users who have given comScore explicit permissions to confidentially capture their Web-wide activities. It is weighted so that the distribution of the demographics matches that of the U.S. Internet user population.

\(^6\)The income and age information is in terms of groups. Consumers are categorized into three income groups and three age groups.
The second data set contains individual-level Kindle purchase records for the years 2008-2012 from comScore. We observe purchase time, Kindle version, price, quantity, and household demographics. Consumers were re-sampled every year, so we cannot distinguish between first-time purchase and upgrading. We take a probabilistic view on the Kindle ownership status in the model. Amazon annually launched new Kindle generations and cut the prices for the existing ones. The price of the most popular Kindle version dropped from $359 in 2008 to $199 in 2012. Kindle sales increased over time, as shown in Table 2.

The third data set contains book characteristics that we collected from Amazon using web scrapers. For each book title in the first data set, we collect its genre information and prices for both paperback and e-book formats. There are 122,068 pieces of title-format information. We group the subgenres into three genres: “lifestyle,” “casual,” and “practical.” Subgenres within the same genre have similar prices, reading purposes, and consumers’ purchasing patterns.\(^7\) We aggregate the book purchase records to form the genre-format-retailer level book quantity choices for each consumer every period in the demand model. We also construct the average prices at genre-format-retailer level using all book titles every year.\(^8\) The prices are in Figure 1. “Casual” books were the cheapest and “practical” books were the most expensive. E-books are lower than paperbacks for all genres. The average prices were stable over time. For 75.2% of the book titles, e-book prices were lower than their paperback prices. A typical paperback cost $17.66, and a typical e-book cost $9.72.

\(^7\)The subgenre definition comes from Amazon. (1) “lifestyle” genre includes “Lifestyle & Home,” “Cooking,” “Travel,” “Fitness & Dieting,” “Crafts, Hobbies & Home,” “Arts & Photography,” “Children’s Book” etc. (2) “casual” genre includes “Fiction,” “Science Fiction,” “Humor,” “Non-Fiction,” “Biographies & Memoirs” etc. (3) “practical” genre includes “Computers & Technology,” “Business & Investing,” “Medical Books,” “Education & Reference” etc. The number of large genres are chosen to keep the richness of the book consumption heterogeneity while avoiding too many zero consumption observations. In the model, consumers make purchasing decisions for each genre given idiosyncratic shocks for each genre-format. A typical consumer in the sample only buys one or two books a year. As the number of defined genres increases, the number of zero consumption decisions and consumers’ book utility increase by construction. We believe that the three genres we define are representative to show the difference across book categories. In particular, “lifestyle” books usually have more pictures. “Casual” books usually serve for entertainment purposes. “Practical” books usually require in-depth reading and taking notes. These features will affect how consumers perceive e-books as substitutes for paperbacks.

\(^8\)The prices used here are historical in the first data set rather than the prices from web scrapers. The sales-weighted and unweighted prices differ by less than 2%. We use the unweighted one in the estimation. See Section 4.2 for more discussion.
We supplement the individual-level data with other relevant information. First, we impute aggregate offline book sales from online and offline retailer market shares (Bowker’s Books & Consumers report, 2012). We get the offline population size from the fraction of consumers who have bought books online (Nielsen Online Shopping Trend report, 2012). Second, we obtain the number of e-books available every year in the Kindle Store from a widely cited blog that takes monthly snapshots of Amazon. The number of available e-books increased from 126,630 in year 2008 to 1,429,500 in year 2012. Finally, we impute Kindle cost and book wholesale prices from industry reports. Kindle cost dropped from $236 to $89.

The data reveal several noteworthy consumption patterns. First, there is considerable heterogeneity in the number of books and genres purchased. Among the book buyers, 13.8% comprised nearly half (46.8%) of the total book purchases. Consumers’ book consumptions

---

Notes: the graph shows the sales-unweighted prices. The curve for practical paperback prices is scaled down by $10 to fit into the same graph.

---

9E-commerce constitutes 25.1%, 35.1%, and 43.8% of the U.S. trade book sales in the years 2010-2012. The rest of the book sales come from offline retailers such as large chain bookstores and independent bookstores. Among book buyers, 44% have purchased books online.

10http://ilmk.wordpress.com/category/analysis/snapshots/.

11First, according to the industry-wide practice, a typical hardcover list price is $29.95, and its e-book list price is $23.95. The publishers sell the books at a wholesale price, 50% of the list price, to the retailers. So the paperback wholesale price is $15 and the e-book wholesale price is $12. Second, we impute the Kindle costs from firms that release teardown reports almost every year (http://www.isuppli.com/Teardowns/News/Pages/Amazon-Kindle-Fire-Costs-$201-70-to-Manufacture.aspx). For years without these reports, we extrapolate by assuming that the cost drops at the same rate as that of computer parts.
across genres are not significantly correlated with the observed household characteristics,\textsuperscript{12} suggesting unobserved heterogeneous genre-specific reading tastes. Second, there is strong substitution between e-books and paperbacks of the same genre. Of the households, 98.66\% bought a particular genre in at most one reading format. For consumers buying e-books, 76.9\% of the time they chose the e-format for “lifestyle” books, 96.4\% for “casual” books, and 61.6\% for “practical” books. “Casual” e-books seem to be the strongest substitute for their paperback counterpart. Finally, e-reading is positively related to “casual” book consumption, suggesting a self-selection process of e-reading based on consumers’ genre-specific tastes.\textsuperscript{13}

3 Model Setup and Demand Estimation

In this section, we first illustrate the pricing incentives using a simple two-period model and then describe the full empirical model.

3.1 A Simple Two-period Model

Consider a two-period model where a firm sells a durable primary good (Kindle) at price $P$ and a complementary good (e-book) at price $p^E$ to a unit mass of consumers. The primary good serves as a gateway product to the complementary good and does not bear any stand-alone value.\textsuperscript{14} Consumers are heterogeneous in their tastes for the complementary good. The value of a unit of the complementary good $v$ is uniformly distributed on $[0, 1]$. The utility of the primary good is a function of the utility from subsequent complementary good consumption $u = \lambda (v - p^E) - P$. The coefficient $\lambda$ is the quantity of the complementary goods.\textsuperscript{15} Consumers and the firm share the same discount factor $\delta$ and live for two periods.

\textsuperscript{12}The correlation coefficient is 0.03 on average.
\textsuperscript{13}The demand estimates in Section 4 reveal that the “casual” category has the largest positive e-format fixed effect, while the “practical” category has a negative fixed effect. The e-format fixed effect contributes to the gain in book flow utilities from adopting e-readers. Consumers who like reading “casual” books gain more and are more likely to buy e-readers.
\textsuperscript{14}This assumption is relaxed in the full empirical model.
\textsuperscript{15}$\lambda$ is a function of the taste parameter $v$ and the book price $p^E$ in the full empirical model. We assume that it is a constant here to keep the analytical solution simple while illustrating the same qualitative results.
The marginal costs are assumed to be zero. Consumers have rational expectations about the firm’s pricing policies and correctly predict the future prices in equilibrium. The firm chooses prices $\bar{p}_1 = \{p^E_1, P_1\}$ and $\bar{p}_2 = \{p^E_2, P_2\}$ in period 1 and 2. Notice that the marginal consumer in period 1, $v^*_1$, is indifferent between buying and waiting:

$$\lambda \left[ v^*_1 - P_1 + \delta (v^*_1 - p^E_2) \right] - P_1 = \delta \left[ \lambda (v^*_1 - p^E_2) - P_2 \right] \geq 0$$

where $v^*_1 = \frac{P_1 - \delta P_2}{\lambda} + p^E_1$. Consumers who buy Kindles and e-books in period 1 are in the range $[v^*_1, 1]$. Similarly, the marginal consumer in period 2 $v^*_2$ satisfies $\lambda \left( v^*_2 - p^E_2 \right) - P_2 = 0$, where $v^*_2 = \frac{P_2}{\lambda} + p^E_2$. Consumers who buy Kindles and e-books 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 as well as new features of IPD with complementary products. In particular, the firm’s target is to first extract the most from high-valuation consumers on the primary good and then appeal to low-valuation consumers, while earning the most from the complementary good 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 changes. A decrease in $P_1$ reduces the demand of the primary product in period 2 and changes consumer expectations of $P_2$. $v^*_1$ 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 are self-selected into buying the primary good based on their heterogeneous tastes for the complementary good. Second, the demands of the two products are interrelated. Consumers trade off between the utility from current Kindle purchase and the value of waiting, both of which further depend on the current and future e-book prices. Third, the firm needs to coordinate the pricing of the two products. $p^E$ affects the profit from a primary product owner,
while $p^E$ and $P$ jointly affect the number of primary product owners. The full model that we study captures all the features of the simple model while allowing for richer heterogeneity and non-linear demand elasticities.

Using backward induction to solve for period 2 and period 1 prices, we get

$$
\bar{p}_2 (\bar{p}_1) = \arg \max_{\bar{p}_2} \pi_2 = (1 - v_2^*) \lambda p_2^E + (v_1^* - v_2^*) P_2
$$

$$
\bar{p}_1 = \arg \max_{\bar{p}_1} \pi_1 + \delta \pi_2 = (1 - v_1^*) (\lambda p_1^E + P_1) + \frac{\lambda \delta}{4} \left( 1 + \frac{1}{\delta} \left( \frac{p_1^E + \frac{P_1}{\lambda} - 1}{\lambda} \right)^2 \right)
$$

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, $P_1^* = \lambda (1 + \delta)$ and $p_1^E = 0$ if $\lambda \geq 1$, and $P_1^* = 0$ and $p_1^E = 1 + \delta$ if $\lambda < 1$.

The optimal strategies with complementary products differ from the harvesting (price-cutting) strategy of the traditional single-product IPD in two ways. First, both harvesting (price-cutting) and investing (price-raising) can be optimal. If $\lambda \geq 1$, it is optimal to harvest on the primary good and invest in the complementary good. The opposite is true for $\lambda < 1$. Second, the firm should 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 we get from the full model echo the results from this simple model.

### 3.2 Consumer Problem

In this subsection, we outline the dynamic demand model of consumer e-reader adoption and upgrading decisions, as well as book purchase decisions. The consumer taste heterogeneity, demand elasticities, and substitution patterns between paperbacks and e-books are critical inputs to the computation of optimal pricing strategies in the next subsection. The

---

16In the full model, we show exactly the same results: for avid readers (high $\lambda$), the firm should harvest on the primary product and invest in the complementary product; the opposite is true for general readers (low $\lambda$).
time period is one year. Every period, consumers make ex-ante dynamic Kindle adoption decisions, and then make book purchase decisions based on their device-adoption statuses; they never drop out of the market. Consumers have persistent book tastes and change their book consumption given time-varying book prices and availability.

The consumer book consumption is modeled at the genre-format-retailer level instead of book title level. First, firms care about aggregate book sales when pricing rather than the sales of a particular book title. Therefore, we focus on predicting book quantity choices instead of predicting which book title would be popular.\(^{17}\) Second, we do not have aggregate sales data for the book titles in the sample, which prevent us from estimating title fixed effects and accounting for price endogeneity issues. The benefit of modeling at the genre-level is that the average genre price is not endogenous to the quality of a particular book title in that genre. Section 4.2 provides more discussion on this.

We make the following assumptions for tractability and data limitation reasons. First, we assume that consumers only read e-books on e-readers and not on other screens, such as PCs and tablets; that is, they need to buy an e-reader before purchasing and reading e-books. We conduct robustness checks by allowing consumers to buy or read on other devices after 2010. The predictions on dynamic pricing are qualitatively robust.\(^{18}\) Second, we assume that consumers only use one Kindle at a time and that there is no resale value for Kindles. We also assume that Amazon offers one Kindle version per period, which is the most popular version of the Kindle for every period observed in the data.\(^{19}\)

\(^{17}\)Meanwhile, modeling at the title-level would require strong assumptions on the books that enter consumers’ choice set. It is not appealing to assume that consumers consider millions of books available. It is also not appealing to assume that consumers consider bestsellers, as 99.94% of the titles were purchased less than ten times in the data. Even the best-selling book had only 67 purchases out of 72,619 total purchases. Book consumption is much more dispersed than that of movies and video games.

\(^{18}\)Survey results show that e-readers are the dominant device used. Consumers using dedicated e-readers contribute the most to e-book consumption (the Book Industry Study Group survey, 2011). As a robustness check, we allow consumers to buy other reading devices in the demand estimation. The estimated Kindle qualities are affected, while the key demand-side results remain the same. We can also account for reading e-books on other devices by adding the relevant book profits to Amazon’s profit function in Section 3.3. The impact is that Amazon has weaker incentive to set low e-book prices to induce Kindle adoption because some consumers already own other devices. However, the joint IPD strategy is not qualitatively changed.

\(^{19}\)In practice, consumers are offered up to two generations of Kindles every year, except for year 2012, when three generations were on the market. The most popular version comprised at least 70% of the sales every year. Also, multi-product firm pricing is computationally prohibitive. Goettler and Gordon (2011)
Book quantity, format, and paperback retail channel choice

Consumers choose current period book quantity and format in three genres: “lifestyle,” “casual,” and “practical,” indexed by \( g = 1, 2, 3 \). A Kindle owner can buy both paperbacks and e-books, while a non-owner can only buy paperbacks. If the consumer buys paperbacks, he decides whether to buy from Amazon.com, other online websites, or offline bookstores. Let superscript \( E \) and \( P \) denote e-books and paperbacks respectively. Let superscript 0 denote Kindle non-owners and 1 denote owners. Let \( \{p_g^P, p_g^E\}_{g=1,2,3} \) and \( \{q_{ig}^{P0}, q_{ig}^{P1}, q_{ig}^E\}_{g=1,2,3} \) denote the prices and quantities of books in paperback/e-book format for non-owners/owners of type \( i \) in genre \( g \). The time subscript is dropped for notational simplicity. A Kindle owner of type \( i \) maximizes his period utility from books:

\[
\max_{\{q_{ig}^{P1}, q_{ig}^E\}_{g=1,2,3}} u_{i}^{book,1} = z + \sum_g \frac{1}{b_i} \left( a_{ig}^P q_{ig}^{P1} + a_{ig}^E q_{ig}^E - \frac{(q_{ig}^{P1} + q_{ig}^E)^2}{2} \right)
\]

s.t. \( \sum_g (p_g^P q_{ig}^{P1} + p_g^E q_{ig}^E) + z \leq y_i \) \( (1) \)

where \( a_{ig}^P \) and \( a_{ig}^E \) are taste parameters. \( b_i \) can be interpreted as the price coefficient (discussed later). \( z \) is the numeraire and \( y_i \) is household income. The numeraire price is normalized to be 1. A Kindle non-owner solves a similar problem but without e-book choices:

\[
\max_{\{q_{ig}^{P0}\}_{g=1,2,3}} u_{i}^{book,0} = z + \sum_g \frac{1}{b_i} \left( a_{ig}^P q_{ig}^{P0} - \frac{(q_{ig}^{P0})^2}{2} \right)
\]

s.t. \( \sum_g p_g^P q_{ig}^{P0} + z \leq y_i \) \( (2) \)

The optimal quantity choices in each genre for an owner and a non-owner are

---

20Economides, Seim, and Viard (2008) adopt a similar quadratic functional form without allowing for substitution. For a nice survey on direct utility models of consumer choice in marketing, see Chandukala, Kim, Otter, Rossi, and Allenby (2008).

---
\[
q_{ig}^{P1*}, q_{ig}^{E*} = \begin{cases} 
0, 0 & \text{if } p_g^P > \frac{a_i^P}{b_i} \text{ and } p_g^E > \frac{a_i^E}{b_i}, \\
(a_i^P - b_i p_g^P, 0) & \text{if } p_g^P < \frac{a_i^P}{b_i} \text{ and } p_g^E > \frac{a_i^E}{b_i}, \\
0, a_i^E - b_i p_g^E & \text{if } p_g^P > \frac{a_i^P}{b_i} \text{ and } p_g^E < \frac{a_i^E}{b_i}, \\
\text{or } p_g^P < \frac{a_i^P}{b_i} \text{ and } p_g^E < \frac{a_i^E}{b_i} \text{ and } a_i^P - b_i p_g^P > a_i^E - b_i p_g^E & \text{if } p_g^P > \frac{a_i^P}{b_i} \text{ and } p_g^E < \frac{a_i^E}{b_i}, \\
\text{or } p_g^P < \frac{a_i^P}{b_i} \text{ and } p_g^E < \frac{a_i^E}{b_i} \text{ and } a_i^P - b_i p_g^P < a_i^E - b_i p_g^E \end{cases}
\]

\[
q_{ig}^{P0*} = \begin{cases} 
(a_i^P - b_i p_g^P) & \text{if } p_g^P < \frac{a_i^P}{b_i}, \\
0 & \text{otherwise} \end{cases}
\]

We use the quadratic utility functional form, instead of the discrete choice logit utility or the CES utility, because it has the following advantages: 1) it allows for multiple-unit and corner solutions (zero consumption), which are common in book purchase patterns; 2) the optimal quantity is a linear function of the prices, so that \(b_i\) can be interpreted as the price coefficient. There is perfect substitution between paperbacks and e-books of the same genre.\(^{21}\) Purchases across genres are independent.

Once a consumer chooses to buy paperbacks, he decides among buying from Amazon.com, other online websites, and offline bookstores. This paperback retail channel choice does not interact with the previous book quantity and format choice.\(^{22}\) It is an ex-post decision only if consumers buy paperbacks. Denote the three channels \(A, B, O\). We model the channel choice using a discrete choice logit structure with i.i.d. error terms. The channel utilities contain genre-specific channel fixed effects and time trends. It is \(u_{Ag}^{channel} = A_{0g} + A_{1g} t + A_2\).

\(^{21}\)Notice that this model specification cannot generate positive numbers of books bought for both formats in the same genre. In the data set, only 1.34% of consumers buy positive quantities of both formats. For these consumers, we assume that there are two shopping occasions in a period. The observation \(\{q^P, q^E\}\) is treated as two independent observations \(\{q^P, 0\}\) and \(\{0, q^E\}\). This would mildly overestimate the substitution between paperbacks and e-books.

\(^{22}\)In the data, consumers who choose different retail channels do not have different book consumption patterns. They share the same format substitution pattern and quantity choice pattern.
\(1 \{\text{owner}\} + \zeta_A^g\) for Amazon and \(u_{Og}^{\text{channel}} = B_{Og} + B_{1g}t + \zeta_O^g\) for other online websites.\(^{23}\) It is \(u_{Og}^{\text{channel}} = \zeta_O^g\) for the baseline choice offline bookstores, whose fixed effects and time trends are normalized to zero. \(\{\zeta_A^g, \zeta_B^g, \zeta_O^g\}\) are i.i.d. logit errors. In particular, we allow Kindle owners to prefer buying paperbacks from Amazon, which is captured by \(A_2\).

**Consumer heterogeneity** Consumers differ in their observed demographics such as income \(D_i^{\text{income}}\) and age \(D_i^{\text{age}}\). They also differ in their unobserved book reading tastes. We parameterize the taste parameters as

\[
a_{ig}^P = \theta_{ig} + \beta_1 D_i^{\text{age}} + \eta_{ig}^P
\]

\[
a_{ig}^E = \theta_{ig} + \beta_1 D_i^{\text{age}} + (\theta_g^E + \beta_2 D_i^{\text{age}} + \beta_3 \log n_E) \cdot 1 + \eta_{ig}^E
\]

where each taste parameter contains a deterministic part and an i.i.d. normally distributed error term \(\eta_{ig}\) with mean zero and standard deviation \(\sigma\).\(^{24}\) The two formats in the same genre share the same baseline taste while e-format has extra taste components. In the baseline taste, \(\theta_{ig}\) is the genre-specific reading taste or the fixed effect for genre \(g\). It is heterogeneous across consumers and is not correlated with observed household demographics. We model it using a finite mixture specification. For the e-format exclusive part, \(\theta_g^E\) is the genre-specific fixed effect. \(n_E\) is the time-varying number of e-books available in the Kindle Store, which is exogenously given from the data.\(^{25}\) Consumer may value the increasing

---

\(^{23}\)During the data sample period, e-commerce has grown steadily as part of the total retail sales (http://www.census.gov/retail/mrts/www/data/pdf/ec_current.pdf). We use a linear time trend to parsimoniously capture it.

\(^{24}\)In a robustness check, we allow \(\eta_{ig}^P\) and \(\eta_{ig}^E\) to be correlated within the same genre. The implied substitution patterns and price elasticities are very robust with respect to this specification change.

\(^{25}\)The number of e-books available is unlikely to be endogenous to the e-reader and e-book pricing in this industry. Publishers decide whether or not to launch the e-book version of the books on a title-by-title basis. These decisions are particularly affected by their perception of e-books as a new reading format. They are concerned about the cannibalization of e-books on print book sales, and change their e-book publishing policies from time to time. In addition, e-book introduction is often not platform-exclusive; publishers will sell e-books on all major e-book retail platforms regardless of their e-reader user base. Amazon.com and Barnesandnoble.com both have around three million e-books available by 2012, but their e-reader userbases differ by five times. This is in contrast to the video game industry where games are often exclusive to a particular platform.
availability of e-books. Both the baseline taste and e-format taste can vary by age. Senior consumers are in general less tech-savvy and can have different e-format utilities as compared to young consumers. Finally, we allow the price coefficients to vary across income groups

\[ b_i = b_0 + b_1 D_{income}^i. \]

Recall that the demand function is \( a_{ig} - b_i p_g \). All the components in the deterministic part of the taste parameter \( a_{ig} \) affect demand directly. The price affects the demand through \( b_i \). In all, consumers differ in their observed types (captured by \( D_{income}^i \) and age \( D_{age}^i \)), unobserved types (captured by \( \theta_{ig} \)), and taste shocks \( \eta_{ig} \). The unobserved types and the observed types are independent.

**Indirect flow utility from books** We can calculate the ex-ante indirect flow utilities from book purchase for Kindle non-owners and owners \( v_{book,0}^i \) and \( v_{book,1}^i \) by substituting the optimal book purchase decisions into the utility function and taking expectations over the error terms \( \eta_{ig} \) in \( a_{ig} \).

\[
v_{book,0}^i = y_i + \sum_g E \left( \frac{(a_{ig}^P - b_i p_g^P)^2}{2b_i} \mid q_{ig}^{P0} > 0 \right) \Pr(q_{ig}^{P0} > 0) \tag{5}\]

\[
v_{book,1}^i = y_i + \sum_{T=\{P,E\}} \sum_g E \left( \frac{(a_{ig}^T - b_i p_g^T)^2}{2b_i} \mid q_{ig}^{T,1} > 0, q_{ig}^{T,1*} = 0 \right) \Pr(q_{ig}^{T,1} > 0, q_{ig}^{T,1*} = 0) \tag{6}\]

**Device adoption decision** At the beginning of each period, consumers decide ex-ante whether to buy or upgrade their Kindles. Consumers are forward-looking and can delay purchase. Consumers have perfect foresight about both the book characteristics and the Kindle prices and qualities.\(^{26}\) For years beyond our sample period 2008-2012, we assume

\(^{26}\)We assume perfect foresight because Amazon changed prices annually in the five-year period, which gives us a short panel. As a robustness check, we also try another rational expectation assumption where consumer exceptions follow an AR(1) process. The coefficients in the AR(1) model are empirically estimated using the observed prices and characteristics. The results are robust to different consumer expectation assumptions. The estimated variance of the error terms in the AR(1) model are small, indicating a small difference between the rational expectation assumption and the perfect foresight assumption.
that these variables stop evolving and stay at the year 2012 level.\textsuperscript{27} Book utilities enter device utilities. Kindle adoption is motivated by a higher book utility and Kindle upgrading is motivated by a higher Kindle quality.\textsuperscript{28}

If the consumer buys/upgrades the device in the current period, he receives utility

\[ u_i = \Gamma v_i^{\text{book}} + Q - \alpha_i P + \epsilon_i \]  \hspace{1cm} (7)

If the consumer chooses to wait, he receives utility

\[ \bar{u}_i = \Gamma v_i^{\text{book}} + \bar{Q}_i + \bar{\epsilon}_i \]  \hspace{1cm} (8)

The flow utility from books $v_i^{\text{book}}$ enters the device utility with a coefficient $\Gamma$. It takes value $v_i^{\text{book}} = v_i^{\text{book},0}$ for a Kindle non-owner and $v_i^{\text{book}} = v_i^{\text{book},1} (p^E)$ for an owner. $\bar{Q}_i$ is the quality of the Kindle that the consumer owns and equals 0 if the consumer does not have a Kindle. $Q$ and $P$ are the quality and the price of the new version Kindle offered in the current period. Qualities are estimated as dummies. The price coefficient $\alpha_i = \alpha_0 + \alpha_1 D_{\text{income}}$ is allowed to vary across income groups. The idiosyncratic shocks $\{\bar{\epsilon}_i, \epsilon_i\}$ are identically and independently distributed extreme value type 1 errors, which are also independent of the error terms on the book side. The variances are normalized to be 1.\textsuperscript{29}

Next we describe the dynamic problem. The type subscript $i$ and time subscript $t$ are dropped for notational simplicity. Let prime denote the next period value. For each type of consumer, the state space contains (1) whether the consumer has a Kindle, and if so,
the device he owns \( \tilde{Q} \); (2) the e-book price \( p^E \), which enters the ex-ante flow utility from books \( v^{book} \);\(^{30}\) (3) the offered Kindle price \( P \) and quality \( Q \); (4) the idiosyncratic shocks on the device side \( \tilde{\varepsilon} \equiv \{ \varepsilon, \varepsilon \} \). Let \( V (\tilde{Q}, Q, P, p^E, \tilde{\varepsilon}) \) denote the value function of a consumer with current device \( \tilde{Q} \) at the beginning of the period. \( d = 1 \) indicates buying/upgrading and \( d = 0 \) indicates waiting. The Bellman equation is

\[
V (\tilde{Q}, Q, P, p^E, \tilde{\varepsilon}) = \max \left\{ \Gamma v^{book} \left( p^E \right) + \tilde{Q} + \delta \left[ V (\tilde{Q}, Q', P', p'^E, \varepsilon') \mid Q, P, p^E, d = 0 \right] + \varepsilon,
\right.
\]

\[
\left. \Gamma v^{book} \left( p^E \right) + Q - \alpha P + \delta \left[ V (Q, Q', P', p'^E, \varepsilon') \mid Q, P, p^E, d = 1 \right] + \varepsilon \right\} \tag{9}
\]

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 \( \tilde{Q} \). Conditional on buying/upgrading, the device adoption status evolves deterministically from \( \tilde{Q} \) to \( Q \). The rest of the state space \( \{ Q, P, p^E \} \) evolves to \( \{ Q', P', p'^E \} \) according to consumers’ expectation about next period values. Let \( EV (\cdot) = \int \varepsilon V (\cdot, \varepsilon) d\varepsilon \) denote the expectation of the value function integrated over \( \varepsilon \). The expected value function equation is

\[
EV (\tilde{Q}, Q, P, p^E) = \ln \left[ \exp \left( \Gamma v^{book} \left( p^E \right) + \tilde{Q} + \delta \left[ V (\tilde{Q}, Q', P', p'^E) \mid Q, P, p^E, d = 0 \right] \right) \right]
\]

\[
+ \exp \left( \Gamma v^{book} \left( p^E \right) + Q - \alpha P + \delta \left[ V (Q, Q', P', p'^E) \mid Q, P, p^E, d = 1 \right] \right) \tag{10}
\]

We need to solve the consumers’ and the firm’s dynamic problems jointly in equilibrium. It is useful to rewrite the above equation as a function of the firm’s state space \( \Delta \), which is a vector that contains the number of Kindle non-owners for each type at the beginning of the current period (similar to \( v^1 \) in the simple two-period model). In equilibrium, the Kindle price \( P = P (\Delta) \) and the e-book price \( p^E = p^E (\Delta) \) are functions of the state space.\(^{31}\) We

\(^{30}\)Paperback prices \( p^P \) and e-book varieties \( n^E \) also enter \( v \). They are taken as exogenous from the data.

\(^{31}\)Consumer’s problem is a function of \( \Delta \) because \( \Delta \) enters the firm’s pricing policy, which in turn enters consumer’s problem. We assume that consumers observe it “merely as a convenient way to impose rational expectations over future prices” (Goettler and Gordon 2011).
can equivalently rewrite the consumer’s problem with $\Delta$ in the state space instead of $P$ and $p^E$:

$$EV(\bar{Q}, \Delta) = \ln \left[ \exp \left( \Gamma v_{\text{book}}^E(\Delta) \right) + \bar{Q} + \delta E \left[ V(\bar{Q}, \Delta') | \Delta, d = 0 \right] \right]$$

$$+ \exp \left( \Gamma v_{\text{book}}^E(\Delta) \right) + Q - \alpha P(\Delta) + \delta E \left[ V(Q, \Delta') | \Delta, d = 1 \right] \right]$$

(11)

Notice that there is a unique expected value function for each type in each period. The probability of buying/upgrading is

$$\phi(d = 1 | \bar{Q}, \Delta) = \frac{A}{A + B}$$

$$A = \exp \left( \Gamma v_{\text{book}}^E(\Delta) \right) + \bar{Q} + \delta E \left[ V(\bar{Q}, \Delta') | \Delta, d = 0 \right]$$

$$B = \exp \left( \Gamma v_{\text{book}}^E(\Delta) \right) + Q - \alpha P(\Delta) + \delta E \left[ V(Q, \Delta') | \Delta, d = 1 \right]$$

Given the consumer adoption probability $\phi$, the state space evolves deterministically as $\Delta' = \Delta [1 - \phi]$.

The key feature of the demand system is that Kindle adoption is driven by usage intensity of books. The usage intensity 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 book purchase affect the Kindle adoption decisions; (2) the Kindle adoption statuses influence the book formats from which consumers can choose. Consumers are motivated to buy Kindles for three reasons: gain from current-period book utility, current device prices and qualities, and the option value of device adoption. To see this, take the difference of the two choice-specific
value functions:

$$\{ [\Gamma v^{book} + Q] - [\Gamma v^{book} + \bar{Q}] \} - \alpha P + \delta \left\{ E[V(Q, \Delta') | \Delta, d = 1] - E[V(\bar{Q}, \Delta') | \Delta, d = 0] \right\}$$

(12)

The first term represents the increase in flow utility from an enlarged book format choice set ($v^{book}$ changes from $v^{book} = v^{book,0}$ to $v = v^{book,1}$). For an upgrader, it represents the benefit from the higher Kindle quality. The second term indicates that consumers will respond to a Kindle price drop. The third term is the option value, or the discounted utility gain from Kindles in the future. Both the current and future gains drive consumers to self-select into buying Kindles. We can expect that consumers who like reading benefit more from having Kindles and will adopt earlier.

### 3.3 Firm Problem

In this subsection, we describe the firm’s optimal Kindle and e-book pricing problem. To keep the model tractable, we make several simplifications from the demand model. First, we do not distinguish between book genres. The firm solves for one e-book price as if it changes book prices uniformly across genres. Second, we abstract from the quality improvements to analyze IPD, as they are often intertwined with pricing. The average quality is used in the model. Consumers have only two device ownership statuses: Kindle owner and non-owner.\(^{32}\)

Third, we restrict the pricing policy to be functions of only the two unobserved types and average over the observed types (demographics). This helps reduce the state space greatly from 36 dimensions to two dimensions, while keeping the major heterogeneity in consumer tastes that drives self-selection.

The firm sets Kindle price $P$ and e-book price $p^E$ to maximize total discounted profits from Kindles, print books, and e-books. The demand system provides two key inputs to the firm’s pricing problem: (a) the probability of adopting Kindles for each consumer type $\phi$; (b)\(^{32}\)The upgraders are modeled in a simplified way. They have proportionally higher book flow utilities than first-time adopters. The proportion is calculated from the average value in the demand system.
the book profits generated per Kindle owner $r^1 = (p^E - w^E) \cdot q^E (p^E) + (p^P - w^P) \cdot q^{P1} (p^E)$ and non-owner $r^0 = (p^P - w^P) \cdot q^{P0}$ for each type, where $\{w^P, w^E\}$ are the wholesale prices paid to publishers. $r_{gain} \equiv r^1 - r^0$ is the difference between the book profits generated by a non-owner and an owner. The Bellman equation of the firm is

$$EW(\Delta) = \max_{P, p^E} \pi (P, p^E, \Delta) + \delta E [W(\Delta') | P, p^E \Delta]$$

$$\pi (P, p^E, \Delta) = \pi^{Kindle} + p^{book1} + p^{book0}$$

where Kindle profits $\pi^{Kindle}$ equals the number of new adopters ($= \Delta \cdot \phi$) times the markup. Book profits from Kindle non-owners $p^{book0}$ equals the cumulative number of non-owners ($= \Delta \cdot (1 - \phi)$) times the book profits generated per non-owner 0. Book profits from Kindle owners $p^{book1}$ is similarly constructed. Notice that $\{p^E, P\}$ jointly affects the size and the mix of owners, while $p^E$ affects per owner book profits. Taking the F.O.C. with respect to the Kindle price yields

$$\Delta * \frac{\partial \phi}{\partial P} [P(\Delta) - c] + \Delta * \phi + \Delta * \frac{\partial \phi}{\partial P} * r_{gain} - \Delta * \frac{\partial \phi}{\partial P} * \delta W'(\Delta') = 0 \quad (13)$$

The first-order condition informs the trade-offs for Kindle pricing. Statically, a higher Kindle price increases the marginal gain on the existing Kindle sales (the first term) at the cost of the gains from new adopters (the second term) and their associated book revenue (the third term). The demand elasticities dictate the magnitudes of these effects. Dynamically, two effects are captured in the fourth term: (1) a higher Kindle price today reduces the market size and changes the mix of the two consumer types tomorrow; (2) the current period prices affect the consumer expectation over future prices. Taking the F.O.C. with respect to the e-book prices yields the following equation with similar static and dynamic
trade-offs ($\Delta_0$ is the initial market size at period 0):

$$\Delta \times \frac{\partial \phi}{\partial p^E}[P(\Delta) - c] + \Delta \times \frac{\partial \phi}{\partial p^E} \times r_{gain} + \left[ \Delta_0 - \Delta + \Delta \times \phi \right] \times \frac{\partial r_{gain}}{\partial p^E} - \Delta \times \frac{\partial \phi}{\partial p^E} \times \delta W'(\Delta') = 0$$

static Kindle profit change  
static book profit change  
dynamic future state change

We consider the pure-strategy Markov-perfect Nash equilibrium (MPNE) where both consumers and the firm are forward-looking. The non-commitment pricing policy is sub-game perfect in that prices are optimal given the state of the market in any period. The setup is similar to the framework in Nair (2007) and Goettler and Gordon (2011). The equilibrium requires that the consumer’s expectation over future state is consistent with the firm’s optimal strategy. The equilibrium is defined as the set $\{V^*, W^*, P^*, p^E^*, \Delta'^*\}$, which includes the equilibrium value functions for the consumers and the firm, the optimal pricing policy functions for both Kindles and e-books, and the beliefs about next period state space.

4 Demand Estimation and Supply Simulation Method

This section constructs the likelihood function based on the demand model, and discusses identification and estimation strategies. We then present the model fit, discuss the parameter estimates, and provide implications of the demand estimates.

4.1 Likelihood Function

The total log likelihood is comprised of probabilities for the individual-level device and

\[33\] We consider non-commitment policies because “policies with commitment are not generally subgame perfect” and “the firm has an incentive to deviate from the announced policy after the initial period passes” (Besanko and Winston (1990). Non-commitment policies, in contrast, are “more managerially relevant” (Nair 2007). Consumer rational expectation assumption is common in dynamic equilibrium models. On the theoretical side, such “relatively simple equilibrium policies are effective in explaining the key qualitative features of the data” (Nair 2007). In practice, price drops of digital durable goods such as iPhones occur in a regular manner. Many online websites (decide.com) also provide consumers with price drop predictions based on historical price data, which further enhances consumers’ ability to predict price change.
book choice observations, as well as aggregate offline book sales \( L = L^{\text{Kindle}} + L^{\text{book}} + L^{\text{aggregate}} \).

On the device side, for each consumer type, solving the dynamic programming problem gives the conditional probability of buying/upgrading conditional on holding Kindle version \( \bar{Q} \). It is \( \phi (d = 1 \mid \bar{Q}, Q, P) \) or \( \phi (d = 1 \mid \bar{Q}, t) \), as \( \{P, Q\} \) are unique per period. The conditional probability of buying/upgrading further imply the probabilities of holding a particular generation of Kindle at each point in time \( \Pr (\bar{Q} \mid t) \). Combining these two probabilities, we can calculate the unconditional probabilities of buying/upgrading every period \( \Pr (d = 1 \mid t) = \sum_{\bar{Q}} \phi (d = 1 \mid Q, t) \Pr (\bar{Q} \mid t) \). The device part of the log likelihood function for each consumer type is

\[
L^{\text{Kindle}} = \sum_{t=2008}^{2012} [n_{1t} \log [\Pr (d = 1 \mid t)] + n_{0t} \log [1 - \Pr (d = 1 \mid t)]]
\]

where \( n_{1t} \) is the Kindle sales at time \( t \) and \( n_{0t} = N_0 - n_{1t} \). \( N_0 \) is the initial number of consumers in the market. Summing over the observed types and integrating over the unobserved types, we get the total device-side log likelihood function. Notice that calculating the device choice probability involves calculating the indirect utilities from books \( v_{i}^{\text{book},1} \) and \( v_{i}^{\text{book},0} \), which contain conditional expectations of a truncated normal error and its quadratic term. For \( v_{i}^{\text{book},1} \), the truncation point is a result of a max operator. We use Gauss-Hermite quadrature with 10 nodes to calculate the conditional expectations. The details are presented in the appendix.

On the book side, the individual-level probabilities of book quantity, format, and paperback retail channel choices are combined to form the likelihood \( L^{\text{book}} \). For instance, a Kindle owner \( i \) buys \( q \) paperbacks from Amazon and zero e-book in genre \( g \). The probability of this observation is

\[
f \left( q_{ig}^{P} = q > 0, q_{ig}^{E} = 0 \right) \Pr (\text{Amazon} \mid q_{ig}^{P} > 0, q_{ig}^{E} = 0)
\]

The paperback retail channel choice probability is a standard discrete choice logit probability.
The quantity-format choice probability is calculated based on the optimal quantity solution 
\( q_{P*} = a_{ig}^P - b_i p_g^P \). In particular, we need the probability for 
\( q = \bar{a}_{ig}^P - b_i p_g^P + \eta_{ig}^P, \bar{a}_{ig}^P - b_i p_g^P + \eta_{ig}^P > \bar{a}_{ig}^E - b_i p_g^E + \eta_{ig}^E \), and 
\( \bar{a}_{ig}^P - b_i p_g^P + \eta_{ig}^P > 0 \), where \( \eta_{ig}^P \) and \( \eta_{ig}^E \) are i.i.d. normal errors. The 
calculation involves the probability of a truncated normal error where the truncation point 
is a result of a maximization operator. There is no closed-form expression for it. We use 
Gauss-Chebychev quadrature with 10 nodes and Gauss-Laguerre quadrature with 10 nodes 
to approximate the integrals. The details are presented in the appendix.

Finally, we match the model-predicted aggregate offline book sales \( \hat{H}_{gt}(\Omega) \) to the observed 
aggregate-level data \( H_{gt} \). For each set of parameter values \( \Omega \), we use the simulated error 
terms in the taste parameter \( \{\eta_{ig}^P, \eta_{ig}^E\} \) and the retail channel choice \( \{\zeta_A^{gt}, \zeta_B^{gt}, \zeta_O^{gt}\} \) to calculate 
the predicted offline book quantity for individual \( i \) in genre \( g \).\footnote{For instance, a Kindle non-owner buys 
\( \hat{q}_{ig} = \max(\bar{a}_{ig}^P + \eta_{ig}^P - b_i p_g^P, 0) \cdot 1 \{u_{channel}^{channel} > u_{Ag}^{channel}, u_{Bg}^{channel}\} \). 
Error terms for each individual are simulated for 10,000 times. The set of simulated error terms is fixed 
throughout the estimation to keep the problem stationary.} We sum over the individual 
quantities to form the predicted aggregate offline book sales \( \hat{H}_{gt}(\Omega) \). The differences between 
the observed and predicted offline sales \( L_{aggregate} = \sum_t \sum_g \left[ \left( \hat{H}_{gt}(\Omega) - H_{gt} \right) / H_{gt} \right]^2 \) serve as 
aggregate conditions and enter the likelihood function.

4.2 Identification

Since the book side and the device side are linked through complementarity, many pa-
rameters are identified from both sides. In particular, Kindle adoption is a self-selection 
process that reveals information about consumers’ book reading tastes. The income-specific 
price coefficients and age-specific taste coefficients are identified both by the observed book 
purchases and the observed Kindle adoption patterns across income-age groups. In this sub-
section, we provide more details about the primary identification sources for the book-side 
and device-side parameters.

The book-side parameters include taste parameters \( \{\theta_{ig}, \theta_g^E, \beta_1, \beta_2, \beta_3, \sigma\} \) and price 
coefficient \( b_i \). The genre fixed effects in the baseline taste \( \theta_{ig} \) are captured by a finite mixture
model and are identified from the genre market shares. The genre fixed effects in the e-format taste $\theta_g^E$ are identified from the substitution patterns between paperbacks and e-books across genres. The coefficient on time-varying e-book availability $\beta_3$ is identified from the time-varying substitution patterns between the two formats. The coefficients on age $\beta_1$ and $\beta_2$ are identified from the consumption patterns across age groups. The price coefficient $b_i$ is identified from the variation in book prices across genres, formats, and over time, although the intertemporal variation is smaller than the cross-sectional variation. In the model, paperbacks and e-books in the same genre share the same persistent genre baseline taste. This helps identify the price coefficient. The ratio of paperback price and e-book price differs across genres, which provide important identification source.

The device-side parameters include $\{\Gamma, \alpha_i, \{Q_t\}_{t=2008}^{2012}\}$. Given the same Kindle offered in each period, the Kindle sales among different types of consumers identify the coefficient on the flow utility from books $\Gamma$. Two sources identify the quality dummies $\{Q_t\}_{t=2008}^{2012}$ for each new Kindle generation offered every year: (1) cross-sectionally, the different adoption/upgrading probabilities of consumers owning different Kindle versions; (2) intertemporally, the adoption/upgrading probabilities of different consumer types. In the model, there is a unique pair of Kindle price and quality $\{P_t, Q_t\}$ every year. The fact that price incurs only once and quality enters flow utility every period helps separately identify price coefficient $\alpha_i$ and quality dummies. To avoid overfitting, we use a log time trend $Q_t = Q_0 + Q_1 \log t$ to capture the quality dummies $\{Q_t\}_{t=2008}^{2012}$.

Upgrading can be identified from the link between book consumption and Kindle purchase. In a world without upgrading, Kindle sales and e-book sales would be one-to-one. Extra Kindle sales directly lead to extra e-book sales. Upgrading is identified if the Kindle sales in later years do not lead to a proportional increase in e-book sales. Furthermore, the demographics of Kindle buyers over time offer clues about returning consumers. The early adopters of Kindles in the data are more likely to be older and have higher income. Without upgrading, the change in the demographic composition should be monotonic as the
consumer pool is exhausted. Upgraders can be identified if the income and age distribution of the Kindle adopters in the later years is similar to that in the early years.

**Price endogeneity** The demand estimation is conducted without imposing pricing optimality conditions. The model setup helps eliminate the price endogeneity issue. First, to account for the fact that Kindle prices might be endogenous to their qualities, we explicitly model and estimate the qualities of different Kindle generations. Second, we model book consumption at the genre level instead of at the book-title level. The price of a particular book title might be endogenous to its quality, which in turn affects its sales. We use the average price of a particular genre in the model, which is not endogenous to the qualities of individual book titles in that genre. In other words, a single book title does not drive the average genre price. The genre qualities are estimated as fixed effects in the model. Finally, prices are not endogenous to demand fluctuations over time.

To ensure that bestsellers do not drive average genre price in the data, we first tabulate the sales of the book titles per year in our data. It turns out that 92.82% of the book titles have only one purchase record, 5.53% have two purchases, and 99.94% have less than 10 purchases. This pattern holds for all genres and reading formats. The title with highest sales was bought 67 times among 14,524 total book transactions in that year. This is comparable to the bestseller sales as a percentage of total trade volume in the industry. In general, book consumption is much more dispersed than, for instance, video games. We conduct robustness checks by estimating the model using both sales-weighted and unweighted prices. The results are quite robust. The sales-weighted and unweighted prices only differ by 2% on average across genres and reading formats, and one of them is not systematically higher or lower than the other.

---

36 In fact, a single book title is also unlikely to drive consumer choices such as Kindle purchase. We can compare the book market to the video game market. First, consumers can always buy a title in paperback format, while they cannot play games without buying the console. Second, e-books are not exclusive to retail platforms. Amazon and Barnesandnoble.com have comparable number of e-books available, while many video games are exclusive to particular consoles.  
37 Thanks to Song Yao for suggesting this robustness check.
4.3 Estimation and Simulation Methods

To estimate the demand model, we use the Nested Fixed Point algorithm (NFXP) proposed by Rust (1987). For each iteration, we solve the dynamic programming problem for each consumer type in the inner loop and use MLE in the outer loop. Given a set of parameter guesses, we first calculate the book-side probabilities and the flow utilities from books. We then feed the flow utilities to the device side and solve the dynamic programming problem using the value function iteration method. The flow utilities and expected value functions are calculated separately for the 36 consumer types (3 age group * 3 income group * 4 unobserved segments on book tastes) every period. In particular, the value functions need to be calculated for consumers holding different generations of Kindles. Given the value functions, we can construct the device-side probabilities and form the total likelihood function.

To solve for the supply-side pricing problem, we solve in an inner loop the firm’s and the consumer’s maximization problem and calculate the next period states. In particular, the pricing policies and the next period states need to be consistent as a fixed point. Given the firm’s optimal pricing functions and the belief about the next period state space, consumers solve the dynamic maximization problem. The firm uses consumers’ choice to update the next period state space, and solves the optimal pricing problem again. The inner loop iterates until the belief and the updated next period state space converge to a fixed point. The outer loop updates the value function guess and iterates until convergence. The computation algorithm is in the appendix.

Function approximations are used in the supply-side simulation. First, two demand-side variables enter the supply-side problem: the probability of adopting Kindles $\phi$ and the book profits generated per Kindle owner $r^1$ and non-owner $r^0$. They are calculated from consumer book demand and indirect book flow utilities, which are functions of Kindle price $P$ and e-book price $p^E$. We evaluate them on a set of grid points for $P$ and $p^E$ given the estimated demand system. We then approximate these variables as functions of
Table 1: Model Fit: Kindle Cumulative Sales Among Book Buyers (%)

<table>
<thead>
<tr>
<th>Year</th>
<th>2008</th>
<th>2009</th>
<th>2010</th>
<th>2011</th>
<th>2012</th>
<th>2013</th>
</tr>
</thead>
<tbody>
<tr>
<td>Observed</td>
<td>0.41</td>
<td>1.36</td>
<td>3.54</td>
<td>6.84</td>
<td>10.34</td>
<td>14.54</td>
</tr>
<tr>
<td>Predicted</td>
<td>0.31</td>
<td>1.39</td>
<td>3.83</td>
<td>7.19</td>
<td>11.04</td>
<td>15.41</td>
</tr>
</tbody>
</table>

$P$ and $p^E$ using splines.\textsuperscript{38} Second, we discretize the state space into 20 grid points along each dimension. The value functions are approximated using cubic splines by interpolating between grid points. This gives us differentiable functions to solve the firm’s first-order-conditions.

### 4.4 Model Fit

Table 1 presents the observed and predicted Kindle cumulative sales as a percentage of total market size over time. Figure 2 compares the observed and predicted book sales for both paperbacks and e-books by book genre and by retail channel over time. The demand model is estimated using data in 2008-2012. Data in 2013 are used as an out-of-sample fit test. The model fits the level and the trend well, indicating that the model is able to recover the values of different book formats and retail channels, as well as the values of device adoption and waiting.
Figure 2: Model Fit: Books

Notes: Observed values are indicated with solid lines and predicted values are indicated with dashed lines. For Graph 2-4, the lines from top to down are sales for Amazon paperbacks, other online retailer paperbacks, and Amazon e-books.
Table 2: Parameter Estimates

<table>
<thead>
<tr>
<th>Book Parameters</th>
<th>Device Parameters</th>
</tr>
</thead>
<tbody>
<tr>
<td>( b_0 )</td>
<td>( \alpha_0 )</td>
</tr>
<tr>
<td>0.2452** (0.0026)</td>
<td>0.0042** (9.818e-4)</td>
</tr>
<tr>
<td>( b_1 )</td>
<td>( \alpha_1 )</td>
</tr>
<tr>
<td>-0.0052** (0.0002)</td>
<td>-0.0002** (6.843e-5)</td>
</tr>
<tr>
<td>( \beta_1 )</td>
<td>( \Gamma )</td>
</tr>
<tr>
<td>0.1679** (0.0062)</td>
<td>11.74** (0.6471)</td>
</tr>
<tr>
<td>( \beta_2 )</td>
<td>( Q_0 )</td>
</tr>
<tr>
<td>-0.0023 (0.0165)</td>
<td>-1.0986** (0.4071)</td>
</tr>
<tr>
<td>( \beta_3 )</td>
<td>( Q_1 )</td>
</tr>
<tr>
<td>0.0056 (0.0040)</td>
<td>0.3402** (0.1194)</td>
</tr>
<tr>
<td>( A_2 )</td>
<td>MLE Obj.</td>
</tr>
<tr>
<td>0.5670** (0.0170)</td>
<td>133,364.4</td>
</tr>
<tr>
<td></td>
<td># Obs</td>
</tr>
<tr>
<td></td>
<td>89,382</td>
</tr>
</tbody>
</table>

### Book parameters by genre

<table>
<thead>
<tr>
<th></th>
<th>Lifestyle: ( g = 1 )</th>
<th>Casual: ( g = 2 )</th>
<th>Practical: ( g = 3 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Heterogeneous reading tastes</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>High ( \theta_{ig} )</td>
<td>10.48** (0.0082)</td>
<td>10.04** (0.0178)</td>
<td>11.46** (0.0175)</td>
</tr>
<tr>
<td>Low ( \theta_{ig} )</td>
<td>1.903** (0.0433)</td>
<td>2.079** (0.0112)</td>
<td>4.489** (0.0041)</td>
</tr>
<tr>
<td>E-format utility fixed effects</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( \theta^E_g )</td>
<td>0.7047** (0.0211)</td>
<td>2.763** (0.0595)</td>
<td>-1.372** (0.0217)</td>
</tr>
<tr>
<td>Paperback retail channel fixed effects and time trends</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( A_{0g} )</td>
<td>-0.3793** (0.0128)</td>
<td>-0.7494** (0.1106)</td>
<td>-0.3507** (0.0071)</td>
</tr>
<tr>
<td>( A_{1g} )</td>
<td>0.1249** (0.0061)</td>
<td>0.1238** (0.0409)</td>
<td>0.2562** (0.0065)</td>
</tr>
<tr>
<td>( B_{0g} )</td>
<td>-1.526** (0.0107)</td>
<td>-1.041** (0.1829)</td>
<td>-1.514** (0.0114)</td>
</tr>
<tr>
<td>( B_{1g} )</td>
<td>0.0186** (0.0093)</td>
<td>0.0011 (0.0815)</td>
<td>0.1355** (0.0084)</td>
</tr>
</tbody>
</table>

### Consumer segment tastes and sizes

<table>
<thead>
<tr>
<th>Segment</th>
<th>Genre tastes</th>
<th>Population mass</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>( \theta^L_1, \theta^H_2, \theta^L_3 )</td>
<td>0.0328** (0.0071)</td>
</tr>
<tr>
<td>2</td>
<td>( \theta^H_1, \theta^L_2, \theta^H_3 )</td>
<td>0.0189** (0.0071)</td>
</tr>
<tr>
<td>3</td>
<td>( \theta^H_1, \theta^H_2, \theta^H_3 )</td>
<td>0.0032 (0.0072)</td>
</tr>
</tbody>
</table>

Notes: quality dummies are captured by a log time trend \( Q_t = Q_0 + Q_1 \log t \). Segment 4 consumers have tastes \( \theta^L_1, \theta^L_2, \theta^L_3 \). Population mass of the four segments sum up to 1.
4.5 Parameter Estimates

**Parameter interpretations** Table 2 reports the parameter estimates. The results show that consumers are highly heterogeneous in their unobserved genre-specific reading tastes, which are captured by a finite mixture structure.\(^{39}\) The data reveal four segments. Segments 1, 2, and 3 represent consumers who have high reading tastes for “casual” books, for “lifestyle” and “practical” books, and for all books respectively. They constitute 5.17% of the total population, or 12.9% of book buyers. Segment 4 represents consumers who have low tastes for all genres. For the remainder of the discussion, we refer to the first three segments as avid readers and the fourth segment as general readers. The estimated genre-specific taste parameters imply that an avid reader buys 9.1 more books than a general reader on average every year. Consumers’ book consumption also differs in their observed demographics. Model estimates show that consumers in higher income groups have lower price elasticities of Kindles and books.

In terms of reading format choice, we find that consumer format preferences vary across book genre. Consumers enjoy extra utilities from reading “lifestyle” and “casual” books in e-format and face disutilities from reading “practical” books in e-format. We can express these utilities in terms of numbers of books. If there were no price differences between the two formats, the same consumer would buy 0.70 more “lifestyle” books, 2.76 more “casual” books, and 1.37 fewer “practical” books in e-format as compared to his paperback consumption. The different e-format utilities also result in different substitution patterns between paperbacks and e-books. We find that “casual” e-books have a lower own-elasticity and a higher cross-elasticity with respect to paperbacks, indicating that they are stronger substitutes for

\(^{38}\) We try both linear splines and cubic splines. It turns out that linear splines with 11 breakpoints provide better approximation. First, the functions are highly linear with little local curvatures. Second, the values of the functions level off as \(P\) and \(p^E\) increase. Cubic splines produce small fluctuations around the steady value, which makes the derivatives inaccurate. The accuracy of the derivatives is important in the supply-side problem when solving for the firm’s first-order-conditions. Therefore, we use linear splines to approximate the functions.

\(^{39}\) We determine the number of segments by incrementally adding segments until one of the segment sizes is not statistically different from zero. Besanko, Dube, and Gupta (2003) and Nair (2007) have taken a similar approach.
paperbacks. For other coefficient estimates, we do not find statistically significant evidence that older people dislike the e-format or that e-book variety affects e-format attractiveness, although the two coefficients have expected signs.

Finally, the paperback retail channel estimates show that book buyers are migrating from offline to online, and from other online websites to Amazon.com. One interesting finding is that there is correlation between Kindle ownership and Amazon channel choice. For instance in 2012, we find that a Kindle owner had 61.1% probability of buying paperbacks from Amazon conditional on buying, while the probability was 47.6% for a Kindle non-owner. This could be either a spillover effect (i.e., state dependence) or brand loyalty (i.e., persistent preference for Amazon).

**Consumer heterogeneity** Consumers are highly heterogeneous in their book and Kindle consumption. We find that although avid readers have lower price elasticities for both Kindles and e-books than general readers, they are relatively more sensitive to e-book prices. Meanwhile, general readers are more sensitive to Kindle prices.\(^40\) This is the key finding that drives the optimal pricing policy in the next section. Notice that the differences are all driven by the differences in their unobserved reading tastes. Avid readers and general readers do not differ in the price coefficient. The intuition is that, as compared to general readers, avid readers spend more on books than on devices and care more about subsequent book purchases when considering buying Kindles. Table 3 compares the book consumption and Kindle adoption behaviors of a typical avid reader and a typical general reader. The book quantities are calculated using the average book prices. First, avid readers are less price sensitive to both Kindles and e-books; they have a much higher penetration rate of Kindles (29%) than general readers (3.4%) by the end of 2012. In terms of book consumption, a typical avid reader buys 10.9 paperbacks per year if he does not have a Kindle. Once he adopts a Kindle, he buys 3.1 paperbacks and 13.5 e-books per year. The quantities are smaller for general readers. Amazon gains more in book profits from Kindle adoption by avid readers.

\(^{40}\)The demand elasticities of books we find are comparable to extant literature on book consumption (e.g., Chevalier and Goolsbee 2003; De los Santos, Hortacsu, and Wildenbeest 2012; Reimers and Waldfogel 2014).
Table 3: Consumer Heterogeneity in Kindle and Book Purchase

<table>
<thead>
<tr>
<th>Segment size</th>
<th>Avid reader</th>
<th>General reader</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>12.9%</td>
<td>87.1%</td>
</tr>
<tr>
<td>Demand elasticity</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Kindle</td>
<td>-1.88</td>
<td>-5.74</td>
</tr>
<tr>
<td>Book</td>
<td>-0.42</td>
<td>-0.79</td>
</tr>
<tr>
<td>Ratio: Kindle/Book</td>
<td>4.5</td>
<td>7.3</td>
</tr>
<tr>
<td>Book consumption (per person)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Kindle non-owner: # paperbacks $q_0^P$</td>
<td>10.9</td>
<td>1.3</td>
</tr>
<tr>
<td>Kindle owner: # paperbacks $q_1^P$</td>
<td>3.1</td>
<td>0.9</td>
</tr>
<tr>
<td>Kindle owner: # e-books $q^E$</td>
<td>13.5</td>
<td>4.4</td>
</tr>
</tbody>
</table>

than by general readers.

5 Supply-side Policy Simulations

Given the estimated demand system, we numerically solve for Amazon’s optimal IPD policies and profitability with complementary products. We begin by discussing how the existence of the complementary product affects the IPD of the primary product. We hold the complementary product price flat and solve for the optimal IPD strategy for the primary product only. This is comparable to the traditional single-product IPD case and helps illustrate the novelty of the complementary product setting. We then allow the firm to conduct IPD on both the primary and the complementary products. We show that the firm can benefit from a joint IPD policy that exploits a new dimension of consumer heterogeneity.\textsuperscript{41}

\textsuperscript{41}To conduct the simulation, we use cost information from industry reports with one adjustment. The book cost, wholesale price, is imputed from the standard pricing approach 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. We use these rules and the observed Amazon paperback price to back out the wholesale price, which is $15 for paperbacks and $12 for e-books. However, the observed Amazon e-book price is $9.72 which is below cost. The reasons might be that there are unobserved factors that change Amazon’s actual marginal cost. For instance, there is spillover effect into other product business, negotiated quantity discounts that are not publicly observed, competition pressure, etc. To get a more realistic marginal cost value, we allow for a spillover effect per book transaction in the simulation. We solve the pricing problem with different magnitudes of this spillover effect. The predictions on dynamic pricing are very robust. This also serves as robustness checks on the value of the marginal cost we choose. We pick one value of the spillover effect so that the simulated e-book price level is comparable to the observed one for just the first period. Notice that this is not matching the entire price path because we still take a normative view on dynamic pricing policy. We keep this spillover
We make two clarifications here. First, we define harvesting and investing based on mark-up because cost drop can drive down prices. Firms adopt a harvesting (investing) strategy if the mark-up decreases (increases) over time. Notice that prices can also drop when firms adopt an investing strategy because costs drop faster than prices. Second, the book price we solve is the average price level. Firms can induce average price change through adjusting individual book title prices.

5.1 IPD on the primary product only

In this subsection, we solve for the optimal IPD policy on the primary product, holding the complementary product price exogenously fixed and flat. Compared to the traditional IPD literature, some results are consistent and some are novel in this setting. In particular, the firm loses more from consumers’ strategic behaviors, although complementarity can enhance the firm’s IPD ability on the primary product.

First, complementarity provides extra investing incentives to the firm. The optimal pricing policy is investing rather than harvesting as in the traditional IPD literature. Given the observed e-book price $9.72, the optimal Kindle price drops from $371 in 2008 to $249 in 2012. The markup increases from $135 to $160, suggesting an investing strategy. The incentive to penetrate the market early outweighs the incentive to skim high-valuation consumers.

Second, complementarity influences the firm’s IPD ability on the primary product by changing the demand elasticities of the primary product. In particular, a lower e-book price increases the difference in the demand elasticities of Kindles between the two consumer types. This enhances the firm’s IPD ability so that the firm can invest less in Kindles. When the

\footnote{effect when reporting the pricing and profitability results. Gentzkow (2007) adopts a similar approach when rationalizing the zero price of online newspapers.}

\footnote{The predicted Kindle and e-book price levels are comparable to the observed prices during the period when Amazon was able to set both Kindle and e-book prices (2008-2010).}
e-book price decreases from $9.94 to $9.63, the optimal Kindle price path shifts up and the average price increases from $292 to $305.

Third, it is noteworthy that the profit loss from consumers’ forward-looking behavior is higher in the complementary product setting. Traditional single-product IPD literatures conclude that profits and prices are higher if consumers are myopic and do not intertemporally arbitrage. For instance, Nair (2007) studies the market for console video games and finds that the profit under myopic consumers is 172.2% higher than that under forward-looking consumers. This number is 284% in our case because the difference in profits comes from not only the primary product sales, but also from the subsequent complementary product sales. The optimal price is higher under myopic consumers. It drops from $470 to $358 from 2008 to 2012 if consumers are myopic, and drops from $371 to $249 if consumers are forward-looking.

5.2 IPD on both the primary and the complementary products

In this subsection, we solve for the optimal joint IPD policies on both products (hereon, the joint IPD case).

We find that the shapes of the optimal pricing functions differ as the penetration rates of different consumer types change. This result is driven by the heterogeneous relative demand elasticities of Kindles and e-books within each consumer type, which is novel in the complementary product setting. Optimal policies are functions of the penetration rates of consumer types. As the penetration rate of avid readers increases, the optimal strategy is to harvest on Kindles and invest in e-books. The opposite is true for general readers.43 In Figure 3, we plot the policy functions over the type-specific penetration rates. For a particular penetration rate of avid readers, we average over the optimal prices for all penetration rates.

43Notice that consumer types are still unobserved to the firm. This result is a characteristic of the policy functions, which are functions of the number of consumers in each type.
of general readers. The optimal Kindle and e-book prices exhibit opposite patterns for avid and general readers. In general, it is optimal to invest in the product with higher relative demand elasticities and harvest on the product with lower relative demand elasticities.\footnote{This result is comparable to the static third-degree price discrimination problem, where consumer types are observable (e.g., Robinson 1933, Schmalensee 1981, and Aguirre et al. 2010). Compared to the non-discriminatory case, prices are lower in the weak market, where the demand is more price sensitive, and higher in the strong market. We study IPD in our paper where consumer types are unobserved. Our results can still qualitatively apply to the scenario with observed consumer types (e.g., from transaction history or surveys). Based on our results, the firm should provide promotions on Kindles to general readers and promotions on e-books to avid readers. Notice that intertemporal price discrimination is no longer necessary in this case: the firm can charge a flat optimal price, rather than vary prices over time.}

Avid readers are more price sensitive to e-books than to Kindles. The opposite is true for general readers. The joint IPD policy exploits this new dimension of consumer heterogeneity. The overall price path trades-off both consumer types and depends on the mix of consumer types in the market.

We use the optimal pricing functions to get predicted price trajectories and compare the results to the scenario where the firm conducts IPD only on Kindles and optimally commits to a fixed e-book price (hereon, the single IPD case).\footnote{We do not examine how Amazon gets commitment power in this paper. In practice, commitment power can come from reputation (e.g., Apple’s strategy on hardware) or contractual arrangement (e.g., resale price regulations).}

The predicted price trajectory...
is investing in Kindles in the single IPD case. It is harvesting on Kindles and investing in e-books in the joint IPD case. We present the predicted market outcomes in Table 4 under specification (i). The joint IPD policy induces faster-declining Kindle prices and lower penetration rates of both avid readers and general readers.

The joint IPD policy, as compared to the single IPD policy, benefits the firm in two ways. First, it offers the firm a better screening device. The firm can use different price path combinations (e.g., harvesting on Kindles and investing in e-books, or investing in Kindles and harvesting on e-books) to induce different consumer types to purchase. Figure 4 plots the percentage of avid readers among Kindle new adopters over time. The joint IPD policy induces a higher fraction of avid readers, especially in the early stage. This is because the overall optimal pricing is qualitatively the same as the optimal strategy for avid readers.

---

Notes: Specification (i) shows the scenario where the share of avid readers in the initial market is at the estimated level 13%. Specification (ii) shows the scenario where the share is 30%.

<table>
<thead>
<tr>
<th>Kindle price path ($)</th>
<th>single IPD</th>
<th>joint IPD</th>
<th>single IPD</th>
<th>joint IPD</th>
</tr>
</thead>
<tbody>
<tr>
<td>2008</td>
<td>374 (138)</td>
<td>405 (169)</td>
<td>295 (59)</td>
<td>383 (147)</td>
</tr>
<tr>
<td>2009</td>
<td>331 (146)</td>
<td>342 (157)</td>
<td>259 (74)</td>
<td>327 (142)</td>
</tr>
<tr>
<td>2010</td>
<td>297 (153)</td>
<td>289 (145)</td>
<td>231 (87)</td>
<td>280 (136)</td>
</tr>
<tr>
<td>2011</td>
<td>271 (158)</td>
<td>252 (139)</td>
<td>211 (98)</td>
<td>239 (126)</td>
</tr>
<tr>
<td>2012</td>
<td>252 (163)</td>
<td>226 (137)</td>
<td>197 (108)</td>
<td>209 (120)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>E-book price path ($)</th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>2009</td>
<td>9.63</td>
<td>10.13</td>
<td>10.53</td>
<td>10.47</td>
</tr>
<tr>
<td>2010</td>
<td>9.63</td>
<td>10.36</td>
<td>10.53</td>
<td>10.64</td>
</tr>
<tr>
<td>2011</td>
<td>9.63</td>
<td>10.59</td>
<td>10.53</td>
<td>10.88</td>
</tr>
<tr>
<td>2012</td>
<td>9.63</td>
<td>10.77</td>
<td>10.53</td>
<td>11.13</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Penetration rate by 2012 (%)</th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>avid readers</td>
<td>27.55%</td>
<td>20.1%</td>
<td>22.99%</td>
<td>19.15%</td>
</tr>
<tr>
<td>general readers</td>
<td>2.77%</td>
<td>2.33%</td>
<td>2.92%</td>
<td>2.34%</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Discounted profits (2008-2012, million $)</th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>total</td>
<td>12,597</td>
<td>11,990</td>
<td>21,800</td>
<td>22,196</td>
</tr>
<tr>
<td>from Kindles</td>
<td>2,234</td>
<td>1,642</td>
<td>2,219</td>
<td>2,763</td>
</tr>
<tr>
<td>from books</td>
<td>10,363</td>
<td>10,348</td>
<td>19,581</td>
<td>19,432</td>
</tr>
</tbody>
</table>

Table 4: IPD on Both Goods vs. IPD on the Primary Product Only
Avid readers respond the most to this strategy. This echoes the importance of modeling self-selection based on consumer heterogeneity.

Second, the joint IPD policy limits consumers’ strategic behaviors by providing conflicting incentives to arbitrage on price changes over time. Instead of delaying purchase as in the traditional literature, consumers adopt Kindles earlier even at higher Kindle prices. The firm not only extracts more Kindle profits, but also earns more book profits from subsequent book sales. Figure 5 plots the Kindle sales by consumer type over time under the single IPD case and the joint IPD case. The total Kindle sales are normalized to 100 for each case, so that we can focus on the consumer arbitrage behavior. In the single IPD case, avid readers delay purchase as the Kindle price drops from $295 in 2008 to $197 in 2012 (markups increase, so still “investing”). Interestingly, they adopt Kindles earlier in the joint IPD case, even though the Kindle price is higher and drops faster from $383 to $209. This is driven by increasing e-book prices over time. General readers are much less responsive. In the traditional IPD literature, forward-looking consumers strategically delay purchase and hurt firm profitability. In the joint IPD case, consumers have two conflicting incentives: they should delay purchase, given price-cutting on Kindles, but should adopt earlier, given price-raising on e-books. This
Figure 5: Kindle Sales Over Time

![Chart showing Kindle Sales Over Time for Avid and General Readers]

Notes: The total Kindle sales are normalized to 100 for each case.

limits their ability to intertemporally arbitrage.

The two mechanisms discussed above contribute to the profitability of jointly conducting IPD on Kindles and e-books. However, profitability is not guaranteed. Given that consumers will arbitrage on price changes, it might be better to commit to a fixed price. This is also true in the traditional single-product IPD. We find that profitability of joint IPD increases when there are more avid readers. If the fraction of avid readers is too low, it is better to commit to a fixed e-book price. Table 4 compares the scenario in which the share of avid readers in the initial market is at the estimated level 13% (specification i) to the scenario in which the share is 30% (specification ii). Committing to a fixed e-book price is better under the current market composition because, using the joint IPD policy, the firm can harvest avid readers using high Kindle prices while keeping Kindles attractive using low e-book prices. Yet low e-book prices cannot effectively keep general readers, who are less price sensitive to books; high Kindle prices substantially discourage them from adopting. The firm needs a large enough share of avid readers in the initial market to induce profitability.\textsuperscript{46} If the

\textsuperscript{46}Another reason is that the degree of harvesting/investing depends on the mix of consumer types. If the fraction of general readers is too high, the firm would adjust the price level to accommodate general readers,
fraction of avid readers is 30%, Amazon will benefit substantially from the joint IPD policy. In particular, Kindle profits increase by 24.5%. Given that the observed Kindle ecosystem produced $12.6 billion in revenue over these years,\textsuperscript{47} the total profit gain is $396 million.

5.3 Discussion: Optimality and Reality

In this subsection, we compare Amazon’s observed pricing strategy to our proposed policy. Notice that the focus of this paper is not to explain or rationalize the observed pricing policy. Instead of assuming optimality in reality, we take a normative view and investigate the optimal dynamic pricing policy for a monopolist selling both e-readers and e-books. The fact that the observed policy differs from the proposed policy does not imply that Amazon behaves sub-optimally. Amazon solves a more complex problem in a more complex environment in reality. Yet the monopolist pricing problem in a simplified setup is important and interesting to address. Our paper provides a basic framework and produces new insights on a firm’s hardware and software pricing coordination incentives, and can be used to explore other settings in the future.

The observed Amazon strategy is to conduct IPD only on Kindles and not on books. Our results imply the same pattern except that we propose an investing strategy on Kindle while Amazon adopts a harvesting strategy. In other words, Amazon is cutting Kindle prices faster over time than the proposed strategy. There are two reasons for this difference. First, in 2010, e-book pricing contract between Amazon and the major U.S. publishers switched from a wholesale contract to an agency contract, which drove down Kindle prices. Our proposed pricing strategy is calculated under the wholesale contract and does not account for the contract switch.\textsuperscript{48} Second, the presence of competitors after 2010 also drove down prices. In resulting in a deviation from the optimal pricing strategy to avid readers.

\textsuperscript{47}http://allthingsd.com/20130812/amazon-to-sell-4-5-billion-worth-of-kindles-this-year-morgan-stanley-says/

\textsuperscript{48}Our model can be extended to accommodate the agency contract by modifying the objective function of the firm. The primary difference between the two contracts is who sets e-book prices. Under the wholesale
addition, Amazon makes innovation and investment decisions along with the pricing decision. We do not have data on R&D and a long enough panel data to model innovation. We take product availability and innovation as given from the data.

Our results shed light on the pricing incentives behind Amazon’s reputation for “pricing-at-cost” for Kindles and e-books. We are able to produce a similar strategy (pricing-at-cost for Kindles, and pricing-below-cost for e-books) using our model. We find that Amazon chooses this strategy because of business spillover effects from e-reading-related consumption. First, there is correlation between Kindle ownership and buying paperbacks from Amazon. Second, there is a spillover effect to other product categories. The Kindle ecosystem comprised 10% of Amazon’s total profits in 2011. To rationalize Amazon’s sale of e-books at a significant loss, we allow for a spillover effect per book transaction to other product categories on Amazon. The calibrated spillover effect is $5.50 per book transaction. A recent survey also found a similar spillover effect: Kindle owners spend 56% more with Amazon as compared to Kindle non-owners (Consumer Intelligence Research Partners, 2012).

6 Conclusion

The paper proposes a novel joint IPD strategy under complementary product setting. Coordinating the pricing of the two products can help the firm better screen consumers based on their heterogeneity in the relative demand elasticities of the two products. It also limits

contract. Amazon sets e-book prices for consumers. Under the agency contract, publishers set e-book prices for consumers. We modify our pricing problem to calculate optimal strategies under the agency contract. Our results indicate that the optimal Kindle prices are about $80 lower under the agency contract, which is comparable to the observed Kindle prices in the agency contract period (2010-2012).

49 Without panel data, we cannot distinguish between spillover effect (i.e., state dependence) and brand loyalty (i.e., persistent preference for Amazon).

50 To validate this number, notice that Kindles with no ads are $30 more expensive than those with ads. A typical Kindle owner buys 2.2 more books than a non-owner and produces $16.6 spillover based on our prediction. Kindle ad revenue is comparable to two-year spillover effects. We also check the IPD results with $5 or $6 spillover effects. The results are robust.

consumer’s ability to intertemporally arbitrage. Our study provides new insights into the traditional pricing approach for books. Traditional book prices are set at standard discounts off list prices and remain flat over time.\textsuperscript{52} We find that retailers can conduct IPD on e-books to improve profitability, especially given that e-book prices are easily adjustable due to their digital nature.

The optimal pricing strategies and profitability results can be applied to other industries, including consoles and video games, Apple TV and digital content on iTunes, razors and blades, printers and cartridges, K-cups and espresso machines, and so on. The key common feature is that the usage intensity of the complementary product drives the adoption of the primary product. This will drive the heterogeneity in the relative demand elasticities between the primary and complementary products, based on which the joint IPD policy is built. One needs to estimate the relative demand elasticities for different consumer types and the consumer mix. The overall price trajectory and profitability depend on the mix of consumer types.

One 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, as the majority of sales would 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. The monopoly model in this paper illustrates the fundamental trade-offs for further competitive analysis. It is challenging to solve for a dynamic competition model with both e-readers and e-books, yet the multi-product setting can lead to potentially interesting competition patterns. We hope to explore this competitive setting in the future.

\textsuperscript{52}The same content is priced again only when new book editions are released. For instance, publishers typically first release hardcover editions at higher prices and later release paperback and mass-market paperback editions at lower prices.
References


Appendix

A. Likelihood Function Construction

The format-quantity choice probability comes from the error terms in the taste parameters. Define the realized error terms given the quantity choice $q_{ig}^P$ as $\eta\left(q_{ig}^P\right) \equiv q_{ig}^P + bp^P - \bar{a}_{ig}^P$ and $\eta\left(q_{ig}^E\right) \equiv q_{ig}^E + bp^E - \bar{a}_{ig}^E$. Define the thresholds of worth buying as $\bar{\eta}_{ig}^P \equiv bp^P - \bar{a}_{ig}^P$ and $\bar{\eta}_{ig}^E \equiv bp^E - \bar{a}_{ig}^E$. To simplify the notation, we drop $i$ and $g$ subscripts for now.

**Case 1:** A Kindle non-owner. The optimal quantity choice requires that the normally distributed error terms satisfy

$$\begin{cases} \eta^P = \eta\left(q^P\right) > \bar{\eta}^P & \text{if } q^P > 0 \\ \eta^P \leq \bar{\eta}^P & \text{if } q^P = 0 \end{cases}$$
The probability of buying \( q^P > 0 \) number of books is \( f \left( \eta^P = \eta \left( q^P \right) \mid \eta^P > \bar{\eta}^P \right) \cdot \Pr \left( \eta^P > \bar{\eta}^P \right) = \frac{1}{\sigma} \phi \left( \frac{\eta(q^P)}{\sigma} \right) \). The probability of buying \( q^P = 0 \) is \( \Pr \left( \eta^P < \bar{\eta}^P \right) = \Phi \left( \frac{\bar{\eta}^P}{\sigma} \right) \). Here \( \phi \left( \cdot \right) \) and \( \Phi \left( \cdot \right) \) are pdf and cdf of the normal distribution. So the contribution to the likelihood function for a non-Kindle owner is \( l_{\text{non-owner}} = \prod_g \left[ 1 \left\{ q_g^P > 0 \right\} \frac{1}{\sigma} \phi \left( \frac{\eta(q^P)}{\sigma} \right) + 1 \left\{ q_g^P = 0 \right\} \Phi \left( \frac{\bar{\eta}^P}{\sigma} \right) \right] \).

**Case 2**: A Kindle owner. Similarly, the optimal quantity choice requires that the error terms satisfy

\[
\begin{align*}
\eta^P < \bar{\eta}^P, \quad &\eta^E < \bar{\eta}^E &\text{if } q^P = 0, q^E = 0 \\
\eta^P = \eta \left( q^P \right) > \max \left\{ \bar{\eta}^P, \eta^E + \left( \bar{\eta}^P - \bar{\eta}^E \right) \right\} &\text{if } q^P > 0, q^E = 0 \\
\eta^E = \eta \left( q^E \right) > \max \left\{ \bar{\eta}^E, \eta^P - \left( \bar{\eta}^P - \bar{\eta}^E \right) \right\} &\text{if } q^P = 0, q^E > 0
\end{align*}
\]

Intuitively, paperbacks are chosen because they are worth buying \( (\eta^P > \bar{\eta}^P) \) and they are better than e-books \( (\eta^E > \bar{\eta}^E) \), vise versa. The probability that \( q^P = 0, q^E = 0 \) is \( \Pr \left( \left\{ 0, 0 \right\} \right) = \Pr \left( \eta^P < \bar{\eta}^P \right) \cdot \Pr \left( \eta^E < \bar{\eta}^E \right) \). The probability that \( q^P > 0, q^E = 0 \) is \( \Pr \left( \left\{ q^P, 0 \right\} \right) = f \left( \eta^P = \eta \left( q^P \right) \mid \eta^P > \max \left\{ \bar{\eta}^P, \eta^E + \left( \bar{\eta}^P - \bar{\eta}^E \right) \right\} \right) \cdot \Pr \left( \eta^E > \max \left\{ \bar{\eta}^E, v^E + \left( \bar{\eta}^P - \bar{\eta}^E \right) \right\} \right) \).

Define the thresholds of worth buying as \( \bar{\eta}^P \equiv bP^g - \tilde{a}^P_i \) and \( \bar{\eta}^E \equiv bE^g - \tilde{a}^E_i \). The consumer type subscript \( i \) and book category subscript \( g \) are dropped for notational simplicity.

**Probability Calculation** We need to calculate \( \Pr \left( \left\{ q^P, 0 \right\} \right) \), or the following density:

\[
f \left( \eta^P = \eta \left( q^P \right) \mid \eta^P > \max \left\{ \bar{\eta}^P, \eta^E + \left( \bar{\eta}^P - \bar{\eta}^E \right) \right\} \right) \cdot \Pr \left( \eta^E > \max \left\{ \bar{\eta}^E, \eta^P + \left( \bar{\eta}^P - \bar{\eta}^E \right) \right\} \right)
\]

where \( \eta^P \) and \( \eta^E \) are i.i.d. normally distributed error terms with mean 0 and variance \( \sigma^2 \). \( \bar{\eta}^P \), and \( \bar{\eta}^E \) are known deterministic parts. Define \( \Lambda \equiv \left( \bar{\eta}^P - \bar{\eta}^E \right) \) to simplify the discussions below. It is easier to start with calculating the CDF instead of the PDF: \( \Pr \left( \eta^P \leq \eta \left( q^P \right) \mid \eta^P > \max \left\{ \eta^P, \eta^E + \Delta \right\} \right) \cdot \Pr \left( \eta^E > \max \left\{ \eta^E, \eta^P + \Delta \right\} \right) \).

Define \( \eta^P \leq \eta \left( q^P \right) \) 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 \( D \). Then the 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,

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

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

For the second component,

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

In all, \( CDF(\eta(q^P)) = \Pr(A \mid B \cap C) \Pr(B \cap C) = \left[ \frac{(a - b)c}{1 - b} + \frac{I_1 - ac}{I_2 - c(1 - b)} \right] (1 - c) \) \( I_2 \), where \( a = \Phi \left( \frac{\eta(q^P)}{\sigma} \right) \), \( b = \Phi \left( \frac{\eta^P}{\sigma} \right) \), \( c = \Phi \left( \frac{\eta^P - \Lambda}{\sigma} \right) \), \( I_1 = \int_{\eta^P}^{\eta(q^P)} \Phi \left( \frac{x - \Lambda}{\sigma} \right) dF_x \) and \( I_2 = \int_{\eta^P}^{+\infty} \Phi \left( \frac{x - \Lambda}{\sigma} \right) dF_x \).

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

\[
f (\eta^T = x \mid \eta^T > \max \{ \bar{\eta}^T, \eta^T + (\bar{\eta}^T - \eta^T) \}) = \begin{cases} \\
\frac{a' c}{1 - b} + \frac{I'_1 - a' c}{I_2 - c(1 - b)} (1 - c) & \text{if } x > \bar{\eta}^T \\
0 & \text{otherwise}
\end{cases}
\]

where \( a' = \frac{1}{2} \Phi \left( \frac{\bar{\eta}^T}{\sigma} \right) \), \( b = \Phi \left( \frac{\eta^T}{\sigma} \right) \), \( c = \Phi \left( \frac{\eta^T - \Lambda}{\sigma} \right) \), \( I'_1 = \Phi \left( \frac{x - \Lambda}{\sigma} \right) f_x(x) \) and \( I_2 = \int_{\eta^P}^{+\infty} \Phi \left( \frac{x - \Lambda}{\sigma} \right) dF_x \).

\( f_x \) and \( F_x \) are pdf and cdf of \( N(0, \sigma^2) \).

Taking into account the fact that book quantities are integers, the ultimate probability to calculate is

\[
\begin{align*}
\Pr(\eta(q^P) \leq \eta^T < \eta(q^P + 1) \mid \eta^T > \max \{ \eta^T, \eta^E + \Lambda \}) \cdot \Pr(\eta^T > \max \{ \eta^T, \eta^E + \Lambda \}) \\
&= CDF(\eta(q^P + 1)) - CDF(\eta(q^P)) \\
&= \left[ \frac{\bar{a} c}{1 - b} + \frac{I_1 - \bar{a} c}{I_2 - c(1 - b)} (1 - c) \right] I_2
\end{align*}
\]
where \( \tilde{a} = \Phi \left( \frac{a(q^p+1)}{\sigma} \right) - \Phi \left( \frac{a(q^p)}{\sigma} \right) \) and \( \tilde{I}_1 = \int_{\eta(q^p)}^{\eta(q^p+1)} \Phi \left( \frac{z-\Lambda}{\sigma} \right) dF_x \). There are two integrals to calculate: \( \tilde{I}_1 = \int_{\eta(q^p)}^{\eta(q^p+1)} \Phi \left( \frac{z-\Lambda}{\sigma} \right) dF_x \) and \( I_2 = \int_{\eta(q^p+1)}^{\eta(q^p+2)} \Phi \left( \frac{z-\Lambda}{\sigma} \right) dF_x \). We 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** A Kindle non-owner has the following ex-ante indirect flow utility from books:

\[
v_i^{\text{book},0} = y_i + \sum_g E \left( \frac{(a_{ig}^P - b_{ig}^P)^2}{2b_i} \right) \mid q_{ig}^P > 0 \text{ Pr} \left( q_{ig}^P > 0 \right)
\]

\[
= y_i + \frac{1}{2b_i} \sum_g E \left( \left( \eta_{ig}^P - \bar{\eta}_{ig}^P \right)^2 \mid \eta_{ig}^P > \bar{\eta}_{ig}^P > 0 \text{ Pr} \left( \eta_{ig}^P - \bar{\eta}_{ig}^P > 0 \right) \right)
\]

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

\[
E \left( X^2 \mid X > 0 \right) = \text{Var} \left( X \mid X > 0 \right) = \left[ E \left( X \mid X > 0 \right) \right]^2 = \sigma^2 \left[ 1 - \lambda(\alpha) (\lambda(\alpha) - \alpha) \right] + [-\bar{\eta}_{ig}^P + \sigma \lambda(\alpha) ]^2
\]

where \( \alpha = \frac{\bar{\eta}_{ig}^P}{\sigma} \) and \( \lambda(\alpha) = \frac{\phi(\alpha)}{1-\Phi(\alpha)} \). This is a closed form solution.

A Kindle owner has the following ex-ante flow utility from books:

\[
v_i^{\text{book},1} = y_i + \sum_g E \left( \frac{(a_{ig}^T - b_{ig}^T)^2}{2b_i} \right) \mid q_{ig}^T > 0, q_{ig}^{-T} = 0 \text{ Pr} \left( q_{ig}^T > 0, q_{ig}^{-T} = 0 \right), T = P, E
\]

\[
= y_i + \frac{1}{2b_i} \sum_g E \left( \left( \eta_{ig}^P - \bar{\eta}_{ig}^P \right)^2 \mid \eta_{ig}^P > \max \left\{ \bar{\eta}_{ig}^P, \eta_{ig}^E + \left( \bar{\eta}_{ig}^P - \bar{\eta}_{ig}^E \right) \right\} \text{ Pr} \left( \eta_{ig}^P > \max \left\{ \bar{\eta}_{ig}^P, \eta_{ig}^E + \left( \bar{\eta}_{ig}^P - \bar{\eta}_{ig}^E \right) \right\} \right) \right)
\]

\[
+E \left( \left( \eta_{ig}^E - \bar{\eta}_{ig}^E \right)^2 \mid \eta_{ig}^E > \max \left\{ \bar{\eta}_{ig}^E, \eta_{ig}^P + \left( \bar{\eta}_{ig}^E - \bar{\eta}_{ig}^P \right) \right\} \text{ Pr} \left( \eta_{ig}^E > \max \left\{ \bar{\eta}_{ig}^E, \eta_{ig}^P + \left( \bar{\eta}_{ig}^E - \bar{\eta}_{ig}^P \right) \right\} \right) \right)
\]

where the probability \( \text{ Pr} \left( \eta_{ig}^P > \max \left\{ \bar{\eta}_{ig}^P, \eta_{ig}^E + \left( \bar{\eta}_{ig}^P - \bar{\eta}_{ig}^E \right) \right\} \right) \) is already calculated in appendix 1. To calculate the two conditional expectations \( E \left( \eta_{ig}^T \mid \eta_{ig}^T > \max \{ \bar{\eta}_{ig}^T, \eta_{ig}^{-T} + \left( \bar{\eta}_{ig}^T - \bar{\eta}_{ig}^{-T} \right) \} \right) \) and \( E \left( \eta_{ig}^T \mid \eta_{ig}^T > \max \{ \bar{\eta}_{ig}^T, \eta_{ig}^{-T} + \left( \bar{\eta}_{ig}^T - \bar{\eta}_{ig}^{-T} \right) \} \right) \), we use the conditional expectation definitions \( E \left[ X \mid \mathcal{H} \right] = \int_{-\infty}^{\infty} x \cdot f \left( x \mid \mathcal{H} \right) dx \) and \( E \left[ X^2 \mid \mathcal{H} \right] = \int_{-\infty}^{\infty} x^2 \cdot f \left( x \mid \mathcal{H} \right) dx \). The conditional density \( f \left( x \mid \mathcal{H} \right) \) is calculated in the last section. Given the conditional density, we calculate the conditional expectations using Gauss-Hermite quadrature with 10 nodes.
C. Computation Algorithm for the Dynamic Pricing Problem

The numerical algorithm is similar to that in Goettler and Gordon (2011). It contains an inner loop and an outer loop. The inner loop solves the firm and consumer maximization problem along with the next period state space given the value function guess. The outer loop updates the value function guess and iterates until convergence.

For the agency model where only the Kindle optimal prices are solved, for each iteration $k = 1, 2, ...,$

1) Guess the value functions for the firm and the consumers $\{V_{k-1}, W_{k-1}\}$.

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 $\Delta'$, which in turn depend on their rational expectations of $\Delta'$, we solve for a fixed point in $\Delta'$ such that consumers’ expectations for $\Delta'$ are realized according to the state space evolution equation. In particular, to solve for the fixed point, first guess the next period state space $\Delta'^{m-1}$ and the firm’s optimal pricing policy $P^{m-1}$, where $m$ is the iteration number for the fixed point of the inner loop. Given the guess, solve the consumers’ device adoption problem to get updated next period state space $\Delta'^m$. Given the updated $\Delta'^m$, solve the firm’s first-order conditions at each state and get the updated optimal pricing policy $P^m$. Check convergence of $|\Delta'^m - \Delta'^{m-1}|$ and $|P^m - P^{m-1}|$. If converged, let $\Delta'^k$ and $P^k$ 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 $|V^k - V^{k-1}|$ and $|W^k - W^{k-1}|$ at the state space grid points $\Delta$. If convergence is not achieved, return to step 1.

The pricing problem for the wholesale model is similar except that an extra first-order condition for e-book pricing is simultaneously solved along with the one for Kindle pricing. The optimal pricing policy for e-book is also updated and converged in the inner loop. Throughout the computation, we discretize the state space evenly into 20 grid points on both dimensions. The range of the state space is between 0 and the initial market size of each type. We use a cubic spline to interpolate between the grid points for the value functions and policy functions. This is because solving the firm’s first-order condition requires differentiable continuation values. The convergence is checked at the grid points.