

The First Sale Doctrine and the Digital Challenge to Public Libraries*

Imke Reimers
Northeastern University

Joel Waldfogel
University of Minnesota, NBER, ZEW

July 1, 2022

Abstract

Libraries have traditionally provided free communal access to books, and the role of libraries in knowledge diffusion is facilitated by the first sale doctrine's (FSD) guarantee that libraries may purchase physical books at consumer prices. Increasingly restrictive ebook access terms may imperil libraries, and we compare the welfare cost of higher ebook prices to the welfare benefit of the FSD's guarantee of low physical book prices for libraries and their patrons. Using data on over 8,000 library systems for 2013-2019, we measure the impacts of physical and electronic holdings on the respective types of circulation. We then build a structural model of the library market, and we rationalize the status quo book holdings with a librarian utility function that attaches higher weights to electronic circulation. While higher counterfactual ebook prices would induce libraries to substitute physical for electronic holdings, this would have little effect on patron CS because of consumer willingness to substitute. By contrast, higher physical book prices, as would prevail absent the first sale doctrine, reduce CS by almost ten times as much as an analogous ebook price increase. Moreover, policies that promote physical holdings would stimulate library visits along with their possible concomitant effects on social capital.

*We thank seminar participants at NUS Singapore, Brandeis University, JLU Gießen, Indiana University, Boston University, the Rotman School at Toronto, Stanford University, the University of East Anglia, the 2022 IIOC, and the 2022 Munich Summer Institute. We also thank Devesh Raval and Avi Collis for providing insightful discussant comments.

Introduction

Free access to information is an important contributor to development and innovation.¹ Historically, libraries have provided this access by making physical books available to patrons in local facilities. This practice has been enabled by the “first sale doctrine,” a feature of US copyright law that allows libraries to lend books purchased at the consumer price.² Effectively, libraries operate as subsidized local book-purchasing collectives.

Digitization has the potential to substantially alter the role of libraries. By facilitating zero-marginal-cost distribution of text, video, and music, digitization could deliver an information “utopia” in which all people have access to all information. Ubiquitous book availability could in principle obviate current-day public libraries: Consumers could access books from the cloud, without intermediation by nearby organizations. But two frictions, respectively involving law and consumer preferences, inhibit digital displacement of libraries. First, copyright law’s first sale doctrine guarantees libraries inexpensive access to physical but not electronic books, and this promotes continued reliance on physical books. Law alone is inadequate to keep libraries viable, however. That requires a second condition, that users continue to find physical books appealing, even as digital book technology diffuses more widely.

While publishers could in principle use ebooks to supplant libraries, in the first 15 years since the emergence of ebooks, publishers have instead embraced libraries as complementary distribution channels. Unlike physical books, ebooks are not subject to the first sale doctrine, so publishers are able to make ebooks available to libraries on terms that are more restrictive for libraries and more favorable to the publishers. Per instance of circulation, ebooks cost

¹See, for example, [Berkes and Nencka \(2020\)](#); [Furman et al. \(2018\)](#); [Gilpin et al. \(2021\)](#) and [Karger \(2021\)](#).

²The first sale doctrine stipulates that “an individual who knowingly purchases a copy of a copyrighted work from the copyright holder receives the right to sell, display or otherwise dispose of that particular copy, notwithstanding the interests of the copyright owner.” See <https://www.justice.gov/archives/jm/criminal-resource-manual-1854-copyright-infringement-first-sale-doctrine>.

libraries on average four times what physical books cost. Even on these terms, ebooks have grown to account for roughly a third of library spending on new holdings.

In the past few years, publishers have stirred controversy with additional restrictions on library ebook access, imposing embargoes after release and raising ebook prices further relative to physical books. This could make libraries less useful to patrons or possibly even obsolete; but this fate depends largely on the substitutability of physical and electronic books to patrons. If consumers view ebooks as sufficiently more appealing than physical books, then limited access to ebooks via libraries will diminish interest in libraries, at least as sources for books. On the other hand, if library patrons are readily willing to substitute physical for electronic books, then libraries' continued ability to purchase and lend physical books under the first sale doctrine will limit the vulnerability of libraries to the pricing of digital products.

Librarians and others have recently emphasized the important role of libraries in building social capital in communities, epitomized by the [Klinenberg \(2018\)](#) designation of libraries as “palaces of the people.” Yet, library visits were waning even before the advent of electronic books: After rising for decades, per-capita US library visits began falling in 2009 and by 2012 had fallen more than 10 percent.³ Because patrons can borrow ebooks without visiting libraries, the arrival of ebooks may have hastened the decline in physical library visits and associated community building. Thus – and particularly if patrons are willing to substitute physical for electronic books – inhibiting library access to ebooks might have the subsidiary effect of stimulating visits and concomitant creation of social capital.

This paper explores the effects of digitization on libraries by asking two broad questions. First, how do libraries' decisions about electronic and physical holdings affect users' tendencies to borrow electronic and physical books, as well as library visits? Related, to what extent

³See https://nces.ed.gov/programs/digest/1999menu_tables.asp and <https://www.ala.org/tools/research/librarystats/public/publiclibraries>.

are patrons willing to substitute between electronic and physical book formats? Second, how vulnerable are patrons to publishers' proposed increases in ebook prices; and, analogously, how much protection do patrons derive from low physical book prices guaranteed by the first sale doctrine?

To answer these questions, we use data on over 8,000 libraries in the US from 2013 to 2019, including the number of volumes held in each format and their respective annual circulations, as well as the annual number of visits to each library. Our study has two parts. The first part is a descriptive analysis measuring the impact of physical and electronic holdings on the circulation of books in the respective formats, as well as visits. We employ two strategies to address the potential endogeneity of holdings. We use library fixed effects to control for time-constant unobservable determinants of circulation; and we also use discontinuous jumps in holdings (which arise frequently for ebooks given bundled purchasing arrangements) to measure causal impacts of holdings on circulation. With both approaches, we find that holdings affect patron demand for circulation: Larger physical or ebook holdings bring about higher physical and ebook circulation, respectively. Moreover, the circulation impact of an additional physical book is much larger than for an ebook. Cross effects, by contrast, are negative, indicating that patrons regard physical and ebooks as substitutes. Finally, physical holdings stimulate library visits, while ebook holdings depress them.

The descriptive findings are indicative of mechanisms that we incorporate into the second part of our analysis, a structural model of consumer demand and library book holdings. We use this model to compare the relative impacts of higher electronic and physical book prices as a means of evaluating alternative availability regimes. On the demand side, consumers decide which types of books to borrow as a function of library format holdings. This model, which we implement as a two-level nested logit, has sufficient flexibility to accommodate different possible degrees of substitution among titles within format as well as across formats. Given consumer demand for physical vs electronic books, the supply side of the market con-

sists of libraries choosing how to allocate their book budgets to holdings in each format. We rationalize observed holdings as those that maximize each library’s “utility function,” which consists of a weighted sum of the physical and electronic circulation that patrons choose, subject to the budget constraint implicit in current library spending on holdings. Despite higher ebook prices, the library-specific utility functions rationalizing current holdings generally value electronic circulation substantially more than physical circulation, per instance. Patrons, by contrast, are more willing than librarians to substitute physical for electronic books.

We implement two counterfactual analyses that entail raising electronic and physical book prices, respectively. First, we evaluate the threat to patron CS from publishers’ proposed increases in ebook prices. A doubling of the ebook price would induce substantial substitution toward physical holdings, increasing library visits by 2.3 percent. Because of patron willingness to substitute across books, patron CS would fall by only 0.4 percent. Second, we evaluate the welfare cost of the higher physical book prices that would prevail without the first sale doctrine. Doubling the physical book price has a larger negative effect on patrons than doubling the ebook price. Libraries would substitute substantially toward electronic holdings, and patron CS would fall by 3.4 percent, or by almost ten times more than the impact of higher ebook prices. Finally, library visits would fall by 12.9 percent with doubled physical book prices.

Because libraries offer books without charge, changes in access terms may disproportionately affect low-income populations; and because ebooks are equally available to remote and nearby patrons, increased ebook prices may disproportionately affect rural patrons. Accordingly, we explore welfare impacts of changed access terms separately by community income and urban/rural status. Despite being free, library usage is concentrated in higher-income communities, so restricting electronic or physical access does not deliver disproportionate harm to low-income communities. And despite facing relatively large selections of ebooks,

rural patrons have low per-capita ebook borrowing rates, limiting the negative welfare effects of higher ebook prices on rural patrons.

The paper proceeds in seven sections. Section 1 provides background on the library and ebook markets, including discussions of the copyright literature, the first sale doctrine, and the perspectives of publishers and libraries on book access. The section also presents a simple microeconomic framework making clear the possible misalignment of library holdings choices and patron preferences for formats, as well as the potentially different impacts of increased electronic vs physical book prices. Section 2 describes the data, and Section 3 provides our estimates of the causal impacts of physical and electronic holdings on circulation and visits. Section 4 introduces our empirical structural model, which combines consumer demand for books and formats at libraries with library demand for physical and electronic books, given consumer preferences. Section 5 presents estimates of the structural library patron demand model and the libraries’ implied preference parameters. We use the estimated model in Section 6, which presents the counterfactual scenarios with different library access to electronic and physical books, discussions of heterogeneous effects by income and urban status, and analyses of robustness of the main results with respect to values of estimated parameters. Finally, Section 7 concludes.

1 Background

1.1 Digitization, ebooks, and libraries

In the past century and a half, US public libraries have played a growing role in providing universal access to information.⁴ Modern public libraries have their origins in the 19th century, with libraries financed by Andrew Carnegie, which were “dedicated to the diffusion

⁴Recent research documents effects of libraries on education and innovation outcomes. See [Gilpin et al. \(2021\)](#) and [Karger \(2021\)](#) on library impacts on education outcomes and [Furman et al. \(2018\)](#) on patenting.

of knowledge.”⁵ By 1920, the US had 3,500 public libraries, just over half of which had been endowed by Carnegie between 1883 and 1929.⁶ A series of federal Library Services Acts in the 1950s and 1960s subsidized library development in both rural and urban areas.⁷ By 2019 the US had 16,548 libraries.⁸ In addition to distributing books, public libraries have come to play “important roles in community building.”⁹ According to the American Library Association, library programs help patrons “locating reliable health information, learning the skills needed to thrive in today’s digital economy, helping children and families succeed in school and life, and purchasing affordable health insurance.”¹⁰

Library sharing of books has long been facilitated by the first sale doctrine, which dictates that libraries can purchase *physical* books at the same prices offered to individuals, even as the libraries lend the books without charge.¹¹ Because an instance of library circulation generates less revenue than a purchase, libraries have long held out the possibility of cannibalizing book sales, but the first sale doctrine prevented publishers from pricing or providing access in ways that might control cannibalization.¹²

Digitization has rapidly and substantially changed both the consumer and library markets for books. The US trade book market had \$16.23 billion in 2019 revenue. Since their emergence in 2008, ebooks have grown to account for 12.4 percent of 2018 revenue in the

⁵See [Berkes and Nencka \(2020\)](#). The diffusion quote comes from the dedication of the 1903 Carnegie Library in Washington DC. See <https://www.npr.org/2013/08/01/207272849/how-andrew-carnegie-turned-his-fortune-into-a-library-legacy>.

⁶See <https://dp.la/exhibitions/history-us-public-libraries/carnegie-libraries>.

⁷See <https://www.presidency.ucsb.edu/documents/statement-the-president-upon-signing-the-library-services-and-construction-act-0>.

⁸See <https://www.ala.org/news/sites/ala.org.news/files/content/2019-soal-report-final-accessible.pdf>.

⁹See <https://www.brookings.edu/blog/up-front/2017/03/30/how-public-libraries-help-build-healthy-communities/> as well as [Klinenberg \(2018\)](#).

¹⁰<https://www.ala.org/news/sites/ala.org.news/files/content/2019-soal-report-final-accessible.pdf>.

¹¹[Shiller \(2013\)](#) models the effect of the first sale doctrine on the market for video games.

¹²The theoretical frameworks relevant to impacts of library lending on purchase mirror those in the literatures on piracy and the sale vs renting of information. See [Varian \(2000\)](#), [Bakos et al. \(1999\)](#), and [Smith and Telang \(2016\)](#). In an empirical study of Japan, [Kanazawa and Kawaguchi \(2019\)](#) find that an instance of library circulation of physical books displaces a quarter to a half of a sale.

US and 13 percent of the 170 million 2019 unit sales.¹³ While ebooks make it possible for libraries to lend books remotely – and, indeed, some planners have envisioned the “bookless library” – libraries have overwhelmingly adopted hybrid approaches, continuing to distribute physical books from their branch locations and distributing ebooks online.¹⁴ Ebooks are now a substantial part of library holdings and acquisitions. US public libraries spent \$400 million on ebooks in 2019 – up from about \$200 million in 2012 – accounting for about a third of library expenditures and a fifth of all ebook revenue.¹⁵ Library ebooks may be read by multiple patrons, and libraries make up a larger share of ebook readership than revenue. Ebook circulation reached about 400 million checkouts in 2019, indicating that library lending is an important means by which consumers get access to ebooks.

1.2 The digital challenges to libraries

The emergence of ebooks has created two conflicting challenges for libraries. First, now that ebooks are popular, some major US publishers have further restricted ebook access terms. During 2019, Hachette and Simon & Schuster replaced perpetual licenses with two-year ebook licenses; and audiobook provider Blackstone Publishing announced a new 90-day

¹³The sales figure is from Nielsen. See <https://publishers.org/data-and-statistics/industry-statistics/>.

¹⁴See <https://www.npr.org/2013/01/15/169412670/a-new-chapter-a-launch-of-the-bookless-library> for a discussion of bookless libraries. During 2020, Brewster Kahle, the developer of the Internet Archive, made 1.4 millions on-copyright books available online, ostensibly to provide access to books users might otherwise have obtained at libraries during the pandemic. The Authors Guild promptly sued, and the “national emergency library program” was scrapped. See [Nagaraj and Reimers \(2021\)](#) for evidence on the effect of the Google Books program on book sales and usage. For information on the national emergency library program, see <https://www.publishersweekly.com/pw/by-topic/industry-news/publisher-news/article/83472-publishers-charge-the-internet-archive-with-copyright-infringement.html> and <https://www.theverge.com/2020/6/14/21290902/internet-archive-emergency-library-lawsuit>.

¹⁵Libraries tend to have not only large ebook holdings; they also hold the most popular titles. While title-level holdings and circulation data are not generally available, the Seattle Public library releases its statistics. Among top titles, ebook holdings are as extensive as their physical holdings: The library held 48 of the 2018 USA Today top 50 titles as ebooks and 47 as print books. The Seattle data are available at https://data.seattle.gov/widgets/tmmm-ytt6?mobile_redirect=true.

embargo on sales to libraries, leading to a boycott.”¹⁶ Macmillan also announced an embargo in 2019. For eight weeks after publication, libraries could purchase only one copy of an ebook.¹⁷ To librarians, these changes amounted to publishers “deliberately sabotaging the library system” to induce people to “buy [ebooks], instead of borrow them.”¹⁸ The American Library Association has opposed efforts “to delay or deny library access to digital content,” maintaining that “[p]roviding perpetual access and reducing the price for a single copy is important to ensuring all people have access to the world’s knowledge” at libraries.¹⁹

The second challenge arising from ebook adoption is that ebook usage may decrease library visits and the sort of community building that libraries foster. The concern is real, as visits have declined steadily since 2009 after growing consistently for decades (see Figure 1).

1.3 Ebook prices, acquisition, and bundles

Due to the first sale doctrine, libraries purchase physical holdings at the same price per copy that consumers pay for the books. By contrast, libraries can obtain ebooks under a variety of different arrangements. For example, OverDrive, which supplies ebooks to 76,000 libraries and schools, provides books in four different ways.²⁰ First, a book can be purchased outright

¹⁶See <https://www.libraryjournal.com/?detailStory=publishers-change-ebook-and-audiobook-models-libraries-look-for-answers>.

¹⁷In a letter to librarians, Macmillan CEO John Sargent explained: “We believe the very rapid increase in the reading of borrowed e-books decreases the perceived economic value of a book. I know that you pay us for these e-books, but to the reader, they are free. In the pre-digital world reading for free from libraries was part of the business model. To borrow a book in those days required transportation, returning the book, and paying those pesky fines when you forgot to get them back on time. In today’s digital world there is no such friction in the market. As the development of apps and extensions continues, and as libraries extend their reach statewide as well as nationally, it is becoming ever easier to borrow rather than buy.” See <https://d1x9nywezhk0w2.cloudfront.net/wp-content/uploads/2019/10/29160131/A-Letter-from-John-Sargent-.pdf>. Also see <https://www.denverlibrary.org/blog/books-research/lauren/libraries-and-ebooks-introduction>.

¹⁸See <https://goodereader.com/blog/digital-library-news/major-publishers-are-sabotaging-the-public-library>.

¹⁹See <https://americanlibrariesmagazine.org/blogs/the-scoop/ala-responds-to-macmillan-letter/>.

²⁰See <https://company.overdrive.com/company-profile/who-we-are/>.

and made available to one user at a time (“one copy/one user”). Second, a book can be licensed to be made available via “metered access,” either a fixed time or a maximum number of checkouts. Third, a library can obtain access to each of various publishers’ catalogs, and then pay a “cost per circulation” when a user borrows a book. Finally, libraries can purchase “simultaneous use” contracts for either particular titles or collections from each publisher, which allow unlimited borrowing of those titles until contract expiration.²¹

Data on library ebook pricing are not systematically available, but one US librarian has prominently disclosed price comparisons.²² Based on an analysis of 645 best-selling titles whose physical editions had suggested retail prices of \$24.78, the average Amazon print price was \$16.77, while the physical library editions averaged \$14.14. Individual book copies on Kindle had an average price of \$12.77. Libraries, by contrast, faced an average price of \$45.75 for ebooks, presumably when purchasing metered access.²³ As we show below, however, the cost per circulation pricing scheme appears predominant, as library ebook expenditures are closely related to ebook circulation rather than ebook holdings.

The way that libraries add ebooks to their holdings can give rise to substantial changes in ebook holdings from one year to the next. Rather than purchasing individual copies, as libraries have traditionally done with physical books, ebook cost-per-circulation arrangements allow libraries access to book bundles, and the library incurs a charge when a patron borrows an included title. We exploit the annual jumps in the numbers of ebooks that libraries offer to consumers (their ebook holdings) to measure the impact of holdings on circulation.

²¹See the “OverDrive Marketplace User Guide,” available online at <https://manualzz.com/download/7173254>.

²²Jennifer Rothschild tweets weekly on the library ebook prices for bestsellers. See <https://www.forbes.com/sites/rachelkramerbussel/2019/09/20/why-this-librarian-is-using-twitter-to-fight-changes-to-library-ebook-pricing-terms/>.

²³See <https://smartbitchestrashybooks.com/2020/09/hold-on-ebooks-cost-how-much-the-inconvenient-truth-about-library-eollections/>.

1.4 Framework

A simple model is helpful for guiding the empirical work below. Given a library’s holdings of physical and electronic books (N^P and N^E), consumers collectively maximize their utilities by choosing how many physical and electronic books to borrow. This results in the patron demand functions $Q^P(N^P, N^E)$ and $Q^E(N^P, N^E)$, and consumer welfare is described by a utility function over holdings $U(N^P, N^E) = U(Q^P(N^P, N^E), Q^E(N^P, N^E))$.

If libraries charged for books and sought to generate profits, it would be natural to model libraries maximizing price-weighted sums of physical and electronic circulation, less marginal costs associated with book lending. In reality, libraries derive no revenue from circulation but nevertheless make expenditures on holdings that attract circulation. We model libraries as choosing their holdings of physical and electronic books N^P and N^E to maximize the notional “welfare” that libraries derive from circulation subject to their budget constraints. That is, the libraries maximize $W(N^P, N^E) = W(Q^P(N^P, N^E), Q^E(N^P, N^E))$ subject to $B = P^P N^P + P^E N^E$, where P^P and P^E are the prices of physical and electronic holdings, respectively, and B is the budget.

It is possible that librarians choose holdings to maximize the welfare of patrons, but given that they make their holdings decisions independent of consumer choices, their holdings decisions need not maximize patron welfare. This potential misalignment is illustrated in Figure 2, which depicts a library’s budget constraint, its holdings choice, and librarian and patron indifference curves. With physical holdings on the vertical axis, the librarian indifference curve is depicted – foreshadowing our results – to be steeper than the patrons’ at any holdings choice (N^P, N^E) because librarians derive higher relative marginal utility from electronic circulation. Relative to the depicted holdings, patrons would prefer more physical books and fewer electronic books. Hence, increases in electronic and physical book prices would have asymmetric implications for patron welfare: Increases in physical book prices, which move holdings farther from patrons’ preferences, harm patrons more than increases

in ebook prices. For what follows, we define $\theta = \frac{MRS_{\text{librarian}}}{MRS_{\text{patron}}}$, where θ is the supply-side behavioral parameter reflecting the relative librarian preference for ebook circulation. We provide estimates of θ in Section 5.

2 Data

The data for this study consist of an annual panel of 8,418 US library systems from 2013 to 2019. For each library system and year, we observe the number of physical and electronic volumes in their collection (N^P and N^E), the physical and electronic circulation (Q^P and Q^E), and the number of visits to the library. The data also include the population of the service area (from which we construct market size M). We obtain the data from the Institute of Museum and Library Services (IMLS).²⁴ Because some libraries are not included in all years, our final dataset consists of 50,722 library system-year observations.

We summarize our main variables of interest in Table 1, separately for ebooks and print books. On average, libraries hold 73,617 electronic volumes and 105,571 print holdings.²⁵ Table 1 also provides information on trends in electronic and physical holdings. Electronic holdings have increased quickly, while physical holdings are, on average, stable or declining. There is variation across libraries in holdings growth, however. The mean (median) annual percentage change in electronic holdings is 26.9 (14.8), while the interquartile range runs from 6.1 to 29.2 percent, or from 849 to 14,477 units. Annual growth in physical holdings is much slower: the mean (median) percentage change is -0.8 (0.2) while the interquartile range runs from -3.4 to 2.8 percent, or from -1,550 to 1,005 units. In each year, some libraries experience substantial changes in physical or especially electronic holdings, and we use these sharp instances of variation for identification of causal impacts of holdings on circulation

²⁴See <https://www.ims.gov/research-evaluation/data-collection/public-libraries-survey>.

²⁵Our electronic holdings are primarily ebooks (73 percent) but also include audiobooks and video holdings. The vast majority (88 percent) of our physical holdings are books.

(and visits) below.

The information in Table 1 shows the relative expense on electronic vs physical circulation. For physical books, the ratio of average expenditure to average circulation is \$0.37, while the ratio is four times as high for ebooks. This is the sense in which ebooks are more expensive to libraries than physical books. However, our data suggest that ebook expenditure depends on circulation, while physical book expenditure depends on holdings. When we regress ebook expenditure on ebook circulation and holdings, the circulation coefficient is seven times higher than the holdings coefficient. This dependence of ebook expenditures on circulation demonstrates the prevalence of the cost per circulation model.²⁶

In Section 4 we introduce a model in which librarians allocate their budgets to holdings. Implementing the model requires annual prices per holding for each format. The ratios of expenditure to holdings, for each format, provide simple measures of these prices, which are \$0.91 for physical books and \$0.61 for ebooks. Because we model librarians choosing holdings, below we rely on prices per holding that obey $P^E = 2/3P^P$. While ebooks are more expensive than physical books per instance of circulation, ebooks cost libraries just two thirds as much annually per holding. This reversal arises both because ebooks circulate less than physical books and because libraries using cost-per-circulation arrangements incur charges for ebooks only when the titles are borrowed.

Before moving on, it is worth noting that Table 1 tells much of the story in this paper. Libraries spend almost a third of their acquisition budgets on electronic materials while these materials account for only a tenth of circulation. This suggests both a high librarian marginal valuation of ebook circulation as well as substantial patron enthusiasm for physical books.

²⁶A comparison with physical books is instructive. Because libraries purchase and possess their physical holdings, we would expect physical expenditures to depend on holdings. A regression of physical expenditures on physical holdings and circulation confirms this, delivering a holdings coefficient more than twice the circulation coefficient.

Our data also include information about patron visits to libraries. On average, libraries experience 173,109 (45,239) visits per year, and the volume has declined by 2.0 (1.3) percent per year during the sample.

2.1 Library usage by patron income and urban/rural status

Table 2 provides holdings, circulation, and visit data by community income and urban/rural status. As the top panel shows, holdings and, especially, circulation rise in income levels. Even relative to population, both physical and electronic circulation rise proportionally more than income. Notwithstanding the free nature of library book lending, these patterns indicate that library book usage behaves like a luxury good, suggesting that less favorable book access terms will have larger effects on higher, rather than lower, income patrons.

The bottom panel shows how physical and electronic book holdings and usage vary across library systems according to their urban/rural status. Both physical and ebook holdings tend to be larger in more urban systems, although ebook holdings are far less sensitive to population density. Indeed, rural libraries offer their patrons the choice among roughly twice as many electronic as physical books, while urban libraries have roughly twice as many physical books. Despite the large relative availability of ebooks to rural patrons, rural library lending is dominated by physical books: 96 percent of rural library circulation is physical, compared with just 90 percent in the most urban systems. This suggests that rural patron well-being is less sensitive to increased ebook prices. We explicitly explore welfare impacts of higher electronic and physical book prices by community income and urban/rural status in Section 6.

3 Descriptive evidence

3.1 Holdings and circulation

We are ultimately interested in the welfare effects of various library resource allocation decisions – e.g. how many physical and electronic books to include in collections. As a first step, we measure the extent to which physical and electronic holdings, respectively, attract library users to consumption and the extent to which patrons regard physical and electronic volumes as substitutes.

To this end, Table 3 presents regressions of physical and ebook circulation on physical and ebook holdings as well as year dummies. In columns (1) and (2), which do not include library fixed effects, coefficients on all holdings variables are positive in both regressions, indicating that libraries with larger holdings have higher circulation. These regressions are vulnerable to a concern that some libraries have populations with greater tastes for borrowing books, so that the coefficients on holdings would not reveal the true impact of holdings on circulation. Columns (3) and (4) include library fixed effects, so that coefficients are identified by within-library system changes in physical and electronic holdings. In these specifications, own effects are positive, while cross effects are negative.

While the fixed effects approach deals with time-constant unobservables that may be correlated with both holdings and circulation, one can be concerned about time-varying unobservables. For example, libraries whose patrons have growing demand for books in a format may respond to this demand growth by increasing their format-specific holdings. This could give rise to positive estimated relationships between holdings and circulation for reasons apart from the causal impact of holdings on circulation.

Circumventing this endogeneity requires some source of exogenous variation in holdings. We use sudden jumps in holdings. Even if holdings and circulation are covarying smoothly, we can treat the change in circulation surrounding a large jump in holdings as an “experiment” for measuring the impact of holdings on circulation.

Between 2013 and 2019, the median annual percentage change in ebook holdings is 14.8

percent, but 50.5 percent of libraries experience at least one annual ebook change over 50 percent. The median percentage change in physical book holdings is 0.2 percent. While very few libraries experience any 50 percent year-to-year growth in physical holdings, 19.6 percent of libraries experience at least one annual instance of 10 percent growth.

For the analysis, we define the top decile of annual percent changes in electronic and physical holdings as jumps. We then treat the jumps above these cutoffs as discontinuities, and we implement this discontinuity idea in two steps. First, we estimate “first stage” regressions showing the relationship between these discontinuous jumps in holdings and the size of respective holdings. Column (5) and (6) show “first stage” regressions of physical and electronic holdings on indicators for their respective jumps. Columns (7) and (8) report “reduced form” regressions of physical and electronic circulation, respectively, on both jump indicators. The last two columns report IV regressions of circulation on the respective types of holdings, instrumenting holdings with the jumps and including only the two years surrounding the jumps.²⁷

The fixed effects and instrumental variables approaches deliver similar estimates. The positive own-format coefficients – indicating, for example, that larger physical holdings attract more physical book circulation – show that patrons value holdings. The negative cross-format coefficients – showing that libraries with more ebook growth experience depressed physical book circulation, and vice versa – indicate that patrons view physical and electronic books as substitutes for one another. The similarity of the IV and the fixed effects estimates lends further credibility to our main causal findings: Main effects are positive and significant, cross effects are negative and significant; and the physical main effect is larger than the ebook main effect.

The relative coefficient sizes in Table 3 also offer some clues about library managers’

²⁷The resulting coefficients are quite similar to the Wald IV estimates calculated as ratios of the coefficients in columns (7) and (8) to those in columns (6) and (7) (see Angrist and Krueger, 1991). The coefficients differ from the direct ratios because we have two rather than one instrument per equation.

supply side motivations. Ignoring cross effects for illustration, an additional physical book raises physical circulation by 0.676 (0.675) units in the fixed effects (instrumental variables) estimates, over six times larger than the ebook own effect of 0.108 (0.067). That is, an additional physical holding engenders much more additional circulation than an additional ebook. Because physical books are only 50 percent more costly to libraries per holding, this pattern of coefficients suggests that libraries attach more value to ebook than physical circulation. In our framework, this would imply a θ above one.

3.2 Library visits

As [Klinenberg \(2018\)](#) emphasizes, libraries fulfill functions apart from simply distributing books. By bringing people together to physical spaces, libraries help to build social capital in communities. Ebooks can be borrowed without a physical visit to a library. Hence, it is possible that the shift of resources from physical to electronic books discourages library visits, with possible impacts on social capital formation.

Table 4 reports regressions of visits on library physical and ebook holdings. Column (1) includes year effects but does not include library fixed effects. Both physical and ebook holdings have positive coefficients, indicating that libraries with more holdings have more visits. Column (2) adds library fixed effects, and their inclusion changes the results: Physical holdings continue to have a positive coefficient, while the coefficient on ebook holdings is negative. An additional physical holding raises visits by 0.248 (se = 0.073) while an additional ebook holding reduces visits by 0.082 (0.009).

The library fixed effects address the possibility that different communities have time-constant different preferences for library patronage; but as with the circulation regressions, one can be concerned about time-varying unobservables affecting both holdings and the tendency for patrons to visit libraries. We use our “holdings jump” IV approach for visits as well. Column (3) reports the reduced form regression of visits on the physical and electronic

jump instruments. A jump in physical holdings raises visits by 3,569 per year while a jump in ebook holdings reduces visits by 3,822 per year. The resulting IV estimates are reported in column (4). Using just the interval from two years before until the year after a jump, an additional physical volume raises visits by 0.39 (0.11), while an additional ebook volume decreases visits by 0.05 (0.02) per year.

How large are the depressing effects of ebook holdings on visits? In 2018 the libraries in the sample had 1.25 billion visits, and these libraries collectively had holdings of 800 million ebooks. Using the ebook coefficients from column (2) and (4), this implies that absent ebooks, library visits would have been 6.0 (4.1) percent higher in 2018 using the fixed effects (instrumental variables) estimate, holding physical holdings constant. Visits declined roughly 10.9 percent between 2013 and 2018, and the growth in ebooks between 2013 and 2018 can explain about a third of this decline.²⁸

4 Model

Our empirical implementation of the model, which embodies the findings above, has two broad parts. First, consumers have demand for borrowing physical and electronic books from libraries. Second, given consumer attitudes toward physical and electronic books, along with the respective prices that libraries pay, libraries choose the sizes of their respective physical and electronic holdings to maximize “library utility.” The library’s welfare function attaches value to circulation of physical and electronic books, respectively. Given these implicit valuations, library holdings and patron demand – circulation – emerge as equilibria. The model delivers these equilibrium outcomes, and we use the model to calculate counterfactual equilibria. We entertain effects of different electronic and physical book access on physical

²⁸The elimination of all ebooks would imply between 4.0 and 6.1 percent more visits in 2018. Actual ebook holdings in 2013 were about a quarter of their level in 2018. Hence, the growth in ebooks between 2013 and 2018 would account for 77 percent times 4.0 to 6.1 percent, or about a third of the 10.9 percent decline in visits.

and ebook circulation and patron and librarian welfare, as well as library visits.

4.1 Demand for borrowing

We specify a two-level nested logit model of demand in which consumers at library l choose between three upper-level alternatives: the outside good, a physical book, or an electronic book. Then, having chosen a physical or an electronic book, the consumer chooses a title. Suppressing the library identifier, consumer i 's utility from borrowing book j is given by

$$u_{ij} = \delta_j + \zeta_i^G + (1 - \sigma_2)\zeta_i^f + (1 - \sigma_1)(1 - \sigma_2)\epsilon_{ij}. \quad (1)$$

In this model, the mean utility of product j is given by $\delta_j = x_j\beta + \xi_j$, and ϵ_{ij} follows an extreme value distribution. Further, ζ_i^G and ζ_i^f follow distributions such that $[\zeta_i^h + (1 - \sigma_1)(1 - \sigma_2)\epsilon_{ij}]$ and $[\zeta_i^G + (1 - \sigma_2)\zeta_i^f + (1 - \sigma_1)(1 - \sigma_2)\epsilon_{ij}]$ are also distributed extreme value. The term G is the set of formats at the library (physical and electronic), and the term f denotes an individual format (physical or electronic). The parameter σ_1 reflects the correlation of utilities that consumers experience for books within a format, while the parameter σ_2 reflects the correlation of utilities across formats.²⁹

Consumer utility maximization results in familiar logit expressions for the shares, which yield quantities of electronic and physical books borrowed from a library when multiplied by market size. Thus,

$$s_j = \frac{e^{\delta_j/(1-\sigma_1)}}{D_f^{(\sigma_1-\sigma_2)/(1-\sigma_2)} D_G^{\sigma_2} (1 + D_G^{1-\sigma_2})}, \quad (2)$$

where $D_f = \left[\sum_{j \in \mathcal{J}_f} \exp\left(\frac{\delta_j}{1-\sigma_1}\right) \right]^{\frac{1-\sigma_1}{1-\sigma_2}}$, \mathcal{J}_f is the set of products in format f , and $D_G = \sum_{f' \in G} \left(D_f^{(1-\sigma_1)/(1-\sigma_2)} \right)$. We can use these logit formulas to express circulation as a function of the holdings: $Q^P = Q^P(N^P, N^E)$ and $Q^E = Q^E(N^P, N^E)$, where N^P and N^E are physical

²⁹See [Berry \(1994\)](#) and <http://www.nathanhmilller.org/nlnotes.pdf>.

and electronic holdings, respectively.

Given the nested logit structure of the utility function, the consumer surplus that patrons obtain from a library's holdings is then given by

$$CS = M \times \ln \left[1 + \left(\sum_{f \in \{P,E\}} D_f \right)^{(1-\sigma_2)} \right]. \quad (3)$$

Because library books are free to consumers, there is no price in the utility function and therefore no price parameter that translates utility into a dollar figure. As a result, we can only quantify proportionate changes in CS across ebook access regimes, not its absolute level.

4.2 Library supply of holdings

Libraries do not collect revenue from users. Nevertheless, they operate as if they attached value to instances of circulation. Accordingly, we presume that each library maximizes a welfare function defined over physical and electronic circulation. We describe this function as $W(Q^P(N^P, N^E), Q^E(N^P, N^E))$, where W reflects the notional value the library obtains from instances of circulation of physical and electronic books.

The library's maximization problem is constrained by its budget: $B = P^P N^P + P^E N^E$, where P^P and P^E are the prices per holdings of physical and electronic volumes, respectively, and B is the library's budget for books. Hence, the Lagrangian for each library's full maximization problem is given by

$$\mathcal{L} = W(Q^P(N^P, N^E), Q^E(N^P, N^E)) + \lambda[B - P^P N^P - P^E N^E]. \quad (4)$$

We define the derivatives of the librarian's welfare function with respect to circulation as θ^P and θ^E (e.g., $\frac{\partial W}{\partial Q^P} \equiv \theta^P$), so that $\frac{\theta^E}{\theta^P} \equiv \theta$ introduced in Section 1.4. Then, suppressing

library subscripts, each library's first-order conditions for this problem are:

$$\begin{aligned}\frac{\partial Q^P}{\partial N^P} + \theta \frac{\partial Q^E}{\partial N^P} - \lambda P^P &= 0, \\ \frac{\partial Q^P}{\partial N^E} + \theta \frac{\partial Q^E}{\partial N^E} - \lambda P^E &= 0, \\ P^P N^P + P^E N^E &= B.\end{aligned}$$

The first two first-order conditions can be rewritten succinctly as

$$\begin{bmatrix} \frac{\partial Q^P}{\partial N^P} & \frac{\partial Q^E}{\partial N^P} \\ \frac{\partial Q^P}{\partial N^E} & \frac{\partial Q^E}{\partial N^E} \end{bmatrix} \begin{bmatrix} 1/\lambda \\ \theta/\lambda \end{bmatrix} = \begin{bmatrix} P^P \\ P^E \end{bmatrix}. \quad (5)$$

Once we have estimated demand, the derivatives of circulation (Q) with respect to volumes (N) are known, as are the prices and the library's budget. We can thus solve for the library's welfare weight θ using equation (5).³⁰ Our inference of library managers' welfare weights determines the librarian indifference curve rationalizing the holdings choice, as in Figure 2, and recalls the estimation of welfare weights that a hypothetical planner attaches to different types of consumers in [Berry et al. \(2016\)](#).

4.3 Equilibrium

The library-specific parameters θ show the relative value that the each library attaches to electronic vs physical circulation. The values of these parameters that we infer from the status quo holdings choices $\{N^P, N^E\}$ provide the baseline solution to the model. We can then calculate counterfactual equilibria arising from different physical and electronic book prices by re-solving the model. We solve the model separately for each library by maximizing the librarian utility with respect to the holdings budget constraint.

³⁰Formally we obtain the welfare weights using: $\begin{bmatrix} 1/\lambda \\ \theta/\lambda \end{bmatrix} = \begin{bmatrix} \frac{\partial Q^P}{\partial N^P} & \frac{\partial Q^E}{\partial N^P} \\ \frac{\partial Q^P}{\partial N^E} & \frac{\partial Q^E}{\partial N^E} \end{bmatrix}^{-1} \begin{bmatrix} P^P \\ P^E \end{bmatrix}.$

5 Estimates

5.1 Demand

Following [Berry \(1994\)](#), we can estimate a two-level nested logit model of demand by estimating the following equation:

$$\ln(s_j) - \ln(s_0) = \beta + \gamma \mathbb{1}_j(\text{ebook}) + \sigma_2 \ln(s_{f|G}) + \sigma_1 \ln(s_{j|f}) + \xi_j \quad (6)$$

Here, $\ln(s_{f|G})$ is a format's – electronic or physical book – share of book lending at a library, and $\ln(s_{j|f})$ is a particular book's share of a format's lending.

Our estimation strategy faces two challenges, one related to data and the other to identification. We have as data the total quantity of lending by library and format. That is, we observe Q^f , so we have a direct measure of $s_{f|G} = \frac{Q^f}{Q^E + Q^P}$. We do not observe the circulation of title j , but we do observe library holdings of physical and electronic books, N^P and N^E , respectively. We assume symmetry within format, so that $s_{j|f} = \frac{1}{N^f}$. Then the overall product-level share is given by $s_j = \frac{Q^f}{N^f M}$.

To identify the substitution parameters, we instrument $s_{f|G}$ with group f 's share of total holdings: $\frac{N^f}{N^E + N^P}$. The term $\ln(s_{j|f})$ would normally require an instrument, and $(1/N^f)$ would be a logical candidate. Here, because $(1/N^f)$ is our direct measure of $s_{j|f}$, it requires no instrument. We also include year and library fixed effects, and we suppress the library index for simplicity.

Table 5 reports estimates of the demand parameters. Column 1 reports the model estimated with library and year fixed effects but without instruments. The uninstrumented σ coefficients are close to 1, presumably because we are – in the case of the top-level share – regressing an expression involving Q^f (i.e., $\ln(\frac{Q^f}{N^f M})$) on another expression involving Q^f (i.e., $\ln(\frac{Q^f}{Q^E + Q^P})$). Column (2) shows the first-stage estimate of $\ln(\frac{Q^f}{Q^E + Q^P})$ on the instruments.

The instrument “works”: For example, physical books make up larger shares of circulation in library years in which physical works make up larger shares of the collections. Column (3) reports our two-level nested logit model of demand. We estimate the parameter σ_1 to be 0.937 (0.006) for substitution among titles within a format, and we estimate σ_2 , reflecting substitution across formats, to be 0.671 (0.03). This indicates that titles are closer substitutes for one another within format than across format.³¹ These are the estimates we use in our simulations, and we explore the sensitivity of our main results to alternative parameter estimates in Section 6.4. Although the model appears parsimonious, it is important to recall that both the format-specific mean utilities and the librarian utility parameter θ are library-specific.

5.2 Supply

Given estimates of parameters σ_1 and σ_2 , we calculate format-specific δ 's for each library-year directly from the data, where for each book j in format f , $\delta_j = \ln(s_j) - \ln(s_0) - \sigma_2 \ln(s_{f|G}) - \sigma_1 \ln(s_{j|f})$. This, in turn, allows us to create expressions for Q^P and Q^E as functions of N^P and N^E . Then, using the holdings (N^P and N^E) and the prices of physical and electronic books (per year), we calculate the library-specific welfare weights θ .³² The distribution of these weights is skewed: The mean (median) θ estimate is 5.34 (1.86), and the inter-quartile range runs from 1.28 to 2.69. Most libraries implicitly attach substantially more value to instances of electronic vs physical circulation: 86.3 percent of libraries have θ estimates above unity.

One might be concerned that libraries' decisions to hold electronic as opposed to physical books are influenced by space constraints. A few facts allay such concerns. First, physical

³¹We also estimate a negative coefficient for ebooks, indicating that ebook holdings have lower mean utility than physical holdings.

³²In our main results, we use a common relationship between P^E and P^P ($P^E = 2/3P^P$). We also solved the model using prices specific to groups of library systems by community income and urban/rural status. Results, which we report in Appendix Section A.1, are nearly identical.

holdings decline over time, both overall and per library square footage. In our sample, the mean (median) physical holdings per square foot fall from 5.34 (4.63) in 2013 to 5.09 (4.31) in 2019. Over the same period, mean (median) absolute physical holdings fall from 125,900 (44,388) to 94,698 (36,692). Hence, libraries are not increasingly space-constrained by books. Second, the shadow price of space used for physical books appears small. For example, using the Seattle circulation data introduced in footnote 15, we estimate that 80 percent of physical holdings do not circulate each year.³³ We infer from this that adding physical volumes to the collection would not crowd out holdings that are much used by patrons.

Solving the model leads to two complications that we explore in the Appendix. First, given our parameter estimates, our model does not solve to observed holdings in baseline simulations for 5.7 percent of libraries. Second, roughly a sixth of libraries solve to corners (with zero electronic or physical holdings) with counterfactual doubling of the respective book prices. Our main results exclude libraries that solve to corners; but Appendix Section A.2 shows that we obtain very similar results, with slightly larger welfare losses, if we include the corner solutions.

6 Counterfactuals

This section uses the model for four counterfactual exercises. First, we ask how increased ebook prices, reflecting publisher restrictions on ebook access, affect patrons, libraries, and visits. Second, we explore how an increase in physical book prices, as would occur absent the first sale doctrine, affects market participants. Third, we ask how the results vary by income and urban/rural status. Finally, we explore the robustness of the results to alternative substitution parameters.

³³We use the Seattle data to calculate the number of book copies checked out per year, and we compare this to the number of holdings in our data. We do not observe circulation separately for each copy but rather for each title. Because a copy can circulate no more than 25 times per year, we estimate the number of circulating copies of a title as $\text{int}(\text{checkouts}/25)+1$.

6.1 Raising the electronic book price

We model publishers' proposed ebook restrictions with increases in the price of ebooks by 10 or 100 percent above its current level while holding the physical book price constant. We report the circulation, holdings, and welfare effects in columns 1 and 3 of Table 6. As the first and second rows show, holding the budget fixed, as the ebook price rises, equilibrium physical holdings rise while ebook holdings decline. If the ebook price were twice its baseline level (column 3), physical holdings would increase by 5.0 percent while ebook holdings would decline by 56.5 percent. Physical circulation would rise by 1.7 percent, while electronic circulation would decline by 14.0 percent. Because baseline physical circulation substantially exceeds baseline ebook circulation, total circulation would fall only slightly (by 0.3 percent) with a doubling of the ebook price.

CS would decline by a similarly small amount (0.4 percent), but librarian utility, which assigns substantial weight to ebook circulation, would fall by 0.9 percent. The small effect of the doubled ebook price on consumers arises for two reasons. First, the negative effect is mitigated by consumers' willingness to substitute across and, especially, within formats. Second, relative consumer preference for physical as opposed to electronic books limits the negative effect of an increase in the ebook price. Finally, library visits would rise by 2.3 percent.

6.2 Raising the physical book price

If the first sale doctrine did not exist, libraries would likely face higher prices of physical books. We simulate this by increasing the price of physical books – by 10 or 100 percent – while holding the ebook price constant. We show the counterfactual outcomes in columns 2 and 4 of Table 6. Doubling the price of physical books (column 4) would reduce physical holdings by 52.8 percent while raising ebook holdings by 14.3 percent. As a result, CS would

fall by 3.4 percent while librarian surplus would fall by 2.4 percent. Visits would fall by 12.9 percent.

An increase in the price of physical books would have a larger negative impact on CS than an equal-sized increase in the ebook price. A ten percent increase in the price of physical books (column 2) would have 8.9 times the CS impact of a ten percent increase in the ebook price (column 1), while the analogous ratio for doubling prices (columns 3 and 4) would be 9.2. The reduction in holdings of consumers' preferred format is more harmful to their CS. The benefit to patrons from stable ebook prices is small compared with the protections to print book prices afforded by the first sale doctrine.

6.3 Heterogeneity by income and urban/rural status

We also examine the relative impacts of increased electronic and physical book prices separately by income as well as urban/rural status. In our baseline approach, we solve the model using a single relationship between P^E and P^P , in particular that $P^E = 2/3P^P$ for all libraries. This is correct on average, but given the information in Table 2, one can see that the relationship varies across communities according to income and rural/urban status. This occurs because ebooks expenditure varies with circulation rather than holdings, and circulation varies across types of communities. We simulate separately by income or urban status groups by solving the model using group-specific relationships $P^E = \kappa P^P$. For income quintiles, κ ranges from 0.31 for the lowest income quintile to 1.31 for the highest. In the urban/rural simulation, κ varies between 0.14 for the rural areas and 1.31 for the most urban.

Panel A of Table 7 presents price-doubling results separately for groups of libraries by income quintile. Doubling the ebook price has a larger negative effect on CS as community income rises, while doubling the physical book price has a larger negative effect on CS as income falls across communities. Consequently, increased physical book prices reduce CS

by 13.2 times as much as increased ebook prices for patrons in the lowest income quintile, compared with 7.8 times in the highest income quintile.

Panel B of Table 7 presents price-doubling results separately for groups of libraries according to urban/rural status. Given the relatively low ebook usage by rural library patrons, increased ebook prices have smaller negative impacts on CS in rural areas, where increased physical book prices have larger negative CS impacts. The relative decrease in CS is 7.3 times higher for physical than for electronic book price increases in large cities, while the ratio is 12.2 for rural library patrons.³⁴ In short, the protections of the first sale doctrine are particularly important to rural and low-income communities.

6.4 Robustness to substitution parameters

In our baseline estimates an increase in the physical book prices leads to larger decreases in consumer surplus than an equivalent increase in ebook prices: An increase in physical book prices reduces CS about nine times as much as an increase in the price of ebooks. These results may depend on the estimated substitution parameters (σ_1 and σ_2), and we explore the robustness of our results to these substitution parameters here.

Figure 3 reports the percentage changes in CS for doubling prices of ebooks (blue dashed line) and print books (red dotted line), with varying degrees of substitutability within (σ_1) and across (σ_2) formats. In particular, we choose σ_1 to be either our estimated value of 0.937 or 0.5, and we vary σ_2 from 0.1 to 0.8, keeping $\sigma_2 < \sigma_1$. The baseline results are indicated by the dots. When $\sigma_1 = 0.937$ – and titles are close substitutes within format – the percentage change in CS with a doubling of the ebook price is roughly constant.

³⁴Ebook expenditure is driven by circulation rather than holdings. The price of ebooks in our model is a price per holdings. We also solved the model allowing different prices per holdings in the different income or rural/urban groups. Allowing this flexibility delivered very similar results, both for the percentage impacts on CS from increased prices, as well as the relative sizes of the change in CS from physical vs electronic book price increases. See Appendix A.1 for evidence that allowing for flexible κ delivers similar overall results.

Whatever the substitutability across formats (σ_2), the Δ CS ratio is about eight.³⁵

The shorter lines describe changes in CS for varying levels of substitutability across formats (σ_2) with lower substitutability within format ($\sigma_1 = 0.5$). Here, because within-format substitution is limited the degree of substitutability across formats matters. When across-format substitutability (σ_2) is low, the CS effect of an increase in book prices of either format is larger. However, the CS ratio of interest (the effect of doubling print book prices compared to ebooks) is almost unchanged, ranging between 7.6 and 8.2. While the absolute effects of price changes on CS depend on our estimated preference correlation coefficients, the relative effects do not.

7 Conclusion

Libraries have traditionally drawn patrons together to make shared use of physical books, a practice facilitated by the first sale doctrine, which allows libraries to purchase books at the same prices as consumers. Library ebook prices, not subject to the first sale doctrine, exceed those charged to consumers. Ebooks have become popular at public libraries; and publishers are now raising library ebook prices. This could undermine the value of public libraries to users, particularly if patrons are unwilling to substitute physical for electronic books.

Our findings on the effects of ebook price increases may provide some comfort to library managers: Because consumers are willing to substitute among books – and continue to attach high value to physical books – restricting ebook access would have limited effects on library patron utility. Moreover, shifting holdings toward physical books would increase library visits. On the other hand, the first sale doctrine provides valuable protections to

³⁵Note these numbers are slightly different from those in Table 6 because here we keep the exact same set of libraries across counterfactuals – all those that *never* solve to a corner solution. In Table 6, we keep more libraries, including some of those that solve to corners in counterfactuals with different σ s.

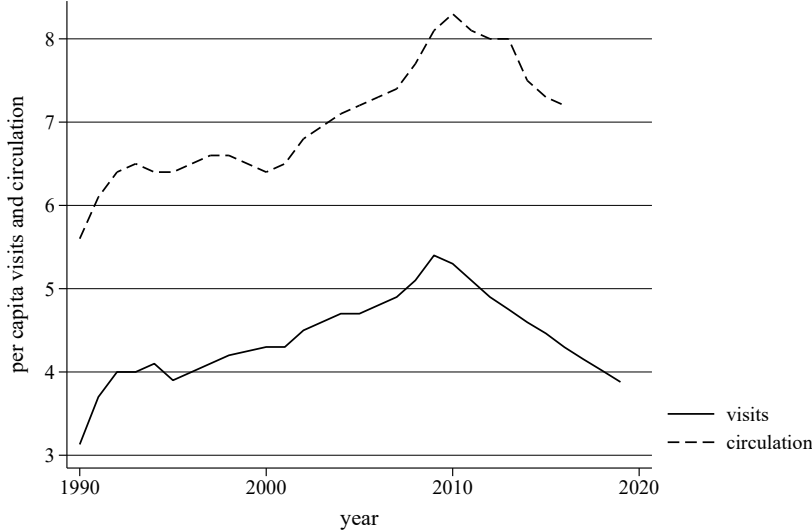
library patrons. In contrast to the negligible percent change in CS arising from a doubling of the ebook price, a doubling of the physical book price would reduce patron CS by 3.4 percent, or by almost ten times the impact of doubled ebook prices. Moreover, the relative vulnerability of patrons to higher physical book prices is greater for rural and for low-income communities. A continuation of favorable library access to physical books appears important as libraries navigate a more digital society.

References

- ANGRIST, J. D. AND A. B. KRUEGER (1991): “Does compulsory school attendance affect schooling and earnings?” *The Quarterly Journal of Economics*, 106, 979–1014.
- BAKOS, Y., E. BRYNJOLFSSON, AND D. LICHTMAN (1999): “Shared information goods,” *The Journal of Law and Economics*, 42, 117–156.
- BERKES, E. AND P. NENCKA (2020): “Knowledge access: The effects of Carnegie libraries on innovation,” *Available at SSRN 3629299*.
- BERRY, S., A. EIZENBERG, AND J. WALDFOGEL (2016): “Fixed costs and the product market treatment of preference minorities,” *The Journal of Industrial Economics*, 64, 466–493.
- BERRY, S. T. (1994): “Estimating discrete-choice models of product differentiation,” *The RAND Journal of Economics*, 242–262.
- FURMAN, J. L., M. NAGLER, AND M. WATZINGER (2018): “Disclosure and subsequent innovation: Evidence from the patent depository library program,” Tech. rep., National Bureau of Economic Research.
- GILPIN, G., E. KARGER, AND P. NENCKA (2021): “The Returns to Public Library Investment,” .
- KANAZAWA, K. AND K. KAWAGUCHI (2019): “Crowding-Out Effects of Public Libraries and the Public Lending Right,” *Available at SSRN 3082016*.
- KARGER, E. (2021): “The Long-Run Effect of Public Libraries on Children: Evidence from the Early 1900s,” .
- KLINENBERG, E. (2018): *Palaces for the people: How social infrastructure can help fight inequality, polarization, and the decline of civic life*, Crown.
- NAGARAJ, A. AND I. REIMERS (2021): “Digitization and the Market for Physical Works: Evidence from the Google Books Project,” *Available at SSRN 3339524*.
- SHILLER, B. R. (2013): “Digital distribution and the prohibition of resale markets for information goods,” *Quantitative Marketing and Economics*, 11, 403–435.
- SMITH, M. D. AND R. TELANG (2016): *Streaming, sharing, stealing: Big data and the future of entertainment*, MIT Press.
- VARIAN, H. R. (2000): “Buying, sharing and renting information goods,” *The Journal of Industrial Economics*, 48, 473–488.

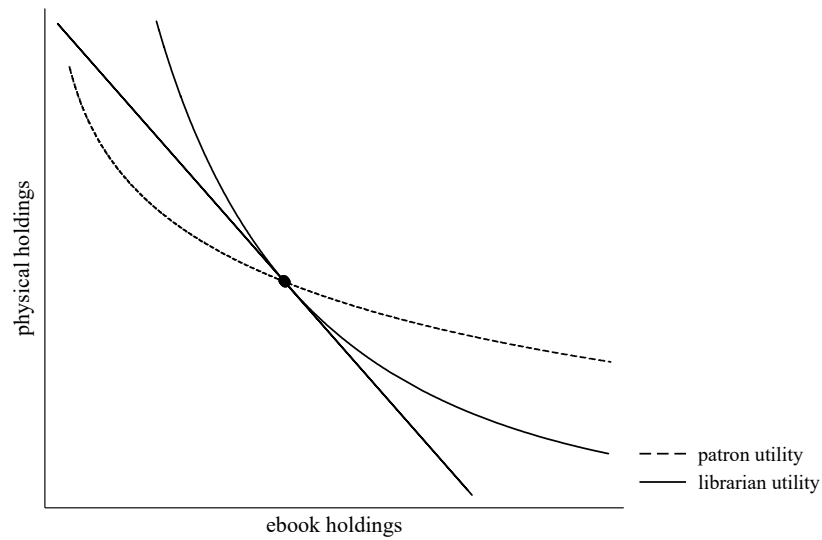
8 Figures and Tables

Figure 1: Library visits and circulation over time



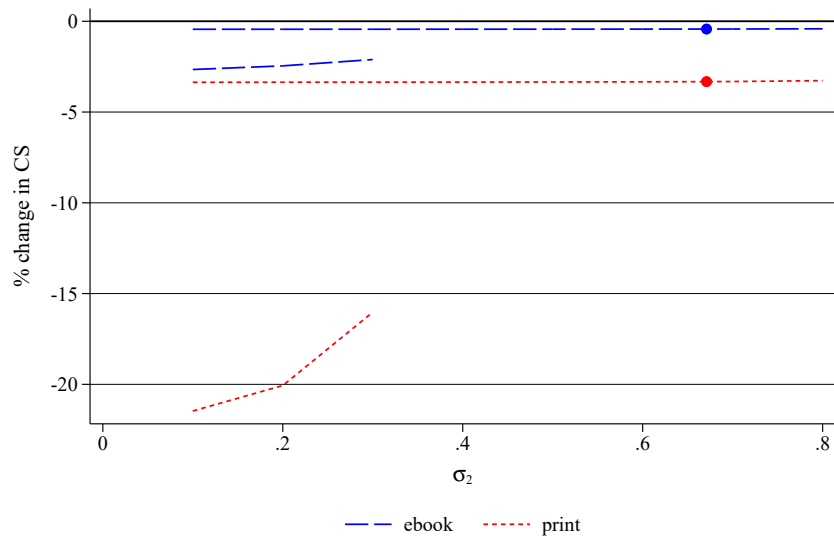
Notes: Data on circulation and visits from selected issues of the Digest of Education Statistics.

Figure 2: Librarian and patron utility and the holdings budget constraint



Notes: The straight line is a library's budget constraint for physical and electronic holdings. The solid curve is a librarian indifference curve over holdings. It reflects both the impact of these holdings on circulation as well as librarian regard for physical and electronic circulation. It is tangent to the budget constraint at the observed library holdings (dot). The dashed curve is the patron indifference curve over physical and electronic holdings (which, in turn, reflects the impacts of respective holdings on circulation) that passes through the librarian holding choice. Here, the patron and librarian indifference curves are drawn to reflect a lower relative patron preference for ebooks. More physical holdings would make patrons better off.

Figure 3: Librarian and patron utility and the holdings budget constraint



Notes: Effects of doubling library book prices on consumer surplus for varying substitution parameters σ_1 and σ_2 . The blue dashed lines describe percent changes in CS when doubling the price of ebooks, and the red dotted lines describe percent changes in CS when doubling the price of print books. The x-axis shows varying levels of σ_2 . The longer lines keep σ_1 constant at its estimated level ($\sigma_1 = 0.937$) and the shorter lines keep σ_1 constant at 0.5. Only libraries that solve to interior solutions for both counterfactuals and for all values of σ_1 and σ_2 are included.

Table 1: Summary statistics

	(1) Mean	SD	Median
<u>Ebooks</u>			
Holdings	73616.6	135825.4	30482.5
% change in holdings	26.9	86.2	14.8
Circulation	30284.4	191922.1	3284.0
Expenditure	44681.6	241931.1	2950.0
<u>Print books</u>			
Holdings	105570.6	405702.5	39284.5
% change in holdings	-0.8	10.4	0.2
Circulation	258391.5	952812.4	50681.0
Expenditure	96050.4	377787.0	23317.0
<u>Visits</u>			
Total	173109.4	584098.5	45238.5
% change in visits	-2.0	20.8	-1.3
Observations	50722		

Notes: Summary statistics at the library-year level, for all libraries from 2013 to 2019. Holdings and circulation describe the number of books of each format, and expenditure is in terms of (nominal) dollars.

Table 2: Summary statistics by income and urban/rural status

Panel A: Income quintile

	Lowest	2 nd	3 rd	4 th	Highest
Median income	19471.5	25678.9	30422.8	36805.0	53602.4
Population	23632.3	30948.2	42189.8	54933.1	79161.2
<u>Ebooks</u>					
Holdings	74610.4	74387.4	77562.0	78986.0	89695.0
Circulation	10250.3	14885.1	33691.5	42379.6	74797.6
<u>Print books</u>					
Holdings	69769.6	81563.7	110561.2	115154.4	220519.2
Circulation	97942.4	159784.3	245936.4	340321.9	636080.1
Total visits	82365.8	119965.1	172271.0	216545.3	400075.0
Observations	39212				

Panel B: Urban/rural

	Large city	Medium city	Town	Rural
Population	97276.0	57316.8	18841.0	9175.7
<u>Ebooks</u>				
Holdings	96609.0	76713.9	70434.1	61339.8
Circulation	82760.2	40832.2	9269.5	4637.4
<u>Print books</u>				
Holdings	237121.0	140696.8	61217.3	33585.2
Circulation	644008.1	379845.9	101915.9	48316.7
Total visits	414132.7	252863.2	83516.9	36629.1
Observations	50427			

Notes: Summary statistics at the library-year level. In the top panel, library systems are grouped by community median income. In the bottom panel library systems are grouped by urban status, where our first group includes metropolitan areas whose principal city has more than 250,000 inhabitants. Our second group includes metropolitan areas whose principal city has fewer than 250,000 inhabitants. The third group includes incorporated, or Census-designated, places outside of metropolitan areas and with populations of at least 2,500. The final group includes Census-designated “rural” places.

Table 3: Holdings and circulation

	OLS		FE		1st stage		Reduced form		IV	
	(1) Phys	(2) Ebook	(3) Phys	(4) Ebook	(5) Phys	(6) Ebook	(7) Phys	(8) Ebook	(9) Phys	(10) Ebook
phys holdings	1.651*** (0.165)	0.285*** (0.0260)	0.676*** (0.152)	-0.229** (0.0959)					0.675*** (0.222)	-0.0219 (0.135)
ebook holdings	0.540*** (0.0524)	0.132*** (0.0145)	-0.164*** (0.0240)	0.108*** (0.0156)					-0.0991** (0.0479)	0.0671** (0.0282)
phys jump					8799.0*** (618.0)		6233.4*** (1629.4)	-281.6 (1016.5)		
ebook jump						56880.8*** (2976.5)	-6826.7*** (2288.9)	3865.7*** (1249.0)		
Observations	50722	50722	50722	50722	20133	20133	20133	20133	20133	20133
$\overline{R^2}$	0.525	0.397	0.980	0.847	0.997	0.665	0.992	0.930		

Notes: The dependent variable in all columns except (5) and (6) is the total circulation of physical books (odd-numbered columns) or electronic books (even-numbered columns). In columns (5) and (6) the dependent variable is holdings. All specifications include year fixed effects. Columns (3) through (10) also include library fixed effects. Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table 4: Physical books, ebooks, and library visits

	(1) OLS	(2) FE	(3) Red form	(4) IV
phys holdings	1.194*** (0.0903)	0.248*** (0.0733)		0.387*** (0.114)
ebook holdings	0.165*** (0.0277)	-0.0815*** (0.00930)		-0.0552** (0.0249)
phys jump			3568.8*** (932.9)	
ebook jump			-3821.8*** (1251.5)	
Observations	50722	50722	20133	20133
$\overline{R^2}$	0.704	0.986	0.992	

Notes: The dependent variable in all columns is the number of visits to a library in a year. All specifications include year fixed effects. Columns (2) through (4) also include library fixed effects.

Table 5: Demand estimates

	(1)	(2)	(3)
	$\ln(s_j) - \ln(S_0)$	$\ln(s_{f \text{inside}})$	$\ln(s_j) - \ln(S_0)$
γ^{ebook}	-0.0450*** (0.00351)	-2.696*** (0.00433)	-0.888*** (0.0894)
σ_2	0.983*** (0.00117)		0.671*** (0.0331)
σ_1	0.995*** (0.000760)		0.937*** (0.00622)
$\ln(N_f/N)$		0.194*** (0.00212)	
Observations	100846	100846	100846
$\overline{R^2}$	0.985	0.829	0.960

Notes: Results from nested logit demand estimations. The dependent variable in columns (1) and (3) is $\ln(s_j) - \ln(S_0)$. Column (2) reports a first stage. All specifications include library and year fixed effects. In column (3) the group share (Q^f/Q) is instrumented using the number of products in the format relative to the number of products overall (N^f/N). Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table 6: Counterfactual prices and % changes in holdings, circulation, and CS

<u>raise:</u>	raise 10%		raise 100%	
	P^e	P^p	P^e	P^p
print holdings	0.73%	-9.74%	5.03%	-52.75%
ebook holdings	-10.82%	1.84%	-56.53%	14.28%
print circulation	0.25%	-0.73%	1.72%	-5.36%
ebook circulation	-2.04%	1.60%	-14.00%	12.08%
total circulation	-0.05%	-0.43%	-0.34%	-3.07%
visits	0.39%	-2.36%	2.31%	-12.85%
librarian utility	-0.13%	-0.35%	-0.91%	-2.44%
CS	-0.05%	-0.48%	-0.37%	-3.41%
Δ CS ratio	8.94		9.15	

Notes: Effect of various changes in electronic or physical book prices on holdings, circulation, and CS. All values are percentage changes from the baseline and based on initial ebook prices related to physical book prices via $P^E = 2/3P^P$. We include all libraries that solved to their observed values for 2018 and solved to interior solutions (positive ebook and physical holdings) given all counterfactual prices.

Table 7: Counterfactual results by income and urban/rural status

Panel A: Income quintile

	Lowest		2 nd		3 rd		4 th		Highest	
<u>double:</u>	<i>P^e</i>	<i>P^p</i>	<i>P^e</i>	<i>P^p</i>	<i>P^e</i>	<i>P^p</i>	<i>P^e</i>	<i>P^p</i>	<i>P^e</i>	<i>P^p</i>
visits	4.67%	-18.33%	3.28%	-14.79%	2.19%	-12.99%	2.00%	-11.39%	1.67%	-11.33%
CS	-0.23%	-3.02%	-0.22%	-2.96%	-0.35%	-2.86%	-0.38%	-2.69%	-0.35%	-2.72%
Δ CS ratio	13.23		13.70		8.20		7.02		7.82	

Panel B: Urban/rural

	Large city		Medium city		Town		Rural	
<u>double:</u>	<i>P^e</i>	<i>P^p</i>	<i>P^e</i>	<i>P^p</i>	<i>P^e</i>	<i>P^p</i>	<i>P^e</i>	<i>P^p</i>
visits	1.62%	-11.74%	2.13%	-11.86%	4.41%	-15.97%	8.40%	-21.75%
CS	-0.45%	-3.31%	-0.30%	-3.49%	-0.26%	-3.56%	-0.29%	-3.49%
Δ CS ratio	7.30		11.75		13.62		12.16	

Notes: Effect of various changes in electronic or physical book availability on visits and CS, for groups of libraries based on median patron income (panel A) and urban status (Panel B). All values are percentage changes from the baseline and based on initial ebook prices calculated as the average expenditure / holdings for the respective library groups reported in Table 2. We include all libraries that solved to their observed values for 2018 and solved to interior solutions (positive ebook and physical holdings) given all counterfactual prices.

A Appendix

This appendix discusses two aspects of the model solution, constancy of the assumed relationship between P^E and P^P , and corner solutions.

A.1 The assumed relationship between P^E and P^P

In our baseline approach, we solve the model using a single relationship between P^E and P^P , in particular that $P^E = 2/3P^P$ for all libraries. This is correct on average, but given the information in Table 2, one can note that the relationship varies across communities according to income and rural/urban status. Accordingly, for the “heterogeneity” simulations, we solve the model using group-specific relationships $P^E = \kappa P^P$. For income quintiles, κ ranges from 0.31 for the lowest income quintile to 1.31 for the highest. In the urban/rural simulation, κ varies between 0.14 for the rural areas and 1.31 for the most urban areas.

In each of these simulations, the choice of the price relationship determines the θ estimates showing librarians’ implied weights on electronic, relative to physical, circulation. Allowing the price relationship to vary across groups of library systems gives rise to both less variation in θ and a lower distribution of values, than in the baseline. While the interquartile range for θ runs from 1.28 to 2.69 in the baseline, the range is between 1.10 and 2.13 in the income simulation and between 1.04 and 2.25 in the urban/rural simulation.

Although prices and implied θ parameters vary across solutions, the results of interest do not vary. In particular, the ratio of the change in CS from a doubling in physical, relative to electronic, book prices is 9.15 in the baseline; and the estimated ratios are 8.29 and 8.92, respectively, in the income and urban/rural simulations. Hence, our baseline results are not driven by the common assumed relationship between P^E and P^P .

A.2 Corner solutions

In our baseline approach (with a common price relationship), the model solves to the observed values of holdings in 94.3 percent of library systems. Of the libraries that solve to baseline observed holdings, 0.27 percent solve to corners (with zero ebook holdings) when doubling the ebook prices, while doubling the physical book price leads to zero physical book corners in 15.92 percent of libraries. For the baseline approach, we exclude all systems that a) do not solve to the observed values in the baseline, and that b) solve to corners with price doubling in either format. This excludes 16.19 percent of library systems.

When we include the systems solving to corners, the results are somewhat larger: for example, rather than reducing CS by 3.41 percent, a doubling of P^P reduces CS by 4.79 percent. As a result, the Δ CS ratio is 9.59 including corners and 9.15 without them.