Cross-channel competition and complementarities in US retail*

Hiroki Saruya Michael Sullivan

Department of Economics, Yale University

June 3, 2023

Click here for the latest version and here for the online appendix.

Abstract

We estimate net effects of offline stores on online spending using US internet-activity and store-location data. These effects are ambiguous in sign: the effect of a retailer's offline stores on its online sales depends on opposing cannibalization and cross-channel complementarity effects. Similarly, effects of offline stores on rivals' online sales depend on opposing business-stealing and showrooming effects. We find that a consumer's spending at multichannel retailer's online store falls (1.1–3.8%, on average) when a rival adds a nearby storefront but rises (7.1–32.3%) when the retailer opens its own storefront. Offline stores often boost Amazon's sales, suggesting showrooming's relevance.

*Email addresses: `hiroki.saruya@yale.edu` and `m.r.sullivan@yale.edu`. We thank Katja Seim and participants at the Yale School of Management Quantitative Marketing internal seminar series for helpful feedback on this project. Michael Sullivan is supported in part by funding from the Social Sciences and Humanities Research Council.

1 Introduction

Contemporary retail markets feature competition between retailers that operate exclusively offline, exclusively online, and through both offline and online channels. The manner in which the offline retail environment affects consumers' online shopping determines, in part, the merits of retailers' channel choices, the extent to which online and offline retail constitute a common market, and the role of offline stores in the digital age. The relationship between offline stores and online sales is multi-dimensional: offline stores could present heterogeneous benefits and harms to online stores across different pairs of offline and online retailers and across different types of products.

Our study estimates effects of offline stores on online spending across retailers and retailing categories. These effects have several components. To begin, a multichannel retailer's offline store may lower that retailer's online sales on account of *cannibalization*. This weakens the retailer's incentive to open brick-and-mortar stores. But various effects that we collectively term *cross-channel complementarities* may offset cannibalization. First, an offline store may boost awareness of—or public opinion toward—the retailer. An offline store may also complement the retailer's online store by offering in-person customer service, accepting in-person returns of items purchased online, by offering pick-up of products purchased online, or by strengthening the retailer's logistical operations — e.g., retailer may operate distributional centres near their offline stores, and these centres may reduce online-order delivery times in their vicinity.

Rival offline stores may also affect a retailer's online sales. This effect may be negative due to standard *competitive effects*, but rival offline stores may also boost a retailer's online sales on account of *showrooming effects*, i.e., benefits that a retailer derives from informative or promotional services offered by rival retailers. To illustrate, a bookstore may invest in informative services by installing product displays, allowing visitors to read unpurchased books in the store, and training staff to educate visitors about books. This investment benefits online book retailers whose product offerings overlap with those of the offline retailer. Online retailers that contain their costs by offering minimal informative services may freeride on an offline store's provision of such services while undercutting this store on price. This behaviour discourages the offline retailer from offering informative services to begin with. Showrooming stems from informative services relevant to products sold by multiple competing retailers (e.g., a store's fitting services for shoes sold by many retailers) rather than from services narrowly affecting a retailer's exclusive product line. The empirical relevance of showrooming is of consequence for retail strategy and policy. Showrooming encourages retailers to develop exclusive product lines for which they can provide informative services that do not apply to products sold by potential freeriders. Showrooming also motivates minimum resale price maintenance, wherein a manufacturer prohibits resale of its products below a certain price and thereby encourages retailers to invest in informative services without fear of being undercut on price.

To estimate effects of offline stores on online spending, we combine a panel of online transactions by US consumers in 2007–2018 with data characterizing the universe of US business locations. We present empirical relationships between the offline stores nearby consumers and online spending, but these relationships do not represent causal effects of the former on the latter. This is because unobservable consumer, retailer, and market characteristics induce spurious correlations between the presence of offline stores and online spending. The foremost endogeneity problem is that the geographical distribution of consumer tastes determines both where retailers open offline stores and patterns of local online shopping, a concern that we call the *location-taste problem.* Bookstores, for example, may choose to open locations near consumers who enjoy reading. Avid readers may also choose to live near bookstores. Either sort of location choice induces a correlation between store locations and consumers' online spending. Our approach to overcoming the location-taste problem involves (i) using a rich set of consumer characteristics to proxy for unobserved tastes and (ii) modelling region-level unobservables using a combination of fixed effects and the local demographic profile, which controls for unobserved shopping tastes of regions' residents. Although the relationship between the offline retail environment and online spending has implications for the structure of retail markets, we leave the study of these implications to future research.

Our first main finding is that rival offline stores generally reduce spending at a retailer's online store. We summarize the effects of rival offline stores on a retailer's online sales using the average percentage change in spending at the retailer's store when a rival store opens within 20km of the consumer. In 2007–2008, these rival effects range from 2.1% to -3.0% across retailing categories when Amazon is included in the average and from -1.1% to -3.8% when Amazon is excluded. In general, Amazon's sales are less negatively affected by rival offline stores than are multichannel retailers' online sales. This suggests showrooming effects: Amazon sells products at prices below those of its multichannel rivals, which could reflect a cost advantage from not offering offline informative services, while freeriding on these services. Offline stores could thereby boost Amazon's sales. This effect does not apply to other multichannel retailers if they do not charge lower prices for online purchases than for in-store purchases. We find that offline bookstores have an especially large positive effect on book sales on Amazon, which aligns with our expectation that the books category is especially prone to showrooming.

A robust finding is that a multichannel retailer's own offline stores boost its online sales. We measure effects of a multichannel retailer's own offline stores on its online sales as the average percentage change in spending at the retailer's online store from placing an additional one of the retailer's own offline stores in the consumer's vicinity. The estimated measures for 2007–2008 range from 7.1% to 32.3% across categories. Results for 2017–2018 are qualitatively similar to those for 2007–2008, but they are less precise in part because of the decreased coverage of our data in the 2017–2018 time period.

1.1 Related literature

We join a literature analyzing the relationship between offline retail and online sales. Earlier studies on cross-channel competition (e.g., Goolsbee 2001, Sinai and Waldfogel 2004, Forman et al. 2009, Brynjolfsson et al. 2009) document evidence for channel substitution while not distinguishing multichannel retailers from single-channel retailers.[1]

[1] Prince [2007] measures the elasticity of demand for computers at online retailers with respect to offline price and argues that the cross-price elasticity increased following the rise in multichannel operations.

More recent papers study substitution between a particular retailer's online and offline retail channels in the context of apparel and home furnishings (Avery et al. 2012, Wang and Goldfarb 2017, Shriver and Bollinger 2022), eyewear (Bell et al. 2018), and groceries (Chintagunta et al. 2012, Pozzi 2013). Our study complements those listed above by analyzing heterogeneity across multiple online stores and product categories. To the best of our knowledge, our work is the first to empirically document heterogeneous effects of offline stores on the sales of own and rival online stores. Our study also complements studies on the effect of online channel on offline sales (e.g., Pozzi 2013, Huang et al. 2023).

Our study also relates to a wider literature on the rise of e-commerce. Examples of articles on online retail's evolution and welfare consequences include Hortaçsu and Syverson 2015, Dolfen et al. 2019, Quan and Williams 2018, and Edgel et al. 2023. Another literature to which our article relates is that on showrooming (e.g., Shin [2007]). Studies on showrooming including Jing [2018], Kuksov and Liao [2018], and Mehra et al. [2018] emphasize the effect of showrooming on offline stores' profits rather than on online sales. Carlton and Chevalier [2001] find evidence that manufacturers internalize freeriding by online retailers in their distribution and pricing strategies. More recently, Goetz et al. [2020] find that bookstore closures in Germany in the 2010s were associated with decreases in overall book sales. Our study complements Goetz et al. [2020] by comparing showrooming effects across categories. See MacKay and Smith [2014] for a discussion of minimum resale price maintenance, which is often rationalized by appeal to showrooming, and for empirical evidence on the effects of resale price maintenance.

2 Data

Our primary data sources are the Comscore Web Behavior Database and the Data Axle business locations database. The Comscore data provide online browsing and transactions records for a panel of US web users in 2007–2008 and 2017–2018. Because there is limited overlap in the data's web users across years, we define a panelist as an individual/year pair. The data feature 147,852 panelists in 2007–2008 and 172,615 panelists in 2017–2018. Variables include panelist characteristics, descriptors of panelists' website visits, and descriptors of panelists' online transactions.[2]

We limit our attention to large cross-category retailers and specialized retailing categories that are well represented in our sample. The large cross-category retailers that we study are Walmart, Costco, Target, and Amazon, and the specialized categories we analyze are books, office supplies, and electronics. Within each category, we analyze retailer-specific sales and store counts for a few large retailers. Table 1 lists these large offline retailers, which we chose based on a consideration of national store counts (see Online Appendix O.1 for details). We analyze offline stores not listed in Table 1 as a grouping of "other" stores within each category. The online stores included in our analysis are Amazon and the online stores associated with each large offline retailer.[3] We only include sales within the product category in question in our analysis — when studying electronics, for instance, we do not include products other than electronics (e.g., computer bags). Table 2 describes our category-specific transactions data for 2007–2008.

We focus on 2007–2008 because the panel's coverage of transactions is higher for that period than for 2017–2018.[4] Coverage may have fallen because internet usage shifted from personal computers, which Comscore tracks, to smartphones and tablets that are not covered by the data. Coverage reductions lower the number of transactions to analyze and consequently reduce precision of estimates obtained using the data. Our findings for 2017–2018 are qualitatively similar to but less precise than those for 2007–2008.

Although our main interest is in online spending, we also construct variables describing web browsing that we use as controls. These panelist-level variables provide the number of times in a year that the panelist visits a website in each of several categories, includ-

[2]See Online Appendix O.1 for a discussion of the Comscore data's representativeness.

[3]In the books category, we do not analyze online stores for Borders and Waldenbooks because our data include no online sales for these retailers in 2007. Similarly, we exclude `booksamillion.com` from our analysis of the 2017–2018 data because these data include no transactions at `booksamillion.com`. We exclude `radioshack.com` from our analysis of the 2017–2018 period for the same reason. Office Depot ceased operating the `officemax.com` online store following its merger with Office Max. This explains why `officemax.com` does not appear in our analysis of the 2017–2018 time period.

[4]See Online Appendix O.1 for evidence of this claim.

Table 1: Large offline retailers by retailing category

Category	Retailers	
	2007–2008	2017–2018
Cross-category	Walmart Target Costco	Walmart Target Costco
Books	Barnes & Noble Books-a-Million Waldenbooks Borders	Barnes & Noble Books-a-Million
Office supplies	Staples Office Depot Office Max	Staples Office Depot Office Max
Electronics	Best Buy Circuit City Radio Shack Apple	Best Buy Radio Shack Apple

Note: Borders, Waldenbooks, and Circuit City each closed all of their brick-and-mortar locations between 2007–2008 and 2017–2018. We therefore exclude these retailers from our analysis of 2017–2018.

Table 2: Summary of consumer panel, 2007–2008

(a) Cross-category retailers

Store	Avg. spend (all)	Positive spending (%)	Avg. spend (pos.)
Amazon	18.51	14.32	129
Costco	3.27	0.54	607
Target	3.84	3.47	111
Walmart	7.16	6.07	118

(b) Bookstores

Store	Avg. spend (all)	Positive spending (%)	Avg. spend (pos.)
Amazon	6.90	8.28	83
Barnes & Noble	1.08	1.83	59
Books-a-Million	0.07	0.13	55

(c) Electronics

Store	Avg. spend (all)	Positive spending (%)	Avg. spend (pos.)
Amazon	4.12	1.83	226
Best Buy	2.75	0.88	311
Circuit City	2.47	0.76	323

(d) Office supplies

Store	Avg. spend (all)	Positive spending (%)	Avg. spend (pos.)
Amazon	0.08	0.10	83
Office Depot	4.36	0.57	768
Office Max	0.39	0.11	350
Staples	5.47	0.84	653

Note: "Avg. spend (all)" reports the mean dollar amount spent at the store across panelists. "Positive spending (%)" reports the share of panelists who spend a positive amount at the indicated store. "Avg. spend (all)" reports the mean dollar amount spent at the store among panelists who make at least one purchase. For panels (b) onward, only Amazons' transactions within the indicated category are included in the analysis. See Table O.5 in the Online Appendix for a version of the table with additional statistics. Table O.6 in the Online Appendix describes panelist spending in 2017–2018.

ing: adult, advert, career, dating, directory, downloads, finance, gaming, government, information, internet/wireless services, malware, media, news, portal, retail, social media, sports, travel, video, weather, and web service.[5] Table O.9 in the Online Appendix reports descriptive statistics for these variables.

We control for panelist characteristics in our primary analysis. These characteristics are: an indicator for household income exceeding $75,000; indicators for the head-of-household's race (white, black, and other); indicators for the head-of-household being less than 40 and between 40 and 54 years old; household size; an indicator for the presence of children in the household; an indicator for the head-of-household being Hispanic; an indicator for broadband internet access; and an indicator for the head-of-household having graduated from college. We additionally construct measures of the demographic profiles of the areas surrounding panelists. These measures are averages of the characteristics enumerated above among panelists within 20km of the panelist in question.[6]

Our other data source is Data Axle, whose database reports the locations of the universe of US businesses at an annual frequency. We use these data to compute, for each Comscore panelist and each retailer, (i) the number of a retailer's locations within 20km of the panelist and (ii) the minimum distance from the panelist to one of the retailer's locations. Table 3 describes these variables for 2007–2008.

3 Descriptive evidence

This section reports empirical relationships between offline stores and online shopping. We argue that the location-taste problem procludes causal interpretations of these relationships, but that controlling for measures of internet browsing and for the demographic profile of consumers' neighbourhoods mitigates the problem.

We assess empirical relationships between offline stores and online spending using Nadaraya-

[5]See Online Appendix O.1 for details on our procedure for categorizing websites.

[6]Recall that the Comscore data reports each panelist's ZIP code of residence; we compute the measures described by the preceding sentence by averaging over Comscore panelists living in a ZIP code tabulation area whose centroid is within 20km of that in which the focal panelist resides.

Table 3: Description of offline retail presence variables (2007–2008)

Table 4: 2007–2008

Category	Store	# stores (20km)		Min. distance		# stores
		Mean	Median	Mean	Median	(2007)
Cross-category	Costco	2.34	0	75.13	25.77	374
Cross-category	Target	7.48	4	33.41	7.44	1446
Cross-category	Walmart	8.33	6	8.66	5.34	3411
Books	Barnes	5.66	2	26.37	11.10	832
Books	Books-a-Million	0.51	0	474.97	134.02	178
Books	Borders	4.35	2	45.23	14.29	660
Electronics	Best Buy	4.95	2	23.75	9.59	851
Electronics	Circuit City	4.16	2	33.75	10.79	685
Electronics	Radio Shack	24.41	12	7.01	3.27	5095
Office Supplies	Office Depot	7.11	4	26.09	9.31	1262
Office Supplies	Office Max	4.73	2	29.18	12.19	982
Office Supplies	Staples	10.63	2	41.76	9.05	1486

Notes: see Table O.7 in the online appendix for a version of the table with additional quantiles and stores, and with figures for 2017–2018.

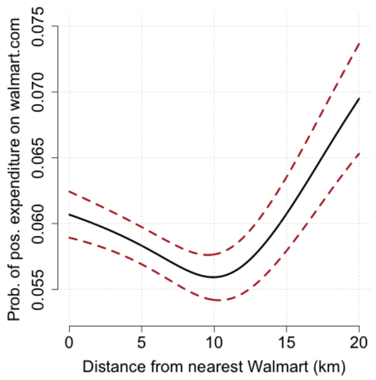
Watson kernel regressions. These regressions estimate the conditional expectation functions $m_{sj}(d_{ij}) = \mathbb{E}[y_{is} \mid d_{ij}]$, where y_{is} is an indicator for the consumer making a transaction at store s and d_{ij} is consumer i's distance from a location of chain retailer j.[7] Figure 1 provides a subset of our results for 2007–2008.[8] First, Figure 1a displays results for a regression of spending at `walmart.com` on the consumer's distance from the nearest brick-and-mortar Walmart store. Among panelists within 10km of a store, panelists who are further away from a store tend to spend less at `walmart.com`. However, we cannot conclude that Walmart storefronts do not cannibalize its online sales, because the relationship plotted in this figure reflects the location-taste problem: Walmart may open stores in areas where consumers enjoy shopping at Walmart through both of the retailer's channels. Although about 75% of panelists live within 10km of a Walmart location (see Table O.7), the fact that our estimated relationship between distance and spending becomes upward sloping after 10km suggests the presence of effects of varying signs that mediate the overall empirical relationship between distance and spending.

Our next regression evaluates the empirical relationship between spending at `walmart.com`

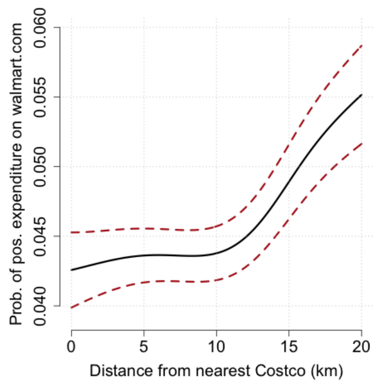
[7]We use a Gaussian kernel with a bandwidth minimizing the sum of squared prediction errors in leave-one-out cross validation.

[8]Online Appendix O.3 provides results for additional pairs of stores and for additional categories.

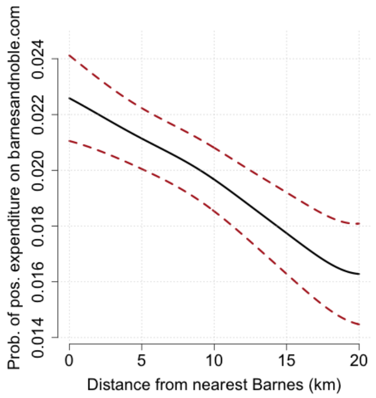
Figure 1: Selected regressions of online spending on distance from retailer (2007–2008)



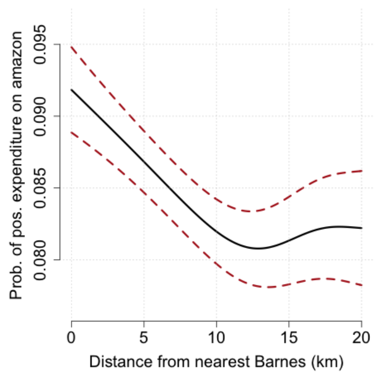
(a) Probability of transaction on `walmart.com` by distance from Walmart



(b) Probability of transaction on `walmart.com` by distance from Costco



(c) Probability of transaction on `barnesandnoble.com` by distance from Barnes & Noble



(d) Probability of transaction on Amazon by distance from Barnes & Noble

Notes: the dotted bands indicate 95% pointwise confidence intervals around our estimates as constructed using the $\hat{v}_{n,1}(x)$ asymptotic variance estimator analyzed in Chu et al. [2020].

and distance from an offline Costco store. Figure 1b displays the result. Consumers further away from a Costco are more likely to purchase from `walmart.com`, which suggests a negative competitive effect of Costco stores on Walmart's online sales. The estimated relationship does not reflect a competitive effect alone, however, but also possibly (i) the location-taste problem and (ii) showrooming effects.

Last, Figures 1c and 1d plot results from regressions of spending on books at Barnes & Noble's online store and at Amazon on distance from a Barnes & Noble store. These figures show a negative relationship between distance from a Barnes & Noble store and each of (i) own and (ii) rival online sales; that is, Amazon boasts higher book sales among consumers nearby a Barnes & Noble store. This contrasts with the relationship plotted in Figure 1b, which shows that consumers living further from a Costco store are more likely to shop at Walmart's online store. The negative relationship of Figure 1d could owe to a strong showrooming effect or the location-taste problem: Barnes & Noble may open offline stores near consumers predisposed to purchase books online.

The empirical relationships discussed above admit various interpretations regarding the contributions of the effects of interest—namely, cross-channel complementarities, showrooming effects, and competitive effects—and the location-taste problem. Inference about the effects of interest therefore requires an alternative approach. Our approach involves controlling for two sets of variables that proxy for unobserved consumer tastes. The first set of variables, which describe the consumer's internet usage, proxy for consumer interests that are not directly related to the consumer's place of residence. The second set of variables describe the consumer's place of residence, and thus capture place-dependent tastes driving both retailer entry decisions and consumer online shopping decisions.

The first set of control variables are the internet usage variables described in Section 2. The idea underlying the use of these controls is that consumers reveal aspects of their interests, aesthetic preferences, and personalities through their choices of visited websites. To fix ideas, consider the location-taste problem that arises when bookstores open in areas with intellectually inclined populations. Intellectual inclination is a component of unobserved (to the econometrician) tastes that may reflect itself in a consumer's web browsing behaviour, e.g., through visits to informational websites. Thus, controlling for variables characterizing internet usage will at least partially control for unobserved tastes including intellectual inclination, thereby allaying the location-taste problem.

To assess the scope for our internet usage variables to address the location-taste problem, we regress indicators for whether a panelist bought a book online on the number of offline bookstores within 20km, the internet usage variables, and consumer characteristics. Table 5 contains the results. The internet usage variables predict online book spending in reasonable ways: visits to websites in the information and news categories—i.e., sites offering informative texts to read, just as books do—predict online book shopping.

Additionally, websites in the adult and gaming categories, which offer relatively little informative and literary content, are associated with less online book purchasing. Table 5 also reports a large positive shift in the estimated coefficient for the number of stores when the internet usage variables are omitted. This suggests that including internet usage variables removes upward bias in the estimates of offline stores' effects on online spending stemming from the fact that bookstores locate near people who like buying books.

Table 5: Regression of online book shopping on nearby stores, internet usage (2007–2008)

		Controls included		Controls excluded	
Variable type	Note	Estimate	t-statistic	Estimate	t-statistic
N. stores (log-transformed)		1.10	6.96	1.18	7.24
N. visits (hundreds)	Adult	-1.10	-10.92		
N. visits (hundreds)	Advert	-1.17	-7.73		
N. visits (hundreds)	Career	-1.11	-2.82		
N. visits (hundreds)	Finance	2.32	19.84		
N. visits (hundreds)	Gaming	-0.44	-7.44		
N. visits (hundreds)	Government	1.12	4.44		
N. visits (hundreds)	Info	4.89	28.46		
N. visits (hundreds)	Malware	-0.33	-7.80		
N. visits (hundreds)	Media	0.40	4.33		
N. visits (hundreds)	Other	0.02	4.47		
N. visits (hundreds)	Portal	0.21	16.75		
N. visits (hundreds)	Retail	3.01	48.34		
N. visits (hundreds)	Social Media	-0.11	-3.65		
N. visits (hundreds)	Video	-0.70	-6.55		
N. visits (hundreds)	Weather	0.12	1.67		
N. visits (hundreds)	Webservice	-0.49	-8.67		
N. visits (hundreds)	Dating	-0.20	-0.93		
N. visits (hundreds)	Internet Wireless	0.51	6.75		
N. visits (hundreds)	News	1.15	13.81		
N. visits (hundreds)	Sports	-0.22	-1.87		
N. visits (hundreds)	Travel	5.53	18.08		
N. visits (hundreds)	Downloads	-0.55	-4.75		
N. visits (hundreds)	Directory	10.02	2.43		
	R^2	0.081		0.012	

Note: The dependent variable is an indicator for whether the panelist ever purchased a book online. "N. stores (log-transformed)" is the number of stores within 20km transformed by $x \mapsto \log(x+1)$. Both regressions also include the panelist characterstics listed in Section 2 and year fixed effects.

Observed consumer characteristics and internet usage variables may not perfectly proxy for the unobserved tastes that give rise to the location-taste problem. Tastes that these variables may fail to capture are those that correlate with the consumer's place of residence. The place that a consumer lives, like the consumer's web browsing behaviour, is an expression of that consumer's tastes. Book lovers, e.g., may prefer to live in regions with highly educated neighbours because they value interactions with such people.[9] We account for region-level unobserved tastes by controlling for the local demographic profile and using regional fixed effects. Controlling for the local demographic profile mitigates the location-taste problem because local demographics explain both online shopping behaviour and store location choice, which implies that a failure to control for the local demographic profile in a regression of online spending on offline stores introduces omitted variable bias. To evince the dependence of online spending on local demographics, we regress spending at each large cross-channel multichannel retailer's online store on panelist characteristics, counts of offline stores within 20km, and the share of the population within 20km of the consumer with household income exceeding \$75,000.[10] Table 6 provides the results, which establish that consumers in higher-income areas spend significantly more at Costco's online store, moderately more at Target's, and less at Walmart's. This is consistent with Costco appealing to consumers of higher socioeconomic status, who are more likely to live in high-income areas condtional on their own income.

Local demographics also correlate with offline store counts, which implies that omitting local demographics from regressions of online spending on offline store counts biases estimates of offline stores' effects. Table 7 reports results from a regression of counts of offline stores within 20km of a consumer on the shares of the population within 20km of the consumer in various demographic groups. The demographic measures predict offline stores counts; the R^2s of the Costco and Target regressions exceed 0.20. Additionally, each retailer has more offline stores in places with more people who have higher incomes.

[9] Preferences may also depend on a consumer's neighbours due to *contextual network effects*, i.e., effects on an individual's behaviour of other individuals' characteristics (see Jullien et al. [2021]). Such effects are relveant when consumption is driven by a desire to impress or fit in with one's peers.

[10] We separately analyze 2007–2008 and 2017–2018. To limit the influence of outliers, we trim observations for which the spending variable exceeding its 98th percentile conditional on positive spending.

Table 6: Regressions of online spending on high-income share (2007–2008)

	Spending		
	costco.com	target.com	walmart.com
	(1)	(2)	(3)
N. stores (log-transformed)	2.551***	0.040	−1.319***
	(0.159)	(0.059)	(0.113)
High income	0.659*	0.620***	−0.237
	(0.352)	(0.165)	(0.248)
High income (average)	2.667***	0.632**	−0.769*
	(0.638)	(0.300)	(0.447)
Mean dep. var.	2.51	3.19	5.70
Observations	147,836	147,749	147,673

Note: "N. stores (log-transformed)" is the log of one plus the number of the retailer's offline stores within 20km. "High income (average)" is the share of people within 20km that have household incomes exceeding $75,000. We include year fixed effects and the consumer characteristics listed in Section 2. (omitted from the table). See Table O.8 in the Online Appendix for results for 2017–2018.

Table 7: Dependence of offline retail environment on local demographics (2007–2008)

	Costco	Target	Walmart
	(1)	(2)	(3)
High income (average)	0.793***	1.015***	0.536***
	(0.008)	(0.011)	(0.009)
Observations	147,852	147,852	147,852
R^2	0.223	0.225	0.137

Notes: the table reports results from a panelist-level regression of the number of a retailer's stores within 20km of a panelist on variables characterizing the demographic profile of the the region within 20km of the panelist's ZIP code of residence. The measures of the demographic profile included are: share of population with household income exceeding $75,000 ("High income (average)"); the share of the population in white and black racial groups; the share of the population under the age of 40 and between the ages of 40 and 54; the average household size; the share with a child in the household; the share that is Hispanic; the share with broadband internet; and the share having graduated from college. All estimates except that for "High income (average)" are omitted from the table.

4 Model of online shopping

As argued in Section 3, empirical relationships between offline stores and online spending do not necessarily represent causal relationships. We formalize this argument in the context of the model on which we base our analysis. This model is described by

$$y_i = h(n_i)'\alpha + z_i'\beta + \omega_i. \tag{1}$$

In (1), y_i is a measure of consumer i's online spending, n_i is a vector of counts of offline stores nearby consumer i, z_i are consumer characteristics, and the function h allows for transformations of store counts n_i. Last, ω_i is an unobservable shifter of spending that depends on unobserved consumer tastes for online shopping.

The unobservable ω_i generates the location-taste problem: it includes unobserved tastes that influence online shopping, and the geographical distribution of these tastes may also affect retailers' location decisions. We model ω_i as a function of (i) unobserved tastes that do not directly relate to place of residence, (ii) region-level tastes, and (iii) a purely idiosyncratic determinant of spending:

$$\omega_i = \xi_i'\psi + \rho_{r(i)} + v_i. \tag{2}$$

The $\xi_i'\psi$ term represents the contribution of taste characteristics that do not directly relate to the consumer's region — examples of ξ_i components include enjoyment of reading, personality traits, and aesthetic preferences. The $\rho_{r(i)}$ term represents tastes specific to the consumer's region $r(i)$. Last, v_i captures disturbances to spending that do not relate to persistent tastes for online shopping — it may reflect, e.g., the timing of consumption needs (e.g., the consumer seeks to replace a broken laptop) or transient shocks to liquidity (e.g., unexpected bills or salary bonuses).

Proxy approach. We address the location-taste problem arising from ξ_i using proxy variables q_i, which we assume depend on unobserved and observed characteristics ξ_i and z_i:

$$q_i = \Pi\xi_i + \Lambda z_i + \eta_i,$$

where η_i is independent of all else. The q_i proxies are thus noisy measurements of ξ_i (conditional on z_i); the assumption that η_i is independent of all else amounts to an assumption of classical measurement error. When Π has full column rank,

$$\xi_i = (\Pi'\Pi)^{-1}\Pi'(q_i - \Lambda z_i) + \tilde{\eta}_i, \tag{3}$$

where $\tilde{\eta}_i = -(\Pi'\Pi)^{-1}\Pi'\eta_i$. Substituting (3) and (2) into (1) yields

$$\begin{aligned} y_i &= h(n_i)'\alpha + z_i'\beta + (q_i - \Lambda z_i)'\Pi(\Pi'\Pi)^{-1}\psi + \tilde{\eta}_i'\psi + \rho_{r(i)} + v_i \\ &= h(n_i)'\alpha + q_i'\gamma + z_i'\tilde{\beta} + \rho_{r(i)} + \tilde{v}_i \end{aligned} \tag{4}$$

for composite parameters $\tilde{\beta} = \beta - \Lambda'\Pi(\Pi'\Pi)^{-1}\psi$ and $\gamma = \Pi(\Pi'\Pi)^{-1}\psi$ and unobservable $\tilde{v}_i = \tilde{\eta}_i'\psi + v_i$. Controlling for q_i is thus similar to directly controlling for ξ_i. Our method to control for ξ_i resembles the proxy (or replacement function) approach (Heckman and Vytlacil 2007) used in estimating production functions (Olley and Pakes 1996, Levinsohn and Petrin 2003, Ackerberg et al. 2015, and Gandhi et al. 2020).

Regional tastes. We now turn to the region-level $\rho_{r(i)}$ unobservables. These unobservables capture differences in tastes across neighbourhoods, which arise when the consumer's choice of neighbourhood correlates with their tastes for online shopping. Conditional on the consumer's own income, for example, a consumer living in a high-income neighbourhood may differ from a consumer living in a low-income area in terms of socioeconomic status and consequently shopping tastes. Differences in tastes across neighbourhoods also arise when consumer tastes are influenced by their neighbours. One solution to the consequent location-taste problem is to specify region fixed effects under the assumption that local taste disturbances do not vary within area known to the researcher. But this approach limits our capacity to use cross-region variation in estimating the effects of n_i on y_i. This limits estimation precision. Indeed, our estimates from regressions with fixed effects for finely defined regions are imprecise, although some results are qualitative similar to our main findings (see Online Appendix O.7). Also, if measurement error accounts for a large portion of time-series variation in store counts, then relying on this variation significantly biases estimates of offline stores' effects. Recall that we observe business locations annually for two years — store counts typically do not change much year-to-year, and hence measurement error may be large relative to true variation in counts. The data's annual frequency also implies measurement error from time aggregation. The possibility of bias from measurement error underlies a common argument against the fixed effects approach to production function estimation (see, e.g., Ackerberg et al. 2007).

We address the endogeneity problem caused by region-level taste unobservables by modelling their dependence on local demographics:

$$\rho_r = \rho_r^{\mathrm{FE}} + g(w_r) \tag{5}$$

where w_r are average demographic characteristics in the region r, and ρ_r^{FE} is a region fixed effect. Note that the regions r need not be the same regions across which the fixed effects vary — in practice, the ρ_r^{FE} vary only across census regions. We specify ρ_r^{FE} to vary only across coarsely defined regions (i.e., census regions) so that there remains ample variation in n_i within these regions for the estimation of α. Substituting (5) into (4) yields

$$y_i = h(n_i)'\alpha + z_i'\tilde{\beta} + q_i'\gamma + \rho_{r(i)}^{\mathrm{FE}} + g(w_{r(i)}) + \tilde{v}_i, \tag{6}$$

which illustrates how we solve the location-taste problem owing to region-level taste unobservables by controlling for w_r.[11]

One mechanism contributing to our estimated effects relates to distributional centres, which retailers may open near their stores for logistical purposes. Distributional centres may also reduce shipping times for nearby consumers. Thus, our estimates capture (i) the effects of offline stores on distributional centre networks and (ii) the effects of centres on spending. If we sought to estimate the effect of an offline store on online spending holding fixed distributional centre networks, simultaneity in retailers' offline store and distributional centre location choices would induce an endogeneity problem. We instead consider distributional centres as a mechanism by which offline stores affect online sales.

Sources of conditional variation in store counts. We address the location-taste problem by controlling for a rich set of consumer and region characteristics. Conditional on these controls, variation in offline store counts reflects search-and-matching frictions in real estate markets. A retailer seeking to open a store, for instance, may choose a location based on the variety of properties contemporaneously listed for sale. Similarly, a consumer

[11] This approach differs from the use of demographics as instruments as described by Berry and Haile [2016]. Although we could instrument n_i with w_r, the associated exclusion restriction is unlikely to hold. This restriction, $\mathbb{E}[\omega_i \mid w_{r(i)}, z_i, q_i] = 0$, is unlikely to hold because consumer i's tastes for online shopping may depend on the characteristics of i's neighbours for reasons noted above, which is precisely the dependency that we address by controlling for w_r.

may decide where to live based on the selection of properties that happen to be currently listed. Orderings of properties in real estate catalogues/websites and prospective buyers' sellers' bargaining strategies may also affect retailers' and consumers' locations. We assume that the timing at which properties are listed and search-and-matching frictions in real estate markets are independent of consumer tastes for online shopping. If so, a reliance on these sources of variation does not introduce an endogeneity problem.

Plausibly exogenous tastes for neighbourhoods also induce variation in proximity to stores conditional on our controls. Consider, e.g., two consumers with the same tastes for online shopping, but who desire to live in different neighbourhoods because of differences in locations of family or differences in valuations of local non-retail amenities. These factors provide variation in tastes for neighbourhoods even conditional on our controls.

5 Estimation details

The primary estimation equation is a linear specification of (4):

$$y_i = \alpha' h(n_i) + z_i'\beta + q_i'\gamma + \rho^{\mathrm{FE}}_{R(i)} + w'_{r(i)}\phi + \varepsilon_i, \tag{7}$$

where $\rho^{\mathrm{FE}}_{R(i)}$ is a fixed effect for i's census region $R(i)$ and $h(n_i)$ is a vector of counts of retailers' locations within 20km of consumer i transformed by $x \mapsto \log(x+1)$. The spending outcomes y_i are the panelist's annual online spending at various retailers and in various categories. The z_i variables are the panelist characteristics of panelist i enumerated in Section 2 in addition to the log of the population within 20km of the panelist. We control for local population because tastes for shopping may vary with population density. In the regressions for specialized retailing categories, we also control for log-transformed counts of Walmart, Target, and Costco stores within 20km. Additionally, q_i includes the internet usage controls, whereas $w_{r(i)}$ includes averages of demographic variables in the consumer's region $r(i)$. Finally, we include year fixed effects in each regression.

We estimate (7) by ordinary least squares. To reduce the dependence of our results on outliers, we trim observations in which the spending outcome exceeds its 98th percentile conditional on positive spending. We log-transform store counts based on our hypothesis that offline stores' effects on online spending are diminishing in the number of stores. These effects are likely to be diminishing because offline stores may affect online spending by creating awareness of their associated retailer and consumers in a market are likely to be aware of a retailer once it has a few offline stores, leaving little scope for additional offline stores to further boost awareness.[12]

5.1 Measures of rival effects and cross-channel complementarities

To facilitate interpretation of our results, we compute scale-free measures of offline stores' effects. The first measure, θ_{js}, is the percentage change in expected spending at online store s when the number of retailer j's nearby offline stores rises from its mean value $\bar{n}_j$ to $\bar{n}_j + 1$, conditional on the mean values of the controls. We also define a store-specific rival effect θ_s^{rival} as the average of θ_{js} across rival multichannel retailers j, weighting each j by that retailer's total number of storefronts. Taking a further average of θ_s^{rival} across online retailers s (weighting each retailer by its total online sales) yields the average rival effect $\bar{\theta}^{\text{rival}}$. Section 6 reports estimates of the average rival effect when Amazon's store-specific rival effect is included in and excluded from the average. Last, we define an average own effect as the average of θ_{ss} across multichannel retailers s, weighting each retailer by its total online sales. Online Appendix O.4 provides further details of these measures.

6 Results

This section presents the results. Throughout this section, we emphasize estimates for 2007–2008. Table 8 summarizes results for regressions with overall (rather than store-specific) spending as outcomes, whereas Table 9 reports measures of average rival and own effects as described by Section 5.1.

[12] Wang and Goldfarb [2017] find, for instance, that awareness largely explains the positive effect of a retailer opening an offline store on that retailer's online sales in areas in which the retailer does not already have a strong presence.

Table 8: Overall spending regressions

(a) 2007–2008

	Cross-category retailers	Bookstores	Electronics	Office supplies
	(1)	(2)	(3)	(4)
N. Stores: Total	-11.754^{***}	0.789^{***}	2.517^{**}	0.468
	(2.476)	(0.190)	(1.199)	(0.771)
Mean dep. var.	187.35	9.14	47.37	12.91
Observations	145,345	146,506	146,404	146,765

(b) 2017–2018

	Cross-category retailers	Bookstores	Electronics	Office supplies
	(1)	(2)	(3)	(4)
N. Stores: Total	-6.894^{***}	0.373^{***}	0.315	0.077
	(1.766)	(0.127)	(0.531)	(0.118)
Mean dep. var.	101.83	4.51	22.28	2.46
Observations	170,169	171,029	170,818	171,131

Note: these tables present estimated coefficients from regressions of the overall spending on offline store counts. The "Mean dep. var" row presents the averages of the dependent variable (expenditures in dollars). Heteroskedasticity-robust standard errors in parentheses.

Cross-category retailers. We begin by studying large cross-category retailers. Column (1) of Table 8a and Table 8b provides results for 2007–2008 and 2017–2018, respectively, of the regressions of overall online spending on the number of stores operated by each retailer. These results suggest that offline stores have a negative effect on online spending, which suggests substitutability between online and offline retail channels.[13] But these estimated aggregate effects conceal heterogeneity across retailers. Table 11a displays estimates for regressions of retailer-specific spending on retailer-specific store counts.[14] The estimated effects of multichannel retailers' offline store counts on their own online sales are generally positive and statistically significant across stores and time periods. Conversely, the estimated effects of rival offline stores on a retailer's own online sales are typically negative. We interpret the negative rival effects as evidence that competitive

[13] In Online Appendices O.6 and O.7, we show that the (average) effect of cross-category stores on online spending remains negative with a different regression specification (Poisson regression), with a different dependent variable (positive spending indicator), with inclusion of heterogeneous effects, and with inclusion of finer fixed effects, although some of these alternative estimates are imprecise.

[14] See Online Appendix Table O.10a for results for 2017–2018.

Table 9: Category-level rival and own effects on expenditures

(a) 2007–2008

	Cross-category retailers (1)	Bookstores (2)	Electronics (3)	Office supplies (4)
Rival	-0.038	-0.031	-0.011	-0.030
	(0.011)	(0.013)	(0.017)	(0.016)
Rival	-0.026	0.021	0.001	-0.030
(incl. amazon)	(0.006)	(0.005)	(0.013)	(0.016)
Own	0.194	0.323	0.071	0.263
	(0.019)	(0.052)	(0.054)	(0.027)

(b) 2017–2018

	Cross-category retailers (1)	Bookstores (2)	Electronics (3)	Office supplies (4)
Rival	-0.020	-0.011	0.048	-0.018
	(0.012)	(0.036)	(0.040)	(0.045)
Rival	-0.013	0.008	0.007	-0.009
(incl. amazon)	(0.004)	(0.008)	(0.012)	(0.013)
Own	0.086	0.108	0.016	0.172
	(0.021)	(0.070)	(0.060)	(0.075)

Note: Each column presents the category-level average rival effect and own effects. Panel 9a displays the results for 2007–2008 and panel 9b displays those for 2017–2018. The "Rival (incl. amazon)" row shows the average rival effects including Amazon. Standard errors are computed by the delta method.

effects generally outstrip showrooming effects for large cross-category retailers.

Table 11b presents estimates of our scale-free measures of rival and own effects for cross-category retailers.[15] For all retailers, the rival effect is negative and the own effect is positive. Additionally, the rival and own effects are heterogeneous across retailers, even across those in the same category. Amazon faces weaker rival effects than multichannel retailers in the same sector. Indeed, Amazon suffers a sales reduction of 1.6% in response to entry of an offline rival store, whereas the second weakest effect is the 3.1% sales reduction suffered by Walmart. This may reflect that Amazon benefitted from showrooming.[16]

[15] See Online Appendix Table O.10b for results for 2017–2018.

[16] Online Appendices O.6 and O.7 show that these patterns are not specific to the choice of the functional form, dependent variable, and fixed effects; we find similar patterns when we estimate Poisson

Books. Now consider spending at online bookstores. Column (2) of Table 8 presents results from regressions of overall books spending on the total number of bookstores. In contrast to the cross-category regressions, the estimated effects are positive and significant for both time periods. Showrooming effects provide one interpretation for these estimates. We also estimate separate effects of own and rival offline stores on each retailer's online sales. Table 12 reports the results.[17] The results are similar to those for cross-category retailers in suggesting that a multichannel retailer's own offline stores raise its online sales. It is less clear, though, that a retailer's online sales suffer from rival offline stores given that we estimate a positive rival effect for Amazon. This suggests that offline bookstore experiences lead consumers to purchase books online due to showrooming effects.

Electronics. Overall spending results for electronics appear in the columns labelled (3) in Table 8, and the store-specific results appear in Table 13.[18] Overall spending positively relates to the total number of offline stores, as in the books category. The estimated store-specific relationships, however, are mixed — Best Buy's and Circuit City's offline stores exhibit positive effects on their respective online sales and negative effects on the other's online sales, but the estimates for Radio Shack and Apple do not fit this pattern. None of the estimates for 2017–2018 are statistically significant. The rival and own effects for 2007–2008 are also qualitatively similar to other product categories.

Office supplies. Last, consider office supplies. The columns labelled (4) in Table 8 present results for regressions of overall office supplies spending on the overall counts of office supplies stores. The estimate of the overall effect of stores on spending is statistically insignificant for each time period. Table 14 and Online Appendix Table O.13 provide results for retailer-specific regressions. These results for 2007–2008 suggest strong own-store effects; for each of Office Max, Office Depot, and Staples, we find that increasing the number of offline stores increases their own online sales. The rival effects are generally negative

regression (rather than linear regressions), models with positive spending indicator (rather than expenditures) as the dependent variable, or models with state rather than census region fixed effects.

[17] See Online Appendix Table O.11 for results for 2017–2018.

[18] See Online Appendix Table O.12 for results for 2017–2018.

between multichannel retailers, with the exception that offline Office Max locations boost Office Depot's online sales. These estimates translate to relatively strong negative rival effects together with relatively strong positive own effects; see Table 14b.

7 Discussion

Sources of cross-channel complemenarities. A robust pattern in the estimates is that a multichannel retailer's offline stores tend to boost its own online sales and lower rivals' online sales. The former result reflects cross-channel complentarities that stem from several sources. We identify the sources of cross-channel complementarities in part based on archived versions of retailers' websites from 2007. The first such source is the option to return items purchased online at one of the retailer's offline stores. All archived retailer websites that we checked—including those for Walmart, Target, Costco, Barnes & Noble, Staples, Office Max, Best Buy, and Circuit City—indicate the acceptance of in-store returns for online purchases. This service makes online ordering more appealing to consumers nearby offline stores.

Another source of cross-channel complementarities is "buy online, pickup in-store" (BOPIS). Retailers often allow consumers to purchase items online for pick-up at an offline store. In addition, retailers often offer to ship items that are not carried by an offline store to that store for pick up without charging the consumer any shipping fees. In March 2007, for instance, Walmart launched its "Site to Store" program—which allows consumers to ship items listed online to offline stores for pick-up without paying shipping fees.

Retailer loyalty programs applying to both online and offline purchases also give rise to cross-channel complementarities. Members of Target's "Red Card" program, for example, qualified for savings on both online and offline purchases in 2007. Staples similarly advertised a cross-channel "Staples Rewards" program then. When a retailer has a greater presence nearby a consumer, the consumer has more opportunities to shop at the retailer. This increases the consumer's benefit from joining the retailer's loyalty program. Joining this program in turn raises the appeal of shopping at the retailer's online store.

Some retailers limit their online services to consumers living near offline stores. Staples, for example, restricted furniture deliveries placed online to consumers living within 20 miles of an offline Staples location in 2007. Office Max similarly restricted free deliveries to consumers within 20 miles of an offline store. Such restrictions also give rise to cross-channel complemenarities. Other sources of cross-channel complementarities include offline stores' function as advertisements for their associated retailers, and their function in providing consumers with information about the fit-and-feel characteristics of products sold exclusively by these retailers.

Cross-channel complemenarities likely inversely relate to the extent that retailers offer online-exclusive items. Costco claimed in 2007 that "most items available on our web site are unique to costco.com" rather than available in offline Costco stores. Such disjointedness of online and offline product lines could reduce the scope for cross-channel complementarities.

Differences in showrooming across categories. Evidence of showrooming effects is strongest for books, followed by electronics. This finding is in accord with the fact that both categories include products sold by multiple retailers that consumers may learn about by visiting offline stores. It is also in accord with the fact that e-commerce is more prominent in the books and electronics categories than others; Hortaçsu and Syverson [2015] note that the share of books and magazines sales accounted for by e-commerce in 2013 was 44.2%, higher than that of electronics and appliances (23.1%) and of office equipment and supplies (17.3%). We expect e-commerce to be more prominent in categories in which it is relatively easy for consumers to resolve uncertainty about products sold online by visiting offline stores,[19] and the categories for which this is possible are those subject to showrooming effects.[20] Also, the extent of product variety can explain why categories

[19] Consumers could possibly learn about some products without visiting offline stores; a consumer can learn about music sold online, for example, by listening to music on the radio.

[20] Showrooming effect could also induce online entry and thus make online retail markets more competitive, leading to larger e-commerce share.

with higher shares of e-commerce are expected to be those which enjoy the larger effect of offline stores on online sales. In a category with many products, offline stores may help consumers discover new varieties and thus bolster consumer interest in the category. Also, consumers may be more likely to make online purchases in categories with extensive product variety because offline stores with limited inventories may not have the exact variety that they seek. Thus, we expect categories with more product variety to have higher online sales near offline stores and also to have greater shares of online sales.[21]

Amazon and showrooming. A conspicuous pattern in Table 9 is that excluding Amazon leads to stronger average rival effects. We explain this pattern by appealing to showrooming effects — Amazon charges lower prices than multichannel retailers while freeriding on its multichannel competitors' informative services to achieve higher sales among consumers living nearby these competitors' offline stores. Table 10 documents price differences between books and electronics stores. For books, the table provides the sales-weighted average ratio of a book's price at Amazon to its price at various multichannel retailers across several best-selling books. The table provides the same average ratio for electronics products including three versions of the PlayStation 3—the 40, 60, and 80 gigabyte versions—and two models of Apple iPods, the iPod Nano and the iPod Shuffle.[22] Table 10 shows that Amazon generally offers lower prices than its competitors for PS3s, iPods, and especially for books. Indeed, Barnes & Noble—Amazon's main competitor—offers a price for best-selling books that is on average 23% higher than Amazon's. Amazon's lower prices mean that consumers who learn about a product at offline stores can generally save by instead purchasing the product on Amazon, which implies that a more offline stores around a consumer may especially benefit Amazon. The fact that rival effects (including Amazon) are most positive in the books category, which we

[21]The substitutability of varieties is important here: a consumer may seek particular book titles or generic clothing items (e.g., a t-shirt of some kind, as opposed to an exact item). Given this consideration, the effective extent of variety seems lower in apparel and health and beauty than in books, music, and videos. Managing an inventory of diverse products may be less costly in a centralized warehouse than in a large network of stores; this could drive online entry in categories with many distinct products.

[22]We consider these electronic devices because they are frequently purchased in the data, which allows us to reliably infer their prices.

Table 10: Price differences across retailers

(a) Books

Retailer	Price ratio with Amazon
barnesandnoble.com	1.23
booksamillion.com	1.08

(b) Electronics

Retailer	Price ratio with Amazon
bestbuy.com	1.08
circuitcity.com	1.06
apple.com	1.08

Notes: This table reports average ratios of products' prices at various multichannel retailers to their prices at Amazon. We compute these averages for the books and electronics categories, and we take the averages over distinct product/year pairs, weighting each by the observed number of corresponding transactions. The books that we include in the analysis are those for which we observe sales in the Comscore data and that were either (i) a *New York Times* best-seller in either fiction or non-fiction for at least one week in 2007 or 2008 or (ii) one of Amazon's top selling books of 2007. This yields 26 book titles for which we observe 1696 transactions collectively. The electronics that we include in our analysis are the iPod Shuffle (1GB), the iPod Nano (4GB), and the 40GB, 60GB, and 80GB versions of the PlayStation 3 (PS3). We observe 355 iPod purchases and 89 PS3 purchases. We obtain a price for each product at each retailer in each year by taking a median over transaction prices for the product in the year in question.

argue is especially prone to showrooming effects, further suggests the empirical relevance of showrooming.

Implications for Retail Strategy. The potential importance of showrooming and cross-channel complementarities has implications for retail strategy, although retailers' decisions on pricing, product offerings and entry/exit also depend on other factors such as cost-side impacts of these choices. Our finding of large cross-channel complementarities suggests that retailers can take advantage of their physical stores to boost online sales, by offering in-person customer services enumerated above. On the other hand, e-commerce retailers may count on informative services provided by rivals and then undercut them on prices, instead of offering in-person services on their own. The optimal strategy on entry/exit further depends on the cost implications of physical stores, e.g., the relative importance of entry and operating costs versus economies of scale and scope. Retailers can also mitigate the concern of rivals' freeriding by developing differentiated products at the cost of developing such products.

Some major retailers' operations illustrate these points. Amazon has become the predominant online retailer despite operating almost entirely online and forgoing the benefit of cross-channel complementarities. This raises the questions about the merits of multichannel retailing. But Amazon has benefitted from economies of scale and scope that have permitted its success despite not realizing cross-channel complementarities; see, e.g., Houde et al. 2022. In addition, Amazon's expansion into brick-and-mortar retail through its acquisition of Whole Foods and its introduction of offline stores (under brands Amazon Go, Fresh, and Style) suggests that it perceives gains from cross-channel complementarities. As studied by Bell et al. 2018, eyewear retailer Warby Parker similarly expanded from a primarily online model to a multichannel model and consequently benefitted from cross-channel complementarities. The multichannel strategy is in stark contrast to the retailer's original business plan to "cut out the middlemen and sell directly to customers" by, e.g., "avoiding landlords and their demands for long, expensive leases." (See King 2023.)

8 Conclusion

In this article, we estimated effects of offline stores on online shopping that vary across retailers and retailing categories. One of our principal findings is that cross-channel complementarities are more empirically relevant than cannibalization: a retailer's own offline stores generally increase its online sales. The offline stores of a multichannel retailer's rivals generally reduce this retailer's online sales, although Amazon often experiences sales increases when its multichannel rivals open offline stores. This latter finding could be explained by showrooming effects that particularly benefit Amazon due to the fact that it generally charges lower prices than do multichannel retailers. The estimates suggest that offline bookstores raise online sales of books, a category that we hypothesize is especially prone to showrooming effects. By contrast, presence of large cross-category stores reduces total online spending. These results additionally suggest a role for offline stores in supporting their associated retailers' online sales, and possibly their competitors' sales.

One direction for future research is the decomposition of cross-channel complementarities into, e.g., in-store returns for items purchased online; "buy online, pickup in-store" services; reward programs; awareness effects; and the value of offline stores in providing about products' fit-and-feel characteristics. Additionally, although we focus on the effects of offline stores on online sales, the online retail environment also affects offline sales. A retailer that invests in an online store, for example, is not only affected by online sales and the costs associated with developing and maintaining its online store, but also the effect of its online store on its offline sales. In general, the market structure of retailing industries will depend on offline-to-online effects and online-to-offline effects; we leave the study of the latter and of these effects' interactions in determining equilibrium market structures to future work.

Table 11: Store-specific cross-category spending in 2007–2008

(a) Coefficients

	Spending			
	amazon	costco.com	target.com	walmart.com
	(1)	(2)	(3)	(4)
N. Stores: Costco	0.466	2.089***	0.212*	0.266
	(0.299)	(0.262)	(0.128)	(0.185)
N. Stores: Target	−0.282	0.122	0.408***	−0.752***
	(0.348)	(0.288)	(0.146)	(0.230)
N. Stores: Walmart	−1.295***	−0.683**	−0.572***	0.640***
	(0.331)	(0.332)	(0.140)	(0.189)
Mean dep. var.	14.10	2.51	3.20	5.71
Observations	146,451	146,857	146,770	146,694
R^2	0.057	0.007	0.017	0.023

(b) Rival effects and own effects

	amazon	costco.com	target.com	walmart.com
	(1)	(2)	(3)	(4)
Rival	-0.016	-0.052	-0.038	-0.031
	(0.006)	(0.030)	(0.018)	(0.013)
Own		0.695	0.063	0.047
		(0.075)	(0.023)	(0.014)

Note: Panel 11a presents the coefficients from the regressions of the expenditures at a given online retailer on the numbers of offline stores of each retailer. The "Mean dep. var" row presents the average expenditures at each online retailer. Heteroskedasticity-robust standard errors in parentheses. Panel 11b displays the scale-free measures of the rival and own effects. Standard errors are computed by the delta method.

Table 12: Store-specific books spending in 2007–2008

(a) Coefficients

	Spending		
	amazon	barnesandnoble.com	booksamillion.com
	(1)	(2)	(3)
N. Stores: Barnes	0.161	0.344***	0.015
	(0.170)	(0.052)	(0.016)
N. Stores: Books-a-Million	0.272*	−0.063	0.056**
	(0.163)	(0.051)	(0.022)
N. Stores: Borders	0.468***	−0.203***	−0.021
	(0.154)	(0.059)	(0.014)
N. Stores: Other	0.555***	−0.036	−0.008
	(0.144)	(0.047)	(0.012)
N. Stores: Waldenbooks	0.008	0.113**	−0.015
	(0.138)	(0.044)	(0.011)
Mean dep. var.	5.53	0.86	0.06
Observations	146,629	146,819	146,869
R^2	0.034	0.008	0.002

(b) Rival and own effects

	amazon	barnesandnoble.com	booksamillion.com
	(1)	(2)	(3)
Rival	0.030	-0.030	-0.052
	(0.006)	(0.013)	(0.042)
Own		0.234	1.579
		(0.035)	(0.582)

Note: See the notes for Table 11.

Table 13: Store-specific electronics spending in 2007–2008

(a) Coefficients

	Spending				
	amazon	apple.com	bestbuy.com	circuitcity.com	radioshack.com
	(1)	(2)	(3)	(4)	(5)
N. Stores: Apple	0.277	−0.162	−0.359*	−0.252	0.047**
	(0.205)	(0.283)	(0.203)	(0.216)	(0.022)
N. Stores: Best Buy	0.031	−0.298	0.589**	−0.530*	0.018
	(0.260)	(0.406)	(0.256)	(0.316)	(0.026)
N. Stores: Circuit City	−0.008	0.530	−0.562**	0.554*	−0.003
	(0.245)	(0.352)	(0.265)	(0.298)	(0.032)
N. Stores: Radio Shack	0.322	0.482	0.196	0.158	−0.059*
	(0.288)	(0.415)	(0.274)	(0.267)	(0.035)
Mean dep. var.	3.22	2.39	2.31	2.13	0.08
Observations	146,819	146,853	146,847	146,850	146,869
R^2	0.011	0.002	0.004	0.004	0.001

(b) Rival and own effects

	amazon	apple.com	bestbuy.com	circuitcity.com	radioshack.com
	(1)	(2)	(3)	(4)	(5)
Rival	0.025	0.048	-0.046	-0.044	0.159
	(0.014)	(0.032)	(0.027)	(0.031)	(0.098)
Own		-0.070	0.146	0.160	-0.254
		(0.123)	(0.063)	(0.086)	(0.140)

Note: See the notes for Table 11.

Table 14: Store-specific office supplies spending in 2007–2008

(a) Coefficients

	Spending			
	amazon	officedepot.com	officemax.com	staples.com
	(1)	(2)	(3)	(4)
N. Stores: Office Depot	0.028	2.190***	−0.209**	−0.278
	(0.019)	(0.370)	(0.104)	(0.438)
N. Stores: Office Max	0.007	0.790**	0.172***	−0.903**
	(0.013)	(0.366)	(0.060)	(0.380)
N. Stores: Other	−0.019	−0.349	−0.045	−0.534
	(0.018)	(0.324)	(0.085)	(0.382)
N. Stores: Staples	0.010	0.005	−0.116	1.875***
	(0.013)	(0.348)	(0.083)	(0.280)
Mean dep. var.	0.07	3.59	0.33	4.54
Observations	146,870	146,856	146,869	146,848
R^2	0.001	0.003	0.001	0.005

(b) Rival and own effects

	amazon	officedepot.com	officemax.com	staples.com
	(1)	(2)	(3)	(4)
Rival	0.041	0.016	-0.162	-0.057
	(0.049)	(0.025)	(0.064)	(0.021)
Own		0.322	0.319	0.212
		(0.050)	(0.100)	(0.029)

Note: See the notes for Table 11.

Bibliography

Daniel Ackerberg, C. Lanier Benkard, Steven Berry, and Ariel Pakes. Chapter 63 econometric tools for analyzing market outcomes. volume 6 of *Handbook of Econometrics*, pages 4171–4276. Elsevier, 2007.

Daniel A. Ackerberg, Kevin Caves, and Garth Frazer. Identification properties of recent production function estimators. *Econometrica*, 83(6):2411–2451, 2015.

Jill Avery, Thomas J. Steenburgh, John Deighton, and Mary Caravella. Adding bricks to clicks: Predicting the patterns of cross-channel elasticities over time. *Journal of Marketing*, 76(3):96–111, 2012.

David R. Bell, Santiago Gallino, and Antonio Moreno. Offline showrooms in omnichannel retail: Demand and operational benefits. *Management Science*, 64(4):1629–1651, 2018.

Steven T. Berry and Philip A. Haile. Identification in differentiated product markets. *Annual Review of Economics*, 8:27–52, 2016.

Erik Brynjolfsson, Yu (Jeffrey) Hu, and Mohammad S. Rahman. Battle of the retail channels: How product selection and geography drive cross-channel competition. *Management Science*, 55(11):1755–1765, 2009.

Dennis W. Carlton and Judith A. Chevalier. Free riding and sales strategies for the internet. *The Journal of Industrial Economics*, 49(4):441–461, 2001.

Pradeep K. Chintagunta, Junhong Chu, and Javier Cebollada. Quantifying transaction costs in online/off-line grocery channel choice. *Marketing Science*, 31(1):96–114, 2012.

Ba M. Chu, David T. Jacho-Chávez, and Oliver B. Linton. Standard errors for nonparametric regression. *Econometric Reviews*, 39(7):674–690, 2020.

Paul Dolfen, Liran Einav, Peter J. Klenow, Benjamin Klopack, Jonathan D. Levin, Laurence Levin, and Wayne Best. Assessing the gains from e-commerce. 2019. NBER working paper 25610.

Danny Edgel, Jean-François Houde, Peter Newberry, and Katja Seim. The impact of distance in retail markets. *AEA Papers and Proceedings*, 113:229–33, May 2023.

Chris Forman, Anindya Ghose, and Avi Goldfarb. Competition between local and electronic markets: How the benefit of buying online depends on where you live. *Management Science*, 55(1):47–57, 2009.

Amit Gandhi, Salvador Navarro, and David A. Rivers. On the identification of gross output production functions. *Journal of Political Economy*, 128(8):2973–3016, 2020.

Georg Goetz, Daniel Herold, Phil-Adrian Klotz, and Jan Thomas Schäfer. The substitutability between brick-and-mortar stores and e-commerce-the case of books. Technical report, Joint Discussion Paper Series in Economics, 2020.

Austan Goolsbee. Competition in the computer industry: Online versus retail. *The Journal of Industrial Economics*, 49(4):487–499, 2001.

James J. Heckman and Edward J. Vytlacil. Chapter 71 econometric evaluation of social programs, part ii: Using the marginal treatment effect to organize alternative econometric estimators to evaluate social programs, and to forecast their effects in new environments. volume 6 of *Handbook of Econometrics*, pages 4875–5143. Elsevier, 2007.

Ali Hortaçsu and Chad Syverson. The ongoing evolution of us retail: A format tug-of-war. *Journal of Economic Perspectives*, 29(4):89–112, November 2015.

Jean-François Houde, Peter Newberry, and Katja Seim. Nexus tax laws and economies of density in e-commerce: A study of amazon's fulfillment center network. *Econometrica*, 2022. Forthcoming.

Yangguang Huang, Chenyang Li, and Si Zuo. Stores going online: Market expansion or cannibalization? 2023. Working paper.

Bing Jing. Showrooming and webrooming: Information externalities between online and offline sellers. *Marketing Science*, 37(3):469–483, 2018.

Bruno Jullien, Alessandro Pavan, and Marc Rysman. Two-sided markets, pricing, and network effects. In Kate Ho, Ali Hortaçsu, and Alessandro Lizzeri, editors, *Handbook of Industrial Organization, Volume 4*, volume 4 of *Handbook of Industrial Organization*, pages 485–592. Elsevier, 2021.

Kate King. Online-only startups adopt a bold new strategy: Opening actual shops. *The Wall Street Journal*, 2023. URL `https://www.wsj.com/articles/`

warby-parker-allbirds-everlane-parachute-collars-co-stores-e4b94623?mod=djem10point.

Dmitri Kuksov and Chenxi Liao. When showrooming increases retailer profit. *Journal of Marketing Research*, 55(4):459–473, 2018.

James Levinsohn and Amil Petrin. Estimating Production Functions Using Inputs to Control for Unobservables. *The Review of Economic Studies*, 70(2):317–341, 04 2003.

Alexander MacKay and David Smith. The empirical effects of minimum resale price maintenance. 2014.

Amit Mehra, Subodha Kumar, and Jagmohan S. Raju. Competitive strategies for brick-and-mortar stores to counter "showrooming". *Management Science*, 64(7):3076–3090, 2018.

G. Steven Olley and Ariel Pakes. The dynamics of productivity in the telecommunications equipment industry. *Econometrica*, 64(6):1263–1297, 1996.

Andrea Pozzi. The effect of internet distribution on brick-and-mortar sales. *The RAND Journal of Economics*, 44(3):569–583, 2013.

Jeffrey T. Prince. The beginning of online/retail competition and its origins: An application to personal computers. *International Journal of Industrial Organization*, 25(1):139–156, 2007.

Thomas W. Quan and Kevin R. Williams. Product variety, across-market demand heterogeneity, and the value of online retail. *The RAND Journal of Economics*, 49(4):877–913, 2018.

Jiwoong Shin. How does free riding on customer service affect competition? *Marketing Science*, 26(4):488–503, 2007.

Scott K. Shriver and Bryan Bollinger. Demand expansion and cannibalization effects from retail store entry: A structural analysis of multichannel demand. *Management Science*, 68(12):8829–8856, 2022.

Todd Sinai and Joel Waldfogel. Geography and the internet: is the internet a substitute or a complement for cities? *Journal of Urban Economics*, 56(1):1–24, 2004.

Kitty Wang and Avi Goldfarb. Can offline stores drive online sales? *Journal of Marketing Research*, 54(5):706–719, 2017.