
Drivers of the Long Tail Phenomenon: An Empirical Analysis

OLIVER HINZ, JOCHEN ECKERT, AND BERND SKIERA

OLIVER HINZ is an assistant professor of e-finance and electronic markets at Goethe-University of Frankfurt (Germany). He studied business administration and information systems at Technischen Universität Darmstadt (Germany) with a focus on marketing, software engineering, and operations research and worked several years as a consultant for business logic in the financial services industry. Dr. Hinz's research has been published or is forthcoming in journals such as the *Journal of Management Information Systems*, *Information Systems Research*, *MIS Quarterly*, *Decision Support Systems*, *Electronic Markets*, *Business & Information Systems Engineering*, and in a number of proceedings (e.g., ICIS, ECIS, PACIS).

JOCHEN ECKERT is customer insights and program manager for eBay Australia and New Zealand. He studied business administration at Goethe-University in Frankfurt (Germany), specializing in marketing and industrial organization, and worked as a project manager for an e-payment provider in Frankfurt. His research focuses on demand changes due to e-commerce and on preference estimation procedures.

BERND SKIERA is chair of electronic commerce at Goethe-University, Frankfurt (Germany), and is a member of the board of the E-Finance Lab. His interests include electronic commerce, online marketing, customer management, pricing, and prediction markets. Dr. Skiera's research has been published in various journals such as *Management Science*, *Marketing Science*, *Journal of Marketing Research*, *Journal of Marketing*, *Journal of Product Innovation Management*, *European Journal of Operational Research*, and *Journal of Interactive Marketing*.

ABSTRACT: The Internet makes it easy to offer large assortments of products, tempting managers to chase the “long tail”—that is, the phenomenon in which niche products gain a significant share of demand among all products. Yet few studies empirically examine the existence and drivers of this long tail phenomenon. This study uses a unique data set with 843,922 purchases from 143,939 customers that a monopolistic video-on-demand operator observed over 111 weeks after its launch of the service. The current analysis centers on the effects of increasing assortment sizes and improved search technologies on measures of the long tail, such as per customer demand, the share of products purchased from the assortment, the distribution of demand across products, and the concentration of demand. Increases in assortment sizes and better assortment quality lead to increases in demand per customer and a longer tail. The length of the tail (i.e., share of purchased products) is also driven by new customers and seasonal effects, such as school vacations, whereas the presence of high-quality blockbuster products shortens the tail. Different search technologies can shift demand toward niche products as well as toward blockbuster products.

KEY WORDS AND PHRASES: electronic commerce, long tail, recommendation systems, search technology, video-on-demand.

THE INTERNET AND RELATED TECHNOLOGIES HAVE VASTLY EXPANDED the variety of products that can be profitably promoted and sold by retailers, leading to a dramatic increase in assortment sizes [8, 12]. Search and recommendation tools support these increases because they reduce customer search costs on the Internet. The book market, for example, has seen a huge increase in the total number of books available (1.5 million books in print in 1990 versus 5.6 million in 2006 [1]), which has increased demand for niche products. For example, books that are not even offered at traditional offline bookstores account for about 30 percent of Amazon.com's sales [7]. Amazon.com alone has an assortment size of 3 million books, approximately 30 times the average assortment size of a traditional bookstore such as Borders. Another example is the online music retailer Rhapsody. Whereas traditional music retailers such as Wal-Mart carry about 55,000 music tracks, Rhapsody's assortment size is 1.5 million, more than 27 times that of the world's largest physical retailer.

As a result of this huge increase in online assortment sizes, Anderson [1] has coined the term *long tail* to describe the phenomenon by which niche products gain a significant share of demand for all products, which then consequently decreases the importance of blockbuster products [6]. This concept assumes an underexploited spectrum of customer tastes that has not been addressed sufficiently or cost-effectively by pre-Internet retailers [39]. However, the low costs of supply and new search technologies enable online retailers to serve these tastes profitably, leading Anderson [1] to predict the end of the blockbuster era.

Yet studies that empirically analyze the effect of increases in assortment size on the distribution of demand across products are scarce (for a review, see [13]). Previous studies of the demand distribution between blockbuster products (hereafter, blockbusters) and niche products (hereafter, niches) [16, 17] do not isolate the specific effects of greater assortment size. Even more worrisome is the nonexistence of research that empirically analyzes how an increasing assortment size affects individual demand in online channels. Such changes in individual demand are far from obvious. On the one hand, larger assortments allow customers to find products that better fit their needs, which should increase customer demand [3, 9, 11]. On the other hand, consumer behavior research shows that too many choices lead to information overload and may cause customers to buy less [5, 9, 14, 24, 25, 29]. In addition, empirical analyses of the effect of search technologies on the distribution of demand across products is rare and offers only mixed results. The few previous studies on the impact of search technology on the demand distribution indicate both decreases [1, 6] and increases [18, 31] in demand diversity. In recent work pertaining to recommender systems, Fleder and Hosanagar [19] determine analytically that these systems favor blockbusters and decrease diversity of demand.

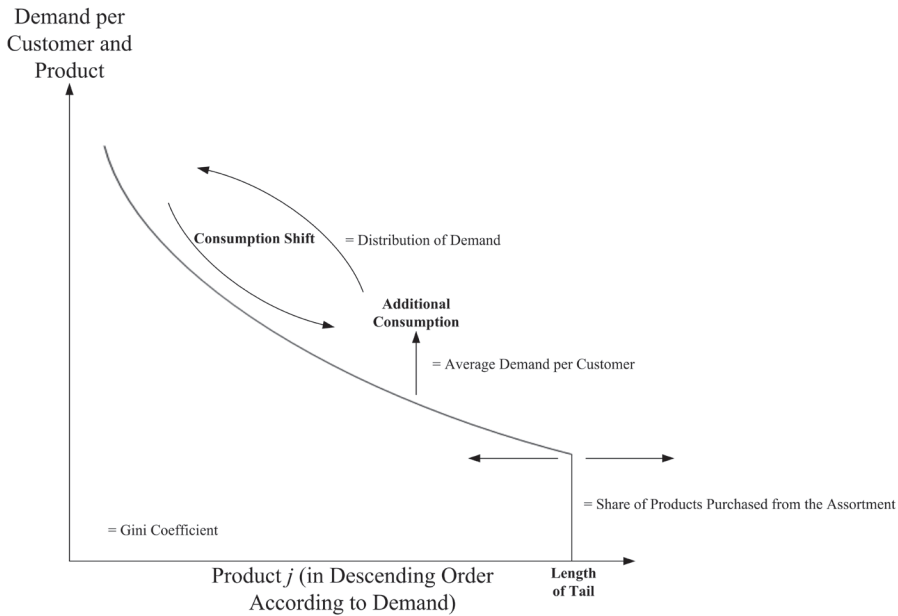


Figure 1. Conceptual Model

Therefore, the current study examines the existence and drivers of the long tail phenomenon to determine why niche products might gain significant market share. Specifically, we analyze the effects of increasing assortment size and improved search technologies on (1) the demand per customer, (2) the share of products purchased from the assortment, and (3) the distribution of demand across products as well as the concentration of demand. In contrast with previous research [6, 16, 17], we examine the effect of the drivers on individual demand, which prevents confounds with changes due to an increasing number of customers.

Figure 1 illustrates the three effects that we analyze. The x -axis in Figure 1 represents the sales ranks of products, ordered according to sales volume from highest (left) to lowest (right). The y -axis in Figure 1 indicates the corresponding sales per product and customer. The first effect describes how the drivers, especially growing assortment size and different search technologies, affect the demand of each customer. If the demand per customer increases, the area below the demand distribution curve grows. Previous studies (e.g., [17]) have focused on aggregated demand and thus examine the combined effects of an increasing number of customers and an increased demand per customer, but we isolate the effect on the latter. Disentangling these effects is feasible because we collect individual transactional data. The second effect shows the impact of the drivers on the share of purchased products. Finally, the third effect describes the influence of the drivers on the distribution of demand across products. We are particularly interested in the effect of increasing assortment sizes. If the curve flattens, demand for additional products comes at the expense of existing products. If the curve remains unchanged and simply lengthens, demand for additional products

actually increases demand per customer. This development can also be captured by concentration measures such as the Gini coefficient.

The remainder of this paper is organized as follows. First, we review existing literature related to the definitions, existence, and drivers of the long tail phenomenon. Second, we describe the data for our empirical study and explain the motivation for our approach to analyzing data. Third, we present the results of the empirical study. Fourth, we conclude with a summary and managerial implications of our findings, as well as suggestions for further research.

Previous Research on the Long Tail Phenomenon

EXISTING LITERATURE IDENTIFIES DIFFERENT DRIVERS of the long tail phenomenon [1, 6, 8, 17, 21]. In addition to aggregating heterogeneous customer preferences more effectively, the Internet enables retailers and manufacturers to increase their assortment sizes and implement new cost-reducing technologies for searching products, such as search filters and recommendation systems. Next, we describe prior studies that analyze assortment size and search technologies as two drivers of the long tail phenomenon.

Assortment Size as a Driver of the Long Tail Phenomenon

Anderson [1] suggests that increases in the assortment size shift demand away from blockbusters and toward niches. However, to the best of our knowledge, only three studies [16, 17, 37] empirically analyze how assortment size increases actually shift demand between blockbusters and niche products. Elberse and Oberholzer-Gee [17] find that an increase in the assortment size in online markets might even shift demand from blockbusters to niches in traditional sales channels, though they also indicate that the demand distribution across blockbusters becomes more concentrated, such that fewer blockbusters generate the majority of the demand. With their data, Elberse and Oberholzer-Gee cannot distinguish between changes in individual demand and changes in the customer base. Elberse [16] instead shows that greater assortment size online does not noticeably change the demand distribution in traditional sales channels. Yet Elberse also does not disentangle the effects of increasing assortment sizes versus search technology on the distribution of demand.

In contrast, Sela et al. [37] analyze how assortment size influences consumers' choices, given the product alternatives available in traditional sales channels. Choosing from large assortments is often more difficult, so customers tend to make easily justifiable product choices. However, Sela et al. do not situate their finding in the distribution of demand between blockbusters and niches. To the best of our knowledge, no study empirically investigates and isolates the effects of increasing assortment size on the distribution of demand per customer across products.

Some literature examines such effects in a nonempirical setting but without addressing empirically how increasing the assortment size influences the demand per customer online. From an online retailer's perspective, increasing assortment sizes

might lead to additional demand per customer and thus additional revenues and profits. However, such increased demand could also reflect just a decrease in demand for existing products. Previous studies cannot disentangle these effects because they cannot control for increases in the number of customers, such that they confound the changes that result from an increasing customer base with those that reflect individual demand changes.

For example, Clemons et al. [12] find evidence that highly dispersed online ratings in online reviews for products positively affect sales growth rates, and they situate their finding in the context of a huge increase in assortment size, which they label *hyperdifferentiation*. Firms that provide highly differentiated assortments thus should experience higher sales growth than firms with less differentiated assortments. Although Clemons et al.'s finding implies that firms can attract additional demand if they increase their assortment size, they cannot distinguish between the two causes because they only have access to aggregated sales data.

A large body of related studies also considers the effects of assortment size on individual demand in traditional sales channels. Several studies suggest that customers might benefit from greater assortment sizes because they can find products that better match their preferences [3, 28, 34], in line with choice theory (for an overview, see [24]). Customers might also value diversity [26], which usually correlates with assortment size. Yet even in traditional sales channels, assortment size can become so excessive that size reductions do not prompt any decrease in demand per customer [4, 5, 15]. Huge assortments seem overwhelming and confusing to customers, which may even lead to consumption avoidance [24, 27]. The potential negative effect of large assortments, due to too much choice, appears in various settings [9, 14, 24, 25, 29].

These effects may be less likely in online channels, though, because search technologies better support navigating through large assortments. Furthermore, supply costs are lower on the Internet and enable online retailers to realize positive margins even for less-popular products. Still, a large share of products with no demand hurts online retailers. Elberse and Oberholzer-Gee [17] show, in the context of video sales, that increases in assortment sizes might substantially increase the number of unsold products in traditional sales channels. Whether this finding holds for online channels, in which customers can search easily for products [2] (which helps extend the search beyond easy-to-find, top-selling products [1, 6]), remains an open question.

Search Technology as a Driver of the Long Tail Phenomenon

Limited literature investigates how different search technologies affect the distribution of demand across products, and virtually no research considers the effects of search technologies on individual demand or whether they might influence shares of purchased products. Search technologies such as recommender systems might decrease the concentration of demand across products and therefore reduce the importance of blockbusters [1, 6], or they might increase the concentration of demand [18, 31].

Fleder and Hosanagar [19] attempt to reconcile these incompatible results by showing that recommender systems generally decrease the diversity of demand by favoring blockbusters. Furthermore, it is possible for individual diversity of demand to increase but aggregated diversity to decrease because recommender systems “push each person to new products, but they often push users toward the same products” [19, p. 698]. These results have not been empirically verified. Brynjolfsson et al. [6] also find evidence that lower search costs for online customers reduce the relative importance of blockbusters because they decrease the relative contribution of the best-selling 20 percent of products to total demand. Brynjolfsson et al. do not specify which search technology drives this result, though, and they cannot determine the impact on demand per customer. Existing literature thus suggests that lower search costs enable shoppers to find more products that better fit their preferences [2, 22, 38], which would mean that improved search technologies should lead to additional demand per customer. However, the lack of empirical studies, despite the crucial managerial implications, leaves unclear whether search technologies lead to additional demand and potentially more profit.

We also expect that search technology affects the share of purchased products. Whereas search technologies, such as “favorite lists,” favor blockbusters and therefore should decrease the share of purchased products, other search technologies, such as search filters, may increase the long tail and increase the share of purchased products. This question is crucial for retailers aiming to manage their assortments, but it has not been addressed in previous studies.

Comparisons with Previous Research

In Table 1, we outline how our study compares with previous research that has also studied the effects of the two main drivers—namely, assortment size and search technology—on the long tail phenomenon, as measured by the distribution of demand, share of purchased products, and demand per customer. We also detail whether prior results were derived analytically, conceptually, or empirically, and we distinguish active search technologies (e.g., search engines, search filters) from passive ones (e.g., recommender systems) [6]. As Table 1 clearly shows, our study closes a gap by empirically examining the effect of both drivers on all measures of the long tail phenomenon.

Setup of Empirical Study

Objectives

STUDIES THAT EMPIRICALLY EXAMINE THE DRIVERS of the long tail phenomenon are scarce. This empirical study investigates the effects of different search technologies and increases in assortment size on demand per customer, share of purchased products, and distribution of demand across products, which represent key measures of the long tail phenomenon in terms of the length, shape, and thickness of the tail.

Table 1. Overview of Previous Research on the Effect of Drivers on the Long Tail Phenomenon

Dependent variable	Distribution of demand		Share of purchased products		Demand per customer	
	Assortment size	Search technology	Assortment size	Search technology	Assortment size	Search technology
Drivers	Active	Passive	Active	Passive	Active	Passive
Anderson [1]	C	C	C	C		
Brynjolfsson et al. [8]	C	C				
Brynjolfsson et al. [6]		A+E				
Elberse [16]	E	E				
Elberse and Oberholzer-Gee [17]	E		E			
Fleder and Hosanagar [18]		A				
Fleder and Hosanagar [19]		A				
Hervas-Drane [22]		A				
Mooney and Roy [31]		A+E				
Hinz and Eckert [23]	E	A				A
This paper		E	E	E	E	E

(continues)

Table 1. Continued

Dependent variable	Distribution of demand		Share of purchased products		Demand per customer	
	Assortment size	Search technology	Assortment size	Search technology	Assortment size	Search technology
Drivers	Active	Passive	Active	Passive	Active	Passive
Related literature with a focus on traditional sales channels						
Bakos [2]					A	
Baumol and Ide [3]					E	
Boatwright and Nunes [4]						
Broniarczyk et al. [5]						
Chernev [9]						
Clemons et al. [12]						
Dhar [14]						
Dreze et al. [15]						
Huffman and Kahn [24]						
Iyengar and Lepper [25]						
Kahn [26]						
Kahn [27]						
Lancaster [28]						
Lehmann [29]						
Quelch and Kenny [34]						
Sela et al. [37]						
Stigler [38]						

Notes: A = analytical, C = conceptual, E = empirical.

Data

To achieve our study objectives, we used transactional data from the video-on-demand market leader in Germany. This monopolistic operator provides access to films from different genres via a platform on the Internet. Once customers have registered on the platform, they pay a fee per film that enables them to watch the film on their computer. The data cover every sale from the start of service in December 2004 through June 2007—that is, all 843,922 purchases made by 143,939 customers. The business model changed in June 2007 such that the provider started allowing customers to store films permanently on their hard drives in addition to renting films for 24 hours. We gathered weekly data but eliminated the first 13 weeks because demand during that period was highly volatile. Thus, our observations start in March 2005. We also eliminated two weeks during which technical problems on the operator’s Web site made the platform frequently inaccessible to customers. Our analysis therefore includes observations from 111 weeks, as we detail in Table 2.

The data exhibit several unique properties. The assortment size changed dramatically over time, growing sixfold to a total of 1,247 films. In addition, the operator changed its search technology three times during the observation period, which represents an interesting natural experiment to analyze the effect of differences in search technology on the long tail phenomenon. In Table 3, we illustrate the functionalities of these three search technologies.

All three search tools in Table 3 offered filters (e.g., customers could search by genre or browse newly added films) and text retrieval for film titles, though otherwise they differed substantially. For example, search tool 1 allowed the indexing of only 300 films, which could be found only through text retrieval or filters. The text search of search tool 1 did not resolve spelling mistakes. In contrast, search tool 2 provided indexes of all the films in the assortment, as well as additional search filters, and was not sensitive to spelling mistakes. Moreover, it offered film recommendations, such that when a customer clicked a link to request more information about a film, the resulting page highlighted seven blockbuster film recommendations. The consumer search costs associated with search tool 2 therefore should be lower than those for search tool 1 because it helps consumers find films faster. Demand for blockbusters also might increase strongly because search tool 2 recommends the most successful films. Finally, the recommendation system with search tool 3 provided seven other films that belonged to the selected film’s genre, which means that it increased the diversity of the films recommended compared with search tool 2 but still favored blockbusters more than search tool 1. Search tool 3 also contained new filters (e.g., by director, actors) to improve the system’s capabilities, and more filters could be used at the same time.

We expect that these various changes all should decrease search costs. However, the introduction of new filters and indexes of all the titles might reduce search costs more for niches than for blockbusters [6], whereas editorial recommendations should increase demand for blockbusters. The difference between the recommendations in search tools 2 and 3 may dilute this effect because recommendations were dispersed across the top blockbusters for all users as well as genre-specific blockbusters.

Table 2. Description of Empirical Data

Variable	Minimum	Mean	Median	Maximum
Weekly sales	3,922	7,602.9	7,835	10,644
Weekly sales top selling product	247	457.3	427	915
Weekly active customers	7,628	12,730.7	12,871	17,039
Weekly assortment size	373	810.3	798	1,247
Weekly Gini coefficient (percent)	55.3	63.7	63.7	73.2
Total number of customers	143,939			
Total sales	843,922			

Table 3. Functionalities of Different Search Tools

Search functionalities	Search tool 1	Search tool 2	Search tool 3
Indexing	Up to 300 films	All films	All films
Text retrieval for film titles	Yes	Yes	Yes
Text retrieval fixes spelling mistakes	No	Yes	Yes
Number of search filters	Low	Medium	High
Filters can be used at the same time	No	No	Yes
Film recommendations	No	Yes (blockbuster recommenda- tions)	Yes (genre-specific blockbuster rec- ommendations)
Application period	Weeks 1–52	Weeks 53–82	Weeks 83–111

Modeling Approach

Our modeling approach consists of three parts in which we model the effects of the two drivers, assortment sizes and search technologies, on (1) the demand per customer, (2) the share of purchased products, and (3) the distribution of demand across products. In addition to the two drivers, we control for additional drivers. In particular, the composition of the customer base might change over time because heavy users are more likely to adopt earlier than light users. We capture this influence by calculating the average historic daily demand per active customers in week t , $ADEC_t$:

$$ADEC_t = \frac{1}{N_{t-1}} \cdot \frac{\sum_{j=1}^{t-1} d_{i,j}}{\sum_{i=1}^{N_{t-1}} AD_{i,t-1}}, \quad (1)$$

where $ADEC_t$ is the average historic daily demand per existing, active customer in week t ; N_{t-1} is the number of active customers in week $t - 1$; $AD_{i,t-1}$ is the duration (in days) of the relationship with active customer i before week $t - 1$; and $d_{i,j}$ is the demand of active customer i in week j .

A higher value indicates that the customer base includes more heavy users. We also control for the number of customers who made their first purchase in week t . They might influence demand because they are likely very curious at the beginning of their relationship with the firm and watch many different films. We therefore distinguish between existing customers (already active in week $t - 1$) and new customers, who purchase for the first time in week t . Thus, we calculate the fraction of new customers for every week t .

Furthermore, we control for changes in the composition of the assortment over time. We collected box office sales for all the films and calculated the sum of the box office sales¹ of all the films in the assortment during week t . A higher value should indicate a higher-quality assortment. The average box office sales for the top 100 films also provides a measure of the appeal of the best-selling films.

To control for other effects, we considered dummy variables that captured seasonality, school vacations, Christmas week, the 2006 soccer World Cup (in Germany, which might have distracted consumers from films), and a trend variable t to capture other unexplained changes over time.

Demand per Customer

With Equation (2), we model the influence of increasing assortment sizes, different search technologies, and the control variables on the demand per customer. We include dummy variables for different search technologies and allow for nonlinear effects of the assortment size by adding a squared assortment size variable:

$$\begin{aligned} dt = & \beta_0 + \beta_1 \cdot ST2_t + \beta_2 \cdot ST3_t + \beta_3 \cdot AssortmentSize_t + \beta_4 \cdot AssortmentSize_t^2 \\ & + \beta_5 \cdot ADEC_t + \beta_6 \cdot FractionNewCustomers_t + \beta_7 \cdot TotalQuality_t + \beta_8 \\ & \cdot QualityBlockbuster_t + \beta_9 \cdot Season_t + \beta_{10} \cdot Christmas_t + \beta_{11} \\ & \cdot SchoolVacations_t + \beta_{12} \cdot WC2006_t + \beta_{13} \cdot t + \varepsilon_t, \end{aligned} \quad (2)$$

where d_t is demand in week t per 1,000 active customers; STk_t is a dummy variable indicating whether search tool k is present ($= 1$) or not ($= 0$) in week t ; $AssortmentSize_t$ is the assortment size in week t ; $ADEC_t$ is the average historic daily demand per existing, active customers in week t ; $FractionNewCustomers_t$ is the fraction of new customers in week t ; $TotalQuality_t$ is the quality of assortment in week t , measured as the sum of box office sales of all films in the assortment in week t ; $QualityBlockbuster_t$ is the average quality of blockbusters (top 100) in week t , measured as the average box office sales of all top 100 blockbusters; $Season_t$ is a dummy variable for the season

in week t (1 = winter, 0 = summer); $Christmas_t$ is a dummy variable for Christmas week (1 = Christmas week, 0 = otherwise); $SchoolVacations_t$ is a dummy variable for school vacations (1 = vacations, 0 = otherwise); $WC2006_t$ is a dummy variable for the 2006 soccer World Cup (1 = WC2006, 0 = otherwise); and t = particular week to capture time trend.

We calculate our dependent variable by dividing the weekly demand for all products by the weekly number of active customers, which yields 111 (weekly) observations. In a noncontractual setting, determining the number of active customers is not trivial because customers have no incentive to announce the end of the relationship. However, marketing studies have developed models to infer the number of active customers from historical purchase behavior [35].

We use Reinartz and Kumar's [35] model to determine the number of active customers in each week t . Specifically, we calculate the probability that a customer is active by calculating T^n , where n is the number of purchases in a given period and T is the time of the last purchase, expressed as a fraction of the observation period. Customers with a purchase probability higher (lower) than 50 percent are classified as active (inactive) customers. We tested the accuracy of this model by verifying its predictions against the observed number of active customers in the future. If the model predicts that a customer is active (inactive) and we find corresponding evidence in our data, we rank it a hit; otherwise, it is a miss. This test is obviously more accurate at the beginning of the observation period because the data are right truncated (only available until August 2007). However, the correlation between the number of active customers per week, as predicted by the applied model and the actual number of active customers (determined by analyzing which existing customers made a purchase in later periods), is 0.85 ($p < 0.01$), and the hit rate of the Reinartz and Kumar model is 78.9 percent. Therefore, the applied model is in line with the actual number of active customers in our right-truncated data and a good proxy for the latent unobservable number of active customers. In Figure 2, we illustrate the number of active customers according to the model by Reinartz and Kumar [35].

Share of Purchased Products

We next examine the influence of the assortment size and different search technologies on the share of purchased products from the assortment by estimating a logit model for our 111 weekly observations. Equation (3) illustrates the general structure of our model, and Equation (4) shows the model specification. We again include dummy variables for different search technologies, allow for nonlinear assortment size effects, and add the previously introduced control variables. Thus,

$$r_t = \frac{z_t}{J_t} = \frac{e^{v_t}}{(e^{v_t} + 1)}, \quad (3)$$

where r_t is the share of purchased products from the assortment in week t ; z_t is the number of purchased products from the assortment in week t ; v_t is a function param-

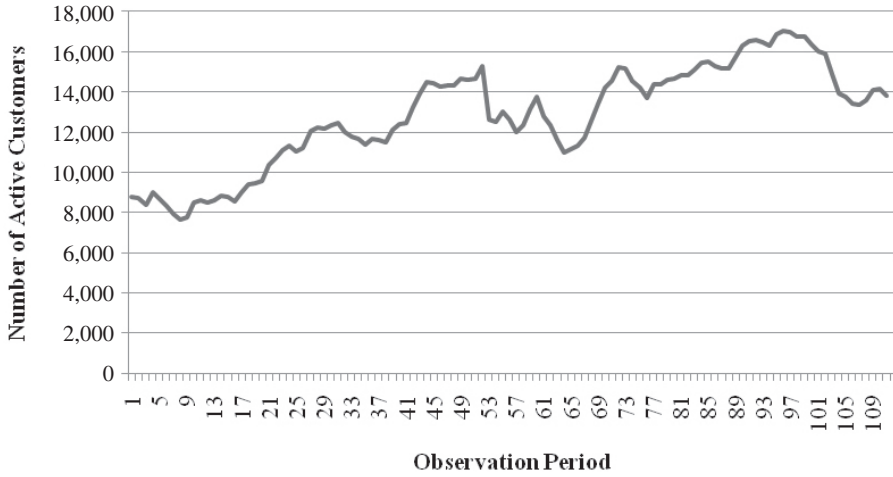


Figure 2. Change in Number of Active Customers over Time

eter of the share of purchased products from the assortment in week t ; and J_t is the assortment size in week t .

$$\begin{aligned}
 v_t = & v_0 + v_1 \cdot ST2_t + v_2 \cdot ST3_t + v_3 \cdot AssortmentSize_t + v_4 \cdot AssortmentSize_t^2 \\
 & + v_5 \cdot ADEC_t + v_6 \cdot FractionNewCustomers_t + v_7 \cdot TotalQuality_t \\
 & + v_8 \cdot QualityBlockbuster_t + v_9 \cdot Season_t + v_{10} \cdot Christmas_t + v_{11} \\
 & \cdot SchoolVacations_t + v_{12} \cdot WC2006_t + v_{13} \cdot t + \varepsilon_t.
 \end{aligned} \quad (4)$$

We estimate our model by taking the logs, according to Equation (5):

$$\begin{aligned}
 \ln(r_t/(1-r_t)) = & v_t = v_0 + v_1 \cdot ST2 + v_2 \cdot ST3 + v_3 \cdot AssortmentSize_t \\
 & + v_4 \cdot AssortmentSize_t^2 + v_5 \cdot ADEC_t + v_6 \cdot FractionNewCustomers_t \\
 & + v_7 \cdot TotalQuality_t + v_8 \cdot QualityBlockbuster_t + v_9 \cdot Season_t + v_{10} \\
 & \cdot Christmas_t + v_{11} \cdot SchoolVacations_t + v_{12} \cdot WC2006_t + v_{13} \cdot t + \varepsilon_t.
 \end{aligned} \quad (5)$$

Distribution of Demand Across Products

We model the effects of increasing assortment sizes and different search technologies on the demand distribution across products. As a first step, we calculate the weekly demand per 1,000 customers for each product and rank the product by descending demand. In line with existing literature [6, 10, 21], we specify a log-linear relationship between the demand for each product and its corresponding demand ranks:

$$\ln(f_{j,t}) = a_t + b_t \cdot \ln(x_{j,t}) + \varepsilon_{j,t}, \quad (6)$$

where $f_{j,t}$ is the demand for product j in week t per 1,000 active customers; $x_{j,t}$ is the demand rank of product j in week t ; a_t, b_t are parameters of the demand distribution function in week t ; $\varepsilon_{j,t}$ is the residual for product j in week t ; and J_t is the number of products in the assortment in week t , that is, assortment size in week t .

The parameter a_t can thus be interpreted as the overall demand in week t , whereas the parameter b_t measures how quickly the demand per product falls as the demand rank increases [6]. We analyze the effect of the drivers on the distribution of demand by linking the parameters a_t and b_t with the drivers described in Equations (7) and (8).

In addition to dummy variables for the different search technologies, we include parameters to account for search tool-specific assortment size effects. Beyond the parallel shifts in the intercepts a_0 (a_1 and a_2) and b_0 (b_1 and b_2), this modeling approach considers possible changes in the slope of the assortment size parameter (a_4 , a_5 , b_4 , and b_5) through differences in the search technologies. We engage in the latter analysis to check if the effects of an increasing assortment size on the demand distribution depend on the particular search technology:

$$\begin{aligned} a_t = & a_0 + a_1 \cdot ST2_t + a_2 \cdot ST3_t + a_3 \cdot AssortmentSize_t + a_4 \cdot ST2_t \\ & \cdot AssortmentSize_t + a_5 \cdot ST3_t \cdot AssortmentSize_t + a_6 \cdot ADEC_t + a_7 \\ & \cdot FractionNewCustomers_t + a_8 \cdot TotalQuality_t + a_9 \cdot QualityBlockbuster_t \quad (7) \\ & + a_{10} \cdot Season_t + a_{11} \cdot Christmas_t + a_{12} \cdot SchoolVacations_t + a_{13} \\ & \cdot WC2006_t + a_{14} \cdot t + \varepsilon_t \end{aligned}$$

$$\begin{aligned} b_t = & b_0 + b_1 \cdot ST2_t + b_2 \cdot ST3_t + b_3 \cdot AssortmentSize_t + b_4 \cdot ST2_t \\ & \cdot AssortmentSize_t + b_5 \cdot ST3_t \cdot AssortmentSize_t + b_6 \cdot ADEC_t + b_7 \\ & \cdot FractionNewCustomers_t + b_8 \cdot TotalQuality_t + b_9 \cdot QualityBlockbuster_t \quad (8) \\ & + b_{10} \cdot Season_t + b_{11} \cdot Christmas_t + b_{12} \cdot SchoolVacations_t + b_{13} \\ & \cdot WC2006_t + b_{14} \cdot t + \varepsilon_t \end{aligned}$$

To determine the parameters a_t and b_t of the demand distribution function, we analyze the demand ranks of all the products across the 111 weeks, which yields 62,260 observations. We estimate all the parameters by generalized least squares after inserting Equations (7) and (8) into Equation (6). Specifically, we estimated Prais-Winsten regressions, which assume AR(1) error structures but no dynamics and which is equivalent to the use of full-information maximum likelihood for an AR(1) model.

Empirical Study Results

Demand per Customer

WHITE'S TEST INDICATES HETEROSKEDASTICITY ($p < 0.05$), so we estimate the model in Equation (2) with robust standard errors and attain the results in Table 4. Specifically, search technology has no significant effect on the demand per customer. Yet we observe positive linear and negative squared assortment size parameters, which indicate rising demand per customer as a result of increasing assortment size, on a diminishing scale. Figure 3 illustrates our results graphically, if we vary only the assortment size variables in our model.

An increasing assortment size leads to increasing demand per customer and, if the products sell at a positive margin, additional profit for retailers. As Figure 3 shows, demand rises from about 600 films per week per 1,000 customers with an assortment

Table 4. Drivers of Weekly Demand per Customer

	Coefficient	Standard error	<i>t</i>	<i>P</i> > <i>t</i>
β_0 [Constant]	155.16	212.14	0.73	0.47
β_1 [ST2]	-32.93	34.27	-0.96	0.34
β_2 [ST3]	-26.70	41.97	-0.64	0.53
β_3 [AssortmentSize]**	0.56	0.27	2.05	0.04
β_4 [AssortmentSize ²]**	-0.00	0.00	-3.00	0.00
β_5 [ADEC]	-0.40	1.51	-0.26	0.79
β_6 [FractionNewCustomers]	849.81	783.63	1.08	0.28
β_7 [TotalQuality]*	6.20e-06	3.54e-06	1.75	0.08
β_8 [QualityBlockbuster]	1.95	1.49	1.30	0.20
β_9 [Season]**	-28.59	13.32	-2.15	0.03
β_{10} [Christmas]	-16.25	24.33	-0.67	0.51
β_{11} [SchoolVacations]	22.18	22.03	1.01	0.32
β_{12} [WC2006]**	-80.80	37.68	-2.14	0.04
β_{13} [<i>t</i>]	-0.62	1.95	-0.32	0.75
Observations	111			
<i>F</i> (14,97)	1,263.95			
<i>R</i> ²	0.45			

** *p* < 0.05; * *p* < 0.1.

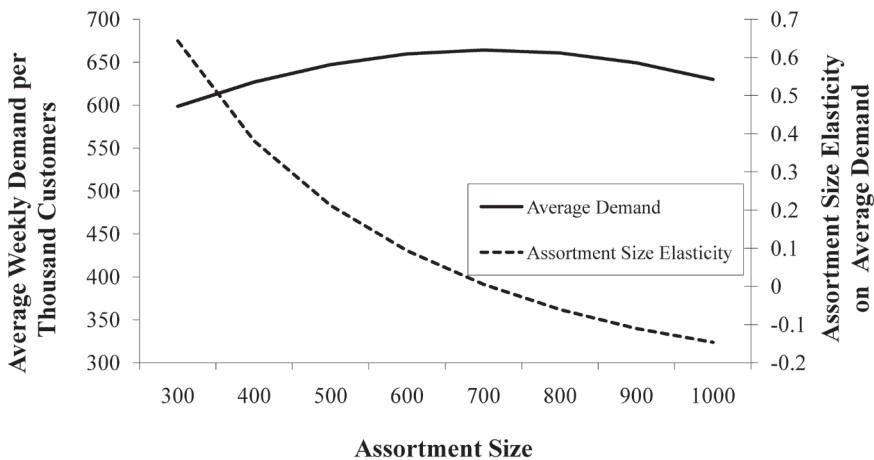


Figure 3. Average Weekly Demand per 1,000 Customers and Assortment Size Elasticity for Different Assortment Sizes

size of 300 films to about 650 films per week per 1,000 customers with an assortment size of 900 films. Furthermore, the positive effect of assortment size on demand per customer decreases with higher assortment size. This assortment size elasticity in Figure 3 reveals that the positive effect of assortment size on demand converges to zero when approximately 700 films are on offer. This convergence also illustrates that increasing the assortment size beyond 700 films generates no additional demand and

therefore no additional profit for the retailer; that is, our results support prior findings that indicate higher assortment sizes lead to additional demand [3, 26, 28, 34]. The decreasing marginal effect of assortment size on demand appears to be driven by saturated consumption levels.

The insignificant value of the parameter that captures the fraction of new customers indicates that a varying composition of the customer base does not influence demand per customer. However, quality of the overall assortment increases demand per customer. An additional film in the assortment that generated approximately \$160 million in box office revenues would increase weekly demand by 1 sale per 1,000 customers. Moreover, the 2006 soccer World Cup (captured by the variable *WC2006*) dramatically lowered demand, and consumption during summer (captured by the variable *Season*) is typically higher than during winter—probably because popular television programming is not original (i.e., repeats) in the summer, which may lead to greater substitution with video-on-demand consumption.

Share of Purchased Products

Again, White's test indicates heteroskedasticity ($p < 0.1$) in the data set, so we use robust standard errors. As we illustrate in Table 5, the results of the model in Equation (4), which analyzes the drivers of the share of purchased products from the assortment, reveal that differences in search technology do not significantly affect the share of purchased products. Increasing the assortment size positively affects this share but on a diminishing scale, as indicated by the significant negative parameter of the squared assortment size. Figures 4 and 5 illustrate our results graphically when we only vary the assortment size. Figure 4 indicates the effect of increasing the assortment, and Figure 5 illustrates the effects on the count of purchased products sold at least once.

In Figure 4, the ratio between the number of purchased products and assortment size first increases and then decreases when the assortment size reaches 1,000 products, whereas Figure 5 suggests that this relation leads to an almost linear relationship between the count of purchased products sold at least once and assortment size. As more products are added to the assortment, additional products face less demand, and this nonmonotonic relationship might explain the varying results in previous studies; that is, adding more products to the assortment can lead to either a longer or a shorter tail, depending on the initial assortment size.

The composition of the customer base also influences the share of the purchased products; a high fraction of new customers increases the probability that products from otherwise unsold niches will encounter demand. Moreover, customers consume more niche products and thus increase the share of purchased products if the overall assortment quality is high (positive *TotalQuality* value). This quality justifies customers' efforts to find less well-known products. On average, search costs seem to decline during school vacations, as indicated by a higher demand for niches during these times. Customers invest more time in finding appropriate niches, which then increases the share of products purchased.

Table 5. Drivers of Weekly Share of Purchased Products at Assortment Size

	Coefficient	Standard error	<i>t</i>	<i>P</i> > <i>t</i>
v_0 [Constant]**	-1.76	0.66	-2.68	0.01
v_1 [ST2]	-0.42	0.03	-1.26	0.21
v_2 [ST3]	-0.29	0.34	-0.85	0.40
v_3 [AssortmentSize]**	0.01	0.00	4.80	0.00
v_4 [AssortmentSize ²]**	-3.07e-06	7.66e-07	-4.01	0.00
v_5 [ADEC]	-0.00	0.00	-0.99	0.33
v_6 [FractionNewCustomers]**	1.27	0.56	2.26	0.03
v_7 [TotalQuality]**	6.09e-08	8.82e-09	6.91	0.00
v_8 [QualityBlockbuster]*	-0.00	0.00	-1.88	0.06
v_9 [Season]	-0.02	0.04	-0.51	0.61
v_{10} [Christmas]	-0.04	0.07	-0.58	0.56
v_{11} [SchoolVacations]**	0.07	0.02	3.51	0.00
v_{12} [WC2006]	-0.09	0.13	-0.72	0.47
v_{13} [<i>t</i>]**	-0.02	0.01	-4.28	0.00
Observations	111			
<i>F</i> (14, 97)	30.50			
<i>R</i> ²	0.49			

** $p < 0.05$; * $p < 0.1$.

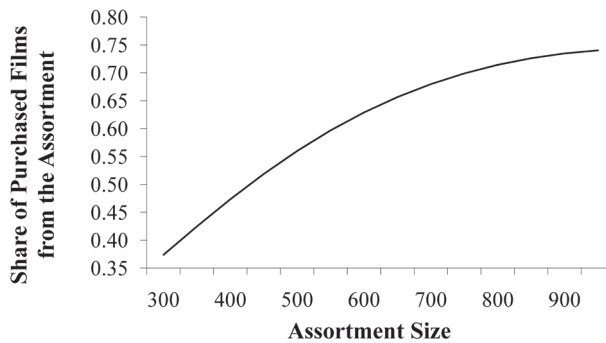


Figure 4. Share of Purchased Products for Different Assortment Sizes

Demand Distribution Across Products

In Table 6, we list the results pertaining to the effects of the drivers on the distribution and concentration of demand, as outlined in the model in Equation (6).

A growing assortment size in the presence of search tool 1 decreases parameter a_i and increases parameter b_i . This result indicates that an increase in assortment size leads to a shift in demand from blockbusters to niches, that is, a fatter and longer tail of the demand distribution: blockbusters lose importance. In Figure 6, we also illustrate the demand distribution if we vary the assortment size in the presence of search tool 1.

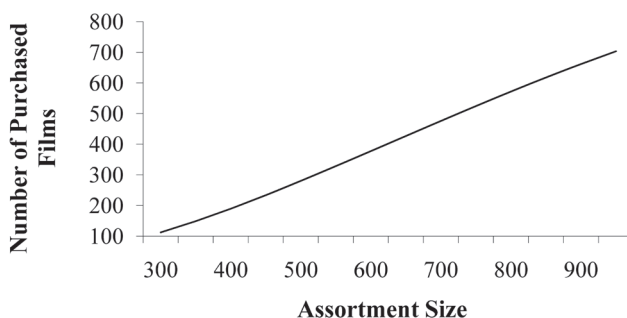


Figure 5. Number of Purchased Products for Different Assortment Sizes

Our results in Table 6 show that changes in search technology significantly affect parameters a_i and b_i , but they do so differently. Search tool 2 decreases parameter b_i through an interaction effect with assortment size, whereas parameter a_i remains unchanged. With a growing assortment size, sales of blockbusters do not drop as much as they do with search tool 1, but demand for products that sell moderately are subject to greater substitution by demand for new products in the assortment. In the presence of search tool 2, customers use the blockbuster recommendations more frequently as the assortment size grows. This finding is as expected: With too much variety, customers simply follow blockbuster recommendations to avoid lengthy search processes. Therefore, blockbusters do not lose as much as products in the middle, and newly added products also confront demand. The negative parameter b_i and the nonsignificant parameter a_i indicate substitution effects. However, search tools do not generate additional consumption.

With Figure 7, we depict the demand distribution function if we vary just the assortment size in the presence of search tool 2. It suggests a typical substitution scenario: products already in the assortment get substituted by newly added products. Because search tool 2 slightly favors blockbusters, blockbuster sales suffer less than those of products in the middle. However, the differences compared with search tool 1 are not visible in the log scale charts in Figure 7.

The effects of search tool 3 compared with search tool 1 (dummy omitted and used as the baseline) are more complex. Search tool 3 has a negative effect on parameter a_i , though the parameter remains positive because of the interaction effect with the assortment size (i.e., the assortment size is always greater than 1,000 in the presence of search tool 3). Therefore, increasing the assortment size with search tool 3 increases a_i more than proportionally, which induces higher demand (if b_i stays constant) because the entire demand distribution curve lifts upward. Yet search tool 3 also has a positive effect on parameter b_i , that is, on the slope of the distribution curve. The negative interaction effect between assortment size and search tool 3 overcompensates for the positive influence of search tool 3 on b_i , so the effect of search tool 3 on the demand distribution is comparable with the effect of search tool 2. Some demand for moderate-selling products shifts toward demand for blockbusters and niches. If the assortment size is great enough, demand for blockbusters can even increase. Our empirical study

Table 6. Drivers of Weekly Demand per Customer and Product

	Coefficient	Standard error	t	$P > t $
a_0 [Constant]**	7.4338	0.2911	25.54	0.00
b_0 [$\ln(x_{jt})$]**	-1.7189	0.0570	-30.16	0.00
a_1 [ST2]	0.3317	0.4626	0.72	0.47
b_1 [ST2 · $\ln(x_{jt})$]	0.0938	0.0863	1.09	0.28
a_2 [ST3]**	-1.4239	0.3645	-3.91	0.00
b_2 [ST3 · $\ln(x_{jt})$]**	0.5153	0.0685	7.53	0.00
a_3 [AssortmentSize]**	-0.0010	0.0003	-2.98	0.00
b_3 [AssortmentSize · $\ln(x_{jt})$]**	0.0004	0.0001	5.19	0.00
a_4 [ST2 · AssortmentSize]	-0.0002	0.0005	-0.33	0.74
b_4 [ST2 · AssortmentSize · $\ln(x_{jt})$]**	-0.0003	0.0001	-2.55	0.01
a_5 [ST3 · AssortmentSize]**	0.0015	0.0004	3.39	0.00
b_5 [ST3 · AssortmentSize · $\ln(x_{jt})$]**	-0.0007	0.0001	-7.74	0.00
a_6 [ADEC]**	-0.0149	0.0021	-7.11	0.00
b_6 [ADEC · $\ln(x_{jt})$]**	0.0021	0.0004	5.57	0.00
a_7 [FractionNewCustomers]**	1.9082	0.2222	8.59	0.00
b_7 [FractionNewCustomers · $\ln(x_{jt})$]**	-6.71e-06	3.16e-06	-2.12	0.03
a_8 [TotalQuality]**	-1.20e-08	3.58e-09	-3.36	0.00
b_8 [TotalQuality · $\ln(x_{jt})$]**	6.52e-09	6.74e-10	9.68	0.00
a_9 [QualityBlockbuster]**	0.0023	0.0010	2.37	0.02
b_9 [QualityBlockbuster · $\ln(x_{jt})$]	-0.0003	0.0002	-1.58	0.12
a_{10} [Season]	-0.0331	0.0251	-1.32	0.19
b_{10} [Season · $\ln(x_{jt})$]**	-0.0101	0.0046	-2.19	0.03
a_{11} [Christmas]**	-0.0953	0.0231	-4.13	0.00
b_{11} [Christmas · $\ln(x_{jt})$]**	0.0247	0.0044	5.62	0.00
a_{12} [SchoolVacations]	0.0143	0.0136	1.05	0.29
b_{12} [SchoolVacations · $\ln(x_{jt})$]**	0.0052	0.0025	2.06	0.04
a_{13} [WC2006]**	-0.2388	0.0348	-6.86	0.00
b_{13} [WC2006 · $\ln(x_{jt})$]	0.0045	0.0068	0.67	0.51
a_{14} [t]	0.0000	0.0026	0.01	0.99
b_{14} [t · $\ln(x_{jt})$]	-0.0003	0.0005	-0.59	0.55
Observations	62,260			
$F(29, 62,230)$	1,941.35			
R^2	0.68			

** $p < 0.05$.

thus supports the analytical findings provided by Fleder and Hosanagar [19], who also showed that search technology increases demand diversity.

As demand becomes more concentrated and blockbusters experience additional demand, many new products also face demand, with simultaneously lengthening tails. In Figure 8, we depict the resulting demand distribution function if we vary assortment

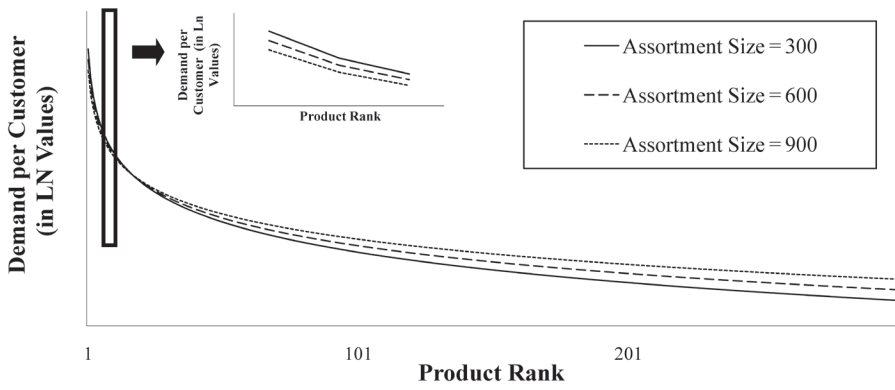


Figure 6. Influence of Assortment Size on Distribution of Demand with Search Tool 1

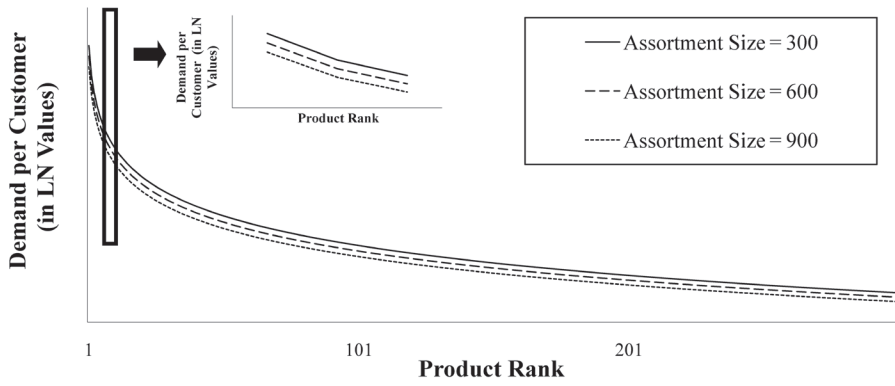


Figure 7. Influence of Assortment Size on Distribution of Demand with Search Tool 2

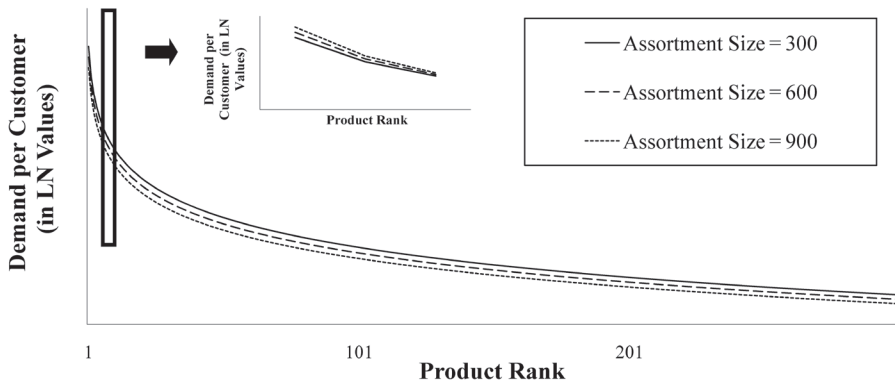


Figure 8. Influence of Assortment Size on Distribution of Demand with Search Tool 3

size in the presence of search tool 3. If customers have more options from which to choose, the product recommendations offered by search tool 3 increase demand for blockbusters.

In summary, we observe two different effects with growing assortment sizes. First, in the presence of search tool 1, we observe the emergence of a typical long tail demand distribution pattern such that demand becomes more evenly distributed and blockbusters lose importance. The share of purchased products indicates that the tail lengthens when additional products are on offer, which increases the average demand per customer, although on a diminishing scale. Too many products may decrease the average demand per customer, though. Additional data with even higher assortment sizes would be required to test whether assortment sizes that are too big eventually lead to decreasing demand in online environments.

Second, slight changes in the search technology do not necessarily minimize the importance of blockbusters. With search tools 2 and 3, customer preferences move toward the extremes: They choose either extreme niches or blockbusters, in line with Dellarocas et al.'s [13] finding that consumers often contribute niche reviews and are highly likely to contribute blockbuster reviews but rarely review films that are not blockbusters or niches.

Products that do not represent either blockbusters or niches suffer lost demand. The first finding might indicate variety-seeking behavior by customers [26] or extremely heterogeneous customer preferences, but the finding that customers tend to stick with blockbusters, regardless of the number of available niches, is in line with the “superstar phenomenon” [20, 32, 33, 36]. As Rosen [36] states, lesser talent is a poor substitute for greater talent, so people have no reason to buy a second-, third-, or 1,000th-best video if the best is available. Furthermore, communication research provides evidence that customers tend to choose their preferred products when the number of alternatives increases [32, 33].

Another explanation for the superstar phenomenon, by Frank and Cook [20], suggests that customers value blockbusters because their consumption facilitates social interactions with other people. Reading a niche book or watching a niche film that has not been consumed by other members of a social network cannot facilitate social interaction like consumption of blockbusters can. Our findings also align with research by Elberse and Oberholzer-Gee [17], who detect a higher concentration of blockbuster demand. Our study illustrates that small changes in search technologies may have significant effects on the distribution of demand.

Heavy users in particular shift demand from blockbusters to niches, in line with Elberse's empirical generalization of McPhee's theory of exposure [30]—namely, “a disproportionately large share of the audience for obscure products consists of relatively heavy consumers” [16, p. 91]. New customers increase demand for blockbusters or buy extreme niches that did not previously face demand, which creates a longer tail. However, moderately selling products get substituted.

In terms of the assortment quality, the results reveal that if the quality of the top 100 films is high (measured by box office sales), demand for blockbusters increases, which seems logical. Yet high quality in the overall assortment actually decreases

blockbuster sales. Perhaps customers are willing to spend more time locating appropriate niches if they expect better-quality niche products, in line with the theory of consumer search behavior [35]. A similar effect during school vacations likely stems from greater available time, which reduces consumers' search costs. Customers search longer, which leads them to niche products, and the demand distribution becomes more evenly distributed, with a longer tail.

During Christmas, there also seems to be greater demand for blockbusters, perhaps because families decide to watch films together and need an option that appeals to the entire family. Finally, during the 2006 World Cup and the winter, the results depict a loss in total demand, as captured by the downward parallel shift of the demand distribution curve.

The Gini Coefficient as a Measure of Concentration of Demand

Existing long tail research [6] often relies on the Gini coefficient to measure and illustrate changes in the distribution of demand. Similarly, we analyze the effects of our two main drivers on the Gini coefficient, a value based on the Lorenz curve that measures the ratio of the area between the line of equality and the Lorenz curve. It ranges from 0 to 1, and a low Gini coefficient indicates a more equal distribution (0 = complete equality), whereas higher Gini coefficients indicate more unequal distributions (1 = complete inequality). However, the Gini coefficient demands careful interpretation because an infinite number of Lorenz curves yield the same Gini coefficients.

In our data set, the weekly Gini coefficients range from 55.3 percent to 73.2 percent, with a mean of 63.7 percent. According to Figure 9, it increases over time, which indicates a higher concentration of demand. Figure 9 also illustrates that the introduction of search tool 2 raises the Gini coefficient, but the changes in the Gini coefficients between search tool 2 and search tool 3 are quite small.

To examine the influence of the drivers on the Gini coefficient further, we consider the 12 weeks before and after the introduction of each search tool and estimate ordinary least square (OLS) regressions with robust standard errors. In Table 7, we provide the effects for the period around the introduction of search tool 2, which indicate that the introduction of the new search tool significantly increased the Gini coefficient. School vacations lower the sales concentration, in support of our findings from the previous analyses. The assortment size does not have any significant effects on the Gini coefficient in this short-term analysis, though, perhaps because it does not change substantially during this short observation period.

We repeat this analysis for search tool 3 (observation period 71–94) and contrast the results with those for search tool 2. Search tool 3 has no significant effect on the Gini coefficient, and if we compare search tool 3 with search tool 1, we observe a significant positive influence on the Gini coefficient ($p < 0.01$). Search tools 2 and 3 thus have similar effects on the concentration of sales. The results from the second observation period also reveal that new customers increase the sales concentration ($p < 0.05$) and seemingly lengthen the tail, though a high fraction of new customers prefers blockbusters, which more than compensates for the effect of sporadic niche

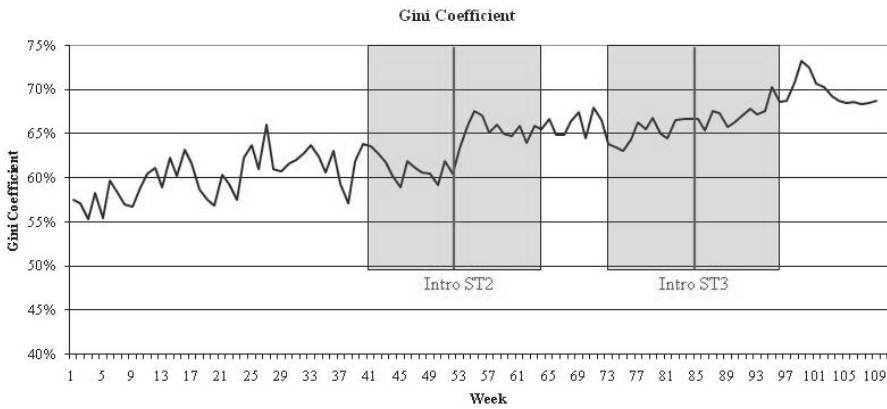


Figure 9. Development of the Gini Coefficient over Time

Table 7. Drivers of the Gini Coefficient (Observation Period 41–64)

	Coefficient	Standard error	<i>t</i>	<i>P</i> > <i>t</i>
g_0 [Constant]**	0.89	0.37	2.41	0.03
g_1 [ST2]**	0.09	0.03	2.57	0.03
g_2 [AssortmentSize]	-0.00	0.00	-0.96	0.36
g_3 [AssortmentSize ²]	5.09e-07	6.19e-07	0.82	0.43
g_4 [ADEC]	0.00	0.00	1.23	0.24
g_5 [FractionNewCustomers]	-0.05	0.20	-0.23	0.83
g_6 [TotalQuality]	-1.64e-09	2.74e-09	-0.60	0.56
g_7 [QualityBlockbuster]	-0.00	0.00	-1.18	0.26
g_8 [Christmas]	0.01	0.01	0.47	0.65
g_9 [SchoolVacations]**	-0.04	0.01	-3.41	0.01
g_{10} [WC2006]	0.01	0.01	0.90	0.39
g_{11} [<i>t</i>]	0.00	0.00	0.28	0.78
Durbin–Watson	2.126098			
Observations	24			
<i>F</i> (10, 12)	10.94			
<i>R</i> ²	0.91			

Notes: The Gini coefficient measures the concentration of demand across products. ** $p < 0.05$.

sales on the Gini coefficient. Changes in assortment size do not influence the Gini coefficient during this short observation period.

Summary and Conclusions

EXISTING LITERATURE CLAIMS THAT ASSORTMENT SIZES and search technologies drive the long tail phenomenon, that is, a scenario in which niche products gain a significant share of overall demand. To address the lack of empirical evidence for this claim, we use a unique data set with 843,922 purchases by 143,939 customers of a monopolistic

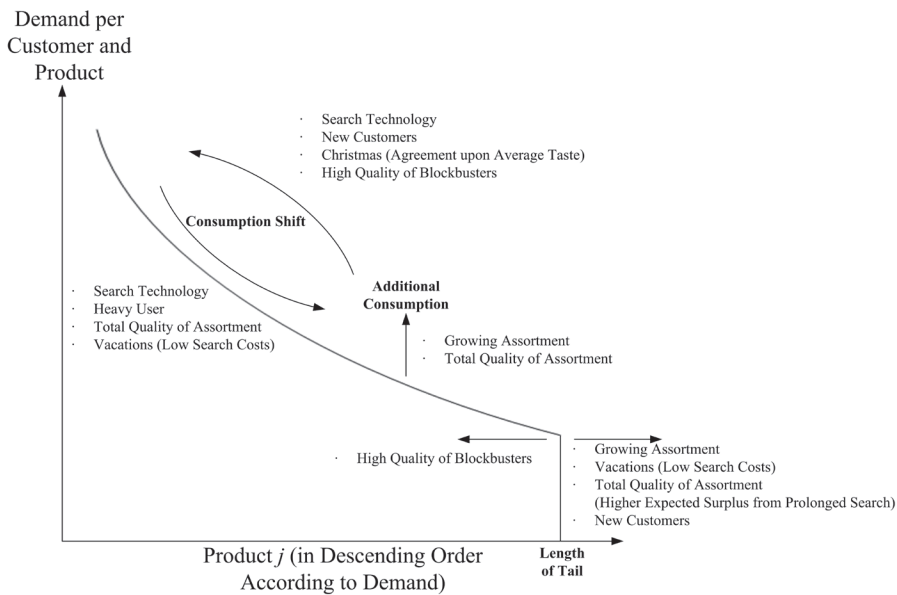


Figure 10. Summary of Empirical Study Results

video-on-demand operator, observed in 111 weeks after the service launch, such that we can measure empirically the effect of both drivers on several measures of the long tail phenomenon, including demand per customer, share of products purchased, distribution of demand across products, and concentration of demand. We summarize our findings in Figure 10.

Our empirical study reveals that a growing assortment leads to greater demand per customer, although on a diminishing scale. This result supports previous findings that have shown increasing assortment sizes allow customers to find more products that fit their preferences [3, 28, 34] and that customers value diversity and exhibit variety-seeking behavior [26]. Retailers can benefit directly by adding more products to their assortment. However, we also find that an assortment beyond a certain size no longer increases demand per customer. The diminishing scale implies substitution between products.

Differences in search technology do not influence the demand per customer, but seasonality and events, such as the 2006 World Cup, do. Not surprisingly, a higher-quality assortment also increases demand.

Our results further reveal that a growing assortment size increases the share of purchased products initially, but once the assortment size reaches 1,000 products, shares decrease. We observe a nearly linear relation between the absolute count of products sold at least once and the assortment size, which suggests that products added to the assortment beyond the maximum number of products observed in our data should experience demand. This result indicates that tails for online businesses lengthen with growing assortments, although we do not find an impact of search technology on the length of the tail. Our results are obviously limited to the technology in our data set,

so further research should consider additional, more sophisticated systems that might incorporate functionalities such as collaborative filtering.

The length of the tail also depends on the proportion of new customers. Compared to existing customers, new customers try obscure products more often. Some of these new customers might even have become customers in order to purchase a particular product that otherwise would have been hard to find. If the quality of the overall assortment is high, customers invest more time searching (in line with information search behavior theory) and may choose a product from distant niches, which leads to a longer tail. However, we observe a shorter tail if the top 100 films offer high quality. Customers come across these blockbusters and decide to purchase them (perhaps because they have heard about their quality) before they invest time in browsing niches.

The effects of the two main drivers on the distribution of demand are as follows: growing assortment size does not necessarily lead to the end of the “blockbuster era,” as predicted by Anderson [1]. There are strong interaction effects between assortment size and search technology. Demand in blockbusters may even increase with assortment size such that blockbusters benefit from this development. At the same time, we find that search technologies that recommend blockbusters (e.g., search tools 2 and 3 in our data set) become more important when the assortment has reached a certain limit. They can slow the substitution of blockbusters (search tool 2) or lead to higher sales of blockbusters (search tool 3) when the assortment is very large and takes substantial time to browse. Products that are not top sellers or niches lose demand, though, and get substituted for blockbusters or very obscure products. If search technology lowers search costs equally and does not favor certain products (i.e., search tool 1), fatter and longer tails occur, along with substitution of blockbusters with growing assortment size, in line with existing analytical research [23].

Our results thus confirm the strong influence of search technologies on the demand distribution. On the one hand, search functionalities, such as additional filters, can lead to a shift in demand from blockbusters to niches. On the other hand, systems based on recommendations may shift demand from niches to blockbusters. Whether search technologies favor niche or blockbuster products, demand might be reversed by minor adjustments to the search technology, as our data illustrate.

Our results also deliver interesting insights regarding the distribution of demand: new customers start with blockbusters or very obscure products, which shifts demand to the extremes of the demand distribution. However, such behavior also leads to a higher overall concentration of blockbuster sales. A shift toward blockbusters is observable if the blockbusters are of high quality as well as during Christmas week, when families frequently watch television together and apparently try to agree on a film that is acceptable to all family members.

Heavy users embrace content on the right side of the demand distribution; we observe a shift toward the right end of the distribution if the customer base consists of many heavy users. This result is in line with Elberse [16], who also finds evidence that niche demand is mainly generated by heavy users. However, our results reveal that heavy users do not drive the length of the demand distribution but, rather, stick to films that were on offer before. We observe a shift of demand toward niche products

if the overall quality of the assortment is high or the search costs are low on average, such as during school vacations. In Figure 10, we summarize these insights.

Overall, the interplay between demand in electronic channels and drivers such as assortment size and search technology is more complex than expected, which reinforces the need for continued research in this area.

Acknowledgments: The authors gratefully acknowledge the many helpful comments from Christine Ebling, Sönke Albers, Christian Barrot, and Michel Clement. In addition, Editor-in-Chief Vladimir Zwass and three anonymous referees provided many helpful and excellent suggestions. This research was supported by the E-Finance Lab Frankfurt.

NOTE

1. These figures came from the following Web sites: www.chart-surfer.de, www.mediabiz.de, and www.the-numbers.com.

REFERENCES

1. Anderson, C. *The Long Tail: Why the Future of Business Is Selling Less of More*. New York: Hyperion, 2006.
2. Bakos, Y. Reducing buyer search costs: Implications for electronic marketplaces. *Management Science*, 43, 12 (1997), 1676–1692.
3. Baumol, W.J., and Ide, E.A. Variety in retailing. *Management Science*, 3, 1 (1956), 93–101.
4. Boatwright, P., and Nunes, C.N. Reducing assortment: An attribute-based approach. *Journal of Marketing*, 65, 3 (2001), 50–63.
5. Broniarczyk, S.M.; Hoyer, W.D.; and McAlister, L. Consumers' perceptions of the assortment offered in a grocery category: The impact of item reduction. *Journal of Marketing Research*, 35, 2 (1998), 166–176.
6. Brynjolfsson, E.; Hu, Y.J.; and Simester, D. Goodbye Pareto principle, hello long tail: The effect of search costs on the concentration of product sales. MIT Working Paper, Cambridge, 2006.
7. Brynjolfsson, E.; Hu, Y.J.; and Smith, M.D. Consumer surplus in the digital economy: Estimating the value of increased product variety at online booksellers. *Management Science*, 49, 11 (2003), 1580–1596.
8. Brynjolfsson, E.; Hu, Y.J.; and Smith, M. From niches to riches: The anatomy of the long tail. *Sloan Management Review*, 47, 4 (2006), 67–71.
9. Chernev, A. When more is less and less is more: The role of ideal point availability and assortment in consumer choice. *Journal of Consumer Research*, 30, 2 (2003), 170–183.
10. Chevalier, J., and Golsbee, A. Measuring prices and price competition online: Amazon and Barnes and Noble. *Quantitative Marketing and Economics*, 1, 2 (2003), 203–222.
11. Clemons, E.K. How information changes consumer behavior and how consumer behavior determines corporate strategy. *Journal of Management Information Systems*, 25, 2 (Fall 2008), 13–40.
12. Clemons, E.K.; Gao, G.G.; and Hitt, L.M. When online reviews meet hyperdifferentiation: A study of the craft beer industry. *Journal of Management Information Systems*, 32, 2 (Fall 2006), 149–171.
13. Dellarocas, C.; Gao, G.; and Narayan, R. Are consumers more likely to contribute online reviews for hit products or niche products? *Journal of Management Information Systems*, 27, 2 (Fall 2010), 127–158.
14. Dhar, R. Consumer preference for a no-choice option. *Journal of Consumer Research*, 24, 2 (1997), 215–231.

15. Dreze, X.; Hoch, S.J.; and Purk, M.E. Shelf management and space elasticity. *Journal of Retailing*, 70, 4 (1994), 301–326.
16. Elberse, A. Should you invest in the long tail? *Harvard Business Review*, 86, 7–8 (2008), 88–96.
17. Elberse, A., and Oberholzer-Gee, F. Superstars and underdogs: An examination of the long tail phenomenon in video sales. Harvard Business School Working Paper Series no. 07-15, Boston, 2006.
18. Fleder, D., and Hosanagar, K. Recommender systems and their impact on sales diversity. In *Proceedings of the 8th ACM Conference on Electronic Commerce*. New York: ACM Press, 2007 (available at <http://portal.acm.org/citation.cfm?id=1250939/>).
19. Fleder, D., and Hosanagar, K. Blockbuster culture's next rise or fall: The impact of recommender systems on sales diversity. *Management Science*, 55, 5 (2009), 697–712.
20. Frank, R.H., and Cook, P.J. *The Winner-Take-All Society*. New York: Free Press, 1995.
21. Ghose, A., and Gu, B. Search costs, demand structure and long tail in electronic markets: Theory and evidence. Net Institute Working Paper no. 06-19, New York, 2006.
22. Hervas-Drane, A. Word of mouth and recommender systems: A theory of the long tail. Net Institute Working Paper no. 07-41, New York, 2007.
23. Hinz, O., and Eckert, J. The impact of search and recommendation systems on sales in electronic commerce. *Business & Information System Engineering*, 2, 2 (2010), 67–77.
24. Huffman, C., and Kahn, B.E. Variety for sale: Mass customization or mass confusion. *Journal of Retailing*, 74, 4 (1998), 491–513.
25. Iyengar, S.S., and Lepper, M.R. When choice is demotivating: Can one desire too much of a good thing? *Journal of Retailing and Consumer Services*, 2, 3 (2000), 139–148.
26. Kahn, B.E. Consumer variety seeking among goods and services. *Journal of Retailing and Consumer Services*, 2, 3 (1995), 139–148.
27. Kahn, B.E. Dynamic relationships with customers: High-variety strategies. *Journal of the Academy of Marketing Science*, 26, 1 (1998), 45–53.
28. Lancaster, K.J. The economics of product variety: A survey. *Marketing Science*, 9, 3 (1990), 189–206.
29. Lehmann, D.R. Customer reaction to variety: Too much of a good thing? *Journal of the Academy of Marketing Science*, 26, 1 (1998), 62–65.
30. McPhee, W.N. *Formal Theories of Mass Behavior*. New York: Free Press of Glencoe, 1963.
31. Mooney, R.J., and Roy, L. Content-based book recommending using learning for text categorization. In *Proceedings of the 5th ACM Conference on Digital Libraries*. New York: ACM Press, 2000, pp. 195–204.
32. Napoli, P.M. Deconstructing the diversity principle. *Journal of Communication*, 49, 4 (1999), 7–34.
33. Neumann, W.R. *The Future of the Mass Audience*. New York: Cambridge University Press, 1991.
34. Quelch, J.A., and Kenny, D. Extend profits, not product lines. *Harvard Business Review*, 72, 5 (1994), 153–160.
35. Reinartz, W.J., and Kumar, V. On the profitability of long lifetime customers: An empirical investigation and implications for marketing. *Journal of Marketing*, 64, 4 (2000), 17–35.
36. Rosen, S. The economics of superstars. *American Economic Review*, 71, 5 (1981), 845–858.
37. Sela, A.; Berger, J.; and Liu, W. Variety, vice, and virtue: How assortment size influences option choice. *Journal of Consumer Research*, 35, 6 (2009), 941–951.
38. Stigler, G.J. The economics of information. *Journal of Political Economy*, 69, 3 (1961), 213–225.
39. Tucker, C., and Zhang, J. Long tail or steep tail? A field investigation into how online popularity information affects the distribution of customer choices. MIT Working Paper no. 4655-07, Cambridge, 2007.

Copyright of Journal of Management Information Systems is the property of M.E. Sharpe Inc. and its content may not be copied or emailed to multiple sites or posted to a listserv without the copyright holder's express written permission. However, users may print, download, or email articles for individual use.