

2007 ASSA Conference Paper
Harvard Business School Working Paper Series, No. 07-015

**Superstars and Underdogs:
An Examination of the Long Tail Phenomenon in Video Sales**

Anita Elberse
Harvard Business School
Boston, MA 02163, USA
Phone: 617 495 6080
Fax: 617 496 5853
Email: aelberse@hbs.edu

Felix Oberholzer-Gee
Harvard Business School
Boston, MA 02163, USA
Phone: 617 495 6770
Fax: 617 495 0355
Email: foberholzer@hbs.edu

(Corresponding Author)

Preliminary Version

September 5, 2006

Superstars and Underdogs: An Examination of the Long Tail Phenomenon in Video Sales

Abstract

The rise of online channels facilitates the distribution of a wide range of products and services. Academics and industry observers agree that online distribution will fundamentally change the number and variety of products that consumers purchase. However, there is sharp disagreement about what type of change will occur. Proponents of the “long tail” idea argue that a sharp increase in the variety of products offered through online channels will fuel a shift in consumption away from hits to a much larger number of lower-selling niche products. While the long-tail view predicts an increase in the heterogeneity of consumption patterns, the well-known superstar effect promises the exact opposite. As consumers have access to their favorite content wherever they are whenever they demand it, consumption patterns will become more, not less uniform, this theory predicts.

To shed light on this debate, we study the distribution of revenues across products in the context of the U.S. home video industry for the 2000 to 2005 period. We find superstar and long-tail effects in home video sales, but each effect comes with a twist. There is a long-tail effect in that the number of titles that sell only a few copies every week increases almost twofold during our study period. But at the same time, the number of non-selling titles rises rapidly; it is now four times as high as in 2000. Many underdogs thus in fact appear to be losers. We also find evidence of a superstar effect. Among the best-performing titles, an ever-smaller number of titles accounts for the bulk of sales. The caveat here is that today’s superstars lack the punch of earlier generations: video sales generally decrease over time across all quantiles of the sales distribution, but this effect is most pronounced among best-selling titles. Our findings have important implications for entertainment companies. Exploiting the tail might prove unprofitable if many titles do not sell at all. At the same time, producing superstars is more difficult than ever. The trends we uncover thus point to significant challenges for the entertainment industry.

Superstars and Underdogs: An Examination of the Long Tail Phenomenon in Video Sales

Industry observers and academic researchers agree that the digitization of information goods and the rise of online distribution channels will fundamentally alter the type and variety of products that consumers purchase. The views starkly differ, however, when it comes to describing what type of change will occur. On the one hand, the theory of superstars predicts that lower distribution and transaction costs will homogenize patterns of consumption (Rosen 1981). Looking for a way to spend their free time, the argument goes, families once upon a time may have considered watching “Sherlock: Undercover Dog” with relatively unknown actor Anthony Simmons. With huge libraries feeding movies on demand, however, they may well revert to superstars Brad Pitt and Angelina Jolie playing assassins in “Mr. and Mrs. Smith,” one of last year’s highest-grossing movies. Because consumers prefer more talented performers, and technology allows these performers to be everywhere at once, a few superstars and their products come to dominate the marketplace (Frank and Cook 1995).

A more recent, alternative theory predicts exactly the opposite. The proliferation of online channels will make consumption more heterogeneous, sharply increasing the variety of products on offer and fueling a shift in consumption away from a relatively small number of “hits” to a much larger number of lower-selling niche products. Dubbed the “long tail” phenomenon by Wired Magazine editor Chris Anderson (2004, 2006), this view holds that niche products will become more prevalent because online retailers are able to catalog and provide a much larger number of products than bricks-and-mortar stores. Moreover, with the help of online search and filtering tools, consumers will find it easier to search for and discover obscure products (Brynjolfsson, Hu and Smith 2003, 2006). Anderson (2006) predicts that the demand for products not available in bricks-and-mortar stores will eventually become as large as the demand for products sold through offline channels.

Which version of the future will emerge is subject to a heated debate. While many companies have started to put more emphasis on niche products, others dismiss ideas about the long tail as “Web utopian fantasies” (The Wall Street Journal 2006a, 2006b). Both superstar and long-tail proponents agree that consumers were constrained in the old world and forced to make suboptimal choices. For example, the limited portability of music meant that many consumers listened to radio stations that sometimes did, and often did not play their favorite songs. In the age of the iPod, this constraint is relaxed considerably, allowing consumers to listen to what they like wherever they are. The superstar and the long-tail view of the world differ sharply, however, in what individuals will consume once they are free to choose from a larger selection. Winner-take-all markets will exist where consumers have similar tastes and agree on what is best. In contrast, a long tail will emerge if niche products and increased variety best satisfy consumer preferences.

The superstar-long-tail debate is critical for producers' and retailers' strategies. With limited space on store shelves, producers were traditionally very discriminating about what they released, and they used intensive marketing, trying to generate a handful of hits. Producers and retailers are now faced with the question whether they need to alter how they manage their assortments.

To shed light on this debate, we study the distribution of revenues across products in the context of the U.S. home video industry for the 2000 to 2005 period. We use a comprehensive data set obtained from Nielsen VideoScan that covers weekly DVD and VHS sales for a random sample of nearly 5,500 unique titles. Video sales are a particularly suitable product category to examine because of the rapid rise of online distribution channels. In 2001, pre-recorded DVD and VHS videos accounted for \$636 million in online sales. By 2005, this number had jumped to \$1.5 billion (Euromonitor 2006). As the long-tail idea would predict, there is anecdotal evidence that this rise in sales went hand in hand with a drastic increase in the number of products on offer. For example, online retailer Amazon is adding 10,000 to 15,000 DVD titles every year, now offering more than 80,000 titles.

Our results show that the distribution of DVD and VHS video sales changed significantly during our study period. In quantile regression models, which control for the changing composition of product characteristics, we observe the emergence of a flatter tail. This flattening is consistent with the idea that studios sell fewer copies of a larger number titles. At the same time, top movies in 2005 clearly did not do as well as the most successful titles in 2000. Collectively, the sales of superstar products declined in every single year since the beginning of our study period. Are the underdogs beating the superstars? To better understand the causes of the flattening tail, we run negative binomial regression models, studying changes in the number of titles that meet particular sales thresholds. In 2005, there were 1.5 times as many titles as in 2000 that sold between 1 and 10 copies per week. Clearly, the long tail is populated by a larger number of products. However, even more strikingly, there was also a dramatic increase in the number of titles that did not sell at all. Compared to 2000, the number of titles with zero weekly sales in 2005 almost doubled, suggesting there are significant business challenges for companies that attempt to benefit from the long tail. Many underdogs, we conclude, are in fact losers. At the superstar end of things, the lower sales are achieved by a significantly smaller number of titles. While best-selling titles do not reach previous sales levels, there is a significant concentration of success on ever fewer titles. In sum, the changes in the distribution of video sales are remarkably similar to what McPhee (1963) called the double jeopardy of niche products in media markets. Increased variety not only fragments the market, with many titles reaching ever smaller audiences, these audiences are also less loyal. When a superstar comes along, they are happy to abandon the niche (Ehrenberg and Goodhardt 1990).

The remainder of the paper is organized as follows. In the first section, we review the theoretical ideas and the literature that documents the phenomena. We describe our data in section

two, outline our empirical approach in section three, and present the empirical results in section four. The fifth section offers concluding remarks.

1. Superstars and Long Tails

Proponents of the long-tail view of the world point out that, across many product categories, brick-and-mortar retailers offer consumers a rather limited variety of product choices. For example, a typical video retailer carries around 5,000 titles. With about 1,500 DVDs, choices at Wal-Mart are even more limited. In contrast, online retailer Amazon, which accounts for over a third of all DVDs sold online, offers nearly 80,000 titles and may add up to 15,000 titles every year.

The limited selection at traditional retailers reflects forces on the supply and the demand side of the market. On the supply side, sellers' transaction costs, the costs involved in distributing goods and collecting payment, are generally lower online (Alba et al 1997, Bakos 1998). Whereas adding additional titles requires more physical space in traditional retail environments, adding variety and communicating the additional choices and prices to customers is much less costly for web-based retailers. Online retailers simply do not face the shelf space constraints that bricks-and-mortar stores are forced to manage, and they can use recommendation engines and other technologies to effectively manage their libraries (Brynjolfsson, Hu and Smith 2003, 2006, Oestreicher-Singer and Sundararajan 2006). These advantages may be particularly important for information goods because these goods can be digitized and distributed almost without cost via the Internet (Bakos 1998, Bakos and Brynjolfsson 2000).

Online retailing also affects the demand side. Search costs, the costs incurred by a buyer to locate an appropriate seller and purchase a product, are lower in online channels (Alba et al 1997, Bakos 1997). Electronic marketplaces in particular lower buyers' costs to find obscure products and obtain information on prices and other product attributes, leading to reduced "fit" costs resulting from consumers making suboptimal product choices (Bakos 1997, 2001, also see Lynch and Ariely 2000).

Because the cost of offering and consuming niche products were substantial both for sellers and for buyers, traditional retail environments offered few such products. Relaxed constraints at the retail level can provide incentives for the producers of content to increase product variety. Traditionally, media providers supplied what limited store shelves could hold, and they used expensive marketing campaigns to generate a handful of hits (The Economist 2005). With increased distribution capacity, the long-tail view holds, it can become more attractive for existing media providers to produce a larger number of lower-selling titles, and for new providers to enter the market with products aimed at smaller market segments. To date, however, not much is known about consumer preferences for long-tail goods and services. As transaction and search costs continue to fall, making

the supply of niche products technically more feasible, consumers' appetite for niche products emerges as the critical factor that will foster, or possibly prevent, the emergence of a lucrative long tail.

Drawing on the large body of marketing research on the way in which consumers respond to product variety, two theoretical arguments seem consistent with the idea that overall consumer demand will surge as the number of product offerings increases. First, with a larger selection of products, customers benefit from being able to identify and purchase products that better match their preferences. A wider assortment increases the likelihood that any one consumer will find exactly what he or she is looking for (Kahn and Lehmann 1991, Kahn 1998, Hoch et al. 1999, Iyengar and Lepper 2000, Chernev 2003, Waldfogel 2003, George and Waldfogel 2006, also see Bakos 1998, Häubl and Trifts 2000, and Lynch and Ariely 2000). The value of such discovery may be substantial: Brynjolffson, Hu and Smith (2003) estimate that the increased variety in online bookstores raised consumer surplus by more than \$700 million annually. Second, variety can also directly contribute to consumer welfare. Due to satiation, curiosity, or fluctuating requirements, an individual consumer will often seek variety within and across consumption occasions (Kahn 1995). The argument implies that a movie buff will prefer a more diverse library of DVDs not only because these titles better match her preferences but also because she values diversity itself.

While the long-tail view is consistent with basic theories of choice and variety-seeking preferences, there are also arguments that the heterogeneity in consumption patterns will decrease in an online world. For example, some recent research has questioned the "variety-is-good" belief (for a survey, see Gourville and Soman 2005). Overwhelmingly large assortments can apparently create confusion or frustration among consumers, and could even lead them to making no choice at all (Huffman and Kahn 1998, Lehmann 1998). A potential negative impact of product assortment on consumer choice, an effect that Gourville and Soman (2005) call "overchoice," has now been documented in numerous settings (e.g., Dhar 1997, Iyengar and Lepper 2000, Boatwright and Nunes 2001, Chernev 2003, Broniarczyk, Hoyer and McAlister 1998).

Moreover, there is a body of research that argues that people tend to converge on the same hit content regardless of the breadth and depth of niche content available. Rosen (1981) explained the "superstar" phenomenon as the result of two factors. First, lesser talent is a poor substitute for greater talent: why would people listen to the world's second-best tenor if the best is also available? Second, with online distribution of digitized information goods, the marginal cost of reproducing and distributing books, records, and videos approaches zero, increasing any cost advantage (or decreasing a disadvantage) that hit products have over content of limited appeal.

Superstars can also emerge as the direct result of consumer preferences. Frank and Cook (1995) attribute the existence of "winner-take-all markets" to the social nature of people, pointing out that people have a keen interest in reading the same books and watching the same movies that others

consume. Similarly, communications research suggests that consumers are more likely to take advantage of the greater availability of their preferred content as the number of content options increases. Instead of facilitating a broad “marketplace of ideas,” increases in the diversity of media content may in fact contribute to polarization—less variety seeking, not more (e.g., Napoli 1997, 1999, 2003, Neuman 1991, Webster 2005). McPhee (1963) was first to describe this phenomenon as the double jeopardy in markets for media products: less popular cultural products not only have smaller audiences, these audiences are also less loyal (see also Ehrenberg and Goodhardt 1990).

As these arguments make clear, whether a long tail will emerge and whether the consumption of niche products is a complement or a substitute for the consumption of “hits” depends on consumer preferences. Because the cost of offering variety was prohibitively high prior to online distribution, not much is known about these preferences. The rise of online channels now affords the opportunity to study consumers’ demand for obscure products and see if this demand increases or reduces the importance of superstar hits. Three questions, illustrated in figure 1, are central to our empirical work:

- 1) Has the mass of the distribution of sales across video titles shifted towards the tail, i.e. toward more obscure titles that each generate a relatively low level of sales?
- 2) Does the emergence of a long tail come at the expense of reduced sales for hits?
- 3) To what extent has the composition of titles in the head and tail of the distribution changed as a result of the proliferation of online channels?

--- Figure 1 ---

2. Data

In our empirical analysis, we study home video market in the period from 2000 to 2005. Our data come from Nielsen VideoScan, a leading provider of information on video sales. We discuss the data source, the industry context, and data limitations in more detail below.

Data Source: Nielsen VideoScan

Nielsen VideoScan tracks weekly sales for video titles sold in the United States. Sales are split by DVD and VHS formats. They can also be broken down by channel: “discount mass stores” (e.g., Kmart, Shopko, and Target), “drug stores” (e.g., CVS, Rite Aid, and Walgreens), “grocery stores” (e.g., Abertson’s, Pathmark, and Safeway), “specialty retail” (e.g., Blockbuster, Hollywood Entertainment,

and Suncoast), and “other mass stores,” which covers both offline (e.g., Best Buy, Circuit City, and Costco) and online stores (e.g., Amazon.com, Borders.com, and Virginmega.com).

We focus our analyses on the period from January 2000 through August 2005, and concentrate on a random sample of titles available in this period. The Nielsen database lists titles alphabetically, presenting information on 50 titles on a single webpage. We created our sample by randomly calling up such pages, recording the sales histories for all 50 titles that appear on a chosen page. The resulting sample includes nearly 5,500 unique titles, 3,700 titles in DVD format and (partly overlapping) 3,000 titles in VHS format.¹ For each title, Nielsen provides a wide variety of characteristics, including the genre, rating, type (e.g., live action versus animated), original language, release language, translation form (e.g. dubbed versus subtitled), runtime, distributor, release date, first release window (e.g., theatrical, television, or direct-to-video), and manufacturer suggested retail price. Table 1 provides descriptive statistics for the variables included in the analyses.

--- Table 1 ---

Industry Context: The Home Video Industry

Because U.S. consumers spend nearly three times as much on videos than on movie theater tickets (Entertainment Merchant Association 2006), home videos are of vital importance for the film industry. Critically, the video window has undergone radical changes in the period under investigation. After DVD technology was introduced in 1997, DVD titles rapidly grew to replace VHS titles as the dominant video format. The number of U.S. households with at least one DVD player jumped from 13 million to 65.4 million from 2000 to 2004 (Standard & Poor’s 2006). While the number of U.S. households that had at least one VCR player also grew, from 88% in 2000 to 99% in 2004, there was a significant shift towards buying and renting DVDs. In our data, only about 24% of unit sales were on DVD in 2000, while this fraction rose to nearly 97% by 2005. Figure 2 displays the trends in DVD and VHS sales, as well as sales by channel.

--- Figure 2 ---

Fueled by the growing popularity of DVDs, the home video market as a whole increased nearly three-fold from \$8B in 2000 to nearly \$24B in 2005. Specifically, Americans spent \$23.84 billion

¹ One “unique title” could include multiple releases of a title, such as a Director’s Cut released a few months after the original version, or a wide-screen and full-screen edition released simultaneously.

on video purchases and rentals in 2005, with sales accounting for over \$16 billion (Video Business 2006). Spending on DVDs amounted to \$22.4 billion, which included \$15.7 billion in sales and \$6.7 billion in rentals (Entertainment Merchant Association 2006, Video Business 2006, Standard & Poor's 2006). Lower prices (most DVDs now sell for \$15 to \$20) coupled with a shortening of the time between theatrical and video releases (on average, DVDs are released for retail sale four months after theatrical release, down from an average of six months a few years ago) helped to significantly grow the video sales market.

Importantly, our study period was characterized by a rapid development of the Internet as a retail channel. According to Euromonitor (2006), the number of online households jumped from just under 45 million in 2000 to over 80 million in 2005. The number of online buyers nearly doubled from 66 million to 123 million between 2001 and 2005, while online shopping expenditures more than doubled in the same period, from \$31 billion to \$77 billion. The central category in this manuscript, pre-recorded DVDs and VHS videos, accounted for \$636 million in sales in 2001 and over \$1.5 billion in 2005. Amazon, which began selling videos in 1998 but offered only about 2,000 DVDs at the time, has been ranked the number one online retailer of VHS videos and DVDs since 1999 (VSDA 2005), and now offers more than 80,000 titles.²

Data Limitations

While the Nielsen VideoScan data are the most comprehensive source of information on industry sales, the available data have a number of limitations. One noteworthy shortcoming is that Wal-Mart does not report its sales to Nielsen (or to any other organization). Another limitation is that VideoScan does not track sales for titles that sell through Amazon.com's Marketplace. Most importantly, however, Nielsen does not break out Internet sales in its statistics. These sales are incorporated in the "other mass and Internet" category shown in Figure 2.³ Hence, we will rely on overall sales to study changes in the distribution of revenue. While it would clearly be desirable to study Internet sales as a separate category, insights into changes in the overall distribution of sales are particularly valuable for two reasons. First, a critical question for media companies is if the sale of niche products crowds out the sales of "hits," an effect that will only be visible in overall sales. Second, there is some reason to believe that the long-tail phenomenon could also change the distribution of sales in regular stores. For example, Barnes & Noble attributes a net increase in special orders in its

² Although video rentals are outside the scope of our study, the growing user base for Netflix, an online retailer with a DVD-by-mail rental subscription model, also illustrates the rapid rise of the Internet as a video distribution channel: Launched in 1998 with 900 titles, Netflix surpassed 1 million subscribers in 2003, and offered roughly 50,000 titles to 4.5 million customers at the end of 2005.

³ While Nielsen VideoScan has the capability to measure online sales as separate category, and even at the level of individual retailers, its contractual arrangements with those retailers prevents it from releasing this information.

brick-and-mortar stores to consumers discovering new books online (Bynjolfsson, Hu and Smith 2003). If these types of spillover effects are important, the total long-tail effect should be measured as changes in the distribution of overall sales.

The data enable us to control for two alternative explanations for a shift in the distribution. First, one could argue that a possible shift in the distribution of sales across video titles could be due to a change in the composition of the DVD buying population. It may well be that early DVD adopters were more likely to be interested in a larger assortment of titles, which would lead to an overestimation of the long tail based on early data. However, by comparing DVD and VHS sales for individual titles, we can account for such a possible changing composition of the user base. Because the two technologies are substitutes and early DVD adopters are likely to be former VHS users, we would expect to see the exact opposite effect for VHS if this explanation were to hold. If we also find a shift towards the tail for VHS sales, that provides even more compelling evidence of the long tail effect. The same applies to our analysis of the possible changing composition of titles in the head and tail of the distribution: comparing, say, genre shifts in DVD versus VHS sales will provide a richer insight into a possible long tail trend.

Second, one could argue that a possible shift in the distribution of sales across titles is simply due to an increase in the supply of titles—either re-releases of old hits or new releases of niche titles, rather than purely a demand effect. However, as we have data on the number of titles that enter the marketplace at any week of the sample period, we can directly investigate whether changes in week-to-week supply of titles fully explain potential shifts in the distribution.

3. Modeling Approach

Our modeling approach consists of three stages. First, we test whether the distribution of revenues across available titles has shifted from year to year using simple nonparametric tests. Second, we estimate a quantile regression model to examine the factors that underlie the shift in the distribution of sales. Third, we estimate a negative binomial regression model to see whether the number of titles reaching particular sales targets has changed over our study period.

Inter-Quartile and Kolmogorov-Smirnov Statistics

As a first step in understanding whether and to what extent the distribution of sales across titles shifts from year to year, we generate various descriptive statistics for these distributions. Specifically, we compare the location ($Q_{0.5}$), scale ($(Q_{0.75}-Q_{0.25})/(Q_{0.75}+Q_{0.25})$), skewness ($(Q_{0.75}+Q_{0.25}-2Q_{0.5})/(Q_{0.75}-Q_{0.25})$), and kurtosis ($(Q_{0.90}-Q_{0.10})/(Q_{0.75}-Q_{0.25})$), where Q_{θ} denotes the sales for the title

in the θ^{th} quantile, with $\theta \in (0,1)$. We also generate two inter-quartile measures, $(Q_{0.50}-Q_{0.25})/Q_{0.50}$ and $(Q_{0.75}-Q_{0.50})/Q_{0.50}$, that capture how the left tail and right tail shift relative to the median, respectively.

We compute the Kolmogorov-Smirnov (KS) statistic to test for any shifts in the mass of the distribution. That is, for pairs of years, we assess whether sales distributions are significantly different from each other. The KS test, which considers the maximum vertical distance between two empirical distribution functions, is sensitive to both the location and shape of the distributions. While it is more sensitive at points near the median of the distribution than near its tails, it allows for a useful initial exploration. We run the test for the full sample and for both VHS and DVD sales separately.

Quantile Regression Model

A weakness of our nonparametric test results is that we cannot be certain if the observed changes in the distribution of sales are due to changes in product characteristics or changes in consumer behavior. We explore this distinction in a series of quantile regression models. Introduced by Koenker and Bassett (1978), and growing in popularity in the fields of finance and economics, quantile regression is rarely used in the field of marketing and management (a study by Taylor and Bunn 1999 is one exception), presumably because most research is concerned with average effects.⁴ In our context, however, it is critical to be able to see how changes in the business environment affect the tails of the distribution of revenues. In a quantile regression model, a specified conditional *quantile* (or percentile) of the outcome variable is expressed as a linear function of observed covariates. (In OLS regression, in contrast, the *mean* of a continuous response variable is expressed as a linear function of a set of independent or predictor variables.) By examining multiple quantiles, it is possible to assess how the entire distribution changes with certain covariates, which means that richer inferences can be drawn from the data through the use of quantile regression than through the use of traditional OLS regression methods alone (e.g., Powell 1991, Koenker and Hallock 2000, Austin et al 2005).⁵ We estimate models of the following general form

$$Q_{\theta}(y|x) = x'\beta(\theta) \tag{1}$$

⁴ Quantile regression models are common in the field of labor economics, for instance (e.g., Chamberlain 1994, Buchinski 1994, 1997, Machado and Mata 2000, 2005), but also have been used in micro-economics (e.g., Eide and Showalter 1998, Koenker and Billias 2001, Levin 2001) and demand analysis (Hendricks and Koenker 1992, Manning, Blumberg and Moulton 1995, Deaton 1997). In finance, Taylor (1999), Chernozhukov and Umantsev (2001), and Bassett and Chen (2001) use the methodology.

⁵ As Koenker and Hallock (2000) explain, quantile regression cannot be achieved by segmenting the response variable into subsets according to its unconditional distribution and then doing least squares fitting on these subsets. Such a form of “truncation on the dependent variable” is doomed to failure for the reasons laid out in Heckman’s (1979) work on sample selection.

where $Q_\theta(y|x)$ denotes the θ^{th} quantile of the distribution of y , the log of video sales, given a vector x of covariates.⁶ The conditional quantile framework provides a full characterization of the conditional distribution of sales in much the same way as ordinary sample quantiles characterize a marginal distribution (Machado and Mata 2005). Quantile regressions are especially informative in our context because we do not expect the heterogeneity in the conditional distribution of y to be captured by location shifts only. To identify the emergence of a long tail in this setting, the covariates in (1) include a set of annual indicators for year t . If a longer tail emerges in our study period, we expect the coefficients on the year effects to decrease in size, $\beta_t(\theta) > \beta_{t+1}(\theta)$, for some quantile θ in the tail end of the distribution. In other words, we should see sales per title fall over time as the tail gets longer.

Negative Binomial Regression Model

Declining sales per title in the tail end of the distribution can be due to changes other than the emergence of markets for niche products. For example, if the longevity of titles declined over time such that once popular titles sell fewer copies once they have been out for a while, this trend too would manifest itself as a flatter tail. To distinguish these types of explanations from the long-tail phenomenon, we analyze changes in the number of titles in the tail. When niche markets become more prominent, average sales will decline but the number of products will increase. As our dependent variable is a count variable, namely the number of titles that meet a particular sales threshold, we estimate negative binomial regression models (Coleman 1964). The sales thresholds we investigate are the 70th, 80th, and 90th percentile of weekly sales. Our controls include characteristics of the population at risk such as the fraction of titles belonging to a particular genre. An advantage of these count models is that we can include a control for the total number of titles that are available in a given year. This allows us to directly test if the long tail is due to release decisions by film producers. As in the quantile regression models, we track changes in the number of titles over time by including year indicators, expecting the coefficients to follow $\beta_t(\theta) < \beta_{t+1}(\theta)$ if there is a growing long tail in video sales.

⁶ This specification assumes linearity of the quantile regression model. As Machado and Mata (2000) argue, additional flexibility can be achieved without sacrificing the semi-parametric nature of the approach, by assuming that the conditional quantile is a given transformation of an affine function of the covariates. We are currently estimating such a Box-Cox quantile regression model.

4. Findings

Our discussion of the findings is organized along the three stages in the modeling approach—inter-quartile statistics and Kolmogorov-Smirnov tests, quantile regression, and negative binomial regression, respectively—but we start with general observations about the shape of the distribution of sales across titles. Table 2 sheds light on the distribution of sales by quantile, for every year of the period under scrutiny, and for VHS and DVD titles.

--- Table 2 ---

The table suggests that the home video sales market has become more concentrated over time. Consider Table 2a, which lists the percentage of sales accounted for by each decile of titles, as well as the 95th and 99th percentile. For VHS titles, for instance, 80% of title-by-week observations account for nearly 3% of sales in 2000, but less than 1% of sales in 2004 and 2005. Similarly, 99% of the observations account for nearly 50% of sales in 2000, and slightly more than half of sales in 2001, but only around a third of sales in 2004 and 2005. The same pattern is visible for DVD titles. Here, 80% of the title-by-week observations account for around 8% of sales in 2000 and 2001, but closer to 4% in 2004 and 2005, while the percentage of sales accounted for by 95% of the observations drops from over 25% to below 20% in the same period. For both formats, the tail appears to have become flatter.

Table 2b and 2c shed light on shifts in an absolute sense. Table 2b reports the median number of titles that exceed a specific *weekly* sales level, while Table 2c does the same for *annual* sales levels. The level of aggregation also differs: Table 2b considers *any version* of a title (e.g. both a Director's Cut and a regular version) as a separate product, while Table 2C counts products at the *title* level, i.e. regards those different versions as being part of the same title.

Across all formats, the number of titles that generate no sales weekly or annually increases. In Table 2b, the number jumps from nearly 500 titles in the year 2000 to over 2,000 titles in 2004 and 2005 for VHS titles, and from less than 100 titles in 2000 to over 1,000 titles in 2005 for DVD titles. The trend is captured in the reported “selling versus non-selling” ratio, which drops from 75% in 2000 to below 50% in 2005. The median number of titles that have sales levels that put them below the 50th percentile increases as well: across both formats, this number jumps from over 1,300 titles in the year 2000 to nearly 2,200 titles in 2004 and nearly 3,300 titles in 2005. This increase is solely due to DVD sales—the opposite pattern is found for VHS. The median number of titles in the 99th quantile also shows the sharpest increase for DVD sales, where it jumps from 5 titles in 2000 to 33 titles in 2005.

Table 2c shows a comparable pattern for annual sales levels. Again, the median number of titles without any sales increases, from 0 in 2000 to over 500 in 2004 and over 1,000 in 2005, and both

formats contribute to that trend. As before, the median number of titles with sales levels below those that correspond to the 50th percentile rises, from nearly 1,400 titles in 2000 to well over 2,000 titles in 2004 and 2005, and DVD titles are largely responsible for this effect. Finally, the median number of titles with an annual sales level (roughly 138,000 in unit sales) that corresponds with the 95th or 99th is relatively stable.

Inter-Quartile and Kolmogorov-Smirnov Statistics

Table 3 reports summary statistics for the distribution in sales.

--- Table 3 ---

Across both formats, the location of ($Q_{0.50}$) decreases, while the scale, skewness and kurtosis measures increase from year to year. This is consistent with a scenario in which the distribution becomes more dispersed, more asymmetrical, and develops a sharper peak and a longer tail over time. The inter-quartile metric ($Q_{0.75}-Q_{0.25}$) sharply decreases in the period from 2000 to 2005. The left-tail inter-quartile measure is either close to one, which is consistent with a very long, low sloping tail, or is missing if $Q_{0.50}$ corresponds to zero sales. The occurrence of many zero sales also explains missing values for the right-tail inter-quartile measure, which otherwise does not reveal a clear trend.

The results of the Kolmogorov-Smirnov tests for each pair of years are displayed in Table 4.

--- Table 4 ---

With one exception (2001 versus 2002), the tests for the full sample reveal that the distributions of weekly sales across titles are significantly different across the years. Looking at the VHS and DVD formats separately, the sales distributions for VHS titles show significant shifts across all years, while those for DVD titles show significant shifts for all comparisons involving the years 2004 and 2005. Substantial shifts in the mass of the distribution of DVD sales thus are a relatively recent phenomenon.

Quantile Regression Model

Many factors can explain the changes documented in Table 4. Perhaps the studios released more comedies in later years, or released more titles directly on video, or simply flooded the market with old material—each of these and other factors could induce a change in the distribution of sales. The quantile regression models in Tables 5a and 5b separate out these types of explanations from a claim that the sales distribution has shifted more generally.

--- Table 5 ---

The tables display estimation results for an OLS model as well as for quantile regression models for the 40th, 50th, 60th, 70th, 80th, 90th, 95th, and 99th quantile. The dependent variable is the log of weekly sales for all titles in our sample, provided the title was available in the channel. A total of 573,753 observations are included in the analysis.⁷ The fit of the models is reasonably good: the OLS model explains about a third of the variation in the data, while the quantile models generate Pseudo *R*-squared values that vary between 0.08 for the 40th quantile (the lowest quantile that can be estimated given the high frequency of zero sales) and 0.26 for the 70th through 90th quantile. Across the specifications in Table 5a and 5b, the effect of a covariate on log sales varies freely from quantile to quantile.

In Table 5a, we study changes in the overall distribution of sales across both formats over time. The OLS results indicate that mean predicted sales declined from year to year. All coefficients on the annual indicators are negative and statistically significant. A similar trend can be observed for the lowest quantile ($\theta=0.40$), but the trend is clearly less pronounced. This decline in sales is consistent with the long-tail idea. As a larger number of lower-selling titles appear in the tail, predicted sales will fall. Looking at the best-selling titles ($\theta=0.95$ and 0.99), Table 5a shows a more drastic decline. For 2005, the coefficient for $\theta=0.99$ is more than five times as large as the coefficient for $\theta=0.40$, indicating there is significant heterogeneity in how sales develop over time. Note that these models control for the large number of product characteristics reported in Table 5a, so these changes are not the result of releasing different types of products. In terms of Figure 1, we observe that the distribution of sales has shifted down in general, but this shift is largest for the better-selling titles. The tail of the distribution has seen a much smaller decrease, implying a shift in the mass towards niche products.

A shortcoming of the models in Table 5a is that they do not allow sales to vary by format and year. Because DVDs became much more popular during our study period, this restriction is important. In Table 5b, we allow the year effects to vary by format. These models control for the same product characteristics as Table 5a. To facilitate the interpretation of the results in Table 5b, Figure 3 provides quantile plots—constructed using the information in the table—for fourteen key covariates.⁸

⁷ Presently, the quantile regression results are based on a randomly drawn 50% sample of our data. We are working on a sparse-matrix approach to estimate these models using our full dataset. Based on comparisons between OLS models for the 50% sample and the full dataset, we expect the results to be qualitatively similar.

⁸ See for example Koenker and Hallock (2000) for an introduction to the interpretation of quantile plots.

--- Figure 3 ---

The quantile figures plot the quantile on the horizontal and the log of sales on the vertical axis. For each of the coefficients, we plot the quantile regression estimates for θ ranging from 0.40 to 0.99 as the solid black curve. The point estimates, shown as dots, can be interpreted as the impact of a one-unit change in the covariate on log sales, holding everything else constant. The gray area depicts the 95% confidence band. The thin straight line in each plot shows the OLS estimate of the conditional mean effect; with the 95% confidence interval.

Consider the quantile plot for “DVD,” an indicator variable for the format. The OLS estimate of 0.85 suggests that, holding all else fixed, DVDs generate higher sales than the alternative format, VHS videos. However, according to the quantile plot, that is an accurate estimate only for titles somewhere around the 70th and 80th quantile. In lower quantiles, DVDs generate relatively higher sales, compared with VHS videos (the coefficient for the 50th quantile is around 1.48), while for the very high quantiles, DVDs actually generate significantly lower sales than VHS videos (the coefficient for the 99th quantile is -0.37). Similarly, the negative OLS estimate for the genre dummy “Documentary” (the coefficient is -0.548) shows that documentaries generate lower sales than other genres, holding all else constant. However, according to the quantile regression results, the disparity is smaller in the lower quantiles of the distribution and larger in the head of the distribution. Again, OLS estimates do a rather poor job of representing these disparities.

The plots thus further underscore the need for a quantile regression approach: note that in virtually all of the panels of Figure 3, the quantile regression estimates mostly lie at some point outside the confidence intervals for the OLS regression, which suggests that the effects of these covariates are not constant across the conditional distribution of the independent variable.

Overall, the quantile regression analyses provide relevant insights into a possible long tail trend, in three ways. First, when we look at the estimates for the “DVD” dummy (which featured in the example above), it is clear that DVDs on average are associated with higher sales, but not so in the higher quantiles of the distribution. That is, among hit titles in the period under investigation, DVDs tend to generate lower sales compared with VHS videos. The pattern is again consistent with a situation in which more, and particularly more obscure, titles are released on DVD, whereas titles for a more mainstream audience tend to appear on VHS, and therefore by definition generate higher sales within the higher quantiles.

Second, we can expand this analysis by assessing temporal dimensions of the sales patterns for each format. The inclusion of the “Year” indicators as well as the “DVD \times Year” terms allows us to assess the time trend in video sales for VHS and DVD separately. To that end, Figure 4 plots the quantile regression point estimates displayed in Table 5 for (a) the “Year” indicators to shed light on

the trend for VHS sales, and (b) the “Year” + “(DVD × Year)” indicators to capture the trend for DVD sales.

--- Figure 4 ---

As we observed in Table 5a, there is a movement toward the tail, and we find that $\beta_t(\theta) > \beta_{t+1}(\theta)$ for year t : both types of titles generate fewer sales from year to year. However, there are important differences in how this sales decrease plays out in both formats. For VHS titles, the sales decline is particularly striking in the head of the distribution. For DVDs, the difference across the years appears slightly more pronounced in the lower quantiles. In other words, DVD titles are selling fewer units from year to year, and the “pain” is felt rather more by the obscure titles, as compared with the hit titles. On a sales-per-title basis, the more obscure titles generate fewer and fewer sales over time.

There are two minor exceptions to this general trend for DVDs. One is the (barely) positive shift from the year 2001 to 2002 in the 40th quantile, the lowest quantile for which the model can be estimated. This suggests that the most obscure titles actually generated higher sales in 2002 than in 2001. The other is the (also barely) positive estimate for the sum of “Year” and “DVD × Year,” in the highest quantiles for 2001 (the 90th quantile and higher) and 2002 (the 99th quantile only). This suggests that compared with 2000, DVDs had higher—not lower—sales levels in 2001 and 2002 in the higher quantiles. However, these exceptions notwithstanding, DVDs have generated lower sales overall since 2003. This again is consistent with a growing number of titles with zero or little sales.

Third, the “Fraction of Sales on DVD” variable allows us to examine whether unobserved characteristics of people switching from the VHS to DVD format play a role in these trends. Suppose it were true that movie buffs are early adopters of the DVD format because they value the better quality. If these movie buffs also like variety, the share of a title’s sales in the DVD format proxies for such tastes, suggesting we should see a negative coefficient on this variable, particularly in the tail of the distribution. The negative OLS coefficient is consistent with this conjecture; titles with a higher share of sales coming from the DVD format sell fewer copies on average. The quantile regression estimates reveal that the coefficient is negative in the lower quantiles, but turns positive in the 90th and 95th quantile, only to dip back into a negative coefficient again in the 99th quantile, although with a confidence band that includes zero. This is in line with the intuition behind the above-discussed results: lower sales levels for DVDs in the lower quantiles, and higher sales for VHS videos in the highest quantile.

The regression analyses further provide a number of insights into the role of the other covariates. Most notably:

- As expected, “Year of Production” has a positive effect on sales throughout the distribution. The effect is strongest in the higher quantiles. Not surprisingly, the tail thus contains slightly older content. Similarly, the number of weeks the title has been available (“Weeks Since Release”) is mostly negatively related to sales, and particularly so in the higher quantiles. This also corresponds to a general decay in the demand for a title over time.
- Video in certain genres that are typically associated with more niche content, like “Foreign” titles, documentaries (as discussed above), and “Adult” titles indeed generate relatively low sales across the range of titles (the coefficients for these variables are negative across all quantiles), and they do generate higher sales among the more obscure titles in the lower quantiles (the coefficients are higher the lower the quantile). However, titles in the “Children” and “Family” genre generate relatively high sales in the higher quantiles (where coefficients are positive), but relatively low sales in the lower quantiles (where coefficients are negative).
- Also usually associated with niche content, foreign-language titles (“Original Language: Non-English” or “Subtitled”) have a relatively strong negative impact on sales in the head. That is not the case for “Dubbed” content, which actually has a positive effect on sales, and more so in the middle quantiles.
- Relative to unrated content, all mainstream ratings (“G,” “PG,” “PG13,” and “R”) have a positive impact on sales, and typically more so in the higher quantiles, while the “adults only” rating “X” is associated with lower sales (with one exception, for the 40th quantile), particularly in the higher quantiles of the distribution.
- Relative to smaller, independent studios, all distributors included in the study have a positive impact on sales across the distribution, but particularly in the higher quantiles. Among all competing studios, Twentieth Century Fox (“Fox”) appears to have had the most successful portfolio of titles, while “Lions Gate” and, to a lesser extent, “Sony” have relatively low sales. Lions Gate is relatively weak in the tail of the distribution.
- Also intuitive, of the three main forms of original release, “Theatrical” outperforms “Direct to Video” and “Television” across all quantiles. Interestingly, television titles actually have a negative association with sales in several higher quantiles—from the 70th to the 95th, to be exact. Also, as far as video releases are concerned, “Rental” contracts mostly outperform “Sell-Through” options, except in the lower quantile reported (where the difference between both coefficients is minimal). While this may seem counter-intuitive (the sell-through model now prevails in the marketplace), it may be related to the underlying formats: the rental model is primarily associated with VHS, while the sell-through model is more prevalent for DVDs.

Negative Binomial Regression Models

Our analyses of the number of titles that meet a particular weekly sales level are presented in Table 6. The five columns correspond to different levels of sales, starting with zero sales, sales below the 70th quantile, sales between the 70th and 80th quantile, sales between the 80th and 90th quantile, and sales above 90th quantile, respectively. To facilitate the interpretation of the results, we transform the estimated coefficients and report incidence rate ratios ($\exp(\beta)$ instead of β). These ratios directly show the factor by which the number of titles changes. For example, an estimate of 2 indicates the number of titles doubles, and estimate of 0.5 implies there were only half as many titles reaching a particular sales threshold.

--- Table 6 ---

Table 6a provides results for a basic model that considers sales across all channels and includes year indicators and the percentage of DVD titles as covariates. The coefficients imply an increase in the number of lower-selling titles for DVDs but also—and more pronounced—a rise in the number of titles with zero sales. As indicated by the estimate of 2.536 reported in the first column of the table, compared to the year 2000, the omitted year, the predicted number of titles with zero sales is over 2.5 times larger in 2001. Similarly, based on the coefficient of 74.631 reported in the first column, it appears that, in 2005, the number of non-selling titles is nearly 75 times as high as in 2000. Compared with VHS, DVD titles are associated with a slightly lower number of titles in this zero sales category (-6.7% ($=100\% \times (0.933-1)$)). Interestingly, the DVD format increases the number of titles with low levels of sales. Although these percentage changes are small ($+2\%$ in model 2 and $+2.8\%$ in model 3), we thus find some evidence in support of the long tail hypothesis. The 90th quantile with titles with weekly sales higher than 125 units hardly qualifies as the “head” of the distribution only—several titles in the sample have weekly sales of over one million units early in their lifecycles—but can be considered as representing both the middle and head. The results for this group, reported in the rightmost column, show that the expected number of titles with weekly sales in the highest decile increases from year to year.

As in our earlier models, the results in Table 6a are susceptible to the concern that studios may have changed their production and release policies during our study period. Table 6b extends the basic model by also including covariates that measure the fraction of titles in a certain genre and with a certain rating. As the results indicate, these controls are important for our results. While we still see a more than fourfold increase in the number of non-selling titles, there is now also clear evidence of a longer tail. The number of titles selling fewer than 10 copies per week steadily increases over time. By 2005, it is 1.816 times as large as in 2000. Controlling for the types of titles that are available also

changes our inferences about the successful titles. We now find that the number of titles that account for relatively high sales levels has decreased in recent years, indicating that the market has become more concentrated. For example, in 2005, compared with 2000, the expected number of titles with weekly sales in the highest decile has declined by over 50%. This evidence complements the quantile regression models in important ways. Whereas the quantile regressions showed a sales decline that was particularly steep at the top end of the market, we now see that the reduced revenues are achieved by only half as many titles. This change is consistent with a superstar effect.

Table 6b captures the central results of this study. Are there important superstar and long-tail effects in U.S. home video sales? The answers turn out to be of the “yes, but...” variety. Yes, there is a long-tail effect in that the number of titles that sell only a few copies every week increases during our study period. But at the same time, the number of non-selling titles also increases substantially; it is now four times as high as in 2000. Many underdogs turn out to be losers. We also find evidence of a superstar effect. Among the best-performing titles, it is an ever-smaller number of films that accounts for the bulk of sales. The caveat here is that today’s superstars lack the punch of earlier years. Video sales generally decrease over time across all quantiles of the sales distribution, but this effect is most pronounced among best-selling titles.

There can be many reasons why a longer tail emerges during our study period. Perhaps studios adjusted their marketing to prolong the life of titles. Another possibility is that more titles are targeted at niche audiences. While qualitative changes such as these are hard to measure, we can easily observe another studio decision: the number of titles that are released. Table 6c controls for this variable. The newly added variable is positive throughout and statistically significant at a 1% level in the zero-sales category as well as the 80th and 90th deciles, but the estimates are very close to 1, indicating there is little economic significance to the number of titles that are on offer.

Nevertheless, controlling for the supply of titles has some impact on the size of the other coefficients in our model. Specifically, accounting for the number of titles available in a given week again reduces the size of the shift toward the tail with low-selling or not selling titles. At the same time, the higher level of concentration in the 90th quantile becomes even more apparent. We conclude that the observed long-tail and superstar effects are only in part due to studios’ supply decisions. What we observe, we believe, is the result of changing consumer decisions as online retailing offers greater variety and lower transaction cost.

5. Conclusion

By examining whether the proliferation of online channels goes hand in hand with a shift in the distribution of sales across products and a changing composition of titles in the head and tail of the

distribution, our study addresses an important debate. We find that online retailing indeed appears to have affected the sales distributions in our study period. In video sales generated over the years from 2000 to 2005, we observe a shift towards the tail of the distribution. In line with the core premise of lower transaction and search costs in online channels, the shift toward the tail becomes somewhat more pronounced in more recent years, and is stronger for DVD titles compared with VHS titles. The shift cannot be fully explained by the changing composition of the user base for DVD versus VHS formats, the number of titles available, as well as characteristics of those titles, such as genres and ratings, which could all be alternative explanations for the observed shift toward the tail.

There is some indication in our data that the popularity of niche titles has gone hand in hand with a significant concentration of success on ever fewer items. Video sales decline across all quantiles of the sales distribution, but the drop is much larger among best-selling titles. While hits as a category generate fewer sales, the role of individual bestsellers is growing over time. In the period from 2000 to 2005, the number of titles in the top 10% of sales drops by more than 50%, an increase in concentration that is familiar from many winner-take-all markets.

As a whole, our findings point to significant business challenges for entertainment companies. At the top end of the distribution, most hits draw smaller audiences, a trend that is particularly worrisome given the economic importance of video sales. At the tail end, we find that there is a rapidly increasing number of titles that never, or very rarely, sell—the long tail appears incredibly flat. We find that the sharp increase in the number of titles that have come onto the market in recent years contributes little to this phenomenon.

With limited space on store shelves, producers traditionally have been very discriminating about what they release, and have focused their marketing resources on a small number of likely hits. Are significant changes warranted in how they manage their assortments? In light of our findings, making radical alterations in resource allocation strategies does not seem advisable. In fact, perhaps above all, our findings imply that it is difficult for content providers to profit from the tail. It is not clear whether the new media environment indeed makes previously unprofitable niche products profitable. While the rise of online channels lowers the barriers to market entry for such products, and thus introduces the possibility that those titles generate some sales, the same phenomenon also likely leads to a flood of products that compete for the attention of consumers. Even if the break-even sales levels for niche titles are low, the intensified competition may make it more difficult, not less to reach these levels.

The insights gained by means of a quantile regression model like the one we estimated here can help producers and retailers better manage their libraries or assortments. Because assessing the value of a producer's content library essentially requires an understanding of the distribution of revenues across titles as well as an understanding of the composition of titles at each point of that

distribution, quantile regression estimates for the title characteristics could provide helpful clues. For example, using the estimates depicted in Figure 3, we can conclude that, controlling for many other factors, children titles tend to generate higher returns than dramas, documentaries, and foreign movies, particularly in the higher quantiles of the distribution. Considering that returns for titles substantially decrease with age, particularly in the higher quantiles of the distribution (as also indicated in Figure 3), a model that includes interaction terms for various title characteristics (such as genres and ratings) with the year of production can be used to more comprehensively establish the value of a library of various types of titles in various stages of their lifecycles.

For online retailers, shelf space constraints do not matter, but problems may be introduced by the relatively large number of titles that do not, or very rarely, sell. Although the margins of long tail products may be higher compared with hit products (which often are used as loss leaders), eliminating the costs of making obscure products available may be critical if the company's objective is to profit from the long tail. A good example of a promising model is Amazon's Marketplace, where third parties pay the costs involved in communicating the availability of a title, and Amazon itself only incurs costs once a customer actually places an order.

Future Research Opportunities

We think three general future research areas are particularly worthwhile. First, while we only considered the home video market, it may be interesting to explore how the distribution of sales across products differs across industries. The more we know about how the proliferation of online channels is affecting sales distributions in various contexts, the closer we may come to an understanding of the factors that potentially drive a shift towards the tail. In light of the possible reasons we discussed for why a strong shift towards the tail may not occur, it appears especially useful to compare trends for products that have a strong social and cultural dimension, such as movies, music and books, to other information products that do not, such as software or market intelligence.

Second, investigating consumption trends at the level of individual consumers and the "baskets" of products they purchase appears particularly promising. As consumers increase the frequency with which they shop online, and as the assortment of products available online continues to grow, will they gradually adopt more content in the tail? Does a select group of customers account for the lion's share of obscure products being sold, or is the appeal of such products more equally distributed? And what factors drive a customer's propensity to adopt obscure products not available in bricks-and-mortar stores? Insights into these questions will help companies in serving their customers more effectively and efficiently.

Third and finally, we cannot escape the notion that the companies that currently garner acclaim for their ability to serve both the head and the tail are mostly retailers—and the most

celebrated examples are large scale, hit retailers. Amazon cost-effectively runs its marketplace for third-party content but also accounts for a large share of online sales for the most popular books, movies and music albums. Google collects revenues not just by placing advertisements for obscure products or firms on equally obscure web pages, but also by selling advertisements on mainstream web pages to large firms. And Netflix couples the appeal of its deep library of over 55,000 titles with high demand for its top titles—the top 50 titles reportedly account for 30% of all rentals (The Wall Street Journal 2006b). In future research, it may be interesting to explore cases of content producers, such as movie studios, record companies, or book publishers, which profit in a fundamental way from the increased product assortments in online channels. Examining whether some firms are likely to suffer in the new environment seems worthwhile as well—for example, established niche providers who now may face strong competition from the growing number of long tail products.

Table 1a: Descriptive Statistics: Dummy Variables at the Unique Title Level

Variable	Full Sample		VHS	DVD
	#	%	%	%
Genre: Action	351	0.06	0.07	0.07
Genre: Adult	397	0.07	0.04	0.09
Genre: Children	514	0.09	0.13	0.06
Genre: Comedy	596	0.11	0.11	0.11
Genre: Documentary	471	0.09	0.09	0.08
Genre: Drama	626	0.11	0.13	0.12
Genre: Family	65	0.01	0.01	0.01
Genre: Foreign	435	0.08	0.06	0.10
Genre: Horror	233	0.04	0.03	0.05
Genre: Japanese Animation	321	0.06	0.05	0.06
Genre: Mystery	179	0.03	0.03	0.04
Genre: Science Fiction	151	0.03	0.03	0.02
Rating: G	51	0.01	0.01	0.01
Rating: PG	139	0.03	0.04	0.03
Rating: PG-13	132	0.02	0.04	0.03
Rating: R	573	0.11	0.12	0.13
Rating: NC-17	4	0.00	0.00	0.00
Rating: X	4	0.00	0.00	0.00
Animation: Animated	724	0.13	0.14	0.12
Animation: Live Action	4,510	0.83	0.82	0.85
Animation: Puppets	41	0.01	0.01	0.00
Original Language: Non English	679	0.12	0.10	0.14
Translation: Dubbed	171	0.03	0.04	0.03
Translation: Subtitled	311	0.06	0.05	0.07
Part of Franchise	1,235	0.23	0.25	0.19
Distributor: Disney	217	0.04	0.06	0.04
Distributor: 20 th Century Fox	167	0.03	0.04	0.02
Distributor: Paramount	171	0.03	0.04	0.02
Distributor: Sony	306	0.06	0.07	0.07
Distributor: Universal	154	0.03	0.04	0.02
Distributor: Warner Bros.	507	0.09	0.12	0.06
Distributor: Lions Gate	221	0.04	0.05	0.04
Original Release: Theatrical	1,405	0.26	0.31	0.27
Original Release: Direct to Video	3,223	0.59	0.47	0.67
Original Release: Television	1,365	0.25	0.27	0.22
Original Release: Rental Re-Price	47	0.01	0.02	0.00
Video Release: Rental	304	0.06	0.05	0.08
Video Release: Sell-Through	592	0.11	0.14	0.16
Video Release: Catalog	2,045	0.37	0.21	0.54
	N=5,455		N=3,067	N=3,717

Table 1b: Descriptive Statistics: Continuous Variables at the Unique Title Level

Full Sample						
	N	Mean	Median	SD	Min	Max
Year Of Production	4,074	1990.51	1998	18.03	1903	2005
Runtime	5,455	106.94	88	177.3	8.00	6000
Suggested Price	5,455	20.56	19.95	18.48	0.99	499.92
VHS						
	N	Mean	Median	SD	Min	Max
Year Of Production	2,307	1988.05	1995	18.09	1914	2005
Runtime	3,067	83.13	85	79.45	16.00	900
Suggested Price	3,067	17.73	14.98	11.79	0.99	199.92
DVD						
	N	Mean	Median	SD	Min	Max
Year Of Production	2,989	1992.78	1999	16.68	1903	2005
Runtime	3,717	125.87	91	205.34	8.00	6000
Suggested Price	3,717	22.55	19.96	20.41	1.00	499.92

Table 1c: Descriptive Statistics: Sales By Year, Channel, and Format

Sales By Channel	2000	2001	2002	2003	2004	- 08/2005
All Formats (N = 1,212,863)						
Drug Stores	1,118,059	262,514	819,342	409,986	459,749	230,331
Grocery Stores	1,823,207	567,061	1,852,020	1,187,498	1,630,999	938,065
Mass Discount Stores	4,684,174	1,776,137	5,414,352	3,624,332	3,208,131	1,436,509
Specialty Retail Stores	5,235,296	4,469,675	6,234,633	6,065,015	5,653,385	3,609,695
Other Mass and Internet Stores	5,315,582	3,701,746	7,699,523	9,552,176	9,912,164	6,471,463
Total	18,176,291	10,777,133	22,019,870	20,839,007	20,864,428	12,686,063
VHS (N = 727,729)						
Drug Stores	1,105,161	255,962	608,420	232,104	153,189	42,282
Grocery Stores	1,792,220	526,035	1,245,910	520,844	347,417	90,978
Mass Discount Stores	4,463,720	1,635,226	2,964,885	875,780	390,910	47,471
Specialty Retail Stores	3,716,796	2,502,929	2,278,921	1,047,714	492,637	176,306
Other Mass and Internet Stores	2,691,209	953,684	1,109,879	457,517	182,239	41,353
Total	13,769,106	5,873,836	8,208,015	3,133,959	1,566,392	398,390
DVD (N = 485,134)						
Drug Stores	12,898	6,552	210,922	177,882	306,560	188,049
Grocery Stores	30,987	41,026	606,110	666,654	1,283,582	847,087
Mass Discount Stores	220,427	140,911	2,449,467	2,748,552	2,817,221	1,389,038
Specialty Retail Stores	1,518,500	1,966,746	3,955,712	5,017,301	5,160,748	3,433,389
Other Mass and Internet Stores	2,624,373	2,748,062	6,589,644	9,094,659	9,729,925	6,430,110
Total	4,407,185	4,903,297	13,811,855	17,705,048	19,298,036	12,287,673

Table 2a: The Distribution of Weekly Sales by Year (N=1,212,863)

% least selling titles	2000 % sales	2001 % sales	2002 % sales	2003 % sales	2004 % sales	- 08/2005 % sales
VHS (N = 727,729)						
10% of titles	0.00	0.00	0.00	0.00	0.00	0.00
20% of titles	0.00	0.00	0.00	0.00	0.00	0.00
30% of titles	0.02	0.00	0.00	0.00	0.00	0.00
40% of titles	0.10	0.02	0.00	0.00	0.00	0.00
50% of titles	0.24	0.11	0.03	0.01	0.00	0.00
60% of titles	0.53	0.32	0.14	0.12	0.01	0.00
70% of titles	1.18	0.82	0.40	0.53	0.12	0.00
80% of titles	2.68	2.20	1.20	1.56	0.82	0.11
90% of titles	7.79	7.31	4.03	5.11	3.37	2.17
95% of titles	17.71	18.99	9.91	12.79	9.33	7.72
99% of titles	49.62	53.92	34.68	44.24	33.79	33.39
DVD (N = 485,134)						
10% of titles	0.00	0.00	0.00	0.00	0.00	0.00
20% of titles	0.03	0.02	0.01	0.01	0.00	0.00
30% of titles	0.14	0.11	0.06	0.07	0.05	0.02
40% of titles	0.40	0.32	0.19	0.21	0.16	0.10
50% of titles	0.93	0.80	0.49	0.48	0.41	0.30
60% of titles	1.99	1.82	1.10	1.02	0.91	0.69
70% of titles	4.01	4.03	2.34	2.08	1.93	1.52
80% of titles	7.81	8.78	5.01	4.43	4.24	3.40
90% of titles	15.89	19.46	11.21	10.44	10.99	9.29
95% of titles	25.15	31.69	19.13	19.44	21.49	19.41
99% of titles	45.57	56.92	40.19	44.25	51.38	49.06

Table 2b: The Number of Titles in Weekly Sales Quantiles by Year

Level of sales	2000 # titles	2001 # titles	2002 # titles	2003 # titles	2004 # titles	- 08/2005 # titles
All Formats (N = 1,212,863)						
Weekly sales = 0	547	946	1446	1959	2749	3474
Weekly sales >0	1576	1799	2154	2776	2841	2973
Q _{0.50} : weekly sales >1 (0)	1312	1488	1779	2272	2264	2973
Q _{0.75} : weekly sales >14 (6)	674	766	948	1059	1057	1572
Q _{0.90} : weekly sales > 87 (44)	257	329	390	412	401	634
Q _{0.95} : weekly sales > 241 (142)	130	168	186	199	199	321
Q _{0.99} : weekly sales > 1478 (1174)	29	32	36	36	36	62
Ratio selling / non-selling	74.2%	65.5%	59.8%	58.6%	50.8%	46.1%
VHS (N = 727,729)						
Weekly sales = 0	480	818	1214	1601	2138	2497
Q _{0.50} : weekly sales >0 (0)	1198	1216	1200	1189	828	538
Q _{0.75} : weekly sales > 5 (0)	680	651	611	508	289	538
Q _{0.90} : weekly sales > 36 (2)	288	291	251	194	112	283
Q _{0.95} : weekly sales > 127 (11)	154	148	121	100	45	136
Q _{0.99} : weekly sales > 1104 (108)	32	32	27	15	3	29
DVD (N = 485,134)						
Weekly sales = 0	68	126	225	356	635	976
Weekly sales >0	374	593	992	1569	1990	2402
Q _{0.50} : weekly sales >7 (3)	246	384	654	952	1110	1675
Q _{0.75} : weekly sales >42 (21)	120	211	348	468	522	847
Q _{0.90} : weekly sales > 174 (114)	44	84	135	176	210	341
Q _{0.95} : weekly sales > 416 (331)	18	38	65	88	108	170
Q _{0.99} : weekly sales > 2196 (2148)	3	5	10	19	20	32

Notes: Any version of a release is counted as a separate product (for example, if there are a Director's Cut and regular version of a particular DVD, they are counted as two products). We report the median number of titles that exceed specific sales numbers. Sales quantiles are based on weekly sales for the 2000 to 2004 period. Sales quantiles for 2005, a year in which we observe sales only through August, are reported in parentheses.

Table 2c: The Number of Unique Titles in Annual Sales Quantiles by Year

Level of sales	2000 # titles	2001 # titles	2002 # titles	2003 # titles	2004 # titles	- 08/2005 # titles
All Formats (N = 1,212,863)						
Annual sales = 0	0	84	191	339	581	1080
Annual sales >0	2275	2733	3459	3923	4474	4703
Q _{0.50} : Annual sales >124 (45)	1373	1415	1779	2135	2322	2282
Q _{0.75} : Annual sales >1113 (485)	693	748	942	1023	1107	983
Q _{0.90} : Annual sales > 7670 (3296)	286	337	390	389	403	388
Q _{0.95} : Annual sales > 25814 (11993)	164	165	192	192	189	162
Q _{0.99} : Annual sales > 138331 (77395)	33	33	39	39	36	21
Ratio selling / non-selling	100%	97%	95%	92%	89%	81%
VHS (N = 727,729)						
Annual sales = 0	0	87	201	358	654	1322
Annual sales >0	2027	2298	2624	2710	2522	1924
Q _{0.50} : Annual sales >42 (1)	1501	1414	1460	1378	965	1525
Q _{0.75} : Annual sales >398 (13)	840	791	752	607	380	796
Q _{0.90} : Annual sales > 3470 (162)	346	344	307	231	120	324
Q _{0.95} : Annual sales > 12140 (620)	179	207	155	97	36	162
Q _{0.99} : Annual sales > 81606 (5037)	50	37	34	10	3	32
DVD (N = 485,134)						
Annual sales = 0	0	3	7	26	54	150
Annual sales >0	646	1010	1729	2344	3168	3831
Q _{0.50} : Annual sales >480 (169)	375	548	919	1229	1418	1988
Q _{0.75} : Annual sales >2784 (994)	200	302	482	592	669	995
Q _{0.90} : Annual sales > 12730 (7071)	81	113	199	232	273	398
Q _{0.95} : Annual sales > 34327 (18912)	40	49	90	122	148	199
Q _{0.99} : Annual sales > 182809 (91439)	9	5	21	28	26	39

Notes: A product is a specific title (for example, unlike Table 2B, if there are a Director's Cut and regular version of a particular DVD available within the same year, they are counted as one and the same product). We report the median number of titles that exceed specific annual sales numbers. Sales quantiles are based on annual sales for the 2000 to 2004 period. Sales quantiles for 2005, a year in which we observe sales only through August, are reported in parentheses.

Table 3: Descriptive Statistics for the Distributions

		2000	2001	2002	2003	2004	- 08/2005
Full Sample (N=1,212,863)							
Location	$Q_{0.5}$	4.00	2.00	2.00	1.00	1.00	0.00
Scale	$(Q_{0.75}-Q_{0.25})/(Q_{0.75}+Q_{0.25})$	0.93	1.00	1.00	1.00	1.00	1.00
Skewness	$(Q_{0.75}+Q_{0.25}-2Q_{0.5})/(Q_{0.75}-Q_{0.25})$	0.76	0.80	0.76	0.83	0.75	1.00
Kurtosis	$(Q_{0.90}-Q_{0.10})/(Q_{0.75}-Q_{0.25})$	5.56	6.10	5.94	6.42	6.88	7.33
Inter-quartile	$Q_{0.75}-Q_{0.25}$	25.00	20.00	17.00	12.00	8.00	6.00
- Left tail	$(Q_{0.50}-Q_{0.25}) / Q_{0.50}$	0.75	1.00	1.00	1.00	1.00	--
- Right tail	$(Q_{0.75}-Q_{0.50}) / Q_{0.50}$	5.50	9.00	7.50	11.00	7.00	--
VHS (N=727,729)							
Location	$Q_{0.5}$	3.00	1.00	1.00	0.00	0.00	0.00
Scale	$(Q_{0.75}-Q_{0.25})/(Q_{0.75}+Q_{0.25})$	1.00	1.00	1.00	1.00	1.00	--
Skewness	$(Q_{0.75}+Q_{0.25}-2Q_{0.5})/(Q_{0.75}-Q_{0.25})$	0.68	0.82	0.67	1.00	1.00	--
Kurtosis	$(Q_{0.90}-Q_{0.10})/(Q_{0.75}-Q_{0.25})$	6.21	6.82	7.00	6.33	6.00	--
Inter-quartile	$Q_{0.75}-Q_{0.25}$	19.00	11.00	6.00	3.00	1.00	0.00
- Left tail	$(Q_{0.50}-Q_{0.25}) / Q_{0.50}$	1.00	1.00	1.00	--	--	--
- Right tail	$(Q_{0.75}-Q_{0.50}) / Q_{0.50}$	5.33	10.00	5.00	--	--	--
DVD (N=485,134)							
Location	$Q_{0.5}$	11.00	10.00	11.00	8.00	5.00	3.00
Scale	$(Q_{0.75}-Q_{0.25})/(Q_{0.75}+Q_{0.25})$	0.94	0.97	0.97	0.95	0.93	1.00
Skewness	$(Q_{0.75}+Q_{0.25}-2Q_{0.5})/(Q_{0.75}-Q_{0.25})$	0.69	0.72	0.66	0.66	0.71	0.71
Kurtosis	$(Q_{0.90}-Q_{0.10})/(Q_{0.75}-Q_{0.25})$	3.53	3.48	3.60	4.07	4.89	5.43
Inter-quartile	$Q_{0.75}-Q_{0.25}$	58.00	64.00	58.00	41.00	28.00	21.00
- Left tail	$(Q_{0.50}-Q_{0.25}) / Q_{0.50}$	0.82	0.90	0.91	0.88	0.80	1.00
- Right tail	$(Q_{0.75}-Q_{0.50}) / Q_{0.50}$	4.45	5.50	4.36	4.25	4.80	6.00

Table 4: Kolmogorov-Smirnov (KS) Two-Sample Tests (Asymptotic)

Years	Full Sample (N=1,212,863)				VHS (N=727,729)				DVD (N=485,134)			
	KS	KSa	D	P>KSa	KS	KSa	D	P>KSa	KS	KSa	D	P>KSa
2000 vs. 2001	0.074	5.146	0.150	0.000	0.090	5.791	0.181	0.000	0.021	0.829	0.044	0.498
2000 vs. 2002	0.082	6.163	0.170	0.000	0.127	8.554	0.257	0.000	0.023	1.058	0.051	0.213
2000 vs. 2003	0.083	6.507	0.174	0.000	0.161	11.162	0.330	0.000	0.022	1.162	0.054	0.134
2000 vs. 2004	0.101	8.373	0.218	0.000	0.233	16.315	0.479	0.000	0.045	2.688	0.122	0.000
2000 vs. 2005	0.136	11.825	0.302	0.000	0.326	22.991	0.672	0.000	0.078	5.087	0.227	0.000
2001 vs. 2002	0.013	1.034	0.027	0.236	0.042	2.970	0.085	0.000	0.011	0.548	0.023	0.925
2001 vs. 2003	0.025	2.084	0.052	0.000	0.078	5.556	0.156	0.000	0.020	1.099	0.043	0.178
2001 vs. 2004	0.042	3.603	0.087	0.000	0.152	10.973	0.307	0.000	0.040	2.527	0.095	0.000
2001 vs. 2005	0.084	7.543	0.179	0.000	0.254	18.500	0.515	0.000	0.077	5.268	0.194	0.000
2002 vs. 2003	0.019	1.626	0.038	0.010	0.044	3.302	0.089	0.000	0.018	1.134	0.037	0.153
2002 vs. 2004	0.034	3.091	0.069	0.000	0.114	8.603	0.229	0.000	0.043	2.899	0.089	0.000
2002 vs. 2005	0.081	7.645	0.166	0.000	0.232	17.573	0.466	0.000	0.086	6.291	0.187	0.000
2003 vs. 2004	0.024	2.294	0.049	0.000	0.078	5.973	0.156	0.000	0.040	2.926	0.082	0.000
2003 vs. 2005	0.066	6.453	0.134	0.000	0.191	14.769	0.383	0.000	0.089	6.842	0.183	0.000
2004 vs. 2005	0.049	4.977	0.099	0.000	0.131	10.193	0.262	0.000	0.057	4.721	0.116	0.000

Notes: The tables report the Kolmogorov-Smirnov (KS) statistic, the asymptotic Kolmogorov-Smirnov (KSa) statistic, the two-sample Kolmogorov-Smirnov test statistic (D), and the p-value for this test. The null hypothesis is that the two empirical distribution functions are identical.

Table 5a: Quantile Regression Analysis—Common Year Effects for DVD and VHS Titles

	log Sales (OLS)	log Sales Q40	log Sales Q50	log Sales Q60	log Sales Q70	log Sales Q80	log Sales Q90	log Sales Q95	log Sales Q99
Fraction of Sales on DVD	-0.000 (0.000)**	-0.002 (0.000)**	-0.002 (0.000)**	-0.001 (0.000)**	-0.001 (0.000)**	-0.000 (0.000)	0.002 (0.000)**	0.003 (0.000)**	0.002 (0.000)**
DVD	1.401 (0.006)**	1.227 (0.002)**	1.633 (0.004)**	1.856 (0.016)**	1.967 (0.027)**	2.007 (0.030)**	1.833 (0.027)**	1.718 (0.011)**	1.483 (0.027)**
Year 2001	-0.353 (0.010)**	-0.231 (0.037)**	-0.488 (0.019)**	-0.535 (0.009)**	-0.520 (0.008)**	-0.501 (0.020)**	-0.593 (0.030)**	-0.627 (0.032)**	-0.803 (0.020)**
Year 2002	-0.602 (0.010)**	-0.269 (0.041)**	-0.611 (0.028)**	-0.821 (0.037)**	-0.973 (0.020)**	-1.072 (0.031)**	-1.189 (0.048)**	-1.273 (0.045)**	-1.414 (0.008)**
Year 2003	-0.843 (0.009)**	-0.305 (0.042)**	-0.692 (0.029)**	-0.968 (0.031)**	-1.257 (0.017)**	-1.511 (0.024)**	-1.771 (0.039)**	-1.845 (0.029)**	-1.975 (0.014)**
Year 2004	-1.146 (0.009)**	-0.388 (0.045)**	-0.847 (0.032)**	-1.180 (0.033)**	-1.520 (0.019)**	-1.858 (0.023)**	-2.142 (0.030)**	-2.311 (0.021)**	-2.383 (0.022)**
Year 2005	-1.384 (0.009)**	-0.463 (0.044)**	-0.981 (0.030)**	-1.346 (0.029)**	-1.699 (0.012)**	-2.073 (0.020)**	-2.445 (0.044)**	-2.644 (0.051)**	-2.735 (0.000)**
Genre: Action	-0.536 (0.010)**	-0.210 (0.006)**	-0.339 (0.003)	-0.436 (0.004)**	-0.512 (0.001)**	-0.594 (0.005)**	-0.572 (0.019)**	-0.678 (0.004)**	-0.961 (0.033)**
Genre: Adult	-0.578 (0.009)**	-0.195 (0.001)**	-0.276 (0.005)**	-0.350 (0.006)**	-0.428 (0.012)**	-0.626 (0.009)**	-0.864 (0.004)**	-1.209 (0.004)**	-2.080 (0.055)**
Genre: Children	-0.298 (0.012)**	-0.123 (0.002)**	-0.206 (0.007)**	-0.247 (0.015)**	-0.207 (0.022)**	-0.079 (0.014)**	0.236 (0.031)**	0.496 (0.028)**	0.523 (0.001)**
Genre: Comedy	-0.363 (0.009)**	-0.131 (0.005)**	-0.212 (0.006)**	-0.223 (0.010)**	-0.213 (0.013)**	-0.209 (0.014)**	-0.192 (0.009)**	-0.290 (0.015)**	-0.631 (0.074)**
Genre: Documentary	-0.553 (0.008)**	-0.132 (0.007)**	-0.234 (0.003)**	-0.317 (0.005)**	-0.425 (0.002)**	-0.577 (0.000)**	-0.755 (0.029)**	-1.051 (0.016)**	-1.714 (0.031)**
Genre: Drama	-0.632 (0.009)**	-0.233 (0.012)**	-0.395 (0.005)**	-0.503 (0.008)**	-0.558 (0.001)**	-0.663 (0.001)**	-0.714 (0.028)**	-0.863 (0.027)**	-1.292 (0.045)**
Genre: Family	-0.099 (0.026)**	-0.104 (0.012)**	0.016 (0.017)	0.030 (0.057)	-0.074 (0.008)**	0.005 (0.004)	0.530 (0.034)**	0.530 (0.024)**	0.686 (0.183)**

Table 5a: Quantile Regression Analysis—Common Year Effects for DVD and VHS Titles (Continued)

	log Sales (OLS)	log Sales Q40	log Sales Q50	log Sales Q60	log Sales Q70	log Sales Q80	log Sales Q90	log Sales Q95	log Sales Q99
Genre: Foreign	-0.484 (0.014)**	-0.198 (0.005)**	-0.321 (0.003)**	-0.443 (0.011)**	-0.503 (0.014)**	-0.636 (0.002)**	-0.832 (0.007)**	-0.902 (0.003)**	-1.643 (0.127)**
Genre: Horror	-0.416 (0.012)**	-0.167 (0.001)**	-0.241 (0.004)**	-0.321 (0.005)**	-0.364 (0.002)**	-0.463 (0.006)**	-0.817 (0.036)**	-1.280 (0.040)**	-2.112 (0.113)**
Genre: Anime	-0.529 (0.021)**	-0.199 (0.001)**	-0.306 (0.004)**	-0.406 (0.016)**	-0.506 (0.045)**	-0.484 (0.016)**	-0.766 (0.032)**	-0.812 (0.005)**	-0.901 (0.142)**
Genre: Mystery	-0.780 (0.014)**	-0.312 (0.007)**	-0.512 (0.000)**	-0.663 (0.004)**	-0.740 (0.009)**	-0.780 (0.004)**	-0.862 (0.033)**	-0.861 (0.006)**	-1.308 (0.046)**
Genre: Sci-fi	-0.784 (0.013)**	-0.298 (0.012)**	-0.522 (0.015)**	-0.642 (0.013)**	-0.749 (0.018)**	-0.876 (0.005)**	-1.022 (0.010)**	-1.034 (0.015)**	-1.626 (0.012)**
Rating: G	0.595 (0.026)**	0.098 (0.027)**	0.237 (0.029)**	0.657 (0.024)**	1.135 (0.024)**	0.912 (0.010)**	0.693 (0.090)**	0.682 (0.084)**	0.352 (0.281)
Rating: PG	0.417 (0.013)**	0.227 (0.001)**	0.410 (0.003)**	0.442 (0.002)**	0.606 (0.012)**	0.758 (0.014)**	0.621 (0.002)**	0.506 (0.019)**	0.627 (0.010)**
Rating: PG13	0.387 (0.015)**	0.233 (0.007)**	0.302 (0.007)**	0.434 (0.003)**	0.591 (0.018)**	0.637 (0.025)**	0.575 (0.004)**	0.730 (0.022)**	1.331 (0.051)**
Rating: R	0.441 (0.009)**	0.217 (0.006)**	0.376 (0.004)**	0.452 (0.000)**	0.546 (0.001)**	0.627 (0.005)**	0.658 (0.009)**	0.789 (0.027)**	1.047 (0.022)**
Rating: NC17	1.894 (0.054)**	1.803 (0.061)**	2.022 (0.057)**	2.002 (0.136)**	2.080 (0.221)**	3.495 (0.282)**	3.775 (0.005)**	3.309 (0.028)**	3.151 (0.062)**
Rating: X	-0.554 (0.048)**	-0.791 (0.003)**	-0.976 (0.018)**	-0.662 (0.059)**	-0.761 (0.078)**	-0.286 (0.005)**	-0.408 (0.053)**	-0.534 (0.009)**	-0.789 (0.109)**
Animation: Animated	0.689 (0.017)**	0.149 (0.007)**	0.281 (0.013)**	0.423 (0.020)**	0.637 (0.027)**	0.854 (0.004)**	1.110 (0.040)**	1.129 (0.025)**	0.619 (0.042)**
Animation: Live Action	0.317 (0.011)**	0.081 (0.002)**	0.162 (0.000)**	0.232 (0.004)**	0.298 (0.013)**	0.411 (0.018)**	0.413 (0.021)**	0.249 (0.018)**	-0.019 (0.032)
Original Language Non- English	-0.090 (0.013)**	0.065 (0.003)**	0.060 (0.014)**	0.053 (0.003)**	0.008 (0.001)**	-0.64 (0.020)**	-0.076 (0.007)**	-0.327 (0.027)**	-0.342 (0.035)**

Table 5a: Quantile Regression Analysis—Common Year Effects for DVD and VHS Titles (Continued)

	log Sales (OLS)	log Sales Q40	log Sales Q50	log Sales Q60	log Sales Q70	log Sales Q80	log Sales Q90	log Sales Q95	log Sales Q99
Translation: Dubbed	0.319 (0.013)**	0.090 (0.004)**	0.207 (0.001)**	0.302 (0.006)**	0.327 (0.001)**	0.357 (0.005)**	0.401 (0.027)**	0.187 (0.012)**	-0.028 (0.151)
Translation: Subtitled	-0.149 (0.014)**	-0.069 (0.005)**	-0.108 (0.014)**	-0.119 (0.002)**	-1.516 (0.002)**	-0.189 (0.005)**	-0.102 (0.010)**	-0.003 (0.045)	-0.268 (0.182)
Part of Franchise	0.490 (0.010)**	0.102 (0.005)**	0.228 (0.006)**	0.339 (0.001)**	0.471 (0.006)**	0.529 (0.001)**	0.387 (0.014)**	0.161 (0.023)**	-0.384 (0.001)**
Year of Production	0.005 (0.000)**	0.000 (0.000)**	0.001 (0.000)**	0.003 (0.000)**	0.005 (0.000)**	0.008 (0.000)**	0.013 (0.000)**	0.015 (0.001)**	0.015 (0.002)**
Runtime	0.001 (0.000)**	0.001 (0.000)**	0.002 (0.000)**	0.002 (0.000)**	0.002 (0.000)**	0.003 (0.000)**	0.003 (0.000)**	0.003 (0.000)**	0.003 (0.000)**
Distributor: Disney	0.608 (0.012)**	0.131 (0.011)**	0.237 (0.006)**	0.333 (0.000)**	0.504 (0.009)**	1.094 (0.016)**	1.409 (0.002)**	1.277 (0.013)**	1.302 (0.046)**
Distributor: 20th Century Fox	0.961 (0.013)**	0.247 (0.006)**	0.676 (0.002)**	1.043 (0.011)**	1.240 (0.011)**	1.363 (0.023)**	1.374 (0.018)**	1.569 (0.003)**	1.325 (0.009)**
Distributor: Lions Gate	0.290 (0.011)**	0.065 (0.007)**	0.124 (0.006)**	0.230 (0.019)**	0.375 (0.024)**	0.507 (0.024)**	0.651 (0.032)**	0.887 (0.000)**	0.875 (0.032)**
Distributor: Paramount	0.884 (0.013)**	0.321 (0.010)**	0.608 (0.005)**	0.775 (0.014)**	0.594 (0.013)**	0.961 (0.017)**	0.970 (0.029)**	0.704 (0.003)**	0.307 (0.144)*
Distributor: Sony	0.691 (0.010)**	0.288 (0.001)**	0.540 (0.001)**	0.735 (0.009)**	0.793 (0.001)**	0.873 (0.010)**	0.855 (0.013)**	0.717 (0.001)**	0.591 (0.023)**
Distributor: Universal	0.644 (0.013)**	0.123 (0.011)**	0.282 (0.020)**	0.428 (0.029)**	0.648 (0.040)**	0.843 (0.052)**	1.024 (0.040)**	1.105 (0.002)**	1.595 (0.006)**
Distributor: Warner Bros.	0.628 (0.008)**	0.161 (0.008)**	0.329 (0.006)**	0.491 (0.004)**	0.647 (0.003)**	0.778 (0.012)**	0.881 (0.020)**	0.869 (0.017)**	0.854 (0.070)**
Original Release: Theatrical	0.523 (0.008)**	0.113 (0.003)**	0.214 (0.001)**	0.301 (0.005)**	0.358 (0.004)**	0.471 (0.003)**	0.761 (0.000)**	0.918 (0.006)**	0.775 (0.055)**
Original Release: Direct to Video	0.122 (0.006)**	-0.042 (0.002)**	-0.045 (0.003)**	-0.011 (0.001)**	0.059 (0.004)**	0.175 (0.002)**	0.458 (0.010)**	0.670 (0.026)**	0.862 (0.023)**

Table 5a: Quantile Regression Analysis—Common Year Effects for DVD and VHS Titles (Continued)

	log Sales (OLS)	log Sales Q40	log Sales Q50	log Sales Q60	log Sales Q70	log Sales Q80	log Sales Q90	log Sales Q95	log Sales Q99
Original Release: Television	-0.063 (0.008)**	0.006 (0.000)**	-0.015 (0.005)**	-0.052 (0.004)**	-0.153 (0.010)**	-0.288 (0.011)**	-0.341 (0.003)**	-0.331 (0.018)**	-0.034 (0.040)
Video Release: Sell-Through	0.582 (0.012)**	0.577 (0.029)**	0.682 (0.030)**	0.730 (0.036)**	0.762 (0.042)**	0.799 (0.038)**	0.939 (0.016)**	1.087 (0.031)**	1.101 (0.042)**
Video Release: Rental	0.580 (0.008)**	0.313 (0.003)**	0.492 (0.018)**	0.617 (0.028)**	0.694 (0.036)**	0.722 (0.010)**	0.613 (0.016)**	0.476 (0.007)**	0.440 (0.011)**
Suggested price	0.001 (0.000)**	-0.001 (0.000)**	-0.001 (0.000)**	-0.000 (0.000)	0.001 (0.000)**	0.001 (0.000)**	-0.001 (0.000)**	0.001 (0.000)**	0.009 (0.001)**
Weeks Since Release	-0.001 (0.000)**	-0.000 (0.000)**	-0.001 (0.000)**	-0.000 (0.000)**	-0.001 (0.000)**	-0.002 (0.000)**	-0.002 (0.000)**	-0.002 (0.000)**	-0.002 (0.000)**
Constant	-9.040 (0.320)**	-0.481 (0.167)**	-1.927 (0.355)**	-4.214 (0.677)**	-8.108 (0.549)**	-13.248 (0.712)**	-21.477 (0.852)**	-24.785 (0.000)**	-23.281 (3.173)**
Observations	573,753	573,753	573,753	573,753	573,753	573,753	573,753	573,753	573,753
(Pseudo) R-squared	0.310	0.081	0.185	0.235	0.255	0.260	0.255	0.251	0.242

Notes: The first column presents OLS estimates, the remainder of the table reports quantile regression model estimates. One asterisk (*) indicates significance at 5% and two (**) significance at 1%. All models omit a dummy for “Year 2000,” and include “Month” fixed effects.

Table 5b: Quantile Regression Analysis—Year Effects Vary by Format

	log Sales (OLS)	log Sales Q40	log Sales Q50	log Sales Q60	log Sales Q70	log Sales Q80	log Sales Q90	log Sales Q95	log Sales Q99
Fraction of Sales on DVD	-0.029 (0.005)**	-0.129 (0.000)**	-0.165 (0.004)**	-0.137 (0.010)**	-0.082 (0.008)**	-0.006 (0.015)	0.154 (0.024)**	0.277 (0.028)**	-0.072 (0.072)
DVD	0.848 (0.013)**	1.609 (0.017)**	1.482 (0.014)**	1.316 (0.011)**	1.060 (0.034)**	0.703 (0.033)**	0.176 (0.012)**	-0.172 (0.026)**	-0.373 (0.018)**
Year 2001	-0.377 (0.008)**	-0.164 (0.004)**	-0.554 (0.015)**	-0.671 (0.007)**	-0.605 (0.016)**	-0.581 (0.013)**	-0.609 (0.021)**	-0.563 (0.025)**	-0.623 (0.036)**
Year 2002	-0.710 (0.008)**	-0.210 (0.004)**	-0.694 (0.019)**	-1.018 (0.012)**	-1.224 (0.010)**	-1.310 (0.008)**	-1.320 (0.012)**	-1.281 (0.007)**	-1.312 (0.011)**
Year 2003	-1.033 (0.007)**	-0.242 (0.005)**	-0.775 (0.021)**	-1.167 (0.019)**	-1.546 (0.020)**	-1.915 (0.017)**	-2.177 (0.023)**	-2.180 (0.010)**	-2.080 (0.011)**
Year 2004	-1.390 (0.007)**	-0.294 (0.004)**	-0.891 (0.017)**	-1.343 (0.015)**	-1.787 (0.020)**	-2.321 (0.019)**	-2.806 (0.035)**	-2.961 (0.055)**	-2.945 (0.026)**
Year 2005	-1.615 (0.007)**	-0.332 (0.002)**	-0.969 (0.016)**	-1.451 (0.016)**	-1.915 (0.018)**	-2.485 (0.013)**	-3.065 (0.024)**	-3.338 (0.020)**	-3.600 (0.019)**
DVD × 2001	0.256 (0.017)**	-0.034 (0.034)	0.395 (0.019)**	0.559 (0.015)**	0.550 (0.046)**	0.552 (0.042)**	0.710 (0.029)**	0.703 (0.052)**	0.749 (0.012)**
DVD × 2002	0.530 (0.016)**	0.083 (0.022)**	0.532 (0.012)**	0.870 (0.015)**	1.108 (0.047)**	1.217 (0.048)**	1.250 (0.036)**	1.268 (0.050)**	1.351 (0.033)**
DVD × 2003	0.619 (0.015)**	-0.121 (0.021)**	0.352 (0.017)**	0.695 (0.018)**	1.083 (0.046)**	1.436 (0.050)**	1.758 (0.020)**	1.871 (0.015)**	1.968 (0.019)**
DVD × 2004	0.695 (0.014)**	-0.489 (0.014)**	0.066 (0.020)**	0.494 (0.020)**	0.990 (0.037)**	1.599 (0.037)**	2.217 (0.028)**	2.424 (0.066)**	2.700 (0.016)**
DVD × 2005	0.666 (0.015)**	-0.758 (0.042)**	-0.185 (0.000)**	0.305 (0.000)**	0.825 (0.028)**	1.495 (0.025)**	2.161 (0.026)**	2.511 (0.051)**	3.087 (0.023)**
Genre: Action	-0.453 (0.007)**	-0.173 (0.000)**	-0.285 (0.004)**	-0.343 (0.001)**	-0.378 (0.005)**	-0.435 (0.001)**	-0.454 (0.000)**	-0.533 (0.007)**	-0.458 (0.025)**
Genre: Adult	-0.544 (0.007)**	-0.196 (0.002)**	-0.285 (0.000)**	-0.342 (0.002)**	-0.417 (0.001)**	-0.529 (0.005)**	-0.765 (0.000)**	-1.103 (0.011)**	-1.749 (0.008)**

Table 5b: Quantile Regression Analysis—Year Effects Vary by Format (Continued)

	log Sales (OLS)	log Sales Q40	log Sales Q50	log Sales Q60	log Sales Q70	log Sales Q80	log Sales Q90	log Sales Q95	log Sales Q99
Genre: Children	-0.129 (0.009)**	-0.088 (0.001)**	-0.149 (0.004)**	-0.158 (0.000)**	-0.057 (0.002)**	0.198 (0.018)**	0.627 (0.006)**	0.835 (0.007)**	1.096 (0.041)**
Genre: Comedy	-0.318 (0.006)**	-0.091 (0.000)**	-0.157 (0.002)**	-0.184 (0.004)**	-0.200 (0.005)**	-0.186 (0.001)**	-0.185 (0.015)**	-0.240 (0.009)**	-0.396 (0.001)**
Genre: Documentary	-0.548 (0.006)**	-0.111 (0.000)**	-0.209 (0.004)**	-0.300 (0.005)**	-0.417 (0.000)**	-0.544 (0.001)**	-0.649 (0.008)**	-0.905 (0.019)**	-1.449 (0.022)**
Genre: Drama	-0.589 (0.006)**	-0.185 (0.004)**	-0.364 (0.008)**	-0.469 (0.005)**	-0.543 (0.005)**	-0.616 (0.000)**	-0.594 (0.009)**	-0.658 (0.007)**	-0.806 (0.000)**
Genre: Family	0.010 (0.020)	-0.035 (0.005)**	0.017 (0.005)**	0.204 (0.000)**	0.545 (0.006)**	0.611 (0.001)**	0.755 (0.023)**	0.643 (0.032)**	0.722 (0.097)**
Genre: Foreign	-0.428 (0.009)**	-0.115 (0.001)**	-0.237 (0.001)**	-0.357 (0.003)**	-0.450 (0.013)**	-0.570 (0.005)**	-0.648 (0.008)**	-0.741 (0.011)**	-0.983 (0.014)**
Genre: Horror	-0.330 (0.008)**	-0.070 (0.002)**	-0.127 (0.007)**	-0.183 (0.008)**	-0.278 (0.017)**	-0.409 (0.017)**	-0.691 (0.020)**	-0.931 (0.025)**	-1.196 (0.068)**
Genre: Anime	-0.474 (0.015)**	-0.105 (0.002)**	-0.216 (0.004)**	-0.351 (0.001)**	-0.446 (0.012)**	-0.492 (0.003)**	-0.677 (0.006)**	-0.950 (0.065)**	-0.646 (0.004)**
Genre: Mystery	-0.652 (0.009)**	-0.204 (0.008)**	-0.352 (0.011)**	-0.469 (0.008)**	-0.562 (0.002)**	-0.699 (0.016)**	-0.814 (0.015)**	-0.898 (0.040)**	-1.080 (0.028)**
Genre: Sci-fi	-0.577 (0.009)**	-0.160 (0.002)**	-0.319 (0.005)**	-0.447 (0.007)**	-0.556 (0.003)**	-0.698 (0.003)**	-0.820 (0.018)**	-0.793 (0.001)**	-1.210 (0.033)**
Rating: G	0.763 (0.018)**	0.533 (0.040)**	0.950 (0.024)**	1.205 (0.005)**	1.075 (0.001)**	0.989 (0.032)**	1.004 (0.031)**	1.040 (0.013)**	1.153 (0.016)**
Rating: PG	0.525 (0.010)**	0.353 (0.025)**	0.602 (0.016)**	0.711 (0.027)**	0.874 (0.024)**	0.935 (0.026)**	0.796 (0.021)**	0.760 (0.038)**	0.726 (0.022)**
Rating: PG13	0.993 (0.012)**	0.968 (0.049)**	1.191 (0.038)**	1.357 (0.031)**	1.520 (0.013)**	1.701 (0.001)**	1.805 (0.029)**	1.960 (0.018)**	1.750 (0.010)**
Rating: R	0.401 (0.006)**	0.139 (0.001)**	0.282 (0.005)**	0.411 (0.014)**	0.521 (0.010)**	0.659 (0.010)**	0.889 (0.004)**	1.071 (0.019)**	1.089 (0.023)**

Table 5b: Quantile Regression Analysis—Year Effects Vary by Format (Continued)

	log Sales (OLS)	log Sales Q40	log Sales Q50	log Sales Q60	log Sales Q70	log Sales Q80	log Sales Q90	log Sales Q95	log Sales Q99
Rating: NC17	1.881 (0.054)**	1.865 (0.000)**	2.112 (0.028)**	2.111 (0.022)**	2.177 (0.154)**	3.517 (0.277)**	3.585 (0.023)**	2.982 (0.013)**	2.381 (0.010)**
Rating: X	-0.161 (0.029)**	0.127 (0.037)**	-0.103 (0.001)**	-0.163 (0.013)**	-0.132 (0.049)**	-0.288 (0.026)**	-0.739 (0.043)**	-0.994 (0.023)**	-1.618 (0.111)**
Animation: Animated	0.546 (0.012)**	0.116 (0.000)**	0.248 (0.001)**	0.410 (0.008)**	0.607 (0.002)**	0.793 (0.009)**	0.943 (0.015)**	1.028 (0.009)**	0.178 (0.063)**
Animation: Live Action	0.248 (0.008)**	0.070 (0.002)**	0.150 (0.006)**	0.209 (0.007)**	0.274 (0.011)**	0.350 (0.019)**	0.353 (0.000)**	0.286 (0.004)**	-0.074 (0.031)**
Original Language Non-English	-0.077 (0.009)**	0.031 (0.000)**	0.024 (0.001)**	0.014 (0.007)**	-0.022 (0.000)**	-0.100 (0.001)**	-0.122 (0.008)**	-0.183 (0.051)**	-0.401 (0.014)**
Translation: Dubbed	0.326 (0.009)**	0.043 (0.001)**	0.136 (0.004)**	0.225 (0.010)**	0.253 (0.013)**	0.361 (0.004)**	0.306 (0.004)**	0.144 (0.008)**	0.028 (0.021)
Translation: Subtitled	-0.221 (0.009)**	-0.091 (0.006)**	-0.122 (0.007)**	-0.144 (0.006)**	-0.180 (0.005)**	-0.185 (0.006)**	-0.182 (0.019)**	-0.123 (0.041)**	-0.115 (0.009)**
Part of Franchise	0.377 (0.007)**	0.064 (0.003)**	0.138 (0.004)**	0.213 (0.005)**	0.264 (0.007)**	0.292 (0.002)**	0.139 (0.022)**	-0.053 (0.039)	-0.170 (0.036)**
Year of Production	0.005 (0.000)**	0.001 (0.000)**	0.002 (0.000)**	0.003 (0.000)**	0.005 (0.000)**	0.007 (0.000)**	0.010 (0.000)**	0.014 (0.000)**	0.018 (0.000)**
Runtime	0.002 (0.000)**	0.001 (0.000)**	0.002 (0.000)**	0.002 (0.000)**	0.003 (0.000)**	0.003 (0.000)**	0.003 (0.000)**	0.003 (0.000)**	0.002 (0.000)**
Distributor: Disney	0.597 (0.009)**	0.119 (0.001)**	0.247 (0.003)**	0.385 (0.006)**	0.595 (0.029)**	0.828 (0.032)**	0.960 (0.002)**	0.894 (0.007)**	0.733 (0.046)**
Distributor: 20th Century Fox	1.083 (0.010)**	0.515 (0.008)**	0.922 (0.029)**	1.162 (0.027)**	1.297 (0.016)**	1.380 (0.013)**	1.347 (0.011)**	1.313 (0.012)**	1.235 (0.031)**
Distributor: Lions Gate	0.176 (0.008)**	-0.014 (0.001)**	0.000 (0.001)	0.064 (0.000)**	0.191 (0.003)**	0.368 (0.006)**	0.485 (0.007)**	0.501 (0.008)**	0.541 (0.026)**
Distributor: Paramount	0.868 (0.010)**	0.182 (0.006)**	0.571 (0.009)**	0.828 (0.011)**	0.978 (0.006)**	1.082 (0.012)**	1.098 (0.013)**	0.912 (0.002)**	0.821 (0.015)**

Table 5b: Quantile Regression Analysis—Year Effects Vary by Format (Continued)

	log Sales (OLS)	log Sales Q40	log Sales Q50	log Sales Q60	log Sales Q70	log Sales Q80	log Sales Q90	log Sales Q95	log Sales Q99
Distributor: Sony	0.668 (0.007)**	0.175 (0.002)**	0.430 (0.014)**	0.645 (0.019)**	0.759 (0.018)**	0.793 (0.013)**	0.683 (0.007)**	0.580 (0.008)**	0.545 (0.021)**
Distributor: Universal	0.747 (0.009)**	0.115 (0.002)**	0.314 (0.007)**	0.546 (0.011)**	0.805 (0.006)**	0.936 (0.028)**	0.877 (0.017)**	0.923 (0.030)**	1.137 (0.066)**
Distributor: Warner Bros.	0.717 (0.005)**	0.143 (0.006)**	0.350 (0.008)**	0.532 (0.003)**	0.702 (0.009)**	0.842 (0.013)**	0.861 (0.032)**	0.903 (0.031)**	0.718 (0.049)**
Original Release: Theatrical	0.534 (0.006)**	0.087 (0.000)**	0.192 (0.001)**	0.268 (0.003)**	0.352 (0.006)**	0.457 (0.003)**	0.678 (0.020)**	0.845 (0.049)**	1.252 (0.057)**
Original Release: Direct to Video	0.140 (0.004)**	-0.031 (0.000)**	-0.032 (0.000)**	0.000 (0.002)	0.072 (0.000)**	0.149 (0.006)**	0.375 (0.008)**	0.598 (0.009)**	1.049 (0.024)**
Original Release: Television	0.019 (0.006)**	0.023 (0.002)**	0.033 (0.000)**	0.018 (0.001)**	-0.024 (0.003)**	-0.103 (0.000)**	-0.113 (0.017)**	-0.043 (0.024)	0.016 (0.029)
Video Release: Sell-Through	0.402 (0.008)**	0.309 (0.025)**	0.362 (0.020)**	0.405 (0.019)**	0.464 (0.026)**	0.473 (0.008)**	0.554 (0.004)**	0.497 (0.013)**	0.323 (0.019)**
Video Release: Rental	0.667 (0.006)**	0.290 (0.002)**	0.507 (0.002)**	0.675 (0.003)**	0.799 (0.002)**	0.836 (0.000)**	0.806 (0.004)**	0.718 (0.015)**	0.688 (0.034)**
Suggested price	-0.004 (0.000)**	-0.002 (0.000)**	-0.003 (0.000)**	-0.003 (0.000)**	-0.003 (0.000)**	-0.003 (0.000)**	-0.003 (0.000)**	-0.002 (0.000)**	0.006 (0.001)**
Weeks Since Release	-0.001 (0.000)**	0.000 (0.000)**	-0.001 (0.000)**	-0.001 (0.000)**	-0.001 (0.000)**	-0.002 (0.000)**	-0.002 (0.000)**	-0.002 (0.000)**	-0.002 (0.000)**
Constant	-9.256 (0.230)	-0.701 (0.043)**	-2.471 (0.034)**	-4.257 (0.050)**	-7.504 (0.029)**	-11.002 (0.006)**	-16.600 (0.266)**	-21.831 (0.781)**	-28.456 (0.970)**
Observations	1,161,405	1,161,405	1,161,405	1,161,405	1,161,405	1,161,405	1,161,405	1,161,405	1,161,405
(Pseudo) R-squared	0.323	0.091	0.193	0.241	0.261	0.270	0.274	0.277	0.270

Notes: The first column presents OLS estimates, the remainder of the table reports quantile regression model estimates. One asterisk (*) indicates significance at 5% and two (**) significance at 1%. All models omit a dummy for “Year 2000,” and include “Month” fixed effects. For the OLS model, we can reject all null hypotheses, $H_0: \text{Year} + (\text{DVD} \times \text{Year}) = 0$, with $p < 0.000$. We also estimate a full set of inter-quantile regressions, comparing the coefficients in each column with the coefficients for Q₅₀: coefficients that are significantly different are printed bold.

Table 6a: Negative Binomial Regression Model for the Number of Titles in Sales Quantiles (I)

	(1) # titles with sales=0	(2) # titles with 0<sales ≤10	(3) # titles with 10<sales≤28	(4) # titles with 28<sales≤125	(5) # titles with sales>125
Year 2001	2.536 (0.111)**	0.997 (0.019)	0.889 (0.023)**	1.048 (0.029)	1.218 (0.038)**
Year 2002	6.360 (0.644)**	0.977 (0.043)	0.877 (0.052)*	1.310 (0.082)**	1.444 (0.103)**
Year 2003	13.639 (2.201)**	1.099 (0.077)	0.883 (0.083)	1.445 (0.144)**	1.593 (0.181)**
Year 2004	30.990 (6.851)**	1.036 (0.099)	0.761 (0.099)*	1.455 (0.198)**	1.702 (0.265)**
Year 2005	74.631 (21.472)**	0.977 (0.122)	0.736 (0.124)	1.461 (0.259)*	2.069 (0.419)**
% Titles on DVD	0.933 (0.007)**	1.020 (0.003)**	1.028 (0.005)**	1.009 (0.005)	0.995 (0.005)
Observations	296	296	296	296	296

Notes: The reported coefficients are incidence rate ratios. The dependent variable is the number of titles that meet a particular weekly sales target. The independent variables include month indicators (which are not reported in the table). The reported sales thresholds of 10, 28 and 125 copies per week correspond to the 70th, 80th, and 90th percentile of weekly sales. Standard errors are in parentheses. One asterisk (*) indicates significance at 5% and two (**) significance at 1% (the latter values are also in bold font).

Table 6b: Negative Binomial Regression Model for the Number of Titles in Sales Quantiles (II)

	(1) # titles with sales=0	(2) # titles with 0<sales ≤10	(3) # titles with 10<sales ≤28	(4) # titles with 28<sales ≤125	(5) # titles with sales>125
Year 2001	1.540 (0.088)**	1.125 (0.043)**	0.973 (0.057)	1.051 (0.060)	0.852 (0.054)*
Year 2002	1.887 (0.228)**	1.245 (0.100)**	0.998 (0.123)	1.269 (0.152)*	0.776 (0.105)
Year 2003	2.345 (0.410)**	1.482 (0.172)**	1.052 (0.187)	1.406 (0.243)*	0.687 (0.134)
Year 2004	3.294 (0.786)**	1.629 (0.258)**	1.114 (0.271)	1.614 (0.381)*	0.576 (0.153)*
Year 2005	4.610 (1.383)**	1.816 (0.362)**	1.231 (0.376)	1.789 (0.530)*	0.494 (0.166)*
% Titles on DVD	1.010 (0.007)	0.993 (0.004)	1.013 (0.007)	1.008 (0.007)	1.019 (0.008)*
% Action	1.499 (0.115)**	0.981 (0.049)	1.026 (0.081)	0.971 (0.075)	0.771 (0.063)**
% Adult	1.046 (0.016)**	1.037 (0.010)**	1.075 (0.017)**	1.051 (0.015)**	1.011 (0.016)
% Children	1.721 (0.063)**	0.954 (0.019)*	1.047 (0.033)	1.057 (0.033)	1.087 (0.037)*
% Comedy	1.181 (0.043)**	0.952 (0.023)*	1.004 (0.038)	0.980 (0.036)	1.091 (0.043)*
% Documentary	1.075 (0.037)*	0.961 (0.023)	0.892 (0.033)**	1.072 (0.038)	1.042 (0.041)
% Drama	1.016 (0.051)	0.974 (0.032)	1.015 (0.051)	0.930 (0.046)	0.947 (0.052)
% Foreign	1.114 (0.041)**	1.088 (0.026)**	1.005 (0.036)	0.988 (0.035)	1.051 (0.043)
% Rating G	0.897 (0.115)	1.017 (0.086)	0.790 (0.105)	1.018 (0.132)	0.969 (0.135)
% Rating PG13	0.766 (0.101)*	1.026 (0.089)	0.810 (0.109)	1.498 (0.195)**	1.218 (0.175)
% Rating X	2.334 (1.304)	1.012 (0.378)	6.883 (4.075)**	7.803 (4.519)**	0.628 (0.390)
Observations	296	296	296	296	296

Table 6c: Negative Binomial Regression Model for the Number of Titles in Sales Quantiles (III)

	(1) # titles with sales=0	(2) # titles with 0<sales ≤10	(3) # titles with 10<sales≤28	(4) # titles with 28<sales≤125	(5) # titles with sales>125
Year 2001	1.357 (0.088)**	1.098 (0.047)*	0.926 (0.061)	0.936 (0.060)	0.794 (0.056)**
Year 2002	1.384 (0.198)*	1.174 (0.110)	0.884 (0.128)	0.957 (0.135)	0.656 (0.101)**
Year 2003	1.462 (0.307)	1.356 (0.186)*	0.876 (0.186)	0.917 (0.189)	0.532 (0.120)**
Year 2004	1.759 (0.500)*	1.447 (0.269)*	0.874 (0.251)	0.914 (0.256)	0.410 (0.126)**
Year 2005	2.019 (0.731)	1.553 (0.368)	0.894 (0.327)	0.846 (0.301)	0.314 (0.123)**
% Titles on DVD	0.996 (0.007)	0.990 (0.005)*	1.007 (0.008)	0.996 (0.007)	1.010 (0.008)
% Action	1.516 (0.115)**	0.984 (0.049)	1.033 (0.081)	0.985 (0.076)	0.774 (0.063)**
% Adult	1.048 (0.015)**	1.037 (0.010)**	1.076 (0.017)**	1.055 (0.015)**	1.012 (0.016)
% Children	1.668 (0.061)**	0.946 (0.020)**	1.029 (0.034)	1.018 (0.033)	1.060 (0.038)
% Comedy	1.212 (0.044)**	0.958 (0.023)	1.015 (0.039)	1.007 (0.037)	1.109 (0.044)*
% Documentary	1.174 (0.048)**	0.978 (0.027)	0.924 (0.040)	1.164 (0.048)**	1.096 (0.050)*
% Drama	1.044 (0.052)	0.979 (0.032)	1.025 (0.052)	0.953 (0.047)	0.963 (0.053)
% Foreign	1.043 (0.041)	1.073 (0.028)**	0.979 (0.039)	0.929 (0.036)	1.009 (0.045)
% Rating G	0.908 (0.115)	1.018 (0.086)	0.791 (0.105)	1.028 (0.133)	0.976 (0.135)
% Rating PG13	0.857 (0.114)	1.052 (0.093)	0.851 (0.118)	1.675 (0.224)**	1.307 (0.192)
% Rating X	1.815 (1.009)	0.965 (0.361)	6.294 (3.743)**	6.228 (3.627)**	0.535 (0.332)
# of Titles Offered	1.000 (0.000)**	1.000 (0.000)	1.000 (0.000)	1.000 (0.000)**	1.000 (0.000)*
Observations	296	296	296	296	296

Figure 1: Graphic Illustration of the Hypotheses

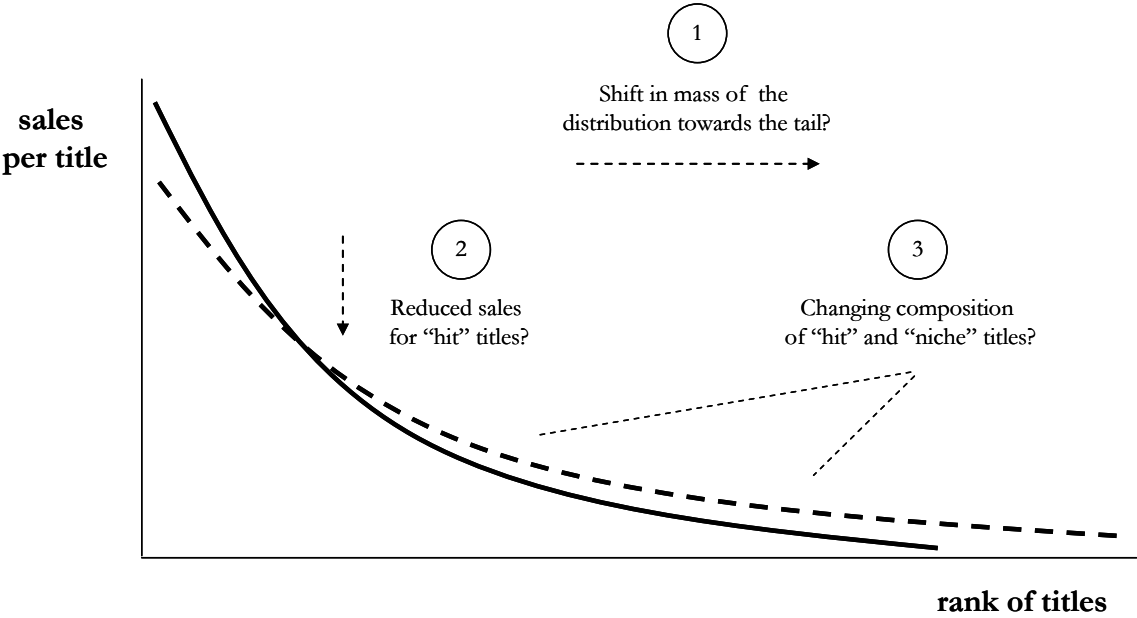
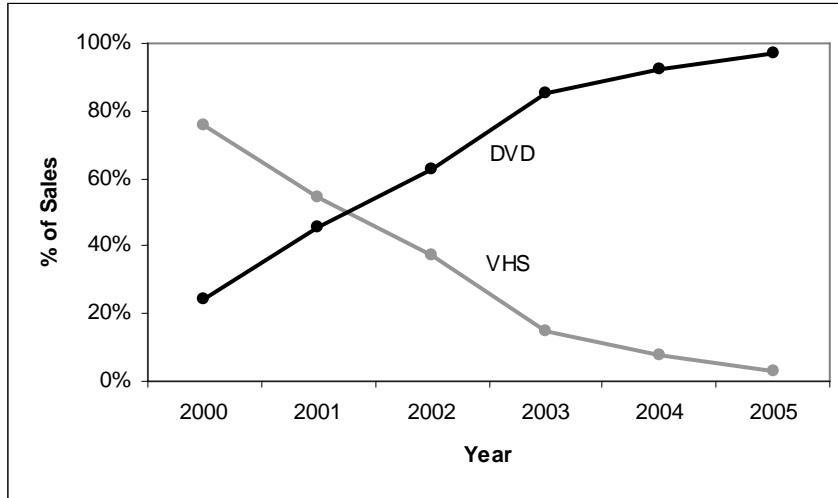


Figure 2: Annual Sales by Format, and Annual Sales by Channel

Video Sales By Year By Format



Video Sales By Year By Channel

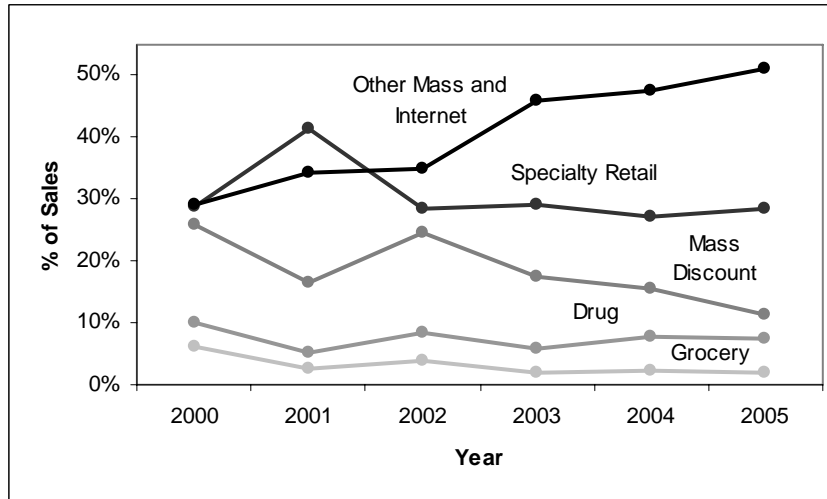


Figure 3: Quantile Plots for Some Key Covariates

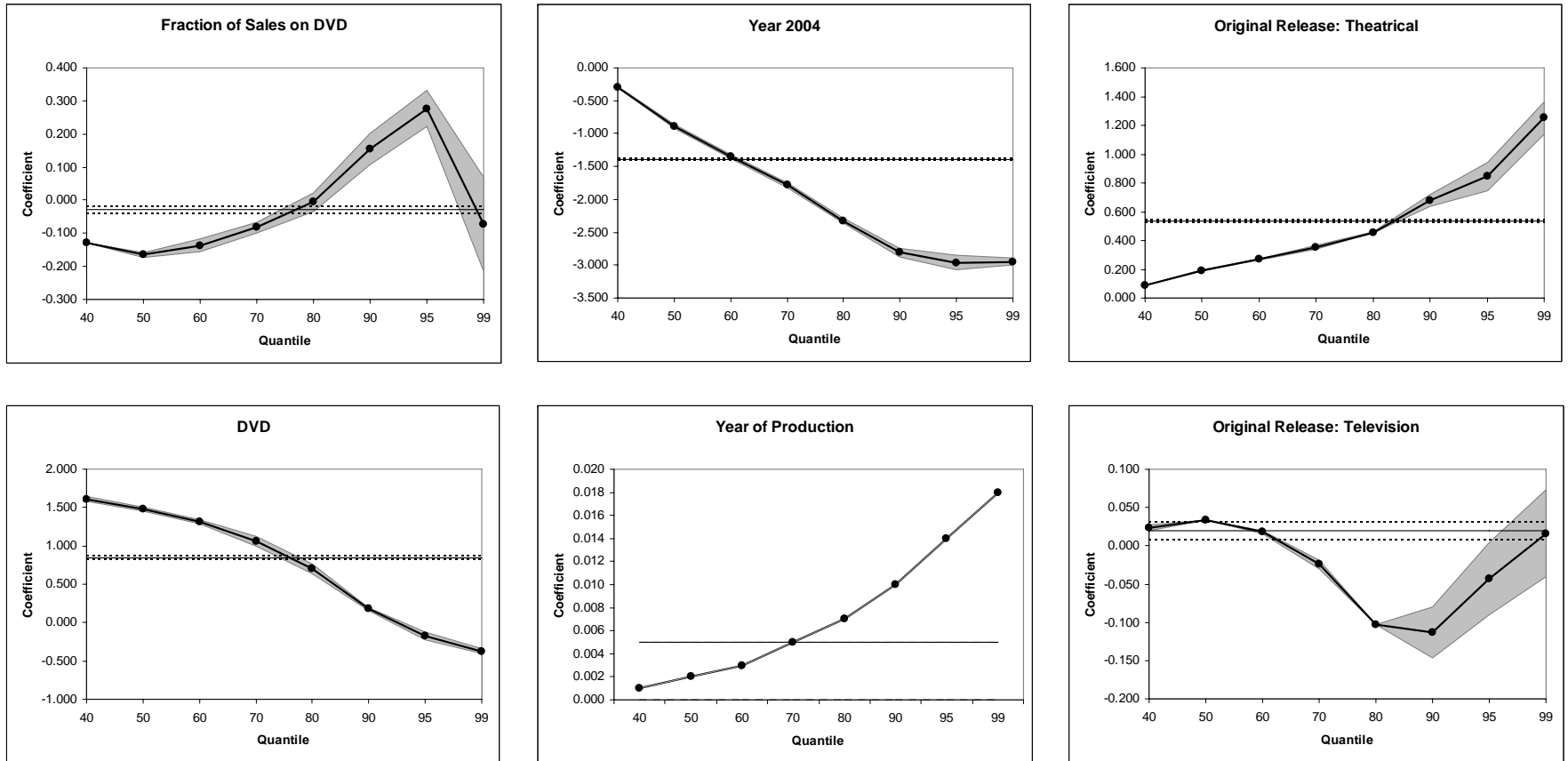


Figure 3: Quantile Plots for Some Key Covariates (Continued)

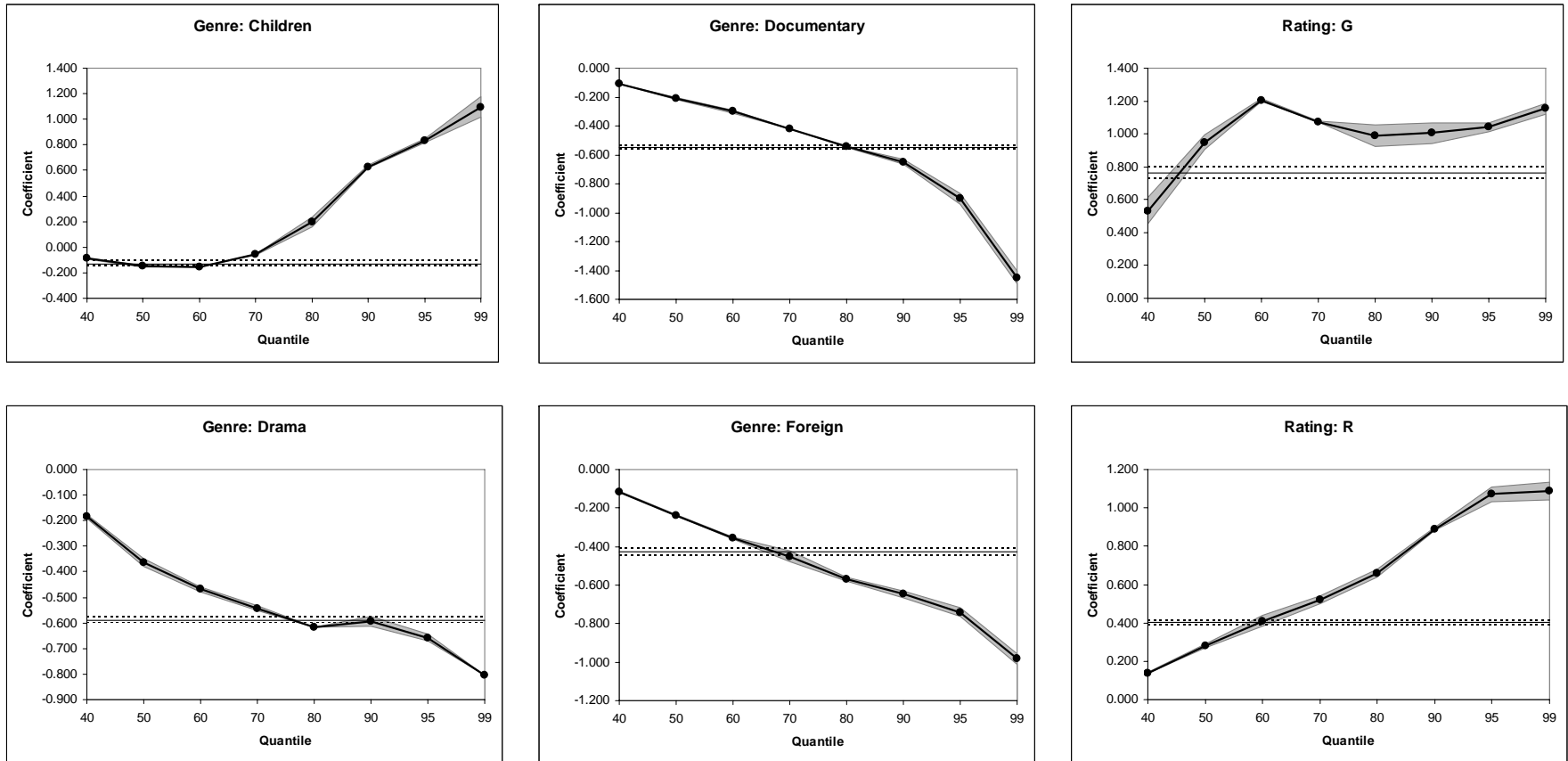
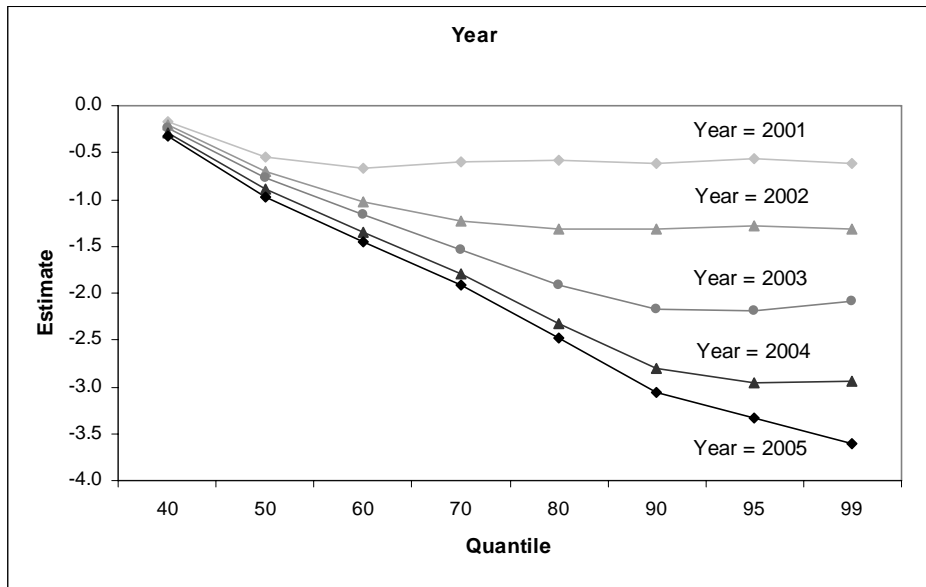
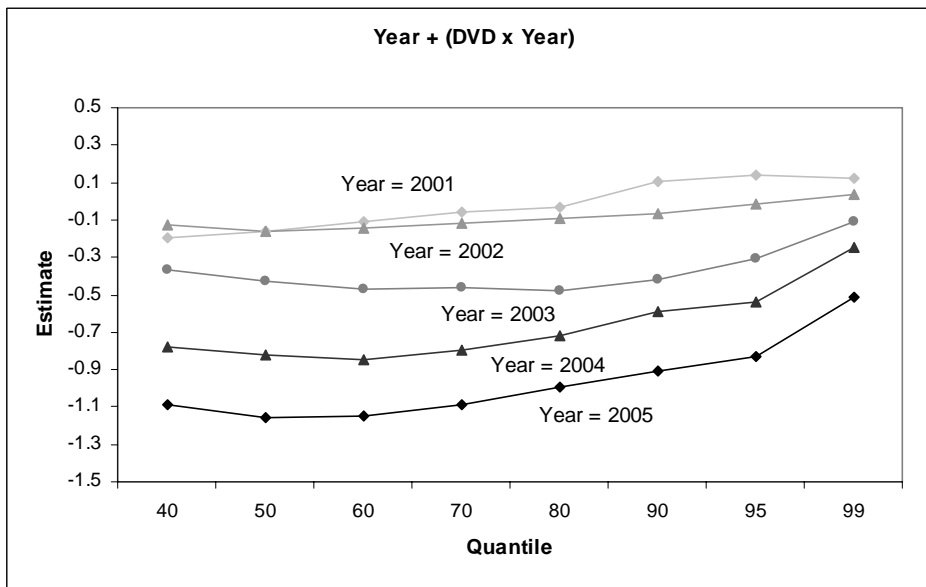


Figure 4: The Time Trend for DVD and VHS, By Quantile

VHS: The Estimate for "Year" By Quantile



DVD: The Estimate of "Year + (DVD x Year) By Quantile



References

- Alba, Joseph, John Lynch, Barton Weitz, Chris Janiszewski, Richard Lutz, Alan Sawyer, and Stacey Wood (1997). Interactive Home Shopping: Incentives for Consumers, Retailers, and Manufacturers to Participate in Electronic Marketplaces. *Journal of Marketing* 61 (July), 38-53.
- Anderson, Chris (2004). The Long Tail. *Wired Magazine*, 12.10 (October).
[www.wired.com/wired/archive/12.10/tail.html]
- Anderson, Chris (2006). *The Long Tail: Why The Future of Business Is Selling Less of More*. New York, NY: Hyperion.
- Austin, Peter C., Jack V. Tu, Paul A. Daly, and David A. Alter (2005). Tutorial in Biostatistics: The Use of Quantile Regression in Health Care Research: A Case Study Examining Gender Differences in the Timeliness of Thrombolytic Therapy. *Statistics in Medicine* 24, 791-816.
- Bakos, J. Yannis (1997). Reducing Buyer Search Costs: Implications for Electronic Marketplaces. *Management Science* 43(12), 1676-1708.
- Bakos, J. Yannis (1998). The Emerging Role of Electronic Marketplaces on the Internet. *Communications of the AMC* 41 (8), 35-42.
- Bakos, J. Yannis (2001). The Emerging Landscape for Retail E-Commerce. *Journal of Economic Perspectives* 15(1), 69-80.
- Bakos, Y. and Brynjolfsson, E. (2000). Bundling and Competition on the Internet. *Marketing Science* 19 (1), 63-82.
- Bassett, Gilbert and Hsiu-Lang Chen. 2001. Quantile Style: Return-Based Attribution Using Regression Quantiles. *Empirical Economics*. March, 26:1, pp. 293-305.
- Boatwright, Peter and Joseph C. Nunes (2001). Reducing Assortment: An Attribute-Based Approach. *Journal of Marketing* 65(July), 50-63.
- Broniarczyk, Hoyer and McAlister (1998). Consumers' Perceptions of the Assortment Offered in a Grocery Category: The Impact of Item Reduction. *Journal of Marketing Research* 35(2), 166-176.
- Brynjolfsson, Erik, Yu (Jeffrey) Hu, and Michael D. Smith (2003). Consumer Surplus in the Digital Economy: Estimating the Value of Increased Product Variety at Online Booksellers. *Management Science*, 49 (11, November), 1580-1596.
- Brynjolfsson, Erik, Yu (Jeffrey) Hu, and Michael D. Smith (2006). From Niches to Riches: Anatomy of the Long Tail. *MIT Sloan Management Review* 47(4), 67-71.
- Buchinsky, Moshe. 1994. Changes in U.S. Wage Structure 1963-1987: An Application of Quantile Regression. *Econometrica*. March, 62:2, pp. 405-58.
- Buchinsky, Moshe. 1997. The Dynamics of Changes in the Female Wage Distribution in the USA: A Quantile Regression Approach. *Journal of Applied Econometrics*. January/February, 13:1, pp. 1-30.
- Chamberlain, Gary. 1994. Quantile Regression, Censoring and the Structure of Wages, in *Advances in Econometrics*. Christopher Sims, ed. New York: Elsevier, pp. 171-209.

- Chernev, Alexander (2003). When More is Less and Less is More: The Role of Ideal Point Availability and Assortment in Consumer Choice. *Journal of Consumer Research* 30(September), 170–183.
- Chernozhukov, V. and L. Umantsev. 2001. Conditional Value-at-Risk: Aspects of Modeling and Estimation. *Empirical Economics*. March, 26:1, pp. 271-92.
- Coleman, J.S. (1964). *Introduction to Mathematical Sociology*. Glencoe: The Free Press.
- Dhar, Ravi (1997). Consumer Preference for a No-Choice Option. *Journal of Consumer Research* 24(September), 215–231.
- Deaton, Angus. 1997. *The Analysis of Household Surveys*. Baltimore: Johns Hopkins.
- Eide, Eric and Mark Showalter. 1998. The Effect of School Quality on Student Performance: A Quantile Regression Approach. *Economics Letters*. March, 58:3, pp. 345-50.
- Ehrenberg, A. S. C. and Goodhardt, G. J., 1990. Double Jeopardy Revisited. *Journal of Marketing* 54 (3), 82.
- Entertainment Merchant Association (2006). *Annual Report on the Home Entertainment Industry*. [www.vgsda.org].
- Euromonitor (2006). *Global Market Information Database: IT and Telecommunications*. Euromonitor International.
- Frank, Robert H., and Cook, Philip J. (1995). *The Winner-Take-All Society*. New York: The Free Press.
- George, Lisa M. and Joel Waldfogel (2006). The New York Times and the Market for Local Newspapers. *American Economic Review* 96 (1), 435-47.
- Gourville, John T. and Dilip Soman (2005). Overchoice and Assortment Type: When and Why Variety Backfires. *Marketing Science*, 24 (3), 382–395.
- Häubl, Gerald and Valerie Trifts (2000). Consumer Decision Making In Online Shopping Environments: The Effects of Interactive Decision Aids. *Marketing Science* 19(1), 4-21.
- Heckman, James (1979). Sample Selection Bias as a Specification Error. *Econometrica* 47 (1), 153-161.
- Hendricks, Wallace and Roger Koenker. 1992. Hierarchical Spline Models for Conditional Quantiles and the Demand for Electricity. *Journal of the American Statistical Association*. March, 87:417, pp. 58-68.
- Hoch, Stephen J., Eric T. Bradlow and Brian Wansink (1999). The Variety of an Assortment. *Marketing Science* 18(4), 527–546.
- Huffman, Cynthia and Barbara E. Kahn (1998). Variety For Sale: Mass Customization or Mass Confusion. *Journal of Retailing* 74(4), 491–513.
- Iyengar, Sheena S. and Mark R. Lepper (2000). When Choice is Demotivating: Can One Desire Too Much of a Good Thing? *Journal of Personality and Social Psychology* 79(6), 995–1006.
- Kahn, Barbara E. (1995). Consumer Variety-Seeking Among Goods and Services. *Journal of Retailing and Consumer Services* 2(3), 139–148.
- Kahn, Barbara E. (1998). Dynamic Relationships with Customers: High Variety Strategies. *Journal of the Academy of Marketing Science*, 26(1) 45–53.
- Kahn, B., D. R. Lehmann (1991). Modeling choice among assortments. *Journal of Retailing* 67(3) 274–299.

- Koenker Roger and Gilbert Bassett (1978). Regression Quantiles. *Econometrica* 46(1), 33-50.
- Koenker, Roger and Kevin F. Hallock (2000). Quantile Regression. *Journal of Economic Perspectives* 15 (4), 143-156.
- Koenker, Roger and Yannis Biliias. 2001. Quantile Regression for Duration Data: A Reappraisal of the Pennsylvania Reemployment Bonus Experiments. *Empirical Economics*. March, 26:1, pp. 199-220.
- Lehmann, Donald R. (1998). Customer Reactions to Variety: Too Much of a Good Thing? *Journal of the Academy of Marketing Science* 26(1), 62–65.
- Levin, Jesse. 2001. For Whom the Reductions Count: A Quantile Regression Analysis of Class Size on Scholastic Achievement. *Empirical Economics*. March, 26:1, pp. 221-46.
- Lynch, John G. Jr. and Dan Ariely (2000). Wine Online: Search Costs Affect Competition of Price, Quality, and Distribution. *Marketing Science* 19 (1), 83-103.
- Machado, José A.F. and José Mata (2000). Box-Cox Quantile Regression and the Distribution of Firm Sizes. *Journal of Applied Econometrics* 15, 253-274.
- Machado, José A.F. and José Mata (2005). Counterfactual Decomposition of Changes in Wage Distributions Using Quantile Regression. *Journal of Applied Econometrics* 20, 445–465.
- Manning, Willard, Linda Blumberg and Lawrence Moulton. 1995. The Demand for Alcohol: The Differential Response to Price. *Journal of Health Economics*. June, 14:2, pp. 123-48.
- McPhee, W. N., 1963. *Formal Theories of Mass Behaviour*, The Free Press of Glencoe, New York.
- Napoli, Philip M. (1997). Rethinking Program Diversity Assessment: An Audience-Centered Approach. *Journal of Media Economics*, 10(4), 59-74.
- Napoli, Philip M. (1999). Deconstructing the Diversity Principle. *Journal of Communication*, 49(4), 7-34.
- Napoli, Philip M. (2003). *Audience Economics: Media Institutions and the Audience Marketplace*. New York: Columbia University Press.
- Neuman, W. Russell (1991). *The Future of the Mass Audience*. New York: Cambridge.
- Oestreicher-Singer, Gal and Arun Sundararajan (2006). Network Structure and the Long Tail of Electronic Commerce. Working Paper, New York University, August 2006.
- Powell, J. (1991). Estimation of Monotonic Regression Models Under Quantile Restrictions. In *Nonparametric and Semiparametric Methods in Econometric and Statistics: Proceedings of the Fifth International Symposium on Economic Theory and Econometrics (357-384)*. W. Barnett, J. Powell and G. Tauchen (Eds.). New York: Cambridge University Press
- Rosen, S. (1981). The Economics of Superstars. *The American Economic Review*, 71(5), 845-858.
- Standard & Poor's (2006). *Industry Surveys: Movies & Home Entertainment*. March 30, 2006.
- Taylor, James W. (1999). A Quantile Regression Approach to Estimating the Distribution of Multiperiod Returns. *Journal of Derivatives*. 7:1, pp. 64-78.
- Taylor, James W. and Derek Bunn (1999). A Quantile Regression Approach to Generating Prediction Intervals. *Management Science* 45(2), 225-237.
- The Economist (2005). *Profiting From Obscurity*. May 5, 2005.
- The New York Times (2006). *On the Contrary: The Long and Short of Sales*. August 20, 2006
- The Wall Street Journal (2006a). *It May Be a Long Time Before the Long Tail Is Wagging the Web*. July 26, 2006, Page B1.

- The Wall Street Journal (2006b). Many Companies Still Cling To Big Hits To Drive Earnings. August 2, 2006, Page B1.
- Video Business (2006). Market Data. [<http://www.videobusiness.com>, accessed August 15, 2006].
- VSDA (2005). A History of Home Video. Accessed December 31, 2005.
(<http://www.idealink.org/Resource.phx/vsda/pressroom/history-of-industry.htx>)
- Waldfogel, Joel (2003). Preference Externalities: An Empirical Study of Who Benefits Whom in Differentiated-Product Markets. *RAND Journal of Economics*, 34 (3), 557-68.
- Webster, J. (2005). Beneath the Veneer of Fragmentation: Television Audience Polarization in a Multichannel World. *Journal of Communication*, 366-382.