Functional Forecasting of Demand Decay Rates using Online Virtual Stock Markets

Wolfgang Jank†, Gareth James‡, Natasha Foutz§

†Robert H Smith School of Business, University of Maryland
‡Marshall School of Business, University of Southern California
§McIntire School of Commerce, University of Virginia

October 21, 2008

Abstract

Forecasting product demand is an important yet challenging planning tool for many industries. It is particularly challenging in industries that feature highly innovative products such as songs or movies in the entertainment industry. The challenge arises from the fact that, since products are highly innovative and thus very distinct, not many comparisons can be drawn from the success of previous products. We propose a novel approach for modeling the decay rate of a product's demand using the predictive power of online virtual stock markets (VSMs). VSMs are online communities that, in a market-like fashion, gather the crowds' opinion about a particular product. We use functional data analysis techniques to capture the pattern of pre-release VSM trading values and hence predict a product's post-release revenue decay. Since both the predictor and response are functional, we first decompose the VSM and decay data using functional shape analysis, and then apply non-linear methods, such as boosting or neural networks, to predict their relationship. We show that this approach provides superior predictive accuracy in comparison to standard methods. Finally, we develop graphical methods to aid in understanding the relationship between a movie's VSM history and its final box office revenue pattern. These plots show that different patterns in the VSM curves can result in very different box office returns.

Key words and phrases: Functional data, smoothing, shape analysis, forecasting, boosting, neural network, virtual markets, movies, Hollywood.

1 Introduction

Successful introductions of new products are crucial for the survival and profitability of major entertainment industries, e.g. motion picture, music, TV, gaming, and publishing. However, such introductions are confronted with enormous investments, short product lifecycles (and hence short revenue windows), and highly uncertain demand. For example, Hollywood studios spend on average \$106.6m to produce and market a film, yet each is screened in theaters for only ten weeks or so, and accrues \$45.7m on average (according to the *U.S Entertainment Industry Market Statistics 2007*).

To increase returns on investment, executives select strategic actions, such as ad spending and release date, in the *months* or even *years prior to* a product's release. As a result, they are keenly interested in accurately forecasting a product's demand pattern over its short lifecycle to guide decision making (Sawhney and Eliashberg, 1996; Bass et al, 2001). For example, in the months leading up to a film's release, executives allocate weekly advertising budgets according to the predicted *rate of demand decay*, i.e. according to whether the film is expected to open big and then decay fast, or whether it opens only moderately but decays very slowly. They rely on similar forecasts to select the best possible release date, one that attracts strongest demand (e.g. a holiday weekend) but weakest competition during the opening weekend and in subsequent weeks (Einav, 2007). Additionally, theater owners use such forecasts to allocate screens and negotiate weekly revenue-sharing with studios.

While early and accurate demand forecasting is crucial, it has long been regarded as one of the most challenging tasks facing marketing academics and practitioners (Bass et al, 2001). Forecasting is particularly challenging for entertainment products due to their unique features. Entertainment products are highly experiential in nature, thus conventional methods that characterize products using tangible attributes and forecast a new product's demand using its attributes relative to the attributes of comparable products do not apply (Bass et al, 2001). Also, demand is driven heavily by consumers' word-of-mouth, which is inherently difficult to measure, and thus rarely accounted for in forecasting models (Mahajan and Wind, 1988; Godes et al, 2005; Godes and Mayzlin, 2004). Moreover, demand is highly heterogeneous across different entertainment products. Take for instance the sample of movie demand patterns in Figure 1. Here we have plotted the log weekly box office revenues for the first ten weeks from the release date for a number of different movies. While revenues for some movies (e.g. *13 GOING ON 30* and *50 FIRST DATES*) decay exponentially over time, revenues for others (e.g. *BEING JULIA*) increase first before decreasing. Even for movies with similar demand patterns (e.g. those on the second row of Figure 1), the speed of decay varies greatly.

Our goal is to develop a flexible and powerful forecasting mechanism that can overcome these challenges and that can provide early and accurate forecasts of demand over an entertainment product's lifecycle. To accomplish this, we propose a powerful forecasting model based on functional data analysis, applied to a novel data source, online virtual stock markets (VSMs). VSMs capitalize on the *wisdom of crowds* (Surowiecki, 2004) and capture consumers' word-of-mouth. In

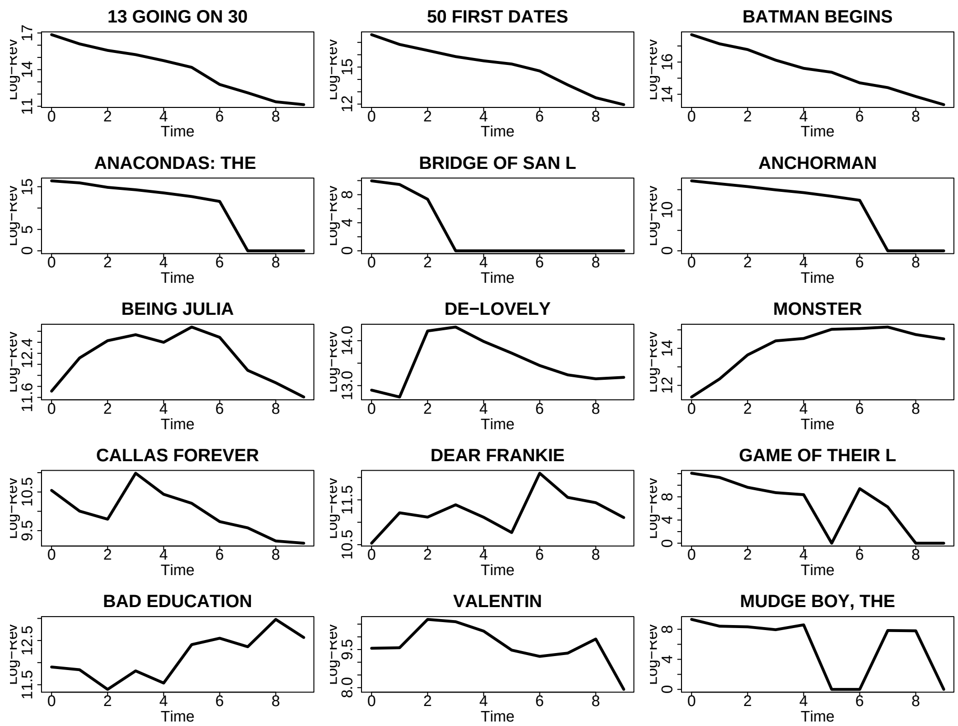

Figure 1: Movie demand decay rates for a sample of movies.

a VSM, participants trade virtual stocks according to their predictions of the outcome of the event represented by the stock (e.g. the demand for an upcoming movie). As a result, VSM trading price histories often give early and reliable demand forecasts. VSMs provide monetary or psychological incentives (e.g. self-pride, entertainment) to participants for discovering information about new products (often widely dispersed across media, e.g. internet, TV, trade magazines) and revealing this information through the embedded trading mechanism (Dahan et al, 2007). VSMs are particularly well suited for forecasting demand of entertainment products as they do not require decomposing a product into tangible attributes. VSMs are especially intriguing from a statistical point of view since the *shape* of the trading prices may reveal additional information such as the speed of information-diffusion which, in turn, can proxy for consumer sentiment and word-of-mouth about a new product (Foutz and Jank, 2008). For instance, a last-moment price spurt may reveal a strengthening hype for a product and may thus be essential in forecasting its demand.

In this paper we use VSM data to predict decay patterns for the box office revenue over the first ten weeks of a movie's lifecycle. Since both the VSM data and the decays are functional, we can view this as a regression problem involving functional predictors (the VSMs) and responses (decay rates). This poses some extra complications over the more standard functional regression frameworks where one more commonly assumes that the predictor or response, but not both, are functional (e.g. Foutz and Jank, 2008). We propose a functional regression model that has three significant advantages. First, by extracting the key features of both the VSM and the decay functions, our model is able to accurately handle functional predictors and responses. Second, most functional regression models assume a linear relationship between the predictor and response. As Figure 1 suggests, demand decay rates are highly heterogeneous and thus hard to model using a linear approach. We instead show how modern non-linear methods such as generalized additive models (Hastie and Tibshirani, 1990), boosting (Friedman et al, 2000), or neural networks, can be incorporated into our functional procedure. These highly non-linear approaches allow us to model much more subtle relationships and we show that, on our data, they produce clear improvements in prediction accuracy. Finally, we address the difficulty of interpreting the results from a model involving functional predictors and responses. Dependence plots are developed which graphically illustrate, for typical VSM shapes, the corresponding predicted decay pattern. These dependence plots provide managerially relevant information as to how particular characteristics of the VSM data will affect the revenue pattern of a given movie.

This paper is organized as follows. In the next section, we provide further background on virtual stock markets in general and our data set in particular. Section 3 describes the process of functional shape analysis and the insights we gain from applying shape analysis to the VSM trading prices as well as to the movie decay patterns. Section 4 discusses how the VSM and decay shapes derived in Section 3 can be used to build our non-linear functional forecasting model. We compare the accuracy of our method to some more standard approaches and also illustrate the insights that can be gained from it using dependence plots. We conclude with further remarks in Section 5.

2 Data

We have two different sources of data. Our input data (i.e. functional predictors) come from the trading histories of an online virtual stock market for movies; our output data (i.e. functional responses) pertain to the weekly demand of those movies. We describe each data source in detail below.

2.1 Online Virtual Stock Markets

Online virtual stock markets (VSMs), also known as prediction markets, idea futures, or betting exchanges, function in ways very similar to real life stock markets except that they are not necessarily based on real currency (i.e. participants often use virtual currency to make trades), and that each stock corresponds to an event or a parameter (rather than a company's shares). Typical events often cover topics of interest to the public, such as economic trends (HedgeStreet), political elections (Iowa Electronic Markets or IEM), sporting events (TradeSports), or the Oscars (Hollywood Stock Exchange or HSX).

If the event has discrete outcomes (e.g. either a democratic or republican candidate wins the 2008 Presidential election), then trading terminates when the event occurs and the final liquidation price of the stock is determined by the actual outcome of the event. Therefore, the current price of a stock can be interpreted as the traders' collective belief of the probability with which the event will occur. For example, if the current price equals 54 cents per share, then the stock can be interpreted as the traders' collective believe that the democratic candidate has a 54% chance of winning. If in fact the democratic candidate wins, then traders holding the democratic candidate's stock will liquidate (or cash in) at \$1 per share; otherwise they receive \$0. Traders can buy or sell stock at any time. If a trader believes the democratic candidate has higher (lower) than 54% chance to win, (s)he will buy (sell) the stock. The more confident (s)he is about this belief, the more shares (s)he will buy (or sell).

Not all events are discrete. Demand for a new product, such as a movie (HSX; IEM), book (MediaPredict), music (HSX), TV show (Inkling; HSX), video game (SimExchange; Yahoo!), or MP3 (Inkling) are all examples of continuous events. In that case, trading commences prior to the new product's introduction and terminates shortly before or after its introduction. The stock's current price then reflects the traders' collective belief about the products future demand; it is liquidated at the actual demand. For example, the current price of a movie stock on the Hollywood Stock Exchange (HSX), e.g. \$50 per share, suggests that traders collectively anticipate \$50 million box office revenues from the movie's first four weeks in theaters. As new information about a movie emerges over time, traders' expectations change. At any time, those who anticipate higher first-four-week revenues, e.g. \$60 million, will buy the stock and those who believe otherwise will sell it.

The source of our data is the Hollywood Stock Exchange (HSX), one of the best known online

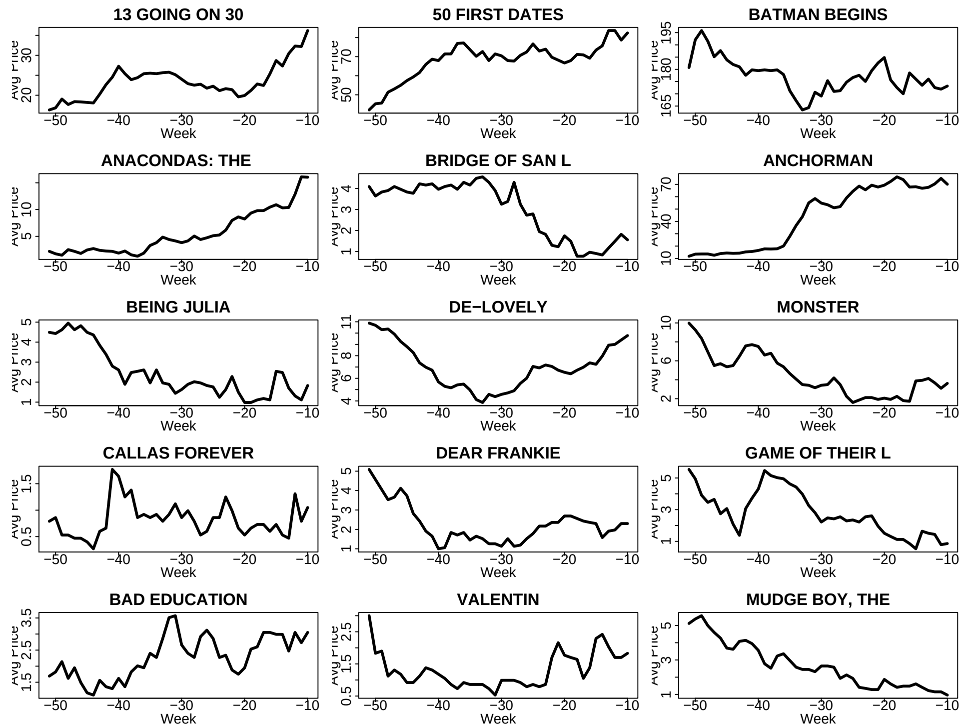

Figure 2: HSX trading histories for the sample of movies from Figure 1.

VSMs. HSX was established in 1996 and has nearly 2 million active participants worldwide. Each trader is initially endowed with $2 million virtual currency and can increase his (her) net worth by strategically selecting and trading movie stocks (i.e. buying-low and selling-high). Traders are further motivated by opportunities to exchange the accrued currency for merchandise and to appear on the daily *Leader Board* that features the most successful traders. Figure 2 shows the sample of HSX trading histories corresponding to the movie demand patterns from Figure 1. Note that since our goal is to accomplish early forecasts, we only consider information prior to ten weeks before a movie's release (i.e. up to time -10 in Figure 2). Ten weeks gives managers ample time to make informed decisions about marketing-mix allocations.

In this paper, we study the HSX trading histories for 262 movies. Table 1 displays basic characteristics for these 262 movies. We can see that, e.g., the average movie-budget equals $45.84 million; the largest fraction (35%) of movies in our sample are comedies, and most (63%) movies are rated P, PG, or PG-13; and movies are distributed rather equally among the 8 major movie studios. Figure 3 shows the distribution of HSX trading prices. The left panel shows the distribution of weekly prices (aggregated across all movies and across all trading periods); we can see that prices are highly right-skewed, with few movies/weeks seeing prices higher than $150. However, our focus is *not* on weekly prices; rather, we are interested in learning from the *trading histories* of individual movies. The right panel shows these histories across all 262 movies. Consistent with the right-skewness in the left panel, we can see that very few movies have trading histories that reach beyond \$100 $\sim$ \$150.

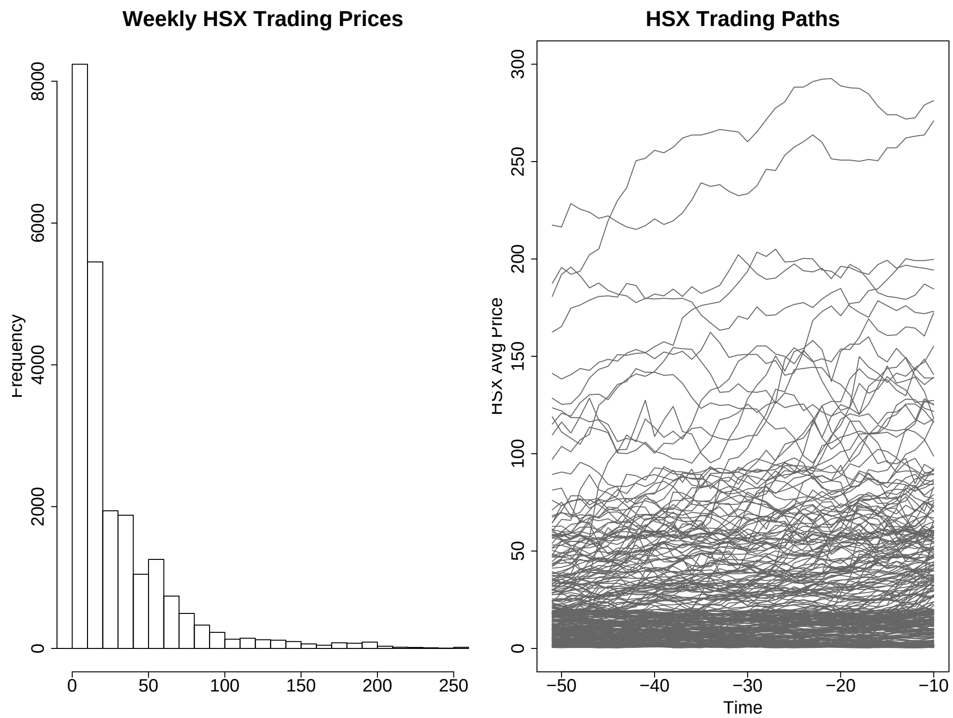

Figure 3: Distribution of HSX weekly trading values and trading paths.

We focus on the entire trading history since there is evidence that the shape of the trading path carries valuable information. For instance, past research in behavioral finance and marketing has documented numerous behavioral anomalies in stock trading, e.g. over- or under-reacting, herding, and information cascading from insiders to less knowledgeable traders (De Bondt and Thaler, 1985; Lo and Mackinlay, 2002; Johnson, Tellis). Many VSMs also limit the maximum number of shares that a participant can trade, resulting in lower liquidity and reduced efficiency in these markets (Lamont and Thaler, 2003). Thus the shape of price histories may indeed contain valuable information, above and beyond the most recent price.

Our functional forecasting model captures differences in shapes of VSM price histories, e.g. trending up or down, concavity vs. convexity, or last-moment spurts. We find that these shapes are predictive of the demand pattern of a product's lifecycle. For example, a rapid increase in early VSM trading prices may suggest a rapid diffusion of awareness among potential adopters and

Sequel: count (%)	30 (11.45%)
Avg. budget ($m)	45.84
Genre: count (%)	
ACTION	22 (8.40%)
ANIMATED	12 (4.58%)
COMEDY	92 (35.11%)
DRAMA	73 (27.86%)
HORROR	13 (4.96%)
SCI-FI	14 (5.34%)
SUSPENSE	22 (8.40%)
OTHERS	14 (5.34%)
Rating: count (%)	
Not Rated	2 (0.76%)
G, PG, PG-13	167 (63.74%)
R, NC-17	93 (35.50%)
Avg. run time (min.)	106.94
Studio: count (%)	
BUENA-VISTA	25 (9.54%)
FOX	25 (9.54%)
MGM-UA	11 (4.20%)
MIRAMAX	17 (6.49%)
PARAMOUNT	21 (8.02%)
SONY	33 (12.60%)
UNIVERSAL	20 (7.63%)
WARNER	25 (9.54%)
OTHERS	85 (32.44%)

Table 1: *Summary of Movie Data.*

strong interest in a product. Thus it can suggest a strong *initial* demand immediately after a new product's introduction to the market place, e.g. a strong opening weekend box office for a movie. Similarly, a new product whose trading prices increase very sharply towards the time of release may be experiencing strong last-moment positive word-of-mouth, which may lead to a reduced decay rate in demand, or increased longevity of a new product.

2.2 Weekly Movie Demand Patterns

Our goal is to predict a movie's demand. Specifically, we want to predict a movie's demand not only for a given week (e.g. at week 1 or week 5), but over its entire lifecycle. A movie's lifecycle lasts typically several weeks, between the opening week-end (week 0) and the time it is removed from theaters. In this work, we focus on the first 10 weeks (between week 0 and week 9) of a movie's lifecycle. Figure 4 shows weekly demand for all 262 movies in our data. The left panel shows the distribution of weekly demand aggregated across all movies and weeks; we can see that (log-) demand is rather symmetric and appears to be bi-modal. It is also interesting to note the large number of zeros in the data; these correspond to movies/weeks with zero demand. (Note that we applied a $\log(\text{Revenue} + 1)$ transformation to our data.) The right panel shows, for each individual movie, the rate at which demand decays over the 10-week period. We can see that while some movies decay gradually, a number have sudden drops, while other movies initially increase after the release week. Our goal is to characterize different demand decay *shapes* and to use the information from VSMs to forecast these shapes.

In the next section, we describe a method to extract shapes from VSM trading price histories and from movie demand patterns. The method is based on the ideas of functional data analysis and we refer to it as *functional shape analysis.*

3 Shape Analysis of Movie Demand and Prediction Markets

In this section, we lay the foundations of our functional forecasting model. Our forecasting model predicts the *shape* of a movie's decay rate, using the *shape* of the market's trading history as input. Thus, as a first step, we explore and characterize the shapes of movies and prediction markets. We do so using *functional shape analysis* (FSA). FSA is done in two steps. First, we recover, from the observed data, the underlying functional object via smoothing techniques. Then, we extract shapes with the help of *functional principal component analysis.* We describe this approach in

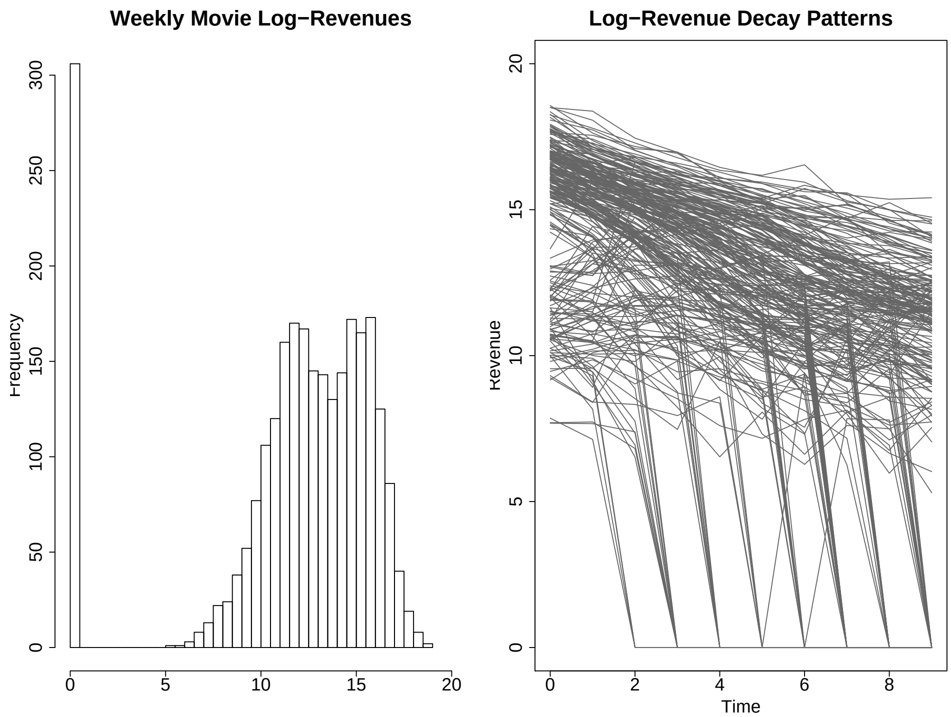

Figure 4: Distribution of movies' weekly demand and demand decay patterns.

detail below, followed by the results of FSA applied to movies' decay rates and markets' trading paths, respectively.

3.1 Shape Analysis via Functional PCA

Functional data analysis operates on a set of *continuous* objects, such as a set of curves describing daily temperature changes (Ramsay and Silverman, 2005), prices in an online auction (Jank and Shmueli, 2006), or market penetration of new products (Sood et al, 2008). Despite their continuous nature, limitations in human perceptions and measurement capabilities allow us to record only discrete observations of these curves. Thus, the first step is to recover, from the observed data, the underlying continuous functional objects by using smoothing techniques. There are a variety of different data smoothers. We use a very flexible and computationally efficient technique, the penalized smoothing spline (Ruppert et al., 2003). The details for implementing smoothing splines are provided in the appendix.

Figure 5 shows the smoothed decay rates from the observed decay patterns in Figure 1. We can see that smoothing captures the main decay patterns while soothing unusual, noisy demand spikes. We apply a similar procedure to the observed HSX trading histories from Figure 2.

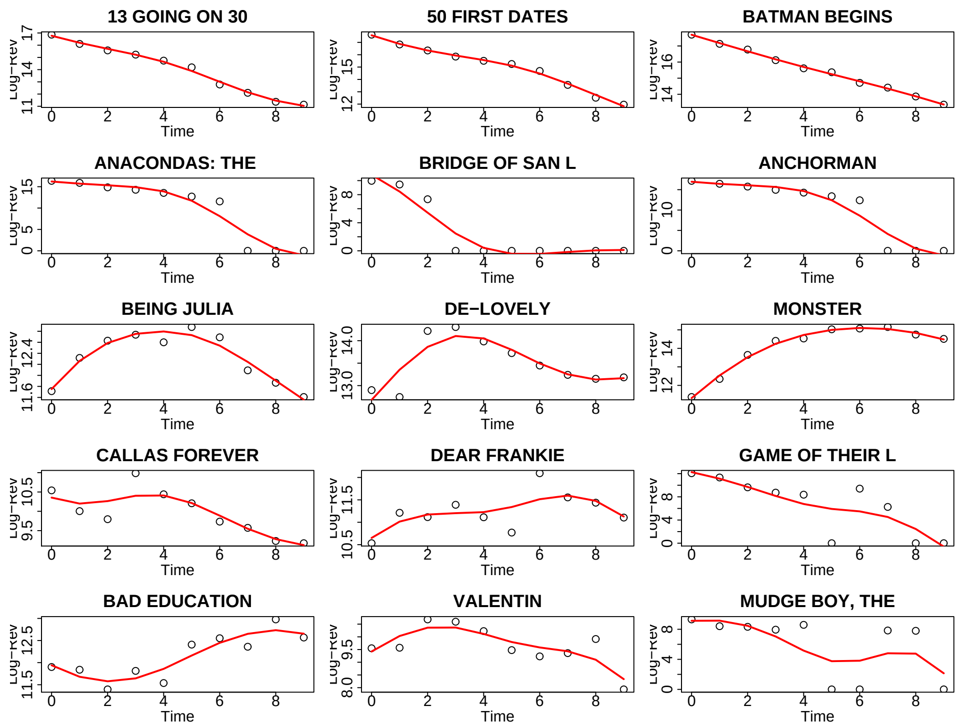

Figure 5: Smooth movie demand decay rates for a sample of movies.

After deriving a smooth decay rate for each individual movie, the next step is to extract typical decay shapes that are common across many movies. Take for instance the 18 movies displayed in Figure 5. We can see that some movies (first panel) decay at a (log-) linear rate; some (second panel) show sudden drops, and some (third panel) show polynomial demand pattern; yet other movies (fourth and fifth panel) display patterns that may be a combination of the first three shapes. In the following, we want to extract the most common decay shapes across all movies. We accomplish this via functional principal component analysis (fPCA).

fPCA is a generalization of ordinary principal component analysis (PCA). Ordinary PCA operates on a set of data-vectors, say, $\mathbf{x}_1, \ldots, \mathbf{x}_n$, where each observation is a p-dimensional data-vector $\mathbf{x}_i = (x_{i1}, \ldots, x_{ip})^T$. The goal of ordinary PCA is to find a projection of $\mathbf{x}_1, \ldots, \mathbf{x}_n$ into a new space which maximizes the variance along each component of the new space and at the same time renders the individual components of the new space orthogonal to one another. In other words, the goal of ordinary PCA is to find a PC vector $\mathbf{e}_1 = (e_{11}, \ldots, e_{1p})^T$ for which the PCS

$$S_{i1} = \sum_j e_{1j} x_{ij} = \mathbf{e}_1^T \mathbf{x}_i \tag{1}$$

maximize $\sum_i S_{i1}^2$ subject to

$$\sum_j e_{1j}^2 = \|\mathbf{e}_1\|^2 = 1. \tag{2}$$

This yields the first PC, $\mathbf{e}_1$. In the next step, we compute the second PC, $\mathbf{e}_2 = (e_{21}, \ldots, e_{2p})^T$, for which, similarly to above, the principal component scores $S_{i2} = \mathbf{e}_2^T \mathbf{x}_i$ maximize $\sum_i S_{i2}^2$ subject to $\|\mathbf{e}_2\|^2 = 1$ and the *additional constraint*

$$\sum_j e_{2j} e_{1j} = \mathbf{e}_2^T \mathbf{e}_1 = 0. \tag{3}$$

This second constraint ensures that that the resulting principal components are orthogonal. This process is repeated for the remaining PC, $\mathbf{e}_3, \ldots, \mathbf{e}_p$.

The functional version of PCA is similar in nature, except that we now operate on a set of continuous curves rather than discrete vectors. As a consequence, summation is replaced by integration. More specifically, assume that we have a set of curves $\mathbf{x}_1(s), \ldots, \mathbf{x}_n(s)$, each measured on a continuous scale indexed by s. The goal is now to find a corresponding set of PC curves, $\mathbf{e}_i(s)$, that, as previously, maximize the variance along each component and are orthogonal to one another. In other words, we first find the PC function, $\mathbf{e}_1(s)$, whose PCS

$$S_{i1} = \int e_1(s) x_i(s) ds \tag{4}$$

maximize $\sum_i S_{i1}^2$ subject to

$$\int e_1^2 ds = \|\mathbf{e}_1\|^2 = 1. \tag{5}$$

Similarly to the discrete case, the next step involves finding $\mathbf{e}_2$ for which the PCS, $S_{i2} = \int e_2(s) x_i(s) ds$ maximize $\sum_i S_{i2}^2$ subject to $\|\mathbf{e}_2\|^2 = 1$ and the additional constraint

$$\int e_2(s) e_1(s) ds = 0. \tag{6}$$

In practice, the integrals in (4) - (6) are approximated either by sampling the predictors, $\mathbf{x}_i(s)$, on a fine grid or, alternatively, by finding a lower dimensional expression for the PC functions $\mathbf{e}_i(s)$ with the help of a basis expansion. For instance, let $\boldsymbol{\phi}(s) = (\phi_1(s), \ldots, \phi_K(s))$ be a suitable basis expansion (Ramsay, 1998), then we can write

$$\mathbf{e}_i(s) = \sum_{k=1}^{K} b_{ik} \phi_k(s) = \boldsymbol{\phi}(s)^T \mathbf{b}_i, \tag{7}$$

for a set of basis coefficients $\mathbf{b} = (b_{i1}, \ldots, b_{iK})$. In that fashion, the integral in e.g. (6) becomes

$$\int e_2(s) e_1(s) ds = \mathbf{b}_1^T \mathbf{W} \mathbf{b}_2, \tag{8}$$

where $\mathbf{W} = \int \boldsymbol{\phi}(s)\boldsymbol{\phi}(s)^T ds$. For more details, see Ramsay and Silverman (2005). In this work, we use the grid-approach.

3.2 Demand Decay Shapes

We apply fPCA to the smooth demand decay rates displayed in Figure 5. The scree plot in Figure 6, which plots the percentage of variance explained by each functional principal component (fPC), shows that the first three fPC's explain more than 98% of the total variation in the data. We thus retain the first three fPC's for further analysis.

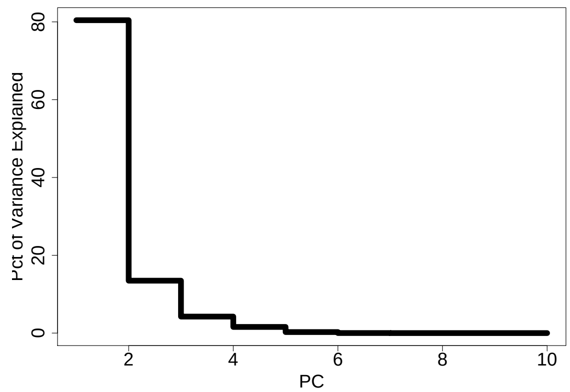

Figure 6: Scree Plot for Revenue.

Figure 7 displays the first three fPC's. We can see that each fPC captures different aspects of a movie's demand decay rate. For instance, the first fPC (solid black line) is almost constant across the first 10 release weeks and puts increasingly more weight on later weeks. In that sense, the first fPC measures a movie's weighted revenue-average over the first 10 release weeks, putting more weight on movies that perform well in later weeks.

Take Figure 8 for illustration. The top panel in Figure 8 shows the revenue decay of a movie; the bottom panel shows the corresponding functional principal component *score* (fPCS), that is, the inner product between the movie's revenue pattern and the corresponding fPC. The black bar in the bottom left panel thus corresponds to the first fPCS for *Batman Begins*. We can see that the first fPCS for *Batman Begins* and for *Monster* are much larger compared to *Anchorman*. Looking at the panel above, we can find an explanation: the average revenue for the two former movies are much larger compared to the latter. In fact, the average (log-) revenue across all movies and all 10 weeks equals 11.59; *Anchorman*'s average is clearly below that. Moreover, while *Monster* starts out slower than *Batman Begins*, it has larger revenues in the later weeks, and thus the corresponding

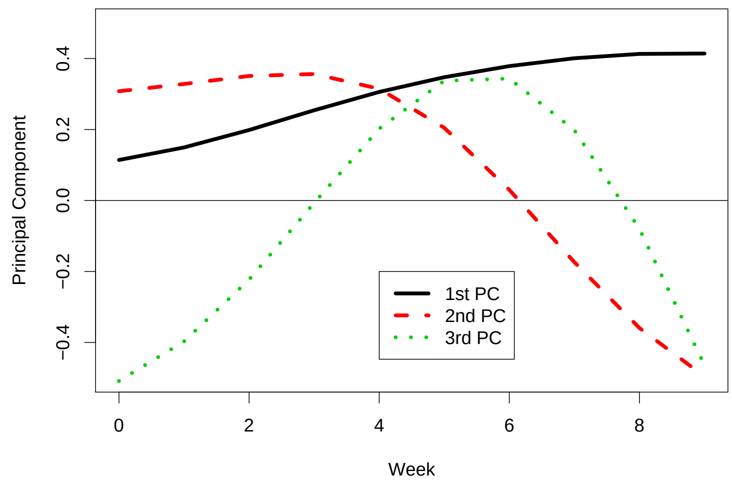

Figure 7: First three functional principal components of revenue.

fPCS are similar.

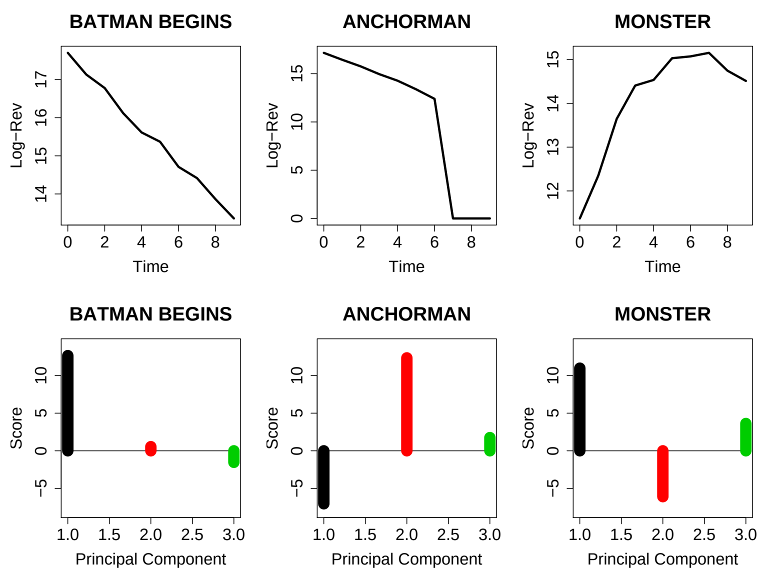

Figure 8: Illustration of revenue decay shapes.

We can interpret the remaining two PCs in Figure 7 in similar fashion. The second fPC (dashed red line) characterizes demand that is steady throughout the first 3 or 4 weeks, and then suddenly drops. An example of such a "sudden dropper" is *Anchorman* in the middle panel of Figure 8. Also note that *Batman Begins*'s second fPCS is average since its demand also drops, albeit not as rapidly as *Anchorman*. In contrast, the second fPCS for *Monster* is lowest since its demand does not drop at all, in fact it increases towards the end of the 10-week period.

The third PCs (green dotted line) in Figure 7 corresponds to movies that develop demand only late. An example of such a "sleeper" movie is *Monster*; its third fPCS (bottom right panel in Figure 8) is highest among all three movies. In contrast, the third fPCS for *Batman Begins* is lowest since its demand drops steadily.

In summary, the three functional principal components (fPC's) reflect three basic movie shapes: fPC1 denotes movies with high average demand over the entire 10-week release period, with special emphasis on movies that hold demand into later weeks ("high late demand"); fPC2 denotes movies whose demand drops suddenly ("sudden droppers"); and fPC3 denotes movies that develop demand only late ("sleepers"). We also want to emphasize that fPC1-fPC3 are not the only possible movie shapes; in fact, they form the basis of many different possible movie shapes (see also Figure 9). In that sense, we can think of fPC1-fPC3 as the basic "DNA" for a wide variety of movie shapes.

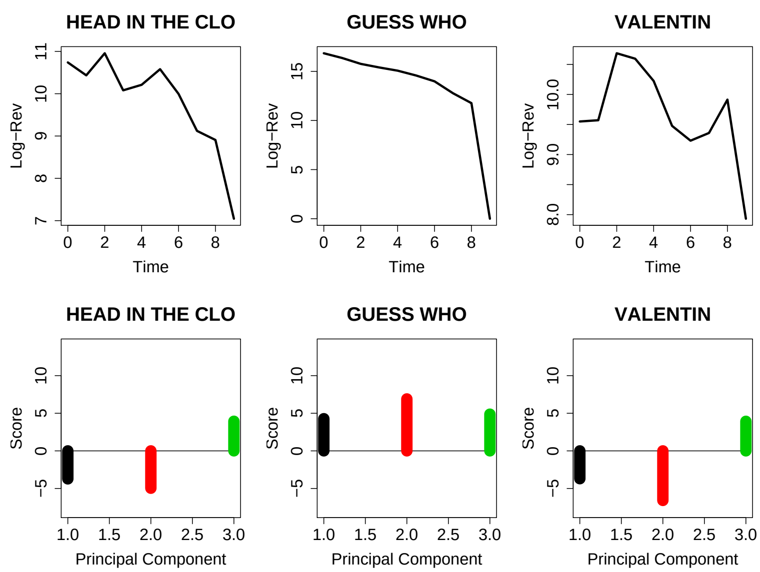

Figure 9: Illustration of other revenue decay shapes.

3.3 Prediction Market Shapes

We apply a similar procedure to the HSX trading histories (Figure 2) by first smoothing the observed trading histories to obtain smooth trading paths and then applying fPCA. The first three fPC's of the trading path explain 99% of the total variation, so we retain these fPC's for further analysis.

Figure 10 shows the first three fPC's of the trading paths. We can interpret them in a similar fashion to Figure 7. The first fPC (black solid line) captures differences in the average trading path. It places almost equal weight on each trading week and thus captures the difference across movies in their week-to-week-averages: a movie with a high average trading path will have a higher fPC score compared to a movie with low average trading path.

In similar fashion, we can interpret the second (red dashed line) and third (green dotted line) fPC. The second fPC captures the difference between early and late prices: a movie with an upward trading path will have a positive fPC score compared to a movie with a flat (or downward) trading path. And the third fPC reflects the difference between mid-term trading activity compared to trading early and late.

In the following, we model demand decay shapes using prediction market shapes as explanatory variables.

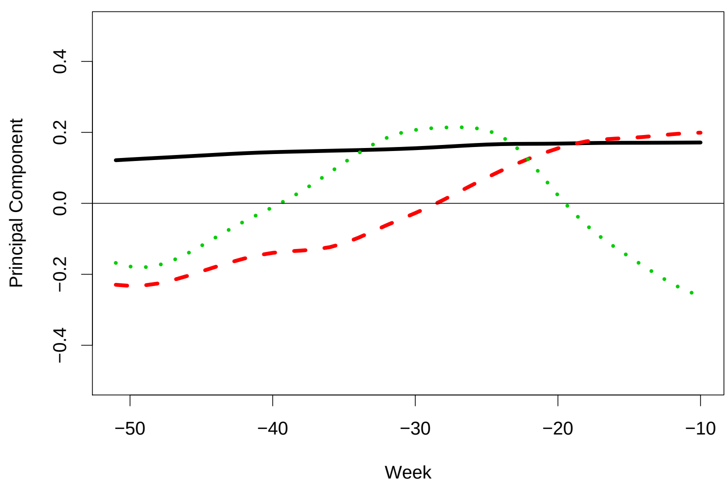

Figure 10: First three functional principal components of HSX trading path.

4 Forecasting Demand Decay Rates

After extracting and characterizing both input (prediction market) and output (movie demand) shapes, we now link these two components using modeling techniques. Our goal is two-fold: on the one hand, we want to find a model that best *predicts* a movie's demand decay rate, given information about the shape of its associated prediction market. On the other hand, given such a model, we also want to obtain new insight into the relationship between prediction markets and movies' success; that is, we want to perform *inference* based on that model. We discuss each of these two goals below.

	Mean	Last Observation	Boost 1	Boost 2	GAM
Week 1	2.061	1.321	1.263	1.026	1.110
Week 2	1.905	1.143	1.153	0.968	1.005
Week 3	1.982	1.379	1.355	1.170	1.240
Week 4	2.425	1.839	1.859	1.772	1.775
Week 5	2.850	2.316	2.336	2.250	2.211
Week 6	3.155	2.649	2.697	2.546	2.534
Week 7	3.531	3.067	3.128	2.919	2.962
Week 8	4.034	3.610	3.653	3.504	3.557
Week 9	4.435	4.007	4.083	3.822	4.009
Week 10	4.772	4.412	4.451	4.070	4.322

Table 2: *Ten-fold cross-validation mean absolute error (on log scale) for five different methods.*

4.1 Prediction

We compare a number of functional and non-functional methods to predict the 10 week box office decay pattern for our 262 movies. One of the advantages of decomposing functional data into a finite dimensional principal component basis is that a large range of standard classical and modern regression methods can be applied to the data by using the principal component scores as predictors. Some of the methods we tested included linear regression, additive models (Hastie and Tibshirani, 1990), neural networks and boosting (Friedman et al, 2000). Linear regression has been frequently applied to functional data and there has been some recent work on extending additive models to the functional domain (Sood et al, 2008; Muller and Yao, 2008). However, we are not aware of previous attempts to apply more modern methods such as neural networks or boosting to such a setting.

Table 2 provides the mean absolute error rates (on the log scale) between predicted and actual box office revenue for two of the non-functional and three of the functional methods that we tested. In all cases we used ten-fold cross-validation to compute the error rates. The first method, "Mean", represents a simple approach where the mean revenue among the training movies, for each of the ten weeks, is used to predict the weekly sales in the holdout sample. This approach does not make any use of the HSX price data, so it represents a baseline to compare other methods against. The second method, "Last Observation", uses the last observed value of the HSX prices after excluding the ten weeks of data prior to release. We form a separate linear regression for each of the ten weeks of observed revenue data with the final HSX price as the predictor. This is the approach that one might expect to provide the best results if all the HSX information was encoded in the final observed price and the shape of the trading history did not provide any useful information.

"Boost 1" was the first of our functional approaches. Here we used the final observed HSX price, as with "Last Observation", and also the first fPC score on the HSX curves, as predictors. We then used the *gbm*() function in R to form ten separate boosting models, one for each of the ten weeks of revenue data. This represented a quasi-functional approach in that we utilized the functional nature of the predictors, by using the fPC score, but did not directly model the fact that the revenue data was also functional. "Boost 2" adopted a fully functional approach. Here we used the same predictors as for Boost 1 with the addition of the second fPC score. However, instead of fitting ten separate models using weekly revenue as the response, we fit only three models using the first three fPC scores on the revenue data as the response. One can combine the predictions on these fPC scores with the estimated decay PC functions to estimate all ten weeks of revenue data. Finally, GAM adopts an identical setup to Boost 2 except that the *gam*() function is used to fit an additive model. Note that we do not report results from neural networks here because we found boosting to be superior on this data.

Not surprisingly, for all five methods, the error rates grew as we attempted to predict revenue further from the release date. Interestingly, the Week 2 error rates dipped slightly from those for Week 1 before increasing again. However, this may have just been a result of lower revenue in Week 2 and hence less variability in the response variable. All four methods that utilized the HSX price data provided considerable improvement over the Mean approach. In comparing Last Observation and Boost 1, the latter method provided a small improvement in Week 1 but otherwise the results were fairly similar. However, much more significant improvements were generated using the fully functional Boost 2 approach (recall the results in Table 2 are all on the log scale). Boost 2 had the lowest error rates among all five methods for each of the ten weeks. This strongly suggests that a non-linear approach that utilizes the functional nature of the HSX and revenue decay data is capable of taking advantage of the fact that the shape of the HSX price data contains useful information beyond that of the final trading price. Finally, GAM outperformed the first three methods but resulted in inferior results relative to Boost 2. GAM also utilized a non-linear regression approach but was restricted to fitting an additive model. The fact that Boost 2 outperformed GAM suggests that there was a non-additive structure in the data which the fully non-linear boosting procedure was able to utilize.

We caution that our attempt here is not to provide a definitive conclusion as to the single best approach in terms of prediction accuracy. Clearly different results could apply using different loss functions and other methodologies. However, the fact that the fully functional boosting approach provided superior mean absolute error rates in comparison to other reasonable approaches demonstrates the merit of the functional methodology on this type of data.

4.2 Inference

The previous section has shown that using a functional shape model can be a powerful approach for forecasting decay rates. However, while the model results in good prediction performance, one downside is that it is conceptually hard to grasp. Both, the model-input (HSX trading paths) as well as the model-output (box-office demand decay rates) arrive in the form of shapes which makes it hard to understand the relationship between response and predictors. In fact, since the model operates on the principal components underlying the input- and output-shapes, it is almost impossible to "see" the relationship between the two. To overcome this drawback, we develop new visualizations that transform the principal components back to their original domain and, as a result, allow for insightful conclusions. The idea is similar in spirit to the "partial dependence plots" described in Hastie et al (2001); however, in contrast to their approach, our graphs take into account the effect of all predictors at the same time (and are hence not "partial"); we thus term our plots "dependence plots."

Recall that our model links input shapes (HSX trading paths) to output shapes (demand decay rates). We operationalize the model by first decomposing both input and output shapes into their first most representative principal components (PCs). Then, we model the relationship between input and output PCs in a multivariate fashion. The top panel in Figure 11 illustrates the process.

In order to make inference with our functional shape model, we have to reverse the process. That is, given an estimated model, we predict the output PCs for a given set of input PCs. We obtain input PCs by assuming an idealized input curve (e.g. linear, exponential or logarithmic), representing an idealized HSX trading path. Using the first PCs of the idealized input curve, we predict the corresponding output PCs. The output PCs correspond to the demand decay rate (i.e. output curve) associated with the idealized HSX trading path (i.e. input curve). The last step consists of reconstructing this output curve from the predicted output PCs. This can be achieved by multiplying each fPC curve by its corresponding predicted PCS. The bottom panel in Figure 11 illustrates this reverse process.

We apply this process to the boosting forecasting model from Section 4.1. Figure 12 shows the

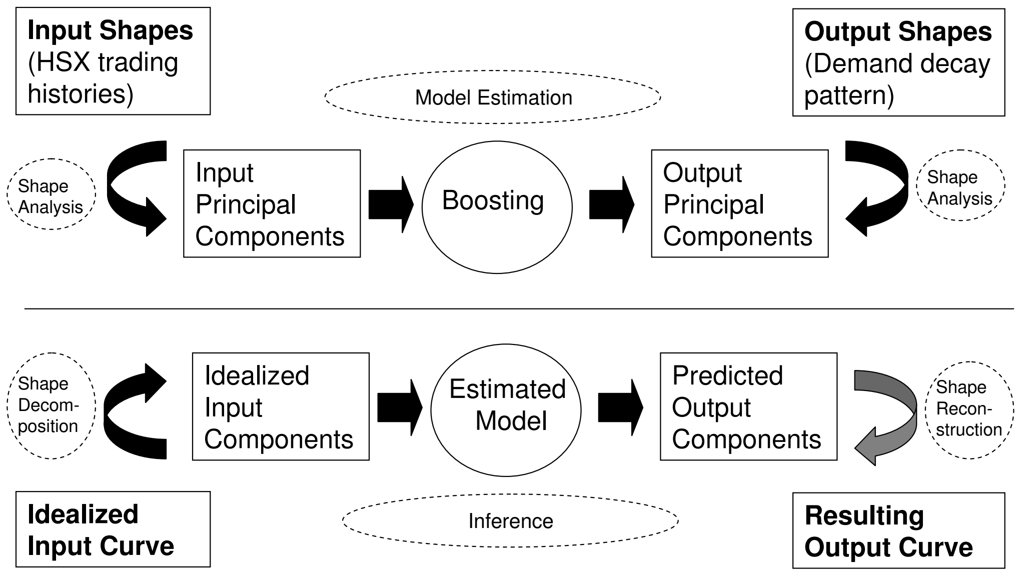

Figure 11: Illustration of reverse-engineering output shapes from an idealized input shape.

results. The left panel shows three idealized HSX trading paths. All three curves start and end at the same values (0 and 100, respectively); their only difference is how they get from the start to the end. The middle curve (dark solid line) grows at a linear rate; the upper and lower curves (light grey dotted and dark grey dashed lines) grow at logarithmic and exponential rates, respectively. In that sense, the three curves represent movies whose market prices either grow at a constant rate (linear rate), or grow fast early but then slow down (logarithmic) or grow slowly early only to increase towards release (exponential). All three curves end at exactly the same market value, so any difference in predicted box office revenue is only due to their difference in shapes.

The right panel of Figure 12 shows the resulting (predicted) demand decay rates. We can see that decay rates differ significantly, both in their initial values as well as in the way that they decay. (Note that decay rates are plotted on the log-scale!) The logarithmic market growth rate (light grey dotted line) results in the highest demand immediately after a movie's release and it remains highest until approx. week 6. On the other hand, while linear market growth (black solid line) results in a similarly high demand early on, its demand drops below that of exponential market growth (dark grey dashed line) after week 3. This is curious because while linear market growth has consistently higher HSX trading values compared to exponential market growth (except for at the very end), its expected movie demand decays much faster. This suggests that the rapid last-moment increase of the exponential shape carries information about a movie's success in later weeks. In other words, movies that experience stronger trading accelerations towards the end are expected to have a slower demand decay. This is different for logarithmic growth. Logarithmic growth builds up momentum very early which results in demand premiums in the beginning stages after a movie's release. However, its demand decays faster compared to exponential growth in later weeks. If one thinks of logarithmic and exponential market shapes as manifestations of early and last-minute hype about a movie, respectively, then our results indicate that early hype is associated with high early demand followed by a rapid decay, and vice versa for last-moment hype.

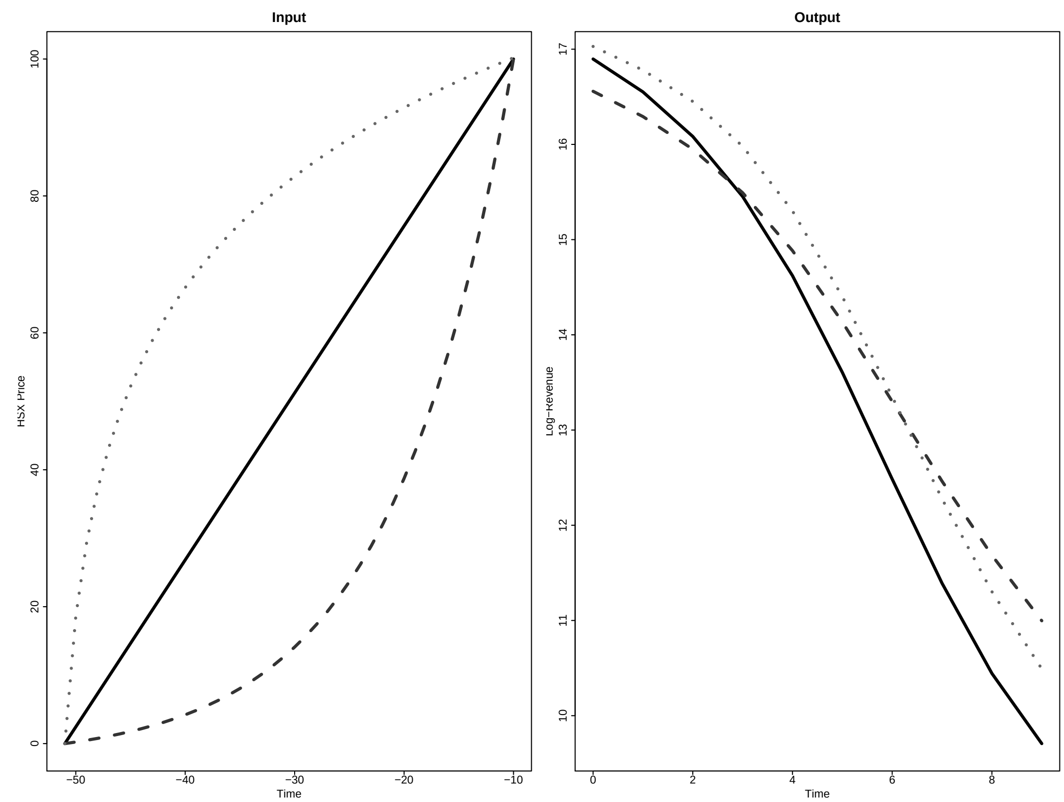

Figure 12: Dependence plots for linear, log and exponential input curves.

Using the approach outlined above, one can investigate additional relationships between input and output shapes. Figures 13 and 14 compare the effects of different *linear* market growth rates. While in Figure 13 all rates *end* in the same market value (but start at different values), the opposite is true for Figure 14. We can observe that the market starting value has an important effect on a movie's expected decay. In particular, Figure 13 shows that while all growth rates end in exactly the same value, the corresponding demand decay rates are vastly different. Conversely, Figure 14 suggests that, given the same market starting value, the ending value has a comparatively smaller effect on demand decay.

Figure 15 investigates a further aspect of market growth. While all growth curves start and end in the same value (similar to Figure 12), growth occurs now in the form of an S-shape (dark grey dashed line) and an inverse S-shape (light grey dotted line), respectively. We can see that both S-shapes result in higher demand (and slower demand decay rates) compared to the linear growth. Comparing one another, the inverse S-shape outperforms the S-shape. The inverse S-shape consists of a first market spurt very early and a second one very late; in other words, it combines the best of both the exponential and the logarithmic growth described in Figure 12. As a result, it has a slower demand decay rate compared to the S-shape, which only has a single spurt and which occurs in the middle of the trading period. This result seems to provide more evidence that early and last-moment hype are important components of a movie's box office performance.

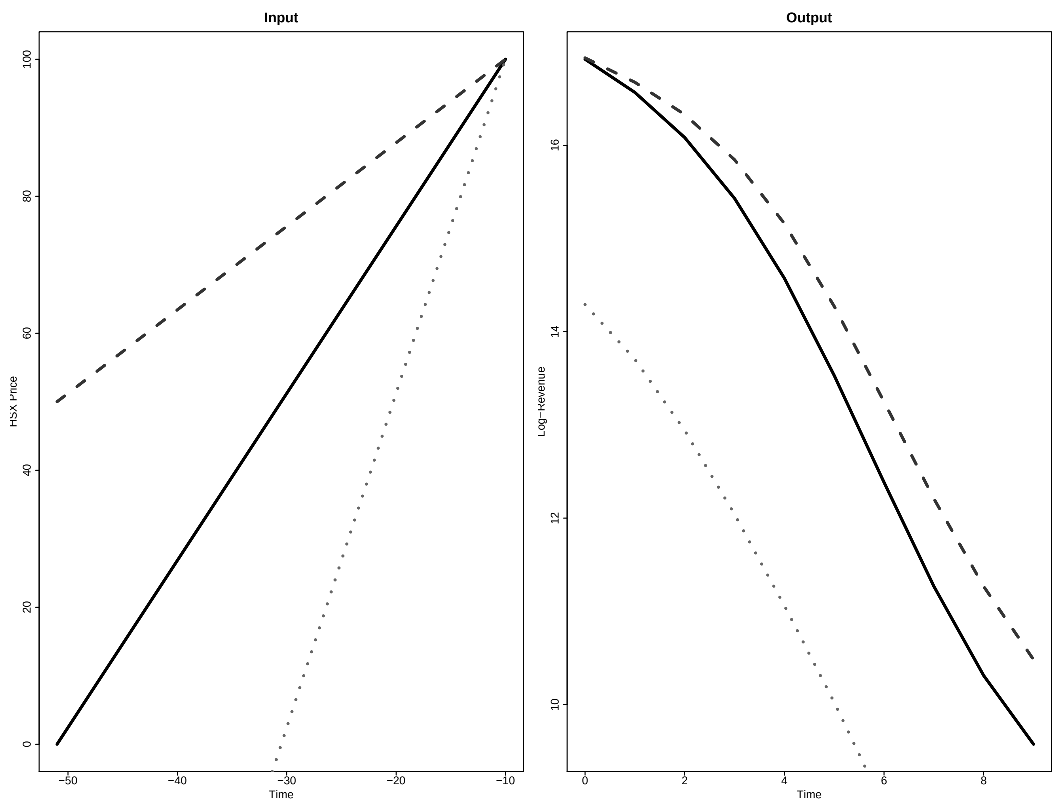

Figure 13: Dependence plots for different starting values.

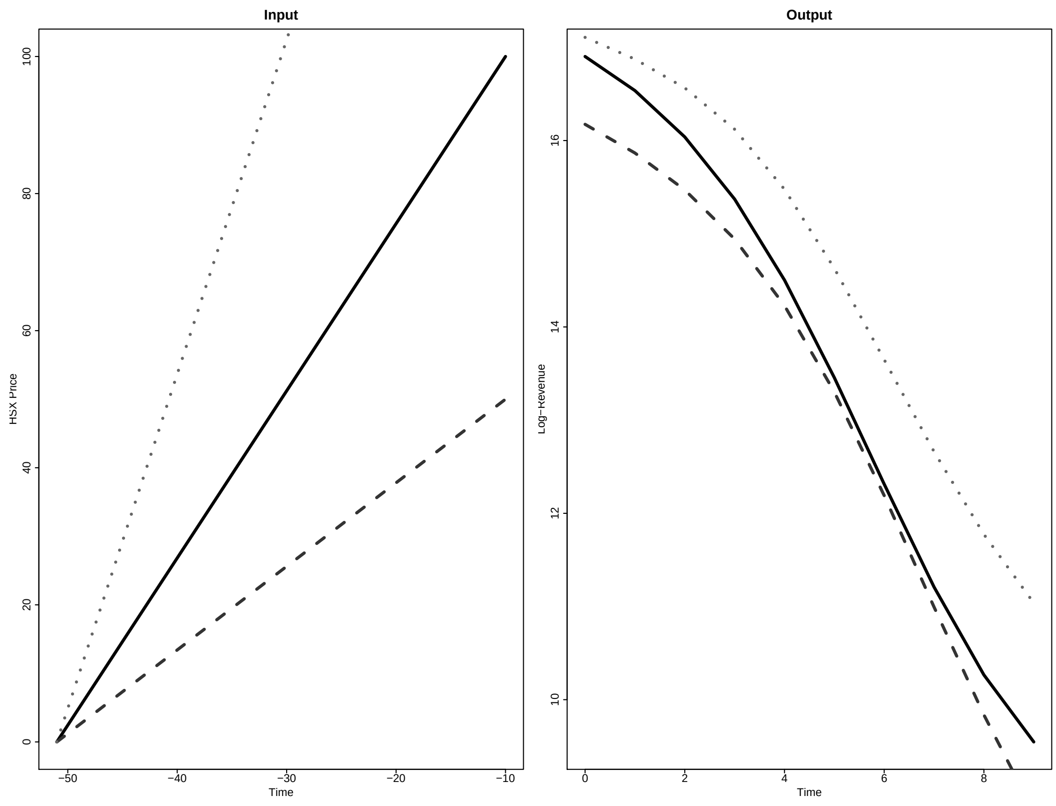

Figure 14: Dependence plots for different ending values.

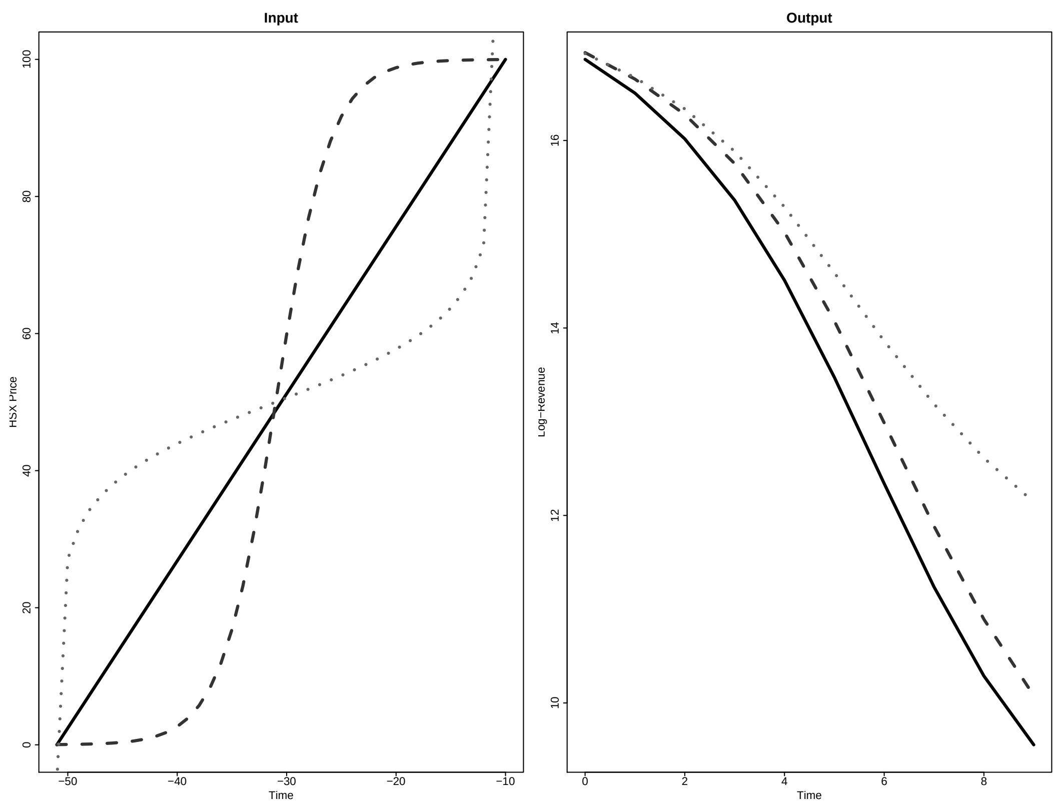

Figure 15: Dependence plots for S and inverse S curves.

5 Conclusion

This paper makes three significant contributions. First, we introduce a new and exciting data source to the statistics community. Online virtual stock markets (VSMs) are novel market-driven mechanisms to capture opinions and valuations of large crowds in a single number. Our work shows that the information captured in VSMs is rich but requires novel statistical ideas to extract all available knowledge. Second, we illustrate the power of functional modeling for analyzing VSMs by developing an innovative functional model that uses shapes both as input and output values and links them using powerful non-linear regression methods. This approach has good predictive accuracy relative to a series of competitor models. Finally, we make our approach practical for inference purposes by developing dependence plots to illustrate the relationship between input and output curves.

One limitation of our approach is that it may only add value in inefficient markets. Inefficient markets are markets where valuable information, above and beyond the information contained final trading price, is captured by the shape of the trading history. As outlined earlier, previous research suggests that VSMs are not fully efficient. Furthermore, the strong predictive accuracy of our functional approach provides further empirical validation for this finding. In addition, our methodology (modeling functional shapes and interpreting their relationship graphically) is applicable beyond just market data. In general it can be used on any regression problem involving functional predictors and responses.

There are several possible future extensions of this work. In terms of the data, we believe there are many other interesting applications of VSM's in different domains (e.g. music, TV, gaming). In addition, our current methodology models the input curves, output curves and the relationship between the two in three separate steps. It is possible that further improvements could be achieved by developing a process that integrated all three steps into one.

A Smoothing Splines

Let $\tau_1, \ldots, \tau_L$ be a set of knots. Then, a polynomial spline of order p is given by

$$f(t) = \beta_0 + \beta_1 t + \beta_2 t^2 + \cdots + \beta_p t^p + \sum_{l=1}^{L} \beta_{pl}(t - \tau_l)_+^p, \tag{9}$$

where $u_+ = uI_{[u \geq 0]}$ denotes the positive part of the function u. Define the roughness penalty

$$\text{PEN}_m(t) = \int \{D^m f(t)\}^2 dt, \tag{10}$$

where $D^m f$, $m = 1, 2, 3, \ldots$, denotes the mth derivative of the function f. The penalized smoothing spline f minimizes the penalized squared error

$$\text{PENSS}_{\lambda,m} = \int \{y(t) - f(t)\}^2 dt + \lambda \ \text{PEN}_m(t), \tag{11}$$

where $y(t)$ denotes the observed data at time t, and the smoothing parameter λ controls the trade-off between data-fit and smoothness of the function f. Using $m = 2$ in (11) leads to the commonly encountered cubic smoothing spline. Other possible smoothers include the use of B-splines or radial basis functions (Ruppert et al., 2003). In this study, we used smoothing splines of order $p = 3$ and a smoothing parameter of $\lambda = 5$ which minimized the average generalized cross validation, averaged across all movies.

References

Bass, F. M., Gordon, K., Ferguson, T. L., and Gith, M. L. (2001). DIRECTV: Forecasting diffusion of a new technology prior to product launch. *Interfaces*, 31(3):S82–S93.

Cohen, N. (2008). Google's lunchtime betting game. *New York Times, January 7, 2008.*

Dahan, E., Lo A. W., Poggio, T., Chan, N., and Kim, A. (2007). Securities trading of concepts (STOC). *Working Paper, MIT.*

De Bondt, W. F. M. and Thaler, R. H. (1985). Does the stock market overreact? *Journal of Finance*, 40:793–805.

Einav, L. (2007). Seasonality in the u.s. motion picture industry. *Rand Journal of Economics*, 38(1):127-145.

Foutz, N. and Jank, W. (2008). The Wisdom of Crowds: Pre-Release Forecasting via Functional Shape Analysis of the Online Virtual Stock Market. *Technical Report*, University of Maryland.

Friedman, J. and Hastie, T. and Tibshirani, R. (2000). Additive logistic regression: A statistical view of boosting. *Annals of Statistics*, 28:337-374.

Godes, D. and Mayzlin, D. (2004). Using online conversations to study word-of-mouth communication. *Marketing Science*, 23(4):545–560.

Godes, D., Mayzlin, D., Chen, Y., Das, S., Dellarocas, C., Pheiffer, B., Libai, B., Sen, S., Shi, M., and Verlegh, P. (2005). The firm's management of social interactions. *Marketing Letters*, 16(3/4):415–428.

Hanson, R. (1999). Decision markets. *IEEE Intelligent Systems*, 143:16–19.

Hastie, T. J. and Tibshirani, R. J. (1990). *Generalized Additive Models*. Chapman and Hall

Hastie, T. J., Tibshirani, R. J. and Friedman, J. (2001). *The Elements of Statistical Learning*. Springer

Jank, W. and Shmueli, G. (2006). Functional data analysis in electronic commerce research. *Statistical Science*, 21:155–166.

Johnson, J., Tellis, G. J., and Macinnis, D. J. (2005). Losers, winners, and biased trades. *Journal of Consumer Research*, 32:324–329.

Lamont, O. A. and Thaler R. H. (2003). Can the market add and subtract? Mispricing in tech stock carve-outs. *Journal of Political Economy*, 111(2):227–268.

Lo, A. W. and Mackinlay, A. C. (2002). *A Non-Random Walk Down Wall Street (Fifth Ed.), 4-47*. Princeton: Princeton University Press, Princeton, NJ.

Mahajan, V. and Wind, Y. (1998). New product forecasting models: direction for research and implementation. *International Journal of Forecasting*, 4(3): 341–358.

Muller, H. and Yao, F. (2008). Functional Additive Models. *Journal of the American Statistical Association*, (To appear).

Plott, C. R. and Chen, K. Y. (2002). Information aggregation mechanisms: concept, design and implementation for a sales forecasting problem. *Working Paper, California Institute of Technology and HP Labs*.

Ramsay, J. O. (1998). Estimating smooth monotone functions. *Journal of the Royal Statistical Society B*, 60(2):365–375.

Ramsay, J. O. and Silverman, B. W. (2005). *Functional data analysis (Second Ed.)*. Springer-Verlag, New York.

Rhode, P. W. and Strumpf, K. S. (2006). Manipulating political stock markets: a field experiment and a century of observational data. *Working Paper, UNC*.

Ruppert, D., Wand, M. P., and Carroll, R. J. (2003). *Semiparametric Regression*. Cambridge University Press, Cambridge.

Sawhney, M. S. and Eliashberg, J. (1996). A parsimonious model for forecasting gross box-office revenues of motion pictures. *Marketing Science*, 152:113-131.

Sood, A., James, G. and Tellis, G. (2008). Functional Regression: A New Model and Approach for Predicting Market Penetration of New Products. *Marketing Science*, (To appear).

Spann, M. and Skiera, B. (2003). Internet-based virtual stock markets for business forecasting. *Management Science*, 4910:1310-1326.

Surowiecki, J. (2004). *The Wisdom of Crowds*. DoubleDay.