# Functional Shape Analysis to Forecast Box-Office Decay using Virtual Stock Exchanges

Wolfgang Jank
RH Smith School of Business
University of Maryland

Jointly with
**Natasha Foutz**, UVA & **Gareth James**, USC

Forecasting Product Demand: Demand for Hollywood Movies

- **Why is forecasting new product demand important?**
  - To make optimal marketing allocation decisions
    - $ of advertising, # of screens, etc
- **Why is it challenging?**
  - Because demand is highly variable
    - High for some products, low for others
      - "The Woodsman" had a budget of $30mill, but made only $54,000 on the first week-end
      - "High Tension" had a budget of only $430,000, and made $1.9 mill on the first week-end
      - "Alexander" made $14 mill on first week-end
  - Because demand changes over life-cycle
    - High early, low later; sometimes vice versa
   - Changing **rate of decay**
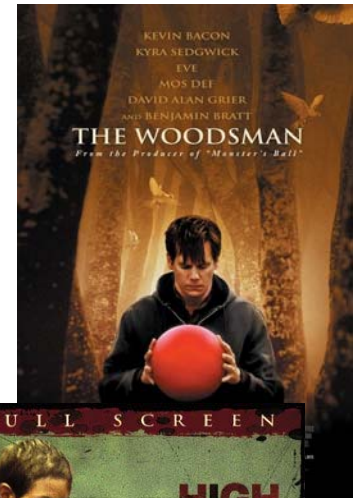
Movie Demand Decay Rates

- **Movie demand decays at varying rates**
  - Some decay at log-linear rate; others have sudden drops
  - Some increase over time ("Sleepers"); others behave very volatile
- **Our goal is to forecast demand decay rates**
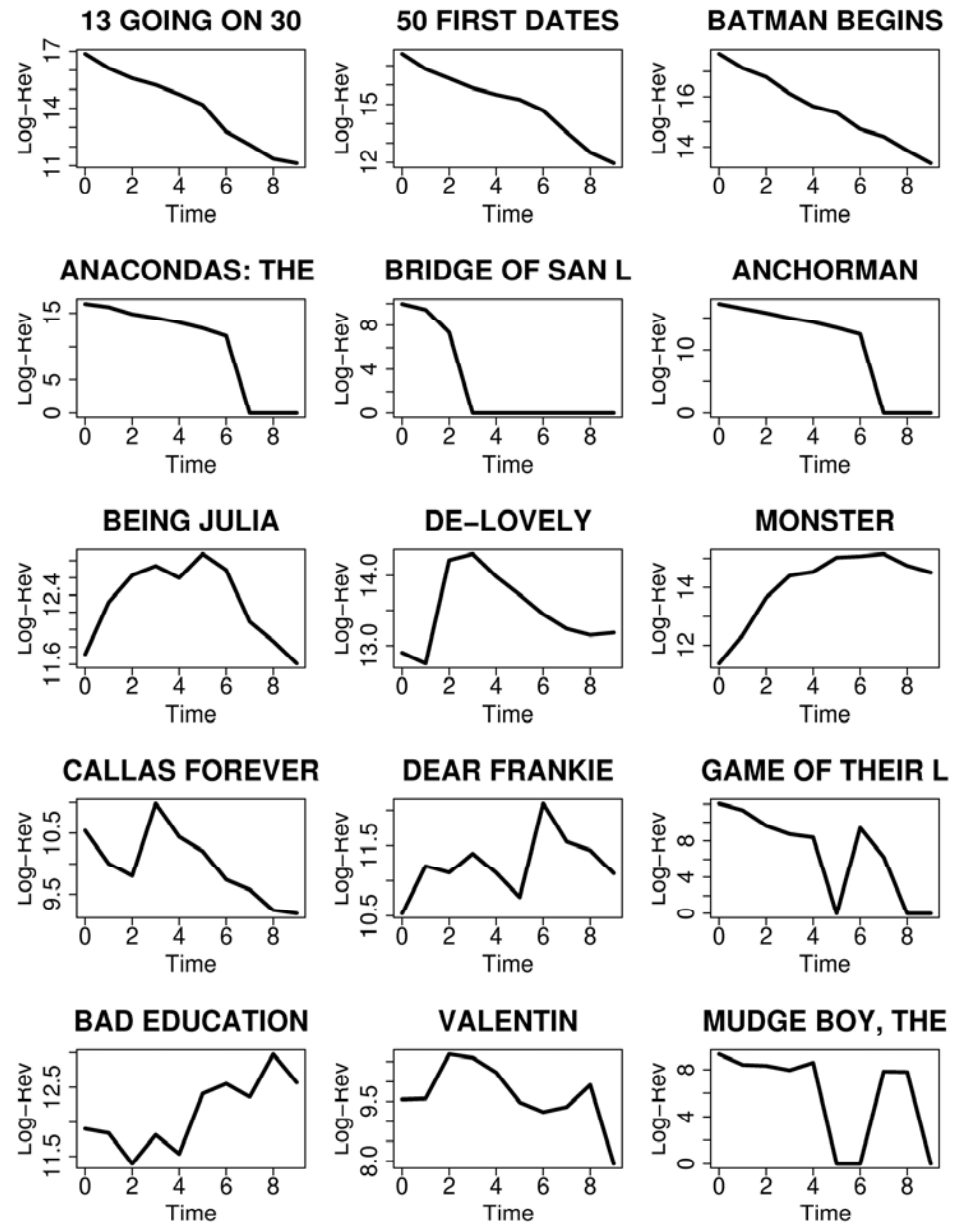
First Reflection

- **The Goal: Forecast entire demand decay pattern**
  - Not just a single value
- **The Challenge**
  - Demand distribution highly irregular
    - Parametric demand models are doomed
  - Want to extract demand **shape** characteristics
- **One more challenge**
  - Want to make **early** forecasts
    - Way before movie is released!
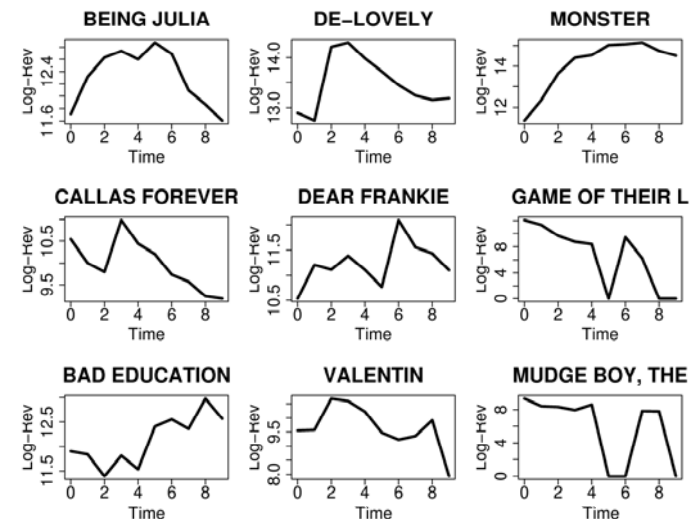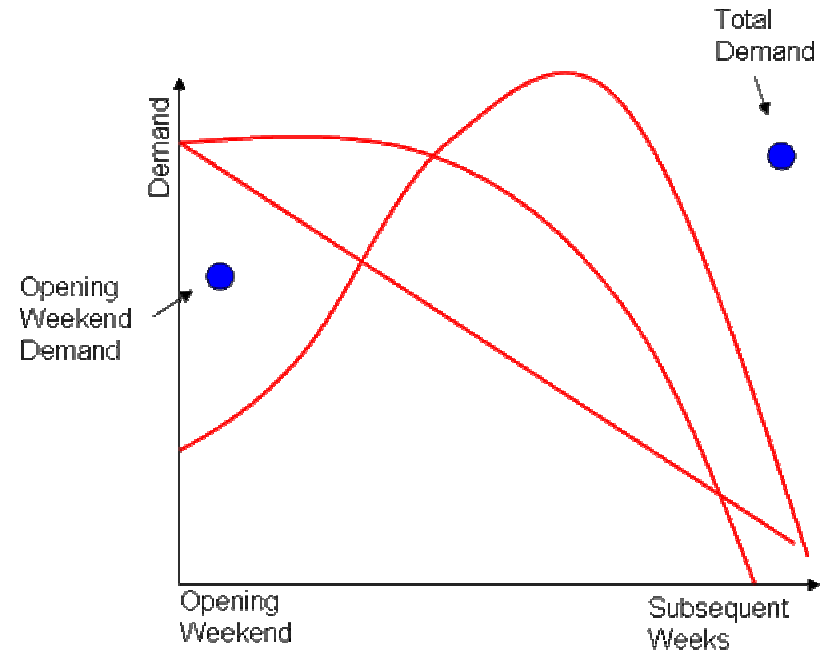
Our only (or main) source of info: Data from a Prediction Market

The Hollywood Stock Exchange (www.hsx.com)

What is a Prediction Market?

- It is a speculative market created for the purpose of making predictions
  - Similar to the stock market
  - Assets' final cash value is tied to an event or parameter
    - Will the next US president be a Democrat?
    - What will be the magnitude of gasoline prices next quarter?
  - Current **market price** can be interpreted as **probability** of the event or the **expected value** of the parameter.
  - Also referred to as information market, decision market, idea future, event derivatives, virtual market, …

Iowa Electronic Market and the 2008 US Presidential Election
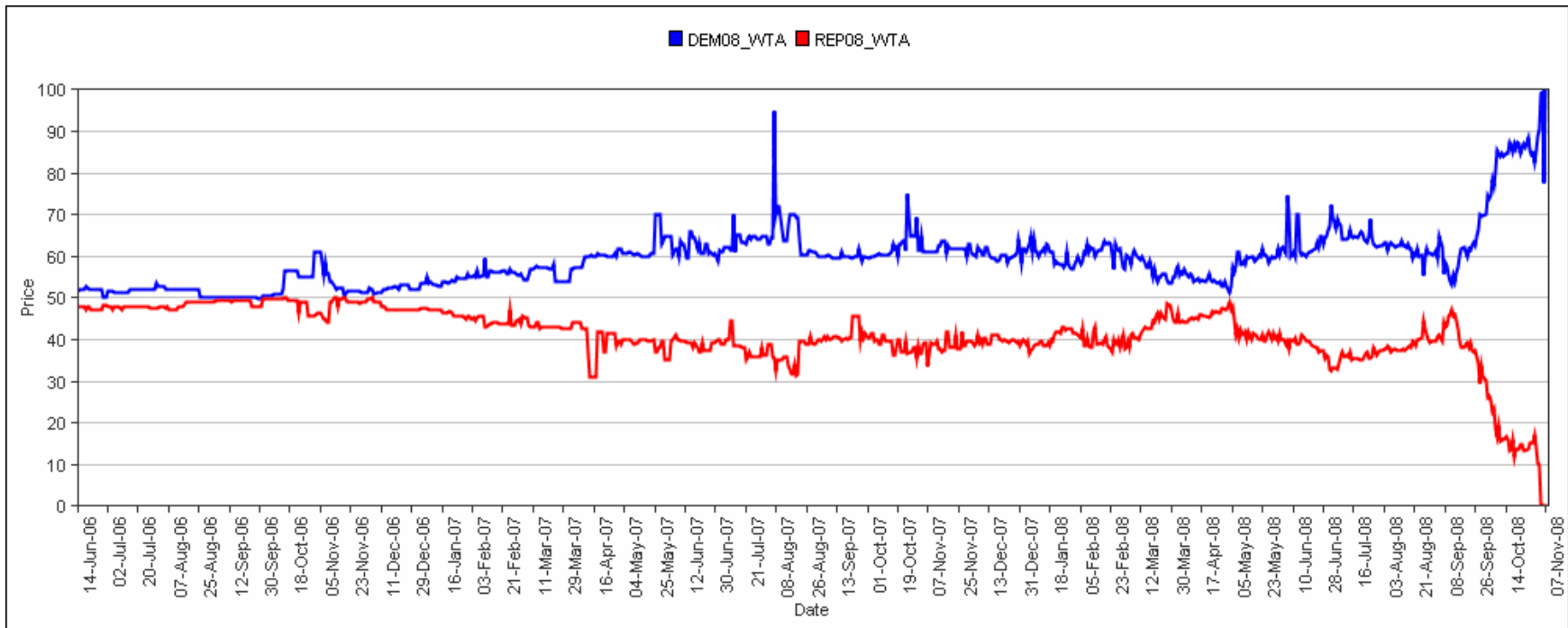
Republican Winner   Democrat Winner

Sports, Politics, Society, Money, News, and much, much more …

Will there be more than 9 named Atlantic Hurricanes in 2008?
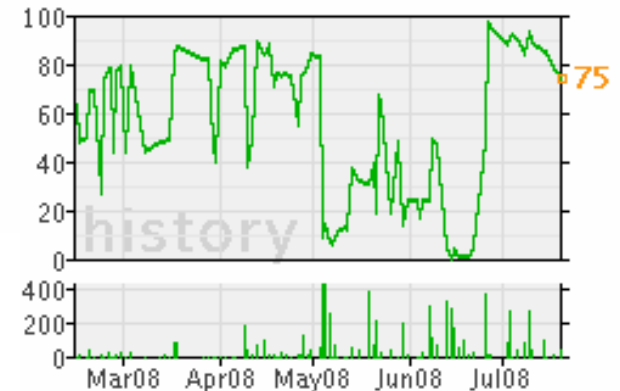

Will The US catch Osama bin Laden while Bush is President?

Will Microsoft buy Yahoo in 2008?

*all courtesy of Newsfutures.com

More major corporations use VSMs

- **Generate ideas**
- **Design new products**
- **Test ad effectiveness**
- **Understand consumer behavior**

We use data from the Hollywood Stock Exchange (HSX)

- On HSX, one can trade movies, actors, and more

MOVIE STOCKS®

COMPLETE LISTING ⊗   ADVANCED SEARCH ⊗
RELEASE SCHEDULE ⊗   WEEKEND OPENERS ⊗

TOP 15 MOST WIDELY HELD

#	Title	Symbol	Price	Change
1	Indiana Jones 4	INDI4	H$226.08	+1 7/8
2	The Simpsons	SIMPS	H$148.01	HALTED
3	Harry Potter and the Order of the Phoenix	HPOT5	H$251.01	-3 1/8
4	The Bourne Ultimatum	BORN3	H$147.03	+2 1/2
5	The Hobbit	HOBIT	H$81.21	+1/2
6	Harry Potter and the Half-Blood Prince	HPOT6	H$213.99	+5/8
7	Austin Powers 4	AUST4	H$59.85	-1/2
8	Terminator 4	TERM4	H$49.21	-3/4

STAR BONDS®

COMPLETE LISTING ⊗   ADVANCED SEARCH ⊗

TOP 15 MOST WIDELY HELD

#	Name	Symbol	Price	Change
1	Johnny Depp	JDEPP	H$107.54	0
2	Angelina Jolie	AJOLI	H$93.53	+1/2
3	Ben Affleck	BAFFL	H$32.32	+1/4
4	Adam Sandler	ASAND	H$101.71	+1 1/4
5	Julia Roberts	JROBE	H$55.50	0
6	Brad Pitt	BPITT	H$111.89	-1 1/4
7	Matt Damon	MDAMO	H$91.62	+1 1/2
8	Samuel L. Jackson	SLJAC	H$25.03	-1/4
9	Nicole Kidman	NKIDM	H$58.99	-3/8

Works just like a Stock Market

What do I get in return?

- **The key challenge for prediction markets are the right incentives:**
  - Why would you submit your wisdom to the market?
  - Some argue that prediction markets based only on play money will not work
  - HSX (and others markets) prove this wrong!

H$hop Redemption

Short on H$? Click here for Cash Only prices.

Summer Special! Receive a free PEN with your first $10 purchase and a free WAND with your first $25.

Sleek and stylish, the HSX twill cap is black with striking white and orange accents with 'BE A PLAYER' on back.

Made of 100% cotton brushed twill with a 6-panel unconstructed crown, self-fabric closure and d-ring slider. This cap is what the cool kids are wearing.

H$10 million + US$12.00

You'll be styling in this 100% pre-shrunk cotton Tee with bold HSX in white and orange accents.

This Navy Tee is perfect for a cool, confident look. Comes in M, L, XL.

H$15 million + US$10.00

Size: Medium

This sporty, casual Tee in Pine is set off nicely with HSX in white and orange accents.

Made with 100% pre-shrunk cotton, this Tee will have your buddies talking. Comes in M, L, XL.

H$15 million + US$10.00

Size: Medium

This black 15 oz ceramic mug will be the envy of your office. Comes with HSX on front and 'BE A PLAYER' on the back.

H$10 million + US$7.50

Your co-workers will be filled with bitter jealousy as you sip your coffee, tea or cocoa.

More incentives…

- The most successful traders are featured on the "leader board"

Leaderboard

LEADER BOARD

	WEEK-TO-DATE			Top 100
rank	screen name		gain	worth
1	MoneyMagnet		30.96 %	H$5,201,661.86
2	Tim		24.84 %	H$3,029,207.37
3	liah katlin		22.05 %	H$2,440,916.59
4	Jeff McDermott		20.01 %	H$2,479,172.70
5	Froggy and the Ho Piggy		19.81 %	H$4,631,992.16
6	Mick Mathews		19.58 %	H$2,996,622.06
7	simon lopez		19.45 %	H$2,388,971.96
8	Wild Bill		18.38 %	H$2,399,967.13
9	Leticia Ford		18.15 %	H$3,281,741.47
10	Gigi^2		16.07 %	H$4,722,570.89

Rankings Current as of 09-JUN-2008

	MONTH-TO-DATE			Top 100
rank	screen name		gain	worth
1	MoneyMagnet		160.08 %	H$5,201,661.86
2	bret shuckerow		75.86 %	H$4,887,319.45
3	Andrew Vytuvis		73.42 %	H$3,589,623.82
4	Murray Hainer		72.67 %	H$3,511,103.52
5	Mike McCarte		72.20 %	H$8,089,736.11
6	Matt Breeden		68.43 %	H$3,368,559.68
7	Shane Spence		68.37 %	H$5,634,247.48
8	Michael Mattioli		65.96 %	H$3,988,725.12
9	Lucas Orner		63.34 %	H$3,266,894.53
10	Marlo MacPoo		62.39 %	H$3,558,930.86

Rankings Current as of 09-JUN-2008

	SEASON-TO-DATE			Top 100
rank	screen name		gain	worth
1	Scott Howard		231.70 %	H$6,633,942.28
2	janira nafarrate		200.32 %	H$8,950,443.45
3	Tigoda		189.12 %	H$8,552,299.73
4	Mike Reyes		188.02 %	H$10,960,805.15
5	Jon Robb		180.63 %	H$5,633,012.60
6	goober saurus		170.11 %	H$7,837,394.54
7	steve harrington		167.01 %	H$7,282,933.26
8	Big Kitty		163.90 %	H$8,773,859.18
9	MoneyMagnet		160.08 %	H$5,201,661.86
10	Shane Spence		154.26 %	H$5,634,247.48

Rankings Current as of 09-JUN-2008

	YEAR-TO-DATE			Top 100
rank	screen name		gain	worth
1	Get Smarter!!! YTD 2008		8972.99 %	H$181,459,747.23
2	(*_*) John McCain n2008		4322.47 %	H$88,449,472.03
3	helmetz		4091.71 %	H$83,834,160.86
4	Decoman YTD 2008		3907.11 %	H$80,142,163.76
5	Group Captain Lionel Mandrake YTD		3659.76 %	H$75,195,191.40
6	maguire		3311.81 %	H$68,236,261.87
7	DragonDrop		3191.06 %	H$65,821,213.29
8	LauK YTD 08 (Reset 2/11)		2792.06 %	H$57,841,232.00
9	NiteKraz YTD		2529.86 %	H$52,597,237.98
10	gav08		2469.58 %	H$51,391,517.03

Rankings Current as of 09-JUN-2008

	LIFETIME			Top 100
rank	screen name		gain	worth
1	fembot		557938.07 %	H$11,160,761,460.23
2	The Full Merino		467657.65 %	H$9,355,153,093.41
3	The Moghul of The Rings		424199.70 %	H$8,485,993,904.89
4	HAL 9000		376020.96 %	H$7,522,419,218.07
5	Trader Jones		324206.05 %	H$6,486,121,059.41
6	HAL 8000		314417.98 %	H$6,290,359,528.98
7	zeitgeist		292655.77 %	H$5,855,115,437.39
8	R A		290873.26 %	H$5,819,465,217.61
9	Pint-Sized Txredd		285621.36 %	H$5,714,427,111.32
10	HSX S&P: MOVIES - BONDS - FUNDS - OPTIONS		279884.56 %	H$5,599,691,244.18

Some curiosities…

- …and you can sell your virtual portfolio on eBay!

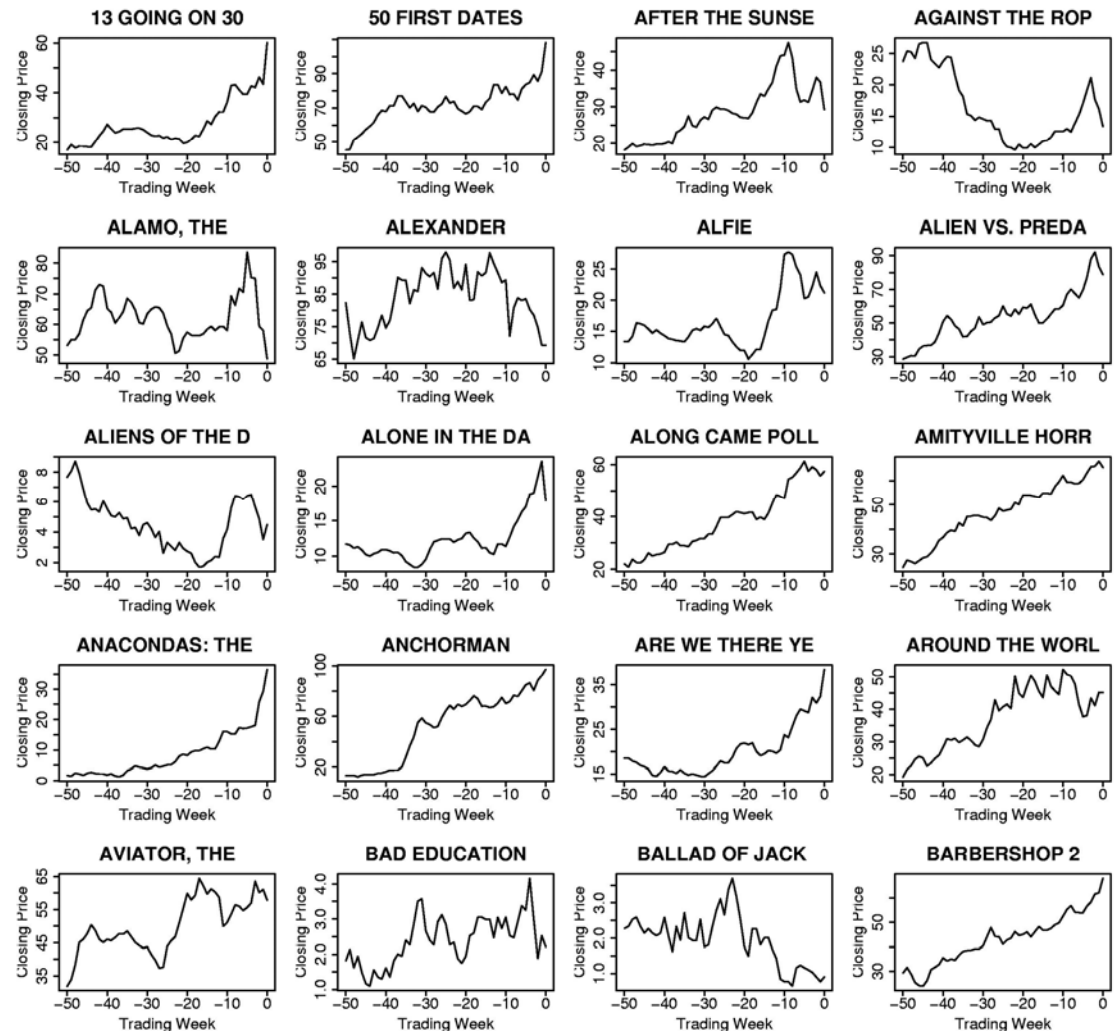Hollywood Stock Exchange Billion Dollar Portfolio

This auction is for my HSX portfolio currently worth just over 1 Billion Dollars. I've gone back to school and I don't have much time to play any more. If you're a member of the HSX movie community and you're scraping along with a few million dollars, now's your chance to save a few years of hard work and buy all the MovieStocks and StarBonds you want to own!
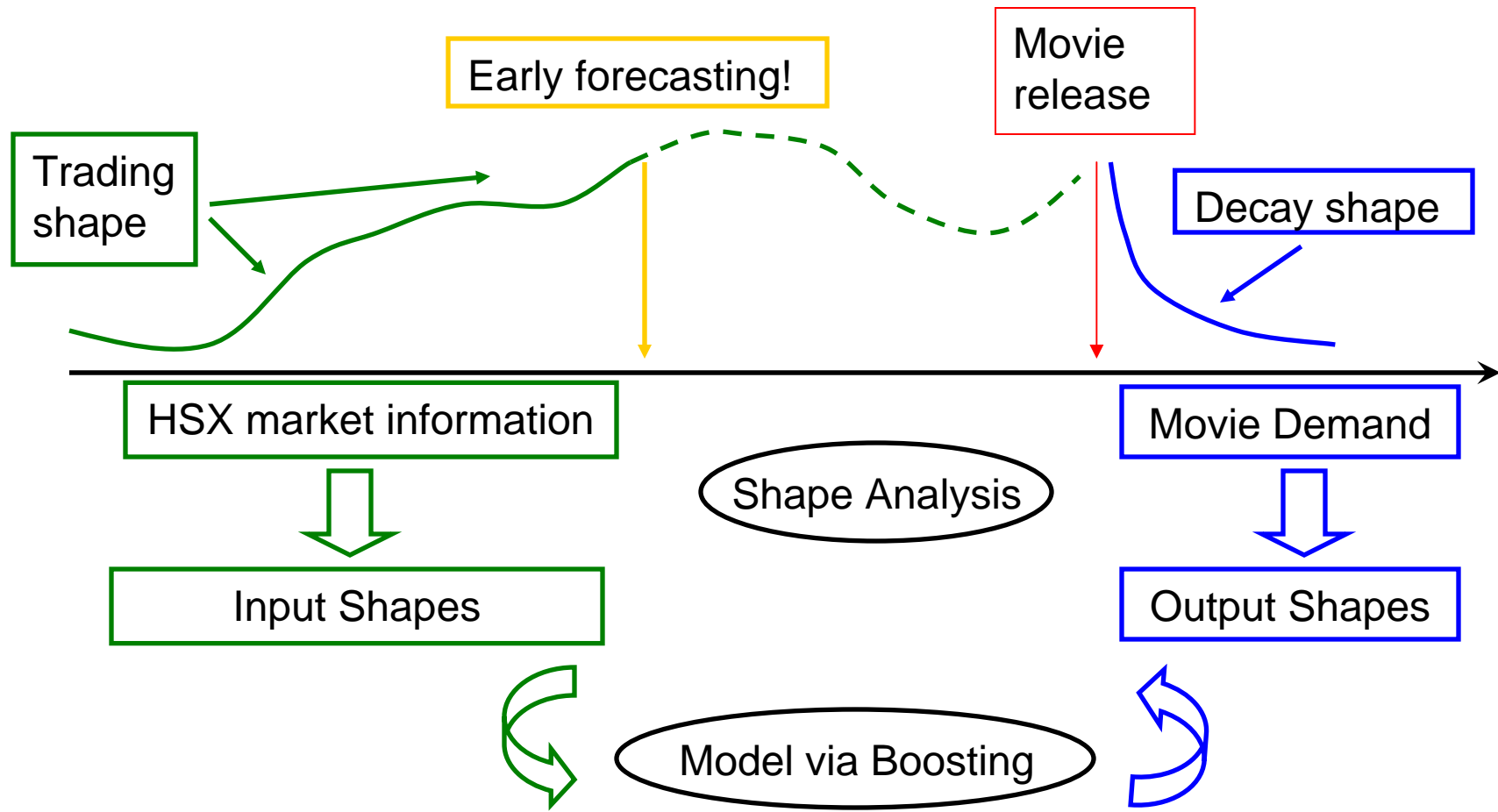
Bidding starts at only One Dollar!

Sample Data from HSX

- We want to use the **shape** of the trading data
  - Concave, convex
  - Linear, nonlinear
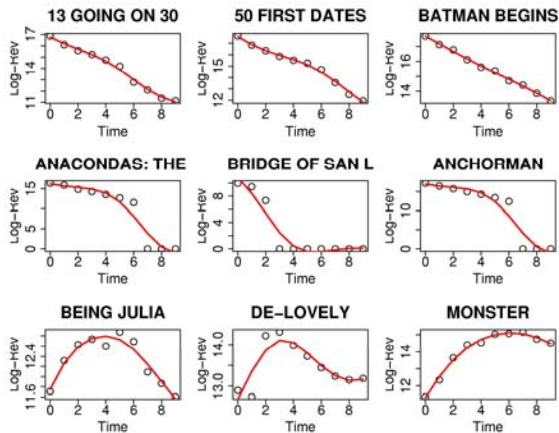 - Volatile, non-volatile

- Shape captures **information cascading**, **herding**

Second Reflection:
Our Full Forecasting Model

Early forecasting!

Movie release

Trading shape

Decay shape

HSX market information

Shape Analysis

Movie Demand

Input Shapes

Output Shapes

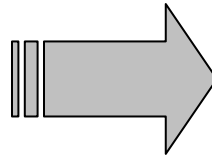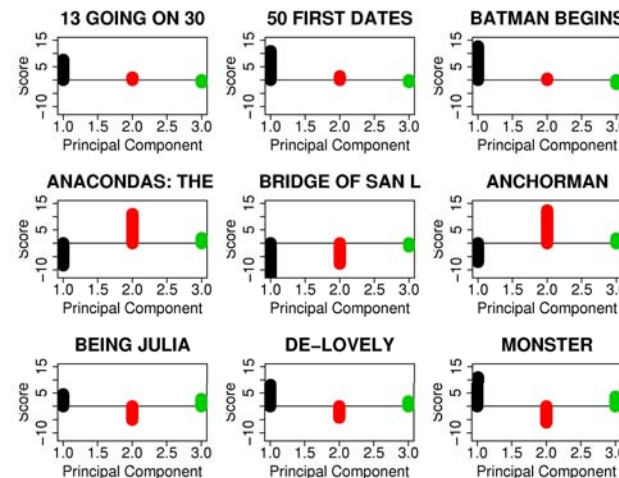Model via Boosting

Shape Analysis: Basic Idea

Smooth observed decays
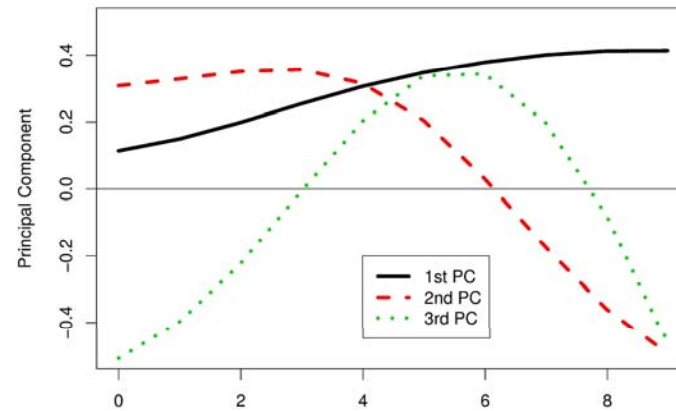
Apply functional PCA

Characterize decay rates by principal component scores
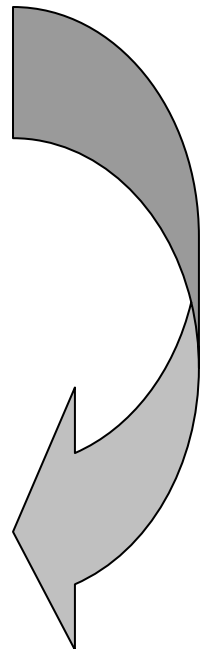
Functional Shape Analysis of Decay

- Functional principal component analysis
  - First 3 PCs explain over 98% of all decay variation
- PC's capture different decay aspects
  - **PC1**: "slow decayers" - movies that hold demand in later weeks
  - **PC2**: "sudden droppers" - movies that decay suddenly after first few weeks
  - **PC3**: "sleeper" movies - movies that gain only very late

Illustration of Decay Shapes
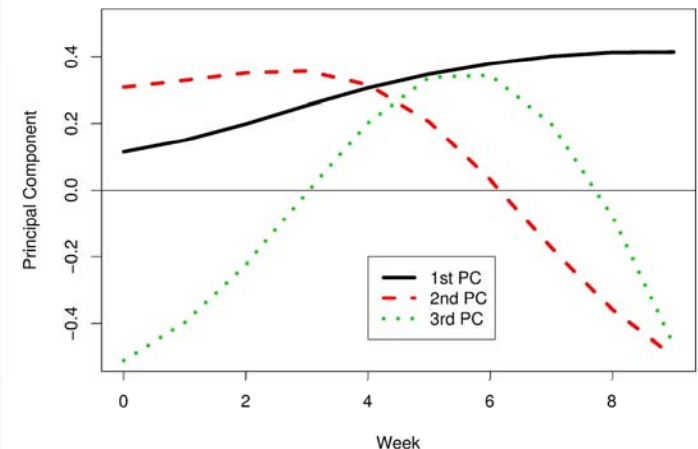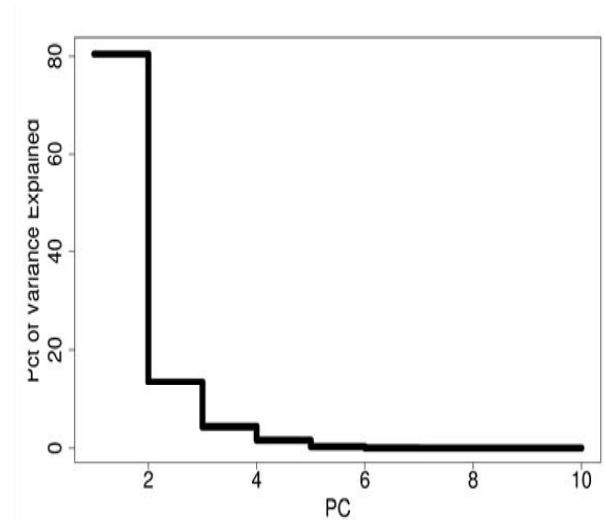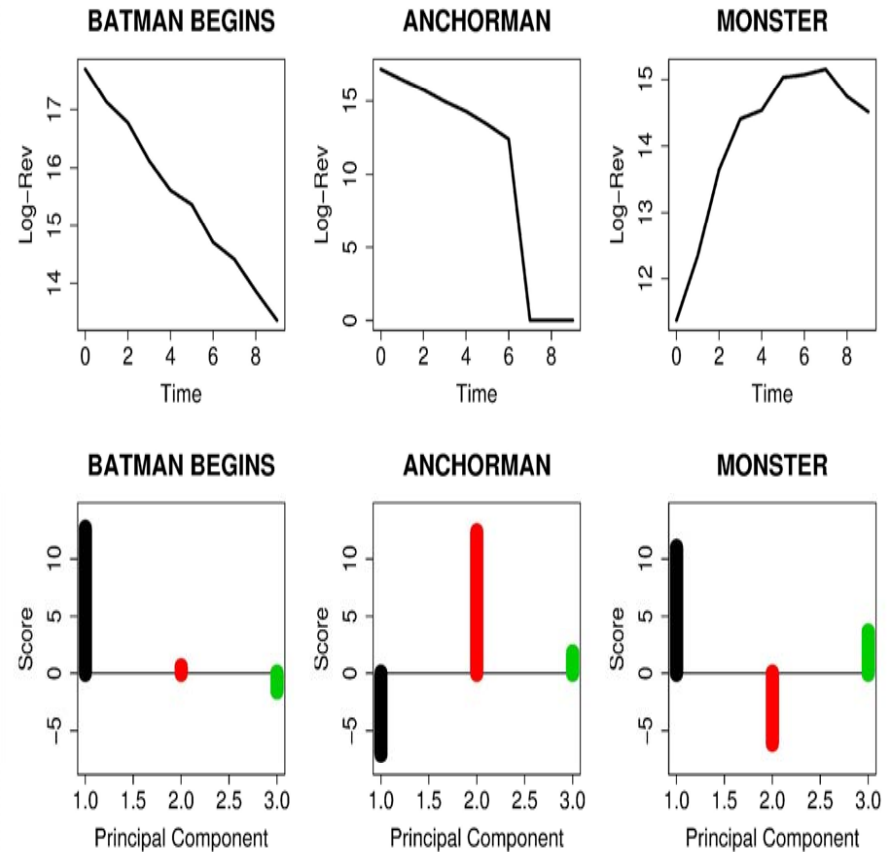
- PC1: Batman and esp. Monster hold demand high in later weeks
- PC2: Anchorman suddenly drops to zero after week 6
- PC3: Monster generates demand only very late
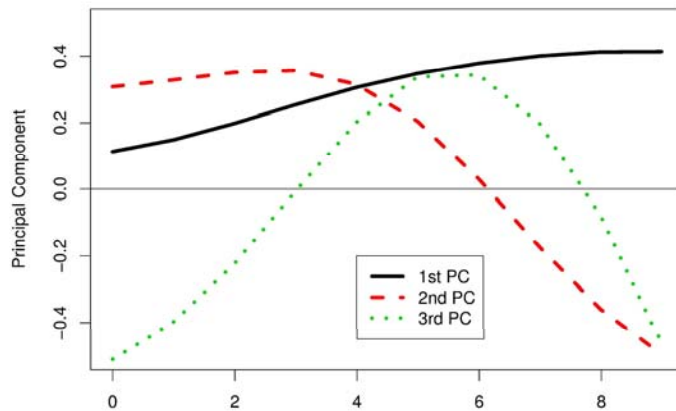
Important: There are more than 3 Shapes -- The "DNA" of Decay

- **3 PCs are only basic building blocks for different decay shapes**
  - ☐ There are many more decay patterns
  - ☐ Movie-specific shapes captured by the corresponding PC-Scores
    - ■ Linear combination of 3 basic PCs
  - ☐ We use similar ideas for capturing shapes of HSX market information

Late-dropper, little-sleeper

Non-dropper, late-sleeper
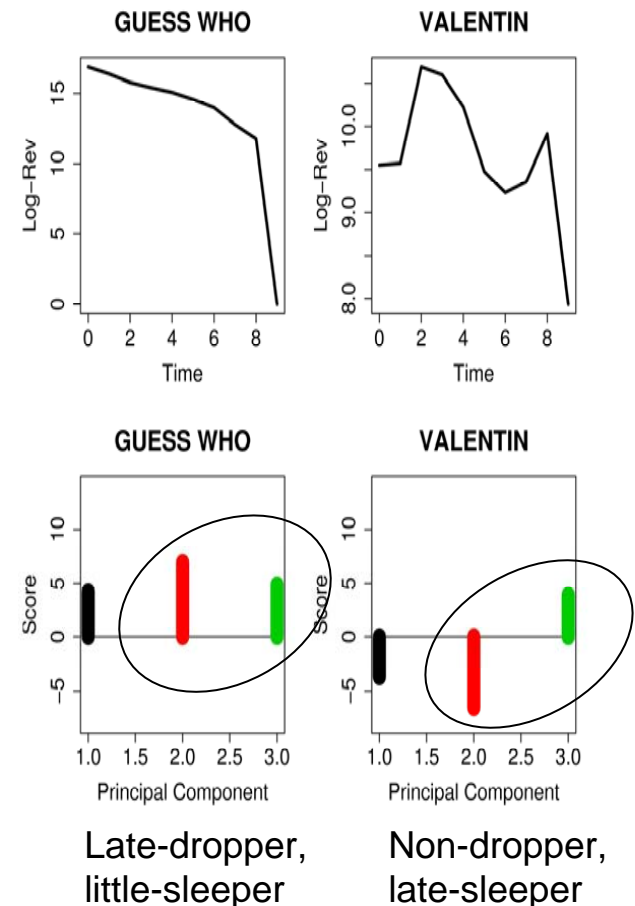
Where we are, where we are going

- **We now successfully have...**
  - ...characterized input and output shapes using the principal component scores
- **We still need to…**
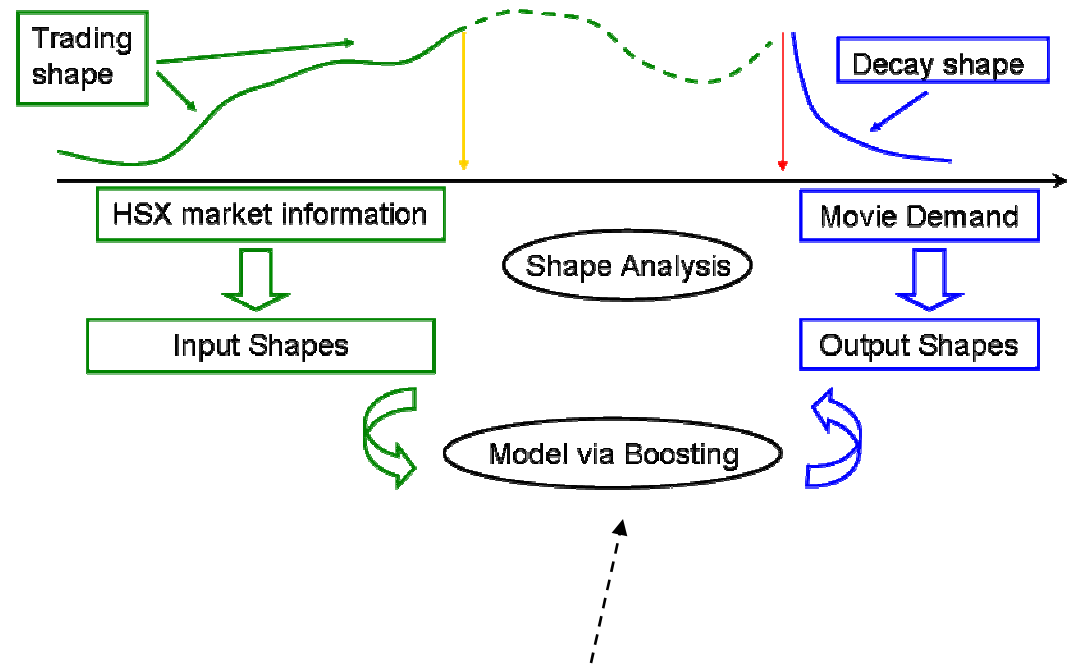 - …model output shapes as a function of input shapes

Trading shape

HSX market information

Shape Analysis

Decay shape

Movie Demand

Input Shapes

Output Shapes

Model via Boosting

Different modeling options: linear, nonparametric, neural networks, boosting, etc

Modeling Input & Output Shapes

- **Different modeling options**
  - Linear Models
    - Simple, straightforward
  - Generalized Additive Models (GAM)
    - Flexible, non-linear relationships between X and Y
  - Boosting
   - Can capture non-additive data structure between X and Y
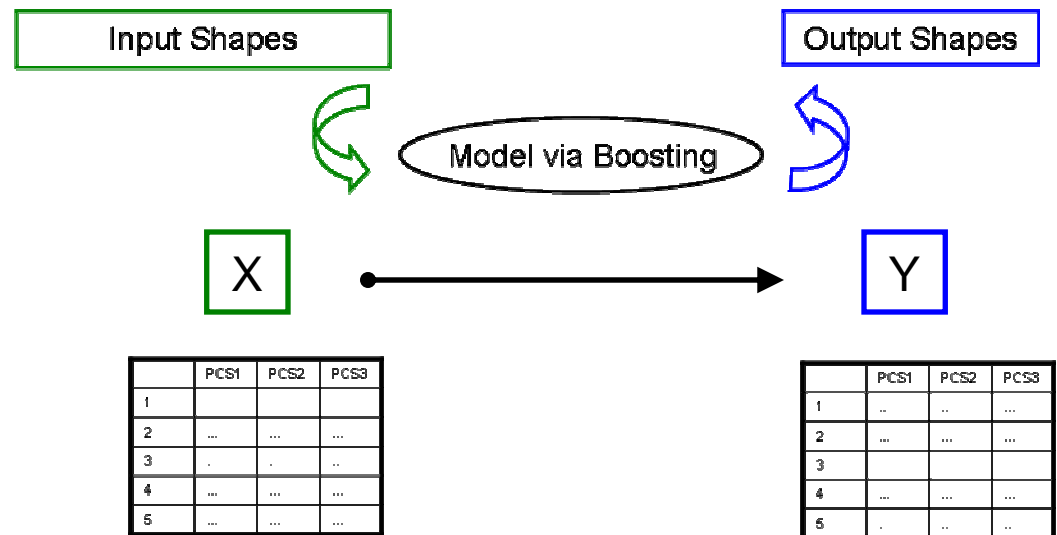
Predictive Accuracy

- **Ten-fold cross-validation to gauge predictive accuracy**
  - Predicting demand for the first 10 weeks after movie release
- **Prediction best for the Boosting Model**
 - Evidence that there data has non-linear, non-additive structure
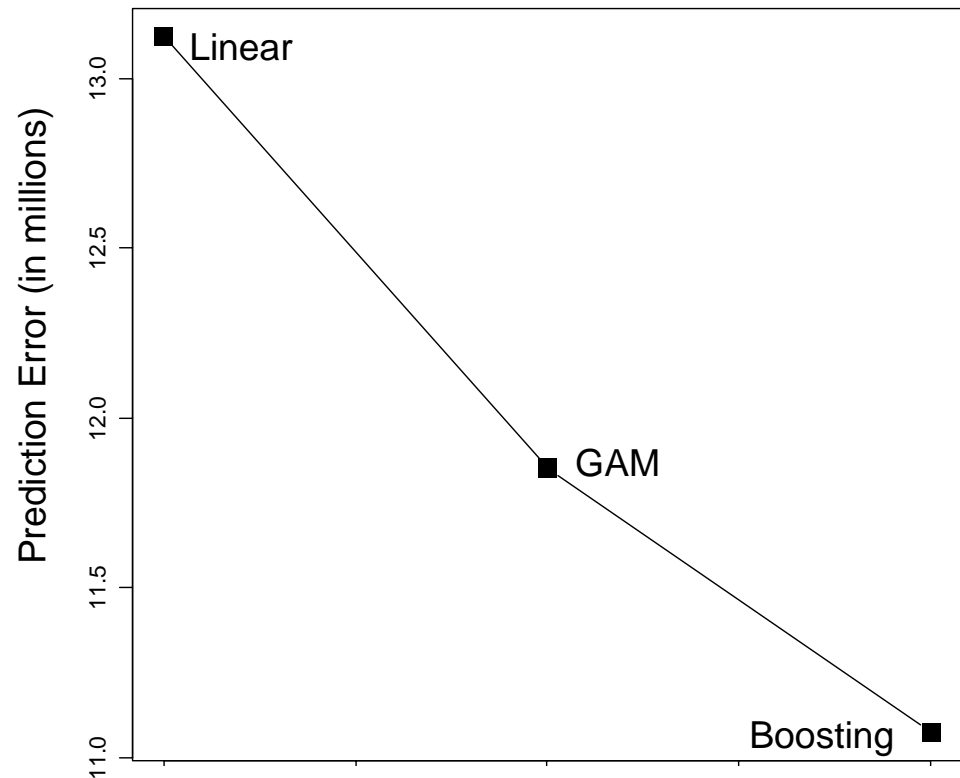
What's left? Inference…

- Functional Shape Analysis:
  - Powerful model for **predicting** demand decay
  - Disadvantage: hard to **interpret**
    - Both model input and output based on principal components
- Solution
  - Reverse-engineer the process
  - Create "dependence plots" that unveil the input-output relationship
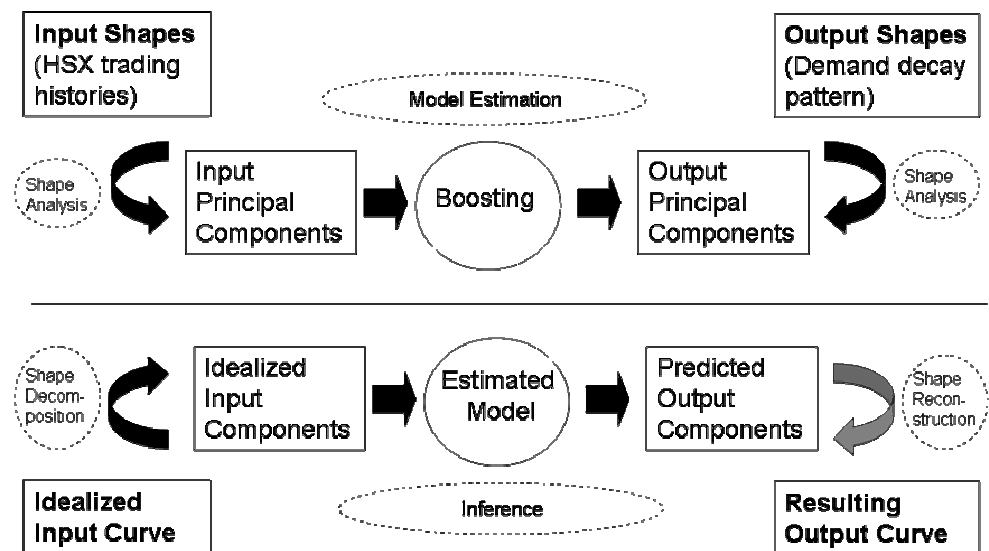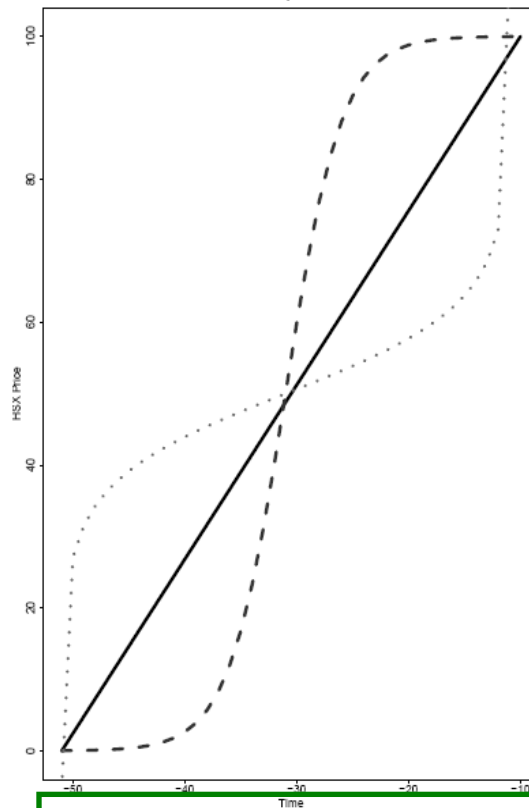
Dependence Plot (1)

- **Input curves measure the effect of different shapes**
  - Curves start and end at same value, only differ in path
  - Early market growth vs late growth
  - Manifestation of early and last-moment **hype** about a movie
- **Resulting demand decay**
  - Early hype leads to high early demand, but rapid decay
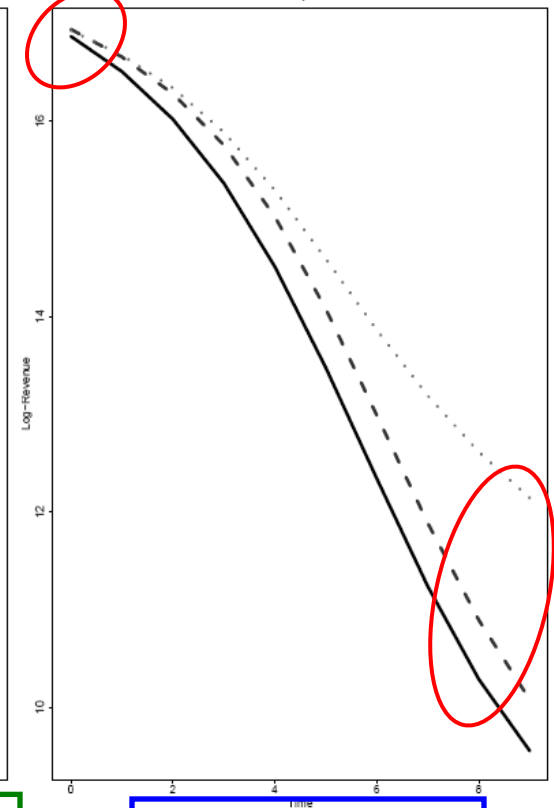 - Last-moment hype leads to lower initial demand, but slower decay

Idealized Input Curve — HSX market information

Resulting Output Curve — Movie Demand
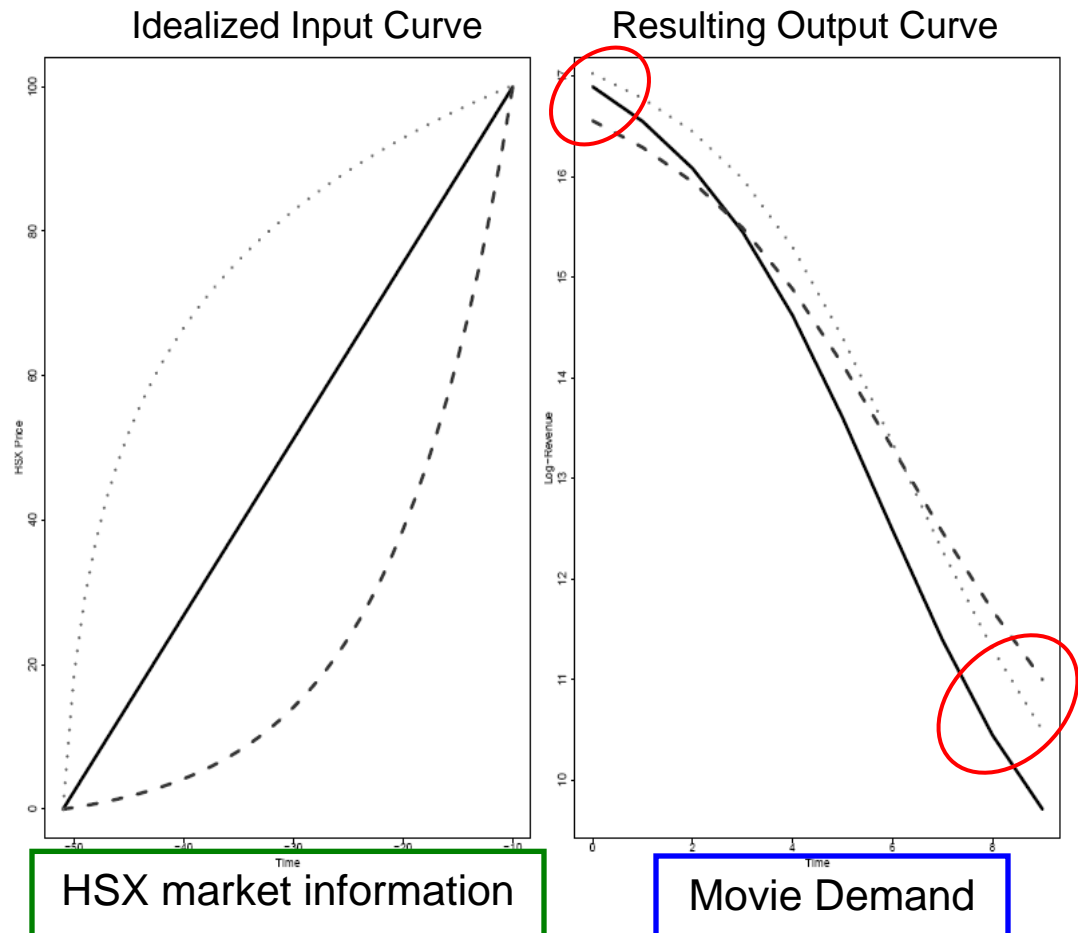
Dependence Plot (2)

- **Input curves**
  - ☐ S-and inverse S-shape
  - ☐ Inverse S-shape combines early and last-moment hype
- **Resulting demand decay**
  - ☐ Both S-shapes result in higher demand
    - ■ Indicates the importance of herding & diffusion dynamics
  - ☐ Inverse-S combines best of both worlds
    - ■ Both early and late diffusion dynamics

Idealized Input Curve

Resulting Output Curve

HSX market information

Movie Demand

Conclusions

- **Market shapes carry valuable information**
  - ☐ Result in more accurate forecasts
  - ☐ Give insight into impact of herding behavior, information cascading, hype, etc
- **Many different methodological challenges**
  - ☐ Modeling decay
  - ☐ Modeling trading shapes
  - ☐ Selecting important market information
  - ☐ Linking response and predictors