Forecasting diffusion with pre-launch online search traffic data

Oliver Schaer
Nikolaos Kourentzes
Robert Fildes

Higher School of Economics
Saint Petersburg

25th May 2016

Lancaster Centre for Forecasting
Lancaster University Management School

LUMS

Lancaster Centre for Forecasting

Current challenges for forecasting product sales

Fast technological lifecycles

High and fast-paced (often global) competition on technological, provider and application level

Abrupt change in consumer and user behaviour

Very limited data availability new product launches

Can search traffic popularity help us to improve sales forecasts?

The video game industry

Worldwide games revenue around £100 billions

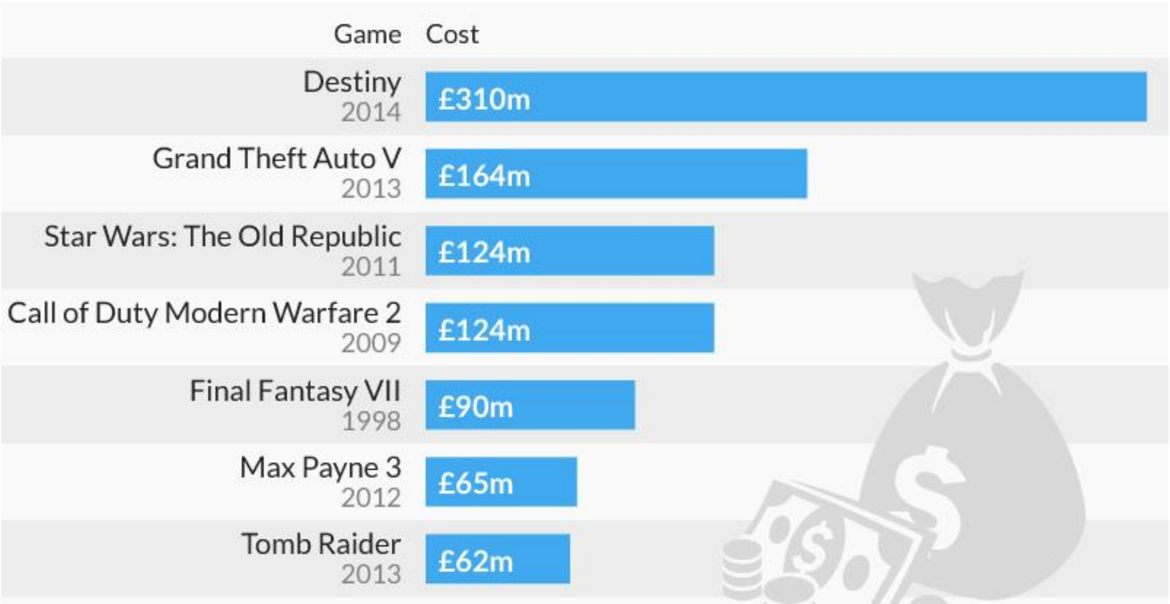
Creates roughly 5% of the global entertainment revenue

U.S. companies alone employ more than 40,000 people

Challenges in the industry

Increasing developing costs

Video games are getting as expensive as movies

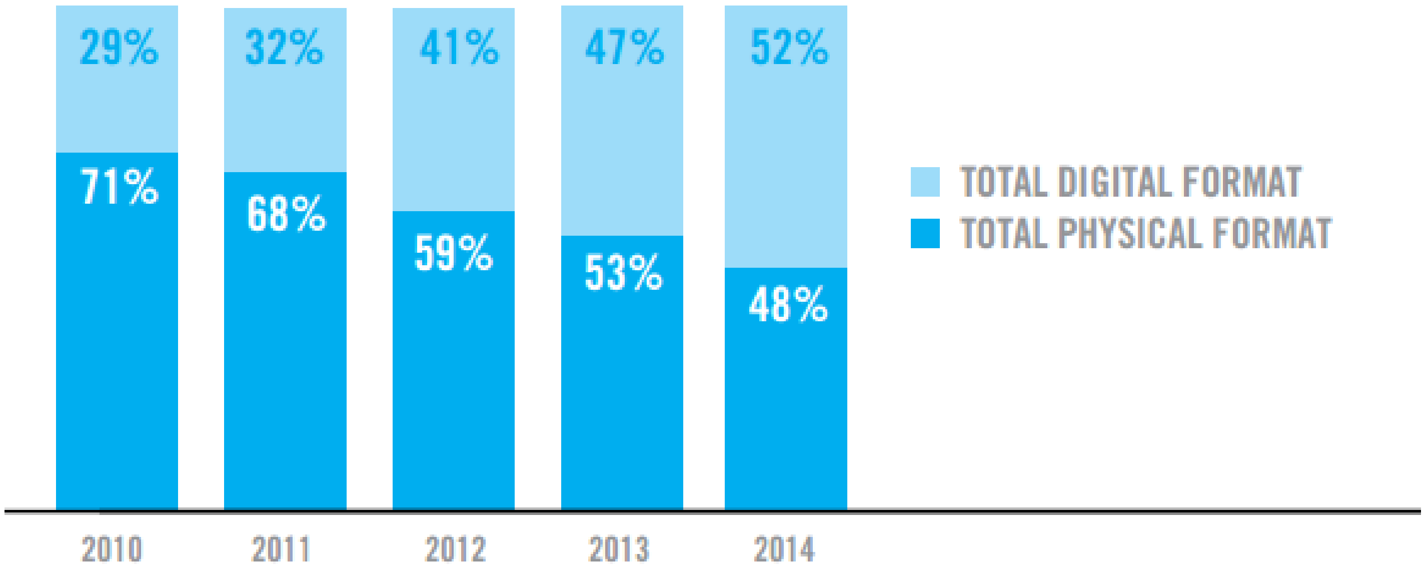


http://www.statista.com/chart/2713/video-games-are-getting-as-expensive-as-movies/

Challenges in the industry

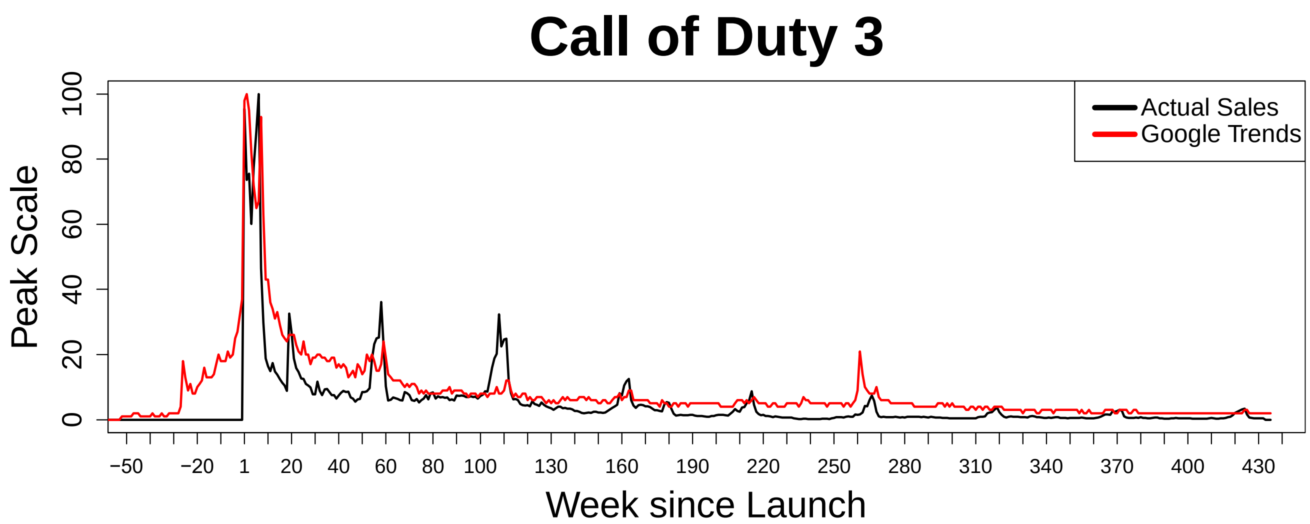
Shift towards the digital format

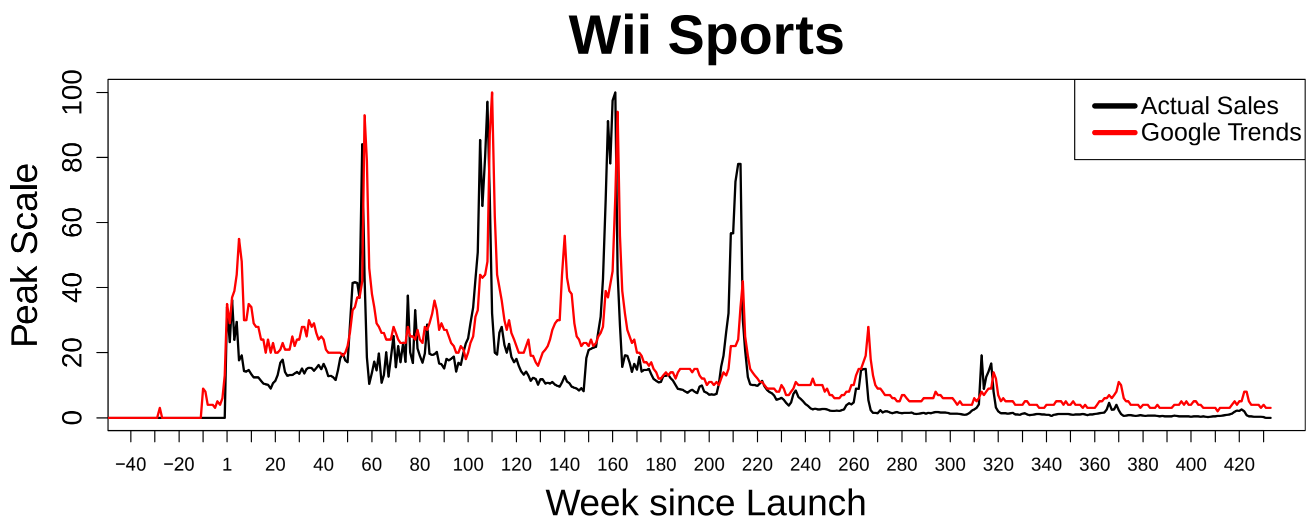
Recent Digital and Physical Sales Information



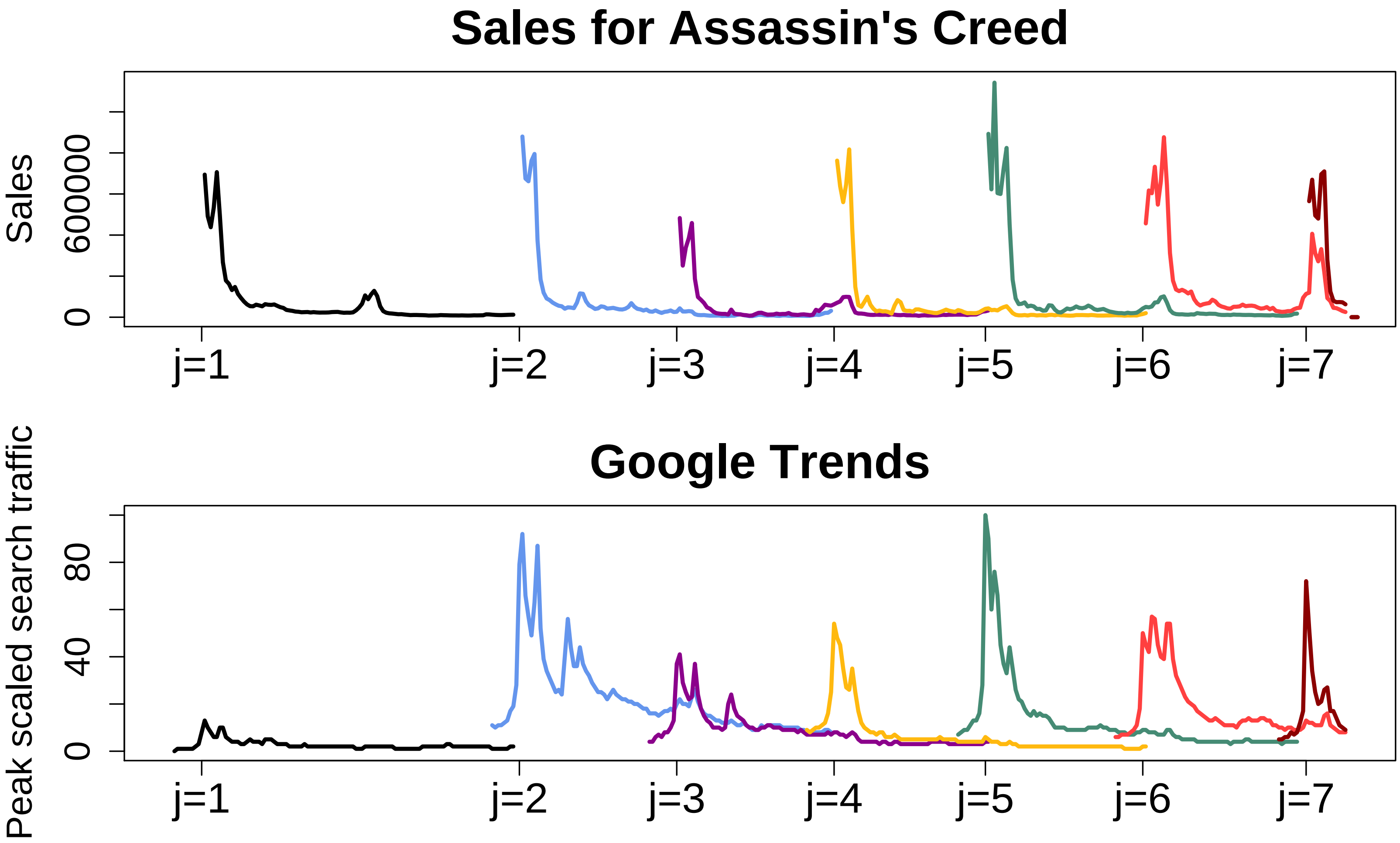
ESA, 2015. Facts, Sales, Demographic and Usage Data

The game life-cycle





Generations of games



Using information from online sources

Using online explanatory variables for pre-launch forecasts

- Forecasting computer game sales using search traffic and social buzz for the opening sales (Xiong & Bharadwaj 2014)
- Box office sales using social network data (Kim et al. 2015) and search traffic (Kulkarni et al. 2012)

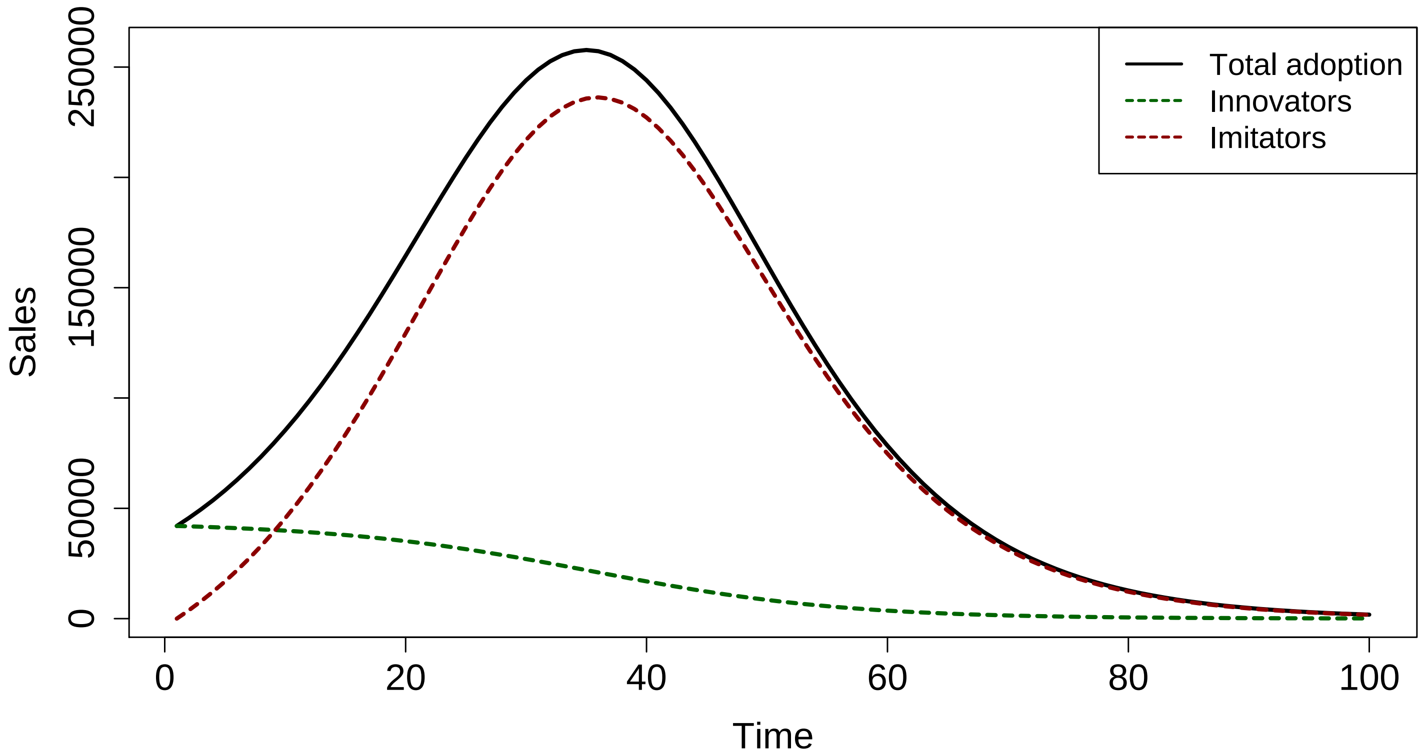
No application to life-cycle forecasting and parameter estimation for diffusion models.

I. Can search traffic data help in estimating diffusion model parameters?

What is a diffusion model

Measuring the adoption of a new product

- A common model in marketing is the Bass (1969) model
- Parameters for Innovators, Imitators and Market size



Ways to obtain model parameters in a pre-launch setting*

By judgement

- However, time consuming and problem with adjustment bias (Fildes et al. 2009)

Forecasting by analogy

- Using parameters from previous or similar products (Kim et al. 2014, Lillien et al. 2000, Norton & Bass 1987)

Market research

- Survey (Bass et al. 2001), Product attributes (Goodwin et al. 2012) or Pre-orders on CD albums (Moe & Fader 2002)

*See Goodwin et al. (2014) for a discussion on challenges with pre-launch forecasting

Experiment – Motivation

Model

- Bass model

Target

- Incorporate search traffic information into the analogy based forecast approach.

Aim

- Estimate market size parameter
- We are also interested in seeing whether there is lead time

The data

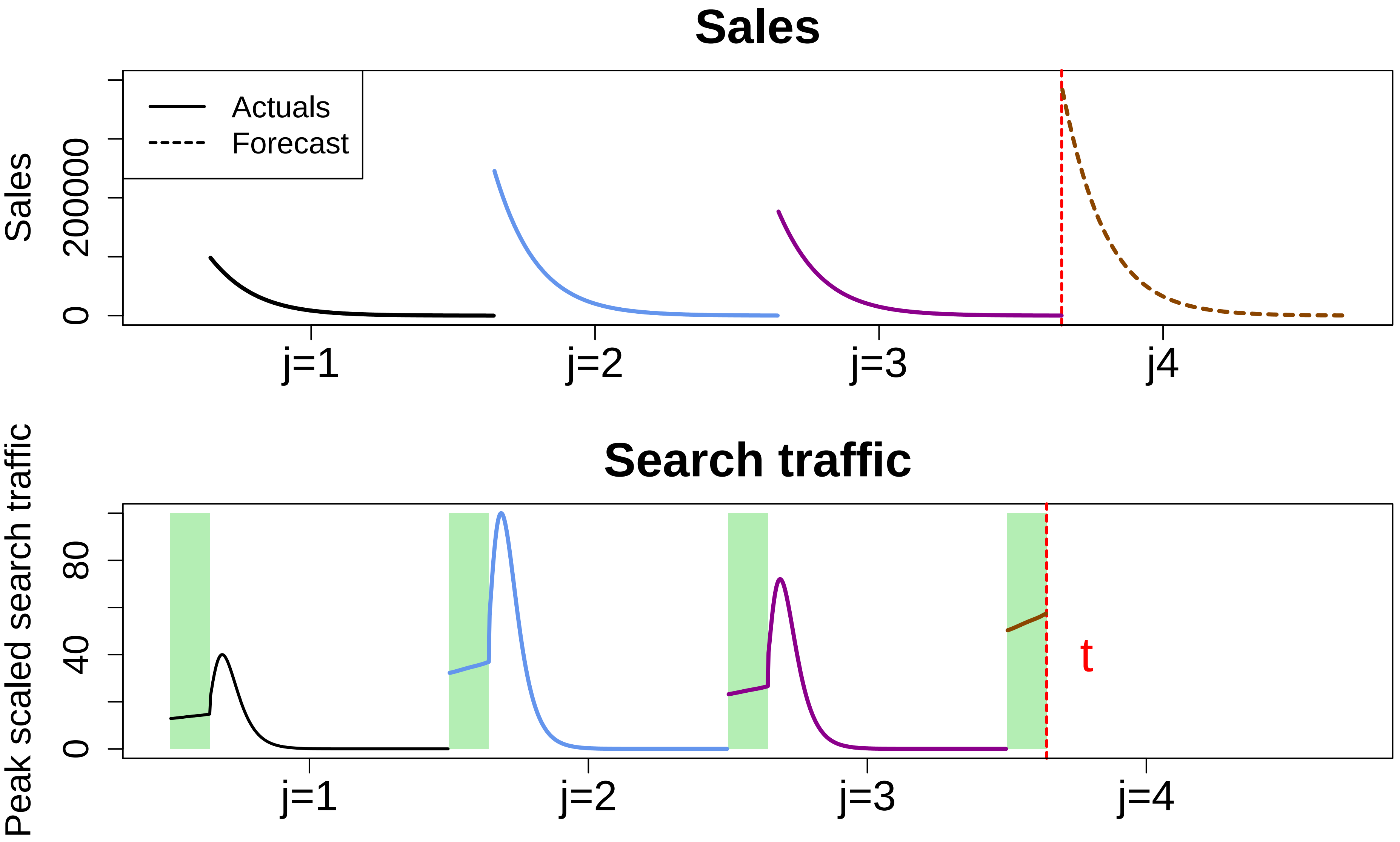
Video game sales from VGchartz

- Global physical sales at weekly frequency
- Using 6 games series with a total of 43 games
- Sales are aggregated across gaming platforms such as PC, Xbox, PS3 or Wii

Search Traffic popularity from Google Trends

- Weekly global search traffic popularity information
- Topic search with game title as keyword

Estimation process



Benchmark models and accuracy measure

Naïve: $m_j = m_{j-1}$

Naïve + Difference: $m_j = m_{j-1} + \Delta m_{j-1}$

Linear Trend: $m_j = \alpha_0 + \alpha_1 j + \epsilon_j$

AR(1): $m_j = \alpha_0 + \alpha_1 m_{j-1} + \epsilon_j$

+ “optimal” fitted Bass model with actuals

Actuals contain two years of data

Numbers of generations needed for model estimation vary

Relative Mean Absolute Error and median across series

Google Trend model selection

Percentage Increase: $$m_j = \frac{\mathrm{GT}_j}{\mathrm{GT}_{j-1}} * m_{j-1}$$

Linear trend: $$m_j = \alpha_0 + \mathrm{GT}_j + \epsilon_j$$

AR(1) + Percentage Increase: $$m_j = \alpha_0 + \alpha_1 m_{j-1} + \beta_1 \frac{\mathrm{GT}_j}{\mathrm{GT}_{j-1}} + \epsilon_j$$

AR(1) + Google Trend: $$m_j = \alpha_0 + \alpha_1 m_{j-1} + \beta_1 \mathrm{GT}_j + \epsilon_j$$

PI	Linear Trend	AR(1) PI	AR(1) GT
1.000	1.079	1.031	1.070

<1 = better

No. Series = 6, Window Size = 6, Lead Time = 1

Performance across series

Lead time 1 week

Naïve	Naïve Diff.	Linear Trend	AR(1)	Optimal
1.020	1.087	1.086	1.070	0.866

<1 = better

No. Series = 6, Window Size = 6

Lead time 6 weeks

Naïve	Naïve Diff.	Linear Trend	AR(1)	Optimal
1.007	1.143	1.088	1.033	0.861

<1 = better

No. Series = 6, Window Size = 6

Conclusion

Fully automated Bass model market size parameter estimation method that includes information from search traffic.

Google trend percentage increase market size estimation method outperformed most benchmark models.

St. Isaac's Cathedral, St. Petersburg

Thank you!

o.schaer@lancaster.ac.uk
@oliverschaer

References

Bass, F.M. (1969). A new product growth for model consumer durables. Management Science, 15, 215-227.

Fildes, R. Goodwin, P., Lawrence, M. and Nikolopoulos, K. (2009). Effective forecasting and judgmental adjustments: an empirical evaluation and strategies for improvement in supply-chain planning. International Journal of Forecasting, 25, 35-47.

Goodwin, P., Dyussekeneva, K. and Meeran, S. (2012). The use of analogies in forecasting the annual sales of new electronics products. IMA Journal of Management Mathematics, 1-16.

Goodwin, P., Meeran, S. and Dyussekeneva, K. (2014). The challenges of pre-launch forecasting of adoption time series for new durable products. International Journal of Forecasting, 30, 1082-1097.

Islam, T. and Meade, N. (1997). The Diffusion of Successive Generations of a Technology: A more General Model. Technological Forecasting and Social Change, 56, 48-60.

References

Jiang, Z. and Jain, D.C. (2012). A Generalized Norton-Bass model for Multigeneration Diffusion. Management Science, 58, 1887-1897.

Kim, T., Hong, J. and Kang, P. (2015). Box office forecasting using machine learning algorithms based on SNS data. International Journal of Forecasting, 31, 365-390

Kulkarni, G. Kannan, P.K. and Moe, W. (2012). Using online search data to forecast new product sales. Decision Support Systems, 52, 604-611.

Lee, H., Kim, S.G., Park, H-W. and Kang. P. (2014). Pre-launch new product demand forecasting using the Bass model: A statistical and machine learning-based approach. Technological Forecasting & Social Change, 86, 4—64.

Meade, N. and Islam, T. (2006). Modelling and forecasting the diffusion of innovation – A 25-year review. International Journal of Forecasting, 22, 519-545.

Norton, J.A. and Bass, F.M. (1987). A Diffusion Theory Model of Adoption and Substitution for Successive Generations of High Technology Products. Management Science, 33, 1069-1086.

Xiong, G. and Bharadwaj, S. (2014). Prerelease Buzz Evolution Patterns and New Product Performance. Marketing Science, 33, 401-421.