## Privacy respecting modeling

Prof. dr. Koen Pauwels Northeastern University BI Norwegian Business School



#### WWW. PHDCOMICS. COM

# Outline

- 1. How do you decide on method?
- 2. Research choices and market situations
- 3. Theory and Empirics
- 4. Cool research questions in Retail Media
- 5. Webrooming example with machine learning

## How do you decide on method ?

# How do you decide on method ?

- What I learned!
- What my advisor said

• What the editor likes

- What the data calls for
- What answers the research question



## Research questions

- Does smoking increase lung cancer chances?
- Is this browser a high or low value customer?
- Are price promotions profitable in long run?
- How will competitors react to my action?
- How much more sales if advertising\$ doubles?
- When can I expect this sales bump to occur?
- Which metrics are Key Performance Indicators
- How important was fake news in US election?
- Where do consumers adopt new products fast

## Privacy respecting modeling

"Econometric time series analysis combines the merits of econometrics, which focuses on the *relationship* between variables, with those of time series analysis, which specifies the dynamics in the model" Franses (1991, p. 240) Sims (1980): 'structural' models often have to place 'incredible' identifying restrictions on how key variables are allowed to behave and influence each other over time

#### Quantitative

Analytical Models Assumes optimality and rationality Method: Game Theory, mathematical proofs (no empirical data)

#### Within Marketing

Structural Models Assumes optimality and rationality Method: creates theoretical model and estimates model parameters based on data using Bayesian estimation Uses policy simulations

#### Market response model

(reduced form models)

Descriptive (statistical relation between observed variables, correlational) Method: Regression, etc. Examines counterfactuals

> **Field experiments, Natural, and Quasi experiments** Method: Experimental, econometric



## In the end, what matters is:

- What do you think ?
- What do your evaluators think ?

What is the purpose of your model/work ?Which model is feasible given obtainable data?How much insight would a more sophisticated model get, would a simpler model give up?

## Research choices matter

- Variables in levels, cumulative or differenced?
- Measured daily, yearly (temporal aggregation)
- For each store? (cross-sectional aggregation)
- Dependent variable: sales, distribution,...?
- Independent variables: actions, competition?
- What to control for: seasonality, trend, ...?
- Include lagged variables for dynamic effects?

## Ex: Return on Marketing Investment



### Key modeling challenge: Dynamic interactions of many variables with unknown feedback loops plus noise.



#### **Return on Marketing Investment scenarios**

	One-shot Marketing	Evolving Marketing (spending more)
Stationary Performance	Business as usual e.g. promotion	Escalation <i>e.g. car rebates in</i> <i>the US</i>
Evolving Performance	Hysteresis e.g. ads induce social postings	Evolving business e.g. <i>AFLAC duck</i>

# Example:



Marketing	Pre-duck	Post-duck
Success !		
Brand awareness	2% (1999)	80 % (2001) 85 % (2006)
Company Revenue	\$ 555 M (1999)	\$ 712 M (2000) \$ 919 M (2001)

## Financial success: RoDuckl ?

- Observable duck lift (full attribution)
  Awareness increase: 3900 %
  Revenue increase: 28 66%
- Very expensive: \$450M (2000-2006)
- Need to continue spending at high level
- Aflac CMO (2007): Shut the duck up !

## Business as usual: Price promotion



## But new brand (on right) adds to base

## Close-up toothpaste Rembrandt (1990s)



### **Car Rebates Decrease GM's Firm Value**



## As They lead to Competitive Escalation



### How to grow Return on Marketing Investment

	One-shot Marketing	Evolving Marketing (spending more)
Stationary Performance	Optimize for profit Look beyond usual for game changer	Allocate \$ to other marketing actions, brands or regions
Evolving Performance	Leverage this now! Learn to repeat Prepare for growth	Profits in long run? Ensure marketing keeps us with sales

## Theory without data is blind



## In marketing

Theory and findings mostly used for

1) Description

2) Prediction

3) Control, i.e. (advice for) policy intervention

## Philip-Hans Franses 2005

- Policy Simulation = "tracking the effects of various doable changes in marketing actions and picking the most preferable one"
- Models should be:
- 1) Descriptive: accurately capture DG patterns
- 2) Predictive: hold-out sample, constant DGP
- 3) Prescriptive: stability and market structure

# Retail Media Back to the Future

Koen Pauwels

Associate Dean of Research, D'Amore-McKim School of Business Distinguished Professor of Marketing, Northeastern University Former Principal Research Scientist at Amazon Ads

# Outline

- 1. How large and how different from offline retail ads?
- 2. Why brands and consumers like retail media
- 3. Lower and upper funnel retail ads for advertiser goals
- 4. How to get started and move to full funnel advertising
- 5. Retail media questions for consumers, brands, and retailers
- 6. Case studies answering some of these questions
- 7. Q&A



#### Retail Media is the digital advertising sold to brands by retailers using their 1st party data as for targeting onsite, instore or on social or programmatic channels

99

Rise of Retail Media: WARC, Sept 2021

#### Top online stores in Turkey in 2022, by e-commerce

(in million U.S. dollars)





## Retail media on the rise

GroupM (biggest media buyer), expects retail media will grow 9.9% to reach \$125.7 billion in 2023 and surpass television revenue in 2028, when it will represent 15.4% of total ad revenue.



# To the delight and \$\$ of retailers

"An efficient, well-run grocery store will *lose* money on the business,

but more than make up for it by selling ads to the companies that make those products

Kroger reported vendor allowances of \$8.7bn, or 7% of revenue"

(Hobart 2023)

#### Global Retailer Operating / EBITDA Margins

amazon	7.1%	Sainsbury's	2.74%
Walmart >	2.4%		4.4%
Carrefour	3.11%	Kroger	3.11%
Migros	3.45%	Ahold Delhaize	4.3%
TESCO	3.23%		2.7%
Woolworths 👩	2.54%	EDEKA	

# What's new, pussycat?

- Data: "Instacart knows consumers' ordering patterns, knows the cross elasticities of demand across different products and categories, and can capture the upside from consumers trading up or trying new brands via ads." This allows online retailers to both offer consumers large and personalized assortments, and offer brands ads targeted at the right consumer in the right buying mood.
- **Consumer ad awareness:** The offline consumer is typically unaware that the store organization and eye-level shelf placements are driven by such manufacturer payments and retail profit calculations. In contrast, *retail media ads are denoted as such to the online consumer*.

## Why brands want it: data is key

#### Motivations for US CPG Brands to Work with Retail Media Networks, 2020 & 2021

% of respondents in each group

	2020	2021
Gain access to a retailer's first-party data	33%	62%
Build a stronger partnership with the retailer	49%	51%
Better understand the shopper at the retailer	19%	51%
Leverage owned and operate retailer inventory	29%	49%
Open co-branding opportunities	38%	37%
Deliver more personalized communications	43%	34%
Access reporting tied to sales (closed loop)	48%	28%
Oblige merchants as part of negotiations	48%	14%
Note: 2020 n=100: 2021 n=100		

Source: Merkle, "The Evolution of Retail Media Networks" conducted by Ugam, Oct 11, 2021



eMarketer | InsiderIntelligence.com

270408

# Why do some consumers like it?

- Content-integrated: fits with the reason consumer visits the site
- 3x more effective than contentseparated (DeHaan et al 2016)
- 1P and 1-click shop convenience

-> More effective than video site, search, social media ads for most

Preference	Global consumers	Marketers	
1.	amazon	🕨 YouTube	
2.	Google	Google	
3.	<b>t</b> ikTok	0	
4.	Ø	<b>J</b> TikTok	
5.	Spotify <sup>.</sup>	Spotify <sup>.</sup>	

## Largest players

Amazon: ~75%

Walmart: ~7%

Instagram: ~2%

Amazon's Advertising business is bigger than Google Display and YouTube Advertising and Global Newspaper Advertising Revenues

Global advertising revenues



## Comparing different retail media networks

	Amazon.com	Walmart.com	Instacart	Kroger	Target
Advertising auction type	2nd price	1st price	2nd price	2nd price	1st price
Placement types	Sponsored Products, Sponsored Brands, Sponsored Display, Amazon DSP (Offsite and Onsite)	Sponsored Products, Brand Amplifier, POV, Marquee, Skyline	Sponsored Products (Search & Browse), Search Keyword Banner (Display)	PLAs (Search & Browse, Savings, Basket Builder) and Targeted Onsite Ads (Display Banner)	Sponsored Products
Cost structure	CPC by product and keyword Amazon DSP is CPM	CPC by product and keyword for all except POV, Marquee, Skyline that are CPM	CPC by product and keyword; Display is CPM	CPC by product and category for everything but TOAs; TOAs are CPM	CPC by SKU
Minimum bid (\$)	0.02c	0.2 automatic 0.3 manual	0.15 CPC for SP, 15 CPM for display	0.5 for S&B, 0.6 for BB,0.30 for Savings, \$28 CPM for TOA	Set by retailer at category level
Targeting options	Keyword, Category, and Individual SKU Targeting	Keyword Targeting	SP: Keyword Targeting SKB: Keyword and Audience	Category Targeting	Keyword and Auto Targeting
Product eligibility	Must be in stock, no minimum, must have positive contribution profit to Amazon	Must be in stock and base item (no variants), no minimum, item must meet relevancy hurdles to appear in-feed	Must be in stock, no minimum, must be loaded up by retailer	Must be in stock, no minimum	Must be in stock, no minimum
Sales attribution	14-day click attribution, same brand	3, 14, or 30-day click attribution, same brand	14-day click attribution, same item, "sales" are based on the order not the delivery	14-day click attribution, same brand, in store and online	Customizable attribution 1, 7, 14, 30-day direct click, 1, 7, 14, 30-day direct view

#### Amazon Advertising's view on its Propositions


### **Sponsored Products**



Placements within Amazon search results and on related detail pages

- Created from product detail page content
- Direct shoppers directly to a product detail page
- Contain native-looking creative elements familiar to
   Amazon
- Reach shoppers with high purchase intent

amazonadvertising

### **Sponsored Brands**

Top placement above the search results, as well as within search results

- Direct shoppers to a page with a collection of products or the brand's Amazon store
- Feature a custom headline, logo and set of products
- Contain native-looking elements familiar to the Amazon shopping experience
- Reach shoppers with high purchase intent



Bai Brasilia Blueberry

Antioxidant Infused

\*\*\*\*\*\*\*\*\*\* 3,025

\$1878 (\$0.09/FLOZ)

More options available

FREE Shipping on elig

\$19.77 vprime (4-5 days)

Beverage...

ine by

cribe & Save

zon Prime

Subscribe & Save Flight

ible for Free Shipping

Free Shipping by Amazo

Bai Cocofusions Variety

\*\*\*\*\*\*\*\* \* 1,618

ave 20% with coupo

\$2099 (\$0.10/FI Oz)

FREE Shipping on eligible

everages.

Pack Antioxidant Infused

amazonadvertising

Pack\_

by Izze Sparkling Juice

\*\*\*\*\*\* \* 2,855

\$1599 (\$0.08/FI Oz) vprime

Sponsored @

Infused...

Bai Bubbles Variety Pack,

\$2003 (\$0.15/FLOz) \$22.99

**Sparkling Antioxidant** 

\*\*\*\*\*\*\* \* 1,498

Save 20% with coupor

More options avai

\$21.08 prime

### **Display ads**

Convey key product attributes or brand differentiators

- Direct to a product detail page, brand website or custom landing page
- Tailor ad creatives
- Use eCommerce creatives with Amazon features
  - for products sold ON Amazon
- Use custom creatives



amazonadvertising

### Offline to Online



### Use the Shopper Funnel



#### Start with Sponsored Search, Move to Sponsored Display



### Think 'Full Funnel'



### **Go Full Funnel**



## Key retail media questions for players

Consumers	Brands	Retailers
<ul> <li>Ad response</li> <li>Optimal exposure</li></ul>	<ul> <li>Retailer fit?</li> <li>Incremental or zero-</li></ul>	<ul> <li>Education vs fatigue</li> <li>Build in house or</li></ul>
frequency <li>Tradeoff privacy,</li>	sum game? Metrics? <li>Budget allocation</li>	outsource? <li>How I fit into the</li>
convenience, price	upper/lower funnel?	brand's ecosystem?

# Higher or lower optimal ad frequency?

- Compared to other settings, are retail consumers more welcoming to ads (content integrated) or do they get annoyed sooner because they are paying more attention (leaning forward vs leaning backward)?
- Inform/Educate: Higher ad frequency may allow brands to increase conversion probabilities, and retailers to enjoy more ad revenue.
- Persuasion fatigue: ads are seen more as manipulative persuasion rather than as information, and should be employed sparingly to balance customer experience with ad revenue for the retailer

# Lower funnel or upper funnel?

- Upper funnel costs more but also has higher click-through rate
- Lower funnel has higher conversion rate: ready-to-buy
- Return on Ad Spend (\$ revenue for \$1 spend) similar: \$6 \$7
- Upper funnel works especially well in high season (Q4 in USA)

Source: Costello 2022 https://skai.io/blog/amazons-advertising-numbers/



# How retail media fits in brand's system

### Making the e-commerce business model more scalable

"Retail Media is a way of making the e-commerce business model more scalable. Let's not forget 80% of all of our sales come from the bricks and mortar. So omnichannel is still going to be where this is won and lost.

"Being able to take rich first-party data, connect it to in-store shoppers, and then look at your technology and advertising in-store to look at how you have different messages in different stores on different days of the week and different times of day, based on the weather."

Alex Crowe, Head of Media Partnerships, Asda UK

#### Revlon's 'omni-ecosystem'



### Expand scope of Amazon Ads research

SALES	Advertising	On-Amazon	Off-Amazon
On-Amazon		Ads halve time for new product to reach search, sales, reviews milestones (Bertozzi et al. 2022)	Needs data from brand managers (clean rooms, Amazon Marketing Cloud, Nielsen, Kantar)
Off-Amazon	I	Webrooming (Zenetti)	Outside vendors

# Quantifying webrooming: Data and Method

**Off-Amazon Amplification:** Customers who research products on Amazon but buy them elsewhere (US Survey in 2/2020)

Analytical approach to compute Amplier Ratio and ROPES

- Survey of Amazon visitors in categories of laptops and smartphones
- Use ML model to predict off-Amazon purchases with Amazon action
- Amplifier Ratio (AR)= webrooming customers / on-Amazon buyers
- ROPES = webrooming customers / on-Amazon category searchers
- Regress ROPES on different Amazon ads to quantify drivers



For every 10 Amazon buyers of Laptops (Smartphones), 16 (21) users research on Amazon.com and purchase elsewhere

# 

### ROPES: Share of off-Amazon buyers to on-Amazon browsers



Total survey sample: 42K Wave 2 (H2) in 2021 Example: The median share of off-Amazon buyers was ~4% vs. on-Amazon buyers in Doll Toys category, with lower quartile ~3%, upper quartile ~5%, min 2% and max 6% for all brands in this category covered by the survey responders

\*\_H2 notation at the end of a brand name indicates that this is the observation from the second wave of surveys

Source: Amazon internal data, Jul-Oct 2021, USA. 50,000 survey responders

### Machine Learning model to predict ROPES

We train a Random Forest model to predict ROPES across categories - based on observed customer and browsing behavior variables related to:

- (1) Customer journey on Amazon in focus product category,
- (2) Past purchase behavior,
- (3) Demographic information.

RF model as good as Gradient Boosting and Deep Learning Neural Nets

- (a) Achieves high level of accuracy of less than 1% error,
- (b) Accurately predicts proportion of off-Amazon buyers in each activity group,\*
- (c) Accurately predicts ROPES across categories, time periods and brands (92.1% no significant difference, average Pearson correlation = 0.83
- (d) Feature importance confirmed by 2SLS Instrumental Variables (Endogeneity)

# Customers who research are more likely to buy on Amazon but also off-Amazon



Total percentage of buyers on Amazon, off-Amazon and non-buyers in different activity groups

Clicks while on product page







■ off-Amazon ■ none ■ on-Amazon

# Upper funnel ads, especially always-on, show most Off-Amazon sales



Source: Amazon internal data, Jul-Oct 2021, USA, 50,000 survey responders. (95% confidence interval - normal approximation)

### Thank You!





# CEO, Dutch SME company Inofec

**"Analytical marketing** is not very common in small

and medium size enterprises in the business-tobusiness sector.

As such, if we had a model or decision support system to enable us to **decide how to allocate resources across channels**, we will have a **huge advantage** compared to our competitors."



### Profit Response: Adwords



Days

### Profit Response: Faxes



Days

### Profit Response: Flyers



Days

### Results

Variable	Profit Effect	Sales Elasticity	Wear-In	Wear-Out	90% Effect
Fax (€)	3.33	0.05	0	6	4
Flyers (€)	0.57	0.04	2	9	5
Adwords (€)	55.72	4.35	1	9	7
eMail (each)	0.71	0.12	2	5	5
Discount (1%)	789	0.75	0	2	1

### Dashboard input and output

### Marketing Dashboard Prototype

Period	Flyers	Faxes	Adwords	Discount
1	0	400	100	7
2	0	0	50	7
3	0	0	50	7
4	4000	0	100	7
5	0	0	100	7
6	0	0	100	7
7	0	400	100	7
8	0	0	100	7
9	0	0	50	7
10	0	0	50	7
11	4000	0	100	7
12	0	0	100	7
13	0	0	100	7
14	0	0	100	7
	80	80	80	80



Total Profits: \$333,068.88

### Comparing profit from saved scenarios

 $\hfile:///C:/Documents$  and S...  $\,\,\times\,\,$ 

С

☆ file:///C:/Documents%20and%20Settings/koenp/My%20Documents/dashboard/wiesel/inofecdashboard/dashjs-v0.0.6a/index.html?load-saved

**Marketing Dashboard Prototype** Profits Graph (Inofec Model) **Saved Plans** З default. double Adwords cut Flyers in half Back ß

**P**-

### Recommendations to the Company

- > Increase spending on faxes
- Decrease spending on flyers

Increase spending on Adwords

*Increase* use of eMail campaigns

### Field Experiment – Design



### Field Experiment – Design



### Field Experiment – Net Profit Changes



### **Experiment Supports Our Assertions**

### AdWords

- Increase raises profits
- -Most of the profit increase shows up in offline orders
- -Higher spending of 17.66 euros per day is more than offset by the higher profits, in either 'Flyer' condition
- Flyers
  - Average daily savings of 137 euros a day are substantially higher than the profits lost from lower sales
  - -Do not panic and invest in flyers if there is budget left

# **Collaboration Phases**



- Define managerial problem– duration: approx. 3 months
- 2
- Leveraging data from existing systems duration: approx. 20 months
- 3
- Employing marketing science approaches duration: approx. 3 months

- 4 Discussing results, designing strategies, and deduct field experiment
  - duration: approx. 7 months and ongoing



#### Ongoing training

- duration: throughout and ongoing
# Impact Statement (1/2)

- Cultural Impact
  - Resources allocation & marketing analytics
  - Substantial discussions (affecting mental model)
  - marketing analytics "department" (institutionalization of approach)
- Methodological Impact
  - Applications of marketing science
  - Internal process (existing information and which other information to collect)

# Impact Statement (2/2)

- > Strategic Impact
  - Determine where generating or losing money and plan appropriate strategic activities.
  - Improved understanding of the role of marketing
  - Strategic change in their focus of activities

# **Generating Ideas from Data**

### Koen Pauwels (Northeastern University) **Rick Staelin (Duke University)**



AMERICAN MARKETING ASSOCIATION

Presented to the JM-AMA Research Workshop February 13-14, 2020, San Diego, CA







### Alternative Research Process: "ET": Empirics $\rightarrow$ Theory New Write New Uncover Propose Marketing theory data patterns paper Questions Owned social media Should brand managers • Own fans defend Battle of the shows negative Ilhan, Kubler delete these posts? brand + cross-post comments by Brand Fans & Pauwels JIM 2018 Does this harm or • 'rally the troops' competing brand fans benefit these brands? with Benefits benefits the brands

## Unique data: Some examples

New types of data

- From Free to Fee → How to market to get customers to pay for services they used to get for free (Pauwels & Weiss, 2008, JM)
- 36 Dutch brands in 15 categories: marketing, sales, brand tracking, online behavior (Pauwels & Van Ewijk, MSI Best Paper, 2013)

Proprieta ry data Pre-Launch Forecasting New Cars (Roberts & Urban 1988, Mgt Sc)

- Pre-experience perceptions, evaluation and choice
- Exposure to word of mouth, test drive, and consumer reports
- Post-experience perceptions, evaluation and choice and updating



## More examples:

New combinations of data	<ul> <li>Facebook, Twitter, YouGov, Press Releases, Media coverage, Firm Performance → How do brand crises develop (target: JM)</li> <li>Retail prices, Consumer purchases, Media Coverage, Stock market → How media coverage affects price wars (Van Heerde, Gijsbrechts and Pauwels, JMR 2015)</li> </ul>
Longitudinal data	<ul> <li>20 years of donation data → variety (Khodakarami et al. 2015, <i>JM</i>)</li> <li>7 years of scanner data → long-term effects 4Ps (<i>JM(R), MKSC</i>)</li> <li>Field experiments over several months (<i>Practice Prize</i> Winner)</li> </ul>



## **Granger Causality**



- Sir Clive Granger: 2004 Nobel Prize
- For a variable X to cause Y in time, changes to X must precede changes to Y
- Forecast of Y based on past of Y + X is better significantly than based on past of Y only
- Limitation: no same-period "causality", so we also consider strong same-period correlations

### For instance, EB Games came up with these metrics:

### Business Objectives

- Quantity of incremental transactions or customers
- Incremental Customer Value (profit per sale)
- Repeat sales rate (reflection on customer experience)
- Defection rates by value segment & tenure

### Marketing's Leading Indicators

- -Consideration
- -Competitive preference
- -Inquiry
- —Trial

### -Initial sales value

- Diagnostics (do not lead business outcomes)
  - Response rates
  - -Cost per contact
  - -Cost per sale
  - -Brand Awareness
  - -Consideration & Preference
  - -Web visits
  - -Downloads
- Long-Term Indicators
  - -Competitive Position on Key Attributes (Brand)
  - -Dissatisfaction Rates (Customer Experience)
  - -Customer equity (future sales potential)

## **Vector AutoRegression MODELS**

<u>Idea:</u> econometric + time-series to describe complex, multiple-equation, dynamic interrelations

$$\begin{bmatrix} S_{t} \\ M_{t} \\ CM_{t} \end{bmatrix} = \begin{bmatrix} \gamma_{11}^{1} & \gamma_{12}^{1} & \gamma_{13}^{1} \\ \gamma_{21}^{1} & \gamma_{22}^{1} & \gamma_{23}^{1} \\ \gamma_{31}^{1} & \gamma_{32}^{1} & \gamma_{33}^{1} \end{bmatrix} \begin{bmatrix} S_{t-1} \\ M_{t-1} \\ CM_{t-1} \end{bmatrix} + \dots + \begin{bmatrix} \gamma_{11}^{I} & \gamma_{12}^{I} & \gamma_{13}^{I} \\ \gamma_{21}^{I} & \gamma_{22}^{I} & \gamma_{23}^{I} \\ \gamma_{31}^{I} & \gamma_{32}^{I} & \gamma_{33}^{I} \end{bmatrix} \begin{bmatrix} S_{t-1} \\ M_{t-1} \\ CM_{t-1} \end{bmatrix} + \begin{bmatrix} u_{S,t} \\ u_{M,t} \\ u_{CM,t} \end{bmatrix}$$

### Order determination:

- Multivariate equivalent of ACF and PACF
- Information Criterion balances power & parsimony

### **Return on Marketing Investment scenarios**

	Stationary Marketing	Evolving Marketing
Stationary performance	Business as usual <i>e.g. beer1</i> <b>VAR</b> in levels	Escalation <i>e.g. car rebates</i> <b>VAR</b> in changes
Evolving Performance	Hysteresis <i>e.g.</i> beer brand X <b>VAR</b> in changes	Evolving business <i>e.g. advertising</i> <b>VAR</b> in changes