Analyzing Multichannel Operations

Anuj Kumar
Hough Faculty Fellow and Assistant Professor
Information Systems Mgt.
337 STZ, Warrington College of Business, Univ. of Florida
Email – akumar1@ufl.edu
Website – https://site.warrington.ufl.edu/kumar/

Education

PhD in Information Systems Management — Carnegie Mellon University, US, May 2011
Master's in Management — IIM Bangalore, India, May 2006

Research Interest

Technology mediated multichannel operations
Online product recommendation networks
Societal impact of IT

Multichannel Sales Operations

- Rapid growth in new technology enabled channels, such as telephone-, web-, and mobile-based channels, through which firm can interact with their customers.

- Opportunities
  - Engage with customers via a variety of channels over different stages of their interactions – more options to advertise and sell products and services.

- Challenges
  - Managing sales and distribution operations on multiple channels is complex – could result in channel conflicts

How do Stores Drive Online Sales? The Less Obvious Effects of Stores on Revenues of a Multi-channel Retailer

Anuj Kumar	Amit Mehra	Subodha Kumar
Univ. of Florida	UT Dallas	Temple Univ.

Motivation

- Retailers increasingly relying on online channel to deliver sales growth and closing down their stores
  - Examples - Gap, Walmart, Nordstrom

- What is the value of physical store channel for retailers?

- Availability of a store at a location should cannibalize online sales there (Forman et al 2009)

- But there are evidence of complementary effect of store openings on the online sales of a multichannel retailer (Bell et al. 2015, Avery et al. 2012, Tang et al. 2016)

Why Complementary Effect?

- To find the effect of stores on online sales, we need variations in availability of stores.
  - Store openings by a multichannel retailer provides variation in store availability
- To find the underlying cause for the effect of store, customer-level transactional data needed
  - Comparison of pre- and post-store opening purchase behavior of a customer can reveal the underlying cause
- Understanding why/how of the complementary effect can help managers design appropriate strategies to optimize the benefit of stores

Research Objective

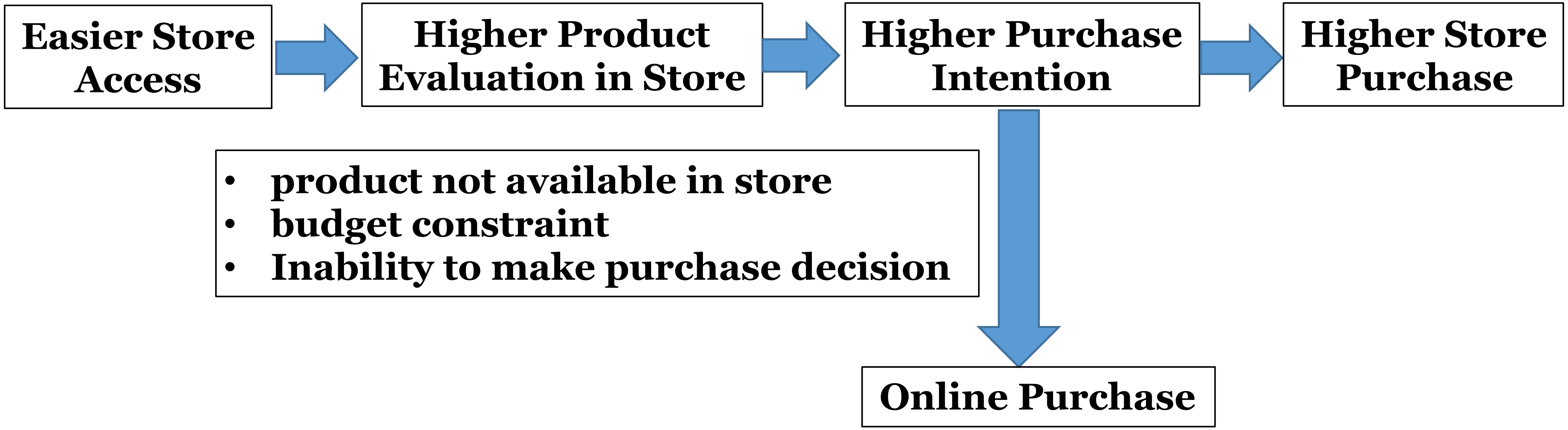
- We collect <u>customer household-level data</u> of a multichannel retailer selling products <u>rich in non-digital attributes</u> to examine

1. **Estimate the causal effect of retailer's store openings on the online purchase behavior of its <u>existing customers</u>.**

1. **Identify the underlying causes for the complementary effect and provide empirical evidence for them.**

Store Engagement Effect

- Stores suited for physical evaluation of non-digital product attributes (Lal and Sarvary 1999)



Association of higher online purchases with higher store interactions → Store Engagement Effect

Store Return Effect

- Higher risk in online purchase of products rich in non-digital attributes
- Easier return policies (return options) results in increased purchase (Wood 2001, Anderson et al. 2009)

Easier Store Access → **Low cost option of store returns** → **Lower risk in online purchase** → **Higher online purchase**

Association of higher online purchases with higher return of online purchases in store → Store Return Effect

Field Setup

- Field study on a large apparel and home goods retailer in the US
- The retailer opened over 30 stores between 1999-2006 and had over 150 stores in 28 states in the US.
- Data on purchase / return transactions on online and store channels by over 1.5 million (10 % random sample) customers from 1999 to 2006.
- Out of 656,949 existing customers as on Jan 2001, distance from the nearest store for 90,326 customers (affected customers) changed due to store openings but not for the remaining 566,623 customers (unaffected customers).

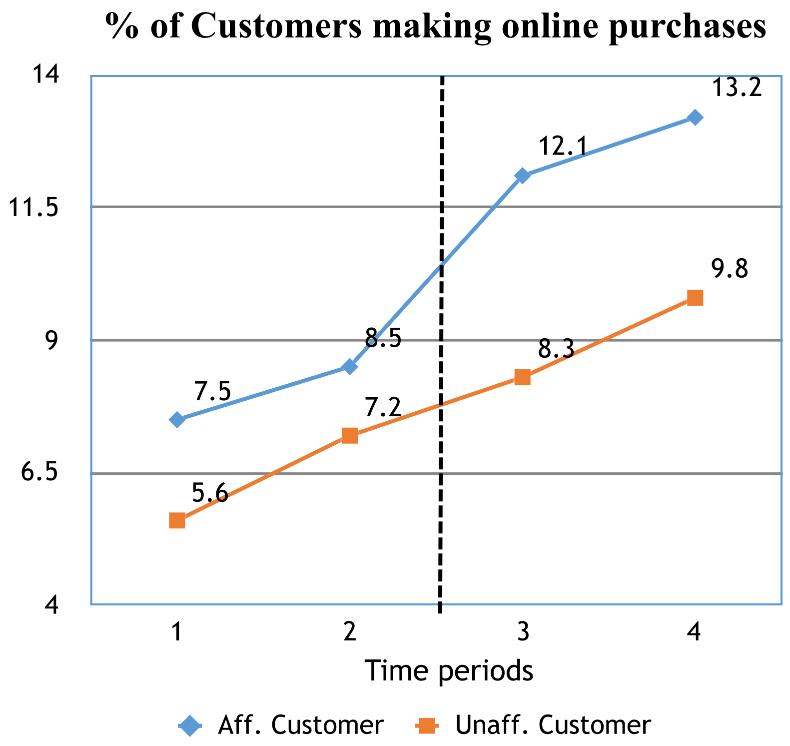
Sample Selection

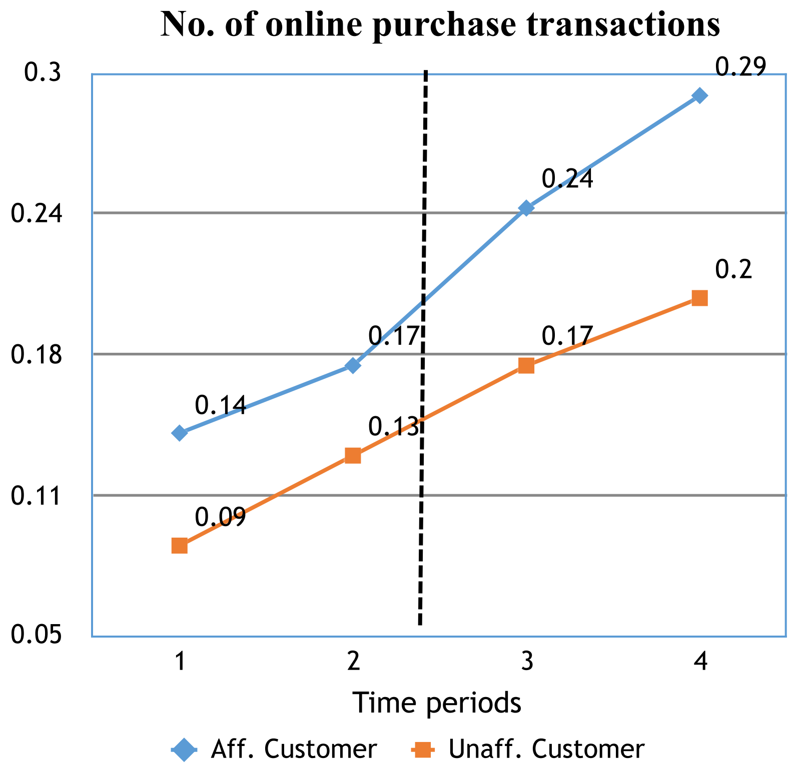
- Select customers
  - whose store distance changed due to 2003 store openings
  - whose tax liability on online sales did not change due to store opening (Anderson et al. 2010)

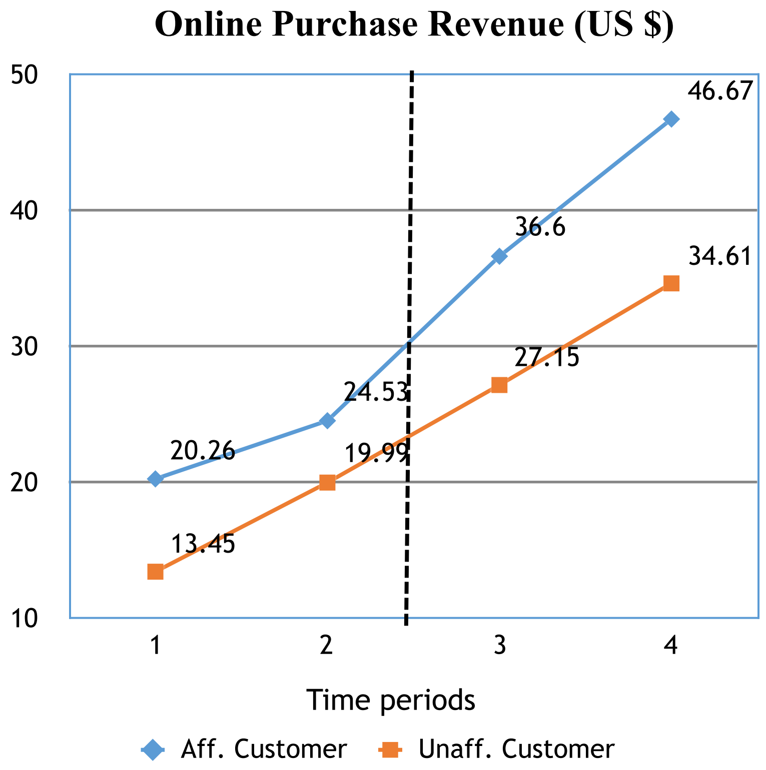
Type of Customers	Store distance in miles	Mean	Std. Dev.	Percentile Values				
				0	25	50	75	100
17277 affected customers	before store opening	148.0	127.4	0.9	22.9	154.1	234.9	473.0
	after store opening	78.0	97.9	0.0	7.9	22.5	144.1	412.0
	Change in store distance	70.0	78.9	0.0	4.7	33.9	114.7	233.5
201096 Unaffected customers	before/after store opening	53.0	73.8	0	6.1	24.9	68.2	794.5

Model Free Evidence

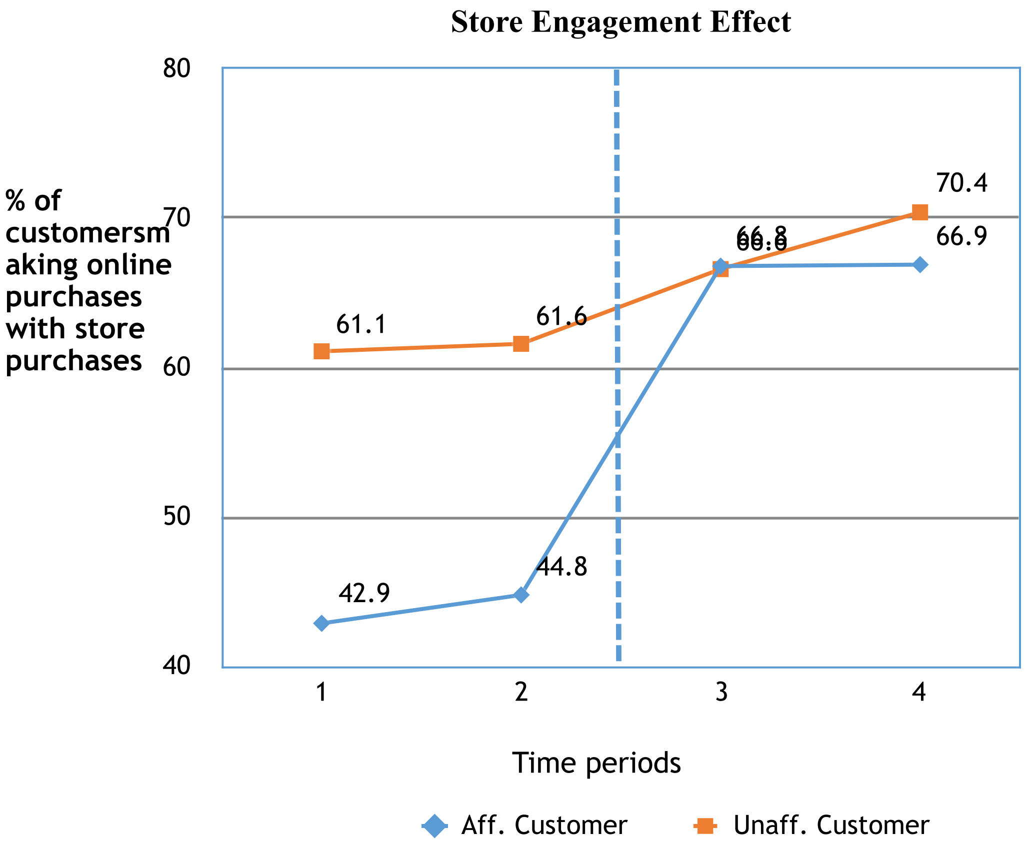
Complementary Effect on Online Purchases

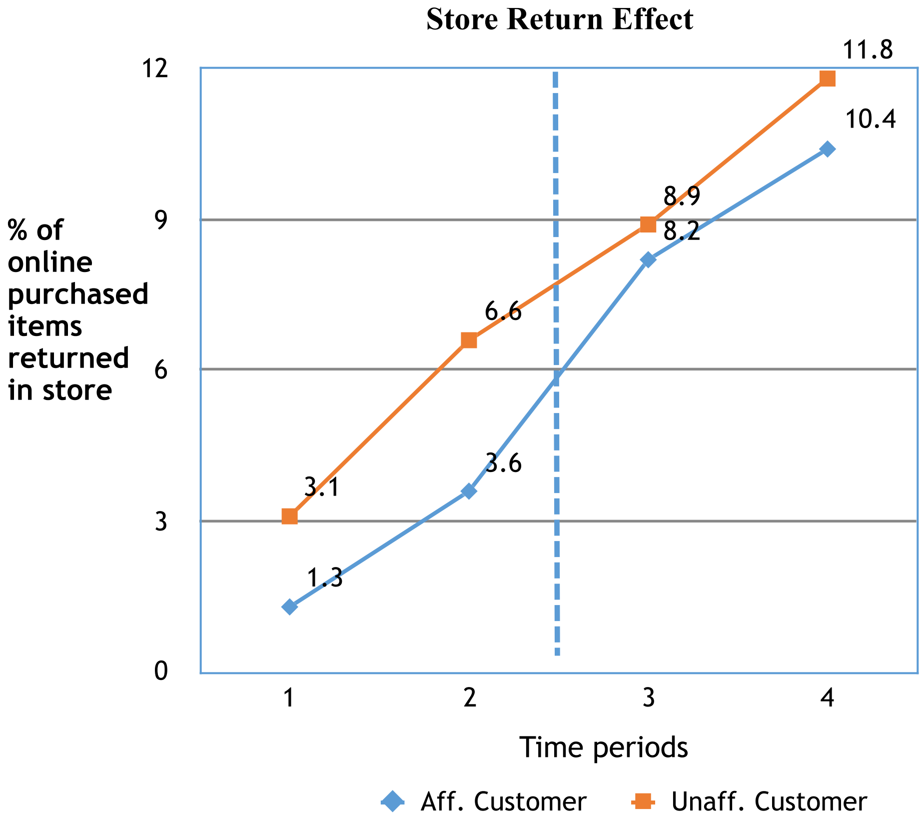






Store Engagement and Return Effects





Estimation Models

Treatment Effect of Store

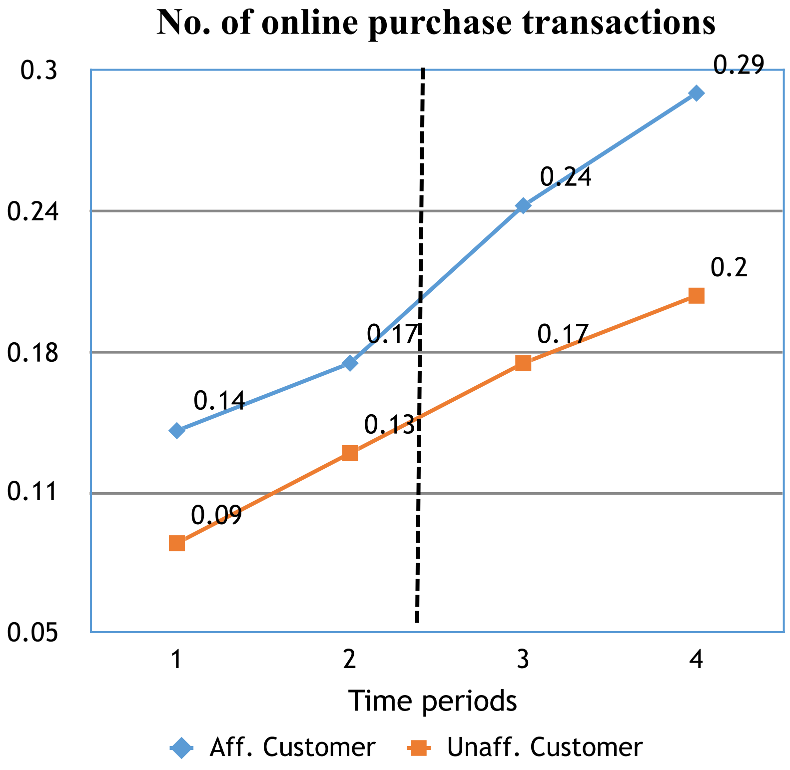
- Treatment effect of store openings on online sales = Online sales with store opening – Counterfactual online sales without store opening

- But sales without store opening is not observed in the area of store opening, so it is inferred from the sales in other areas with no store opening

  Treatment effect = Online sales where store opens – Online sales where no store opens

- As retailers select area to open store based on its socio-economic-demographic characteristics, purchase behavior of customers in area of store opening may be different from area with no stores
  - Statistical methods needed to infer the correct counterfactual online sales

Diff-in-Diff Design



Store opening

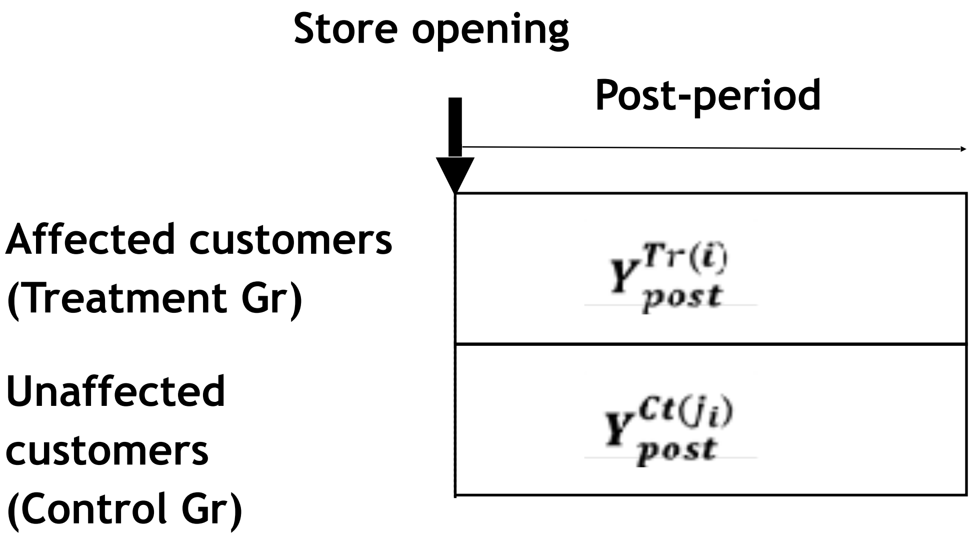
Prior-period | Post-period

	Prior-period	Post-period
Affected customers (Treatment Gr)	Y_{prior}^{Tr}	Y_{post}^{Tr}
Unaffected customers (Control Gr)	Y_{prior}^{Ct}	Y_{post}^{Ct}

$$ATT = \left(Y_{post}^{Tr} - Y_{prior}^{Tr}\right) - \left(Y_{post}^{Ct} - Y_{prior}^{Ct}\right)$$

- Taking the difference of change in purchase behavior from prior- to post-period between treated and control customers accounts for other factors that could affect the purchase behavior of the two groups of customers.

Matching Estimator



- The counterfactual purchase behavior of an affected customer (i) is inferred from the behavior of a sample of unaffected customers (j_i) having similar purchase behavior in the pre-treatment period.

$$ATT = \frac{1}{N_i}\sum_{i}\left[\left(Y_{post}^{Tr(i)}\right) - \frac{1}{N_{j_i}}\sum_{j_i}\left(Y_{post}^{Ct(j_i)}\right)\right]$$

How to Match Customers?

Who are the appropriate customers for inferring counterfactual behavior of affected customers in absence of store

1. Customers who exhibit similar purchase behavior prior to store openings
    1. RFM measures for purchase
    2. Similar purchase trends
2. Customers who have similar socio-demographic variables
    1. Household income, age category, and distance from store
3. Customers from locations with similar zip code-level aggregate socio-economic-demographic variables
    1. Population, median age & income, computer/internet usage, apparel preference index

Experimental Design

	Prior-period	Store opening period	Post-period
Affected customers (Treatment Gr)	$Y_{prior}^{Tr(i)}$		$Y_{post}^{Tr(i)}$
Unaffected customers (Control Gr)	$Y_{prior}^{Ct(j_i)}$		$Y_{post}^{Ct(j_i)}$

$$ATT = \frac{1}{N_i}\sum_{i}\left[\left(Y_{post}^{Tr(i)} - Y_{prior}^{Tr(i)}\right) - \frac{1}{N_{j_i}}\sum_{j_i}\left(Y_{post}^{Ct(j_i)} - Y_{prior}^{Ct(j_i)}\right)\right]$$

Six months period around the date of store opening removed to get rid of the effect of increased awareness about retailer during this period.

Parametric Matching Estimators

- Inverse propensity score weighted least square estimators (Hirano and Imbens 2001)
  - Propensity Score $p\ (X_i)$ = Probability of treatment given matching variables X_i for customer i in pre-store opening period.
  - $Weight_i = [\{Treat_i / p\ (X_i)\} + \{(1\text{-}Treat_i) / (1\text{-}\ p\ (X_i))\}]$, where $Treat_i$ is an indicator variable equal to one if customer i is treated (affected) customer.
  - Estimate the weighted ordinary least square regression which weights observation i with inverse propensity score weight ($Weight_i$)

$$\boldsymbol{Y_{it} = \beta_0 + \beta_1\ Treat_i + \ \beta_2\ Post_t + \ \beta_3\ Treat\ \times Post_{it} + \varepsilon_{it}}$$

  - Coefficient $\boldsymbol{\beta_3}$ estimates the treatment effect

Nonparametric Matching Estimators

- Coarsened exact matching estimator (Iacus et al. 2009)
  - Create bins for each pre-store opening period covariates *X*.
  - Match an affected customer with unaffected customers so that they fall in the same bin for each covariate *X*.
  - Vary the bin size of each covariate such that we have reasonable number of matches for each affected customer
  - Estimate the treatment effect

$$ATT = \frac{1}{N_i}\sum_{i}\left[\left(Y_{post}^{Tr(i)} - Y_{prior}^{Tr(i)}\right) - \frac{1}{N_{j_i}}\sum_{j_i}\left(Y_{post}^{Ct(j_i)} - Y_{prior}^{Ct(j_i)}\right)\right]$$

Results

Estimation Results for Complementary Effect

Variables	Online purchase probability		Online purchase quantity		Online purchase revenue	
	Coeff. Est.	Std. Err	Coeff. Est.	Std. Err	Coeff. Est.	Std. Err
Treat*post	**0.029****	0.014	**0.056****	0.025	**6.59****	3.03
N	873492					

Increase in online purchase revenue by $ 6.59 – a 29 percent increase over the pre-store opening period mean value of $ 22.4

Estimation Results for Store Engagement and Return Effects

Variables	Store Engagement Effect		Store Return Effect	
	Prob. (store purchase |online purchase)		% Online purchase Qty returned in store |online purchase	
	Coeff. Est.	Std. Err.	Coeff. Est.	Std. Err.
Treat*post	**0.110****	0.05	**2.76****	1.34
N	69303		69303	

- **Higher probability of online purchase with store interaction → Store engagement effect**
- **Higher proportion of online purchases returned in store → Store return effect**

Conclusions

- Store opening results in
  - 29 % increase in annual online revenue from existing customers.
  - Higher association of online purchases with store interactions
  - Higher percentage of online purchases returned in store
  - $4.0 increase in net online sales (29% increase) per customer
  - $ 48.35 increase in net store sales (8% increase) per customer

- Overall effect of store opening
  - $ 7.25 million gain in store sales from existing customers
  - $0.6 million gain in online sales from existing customers

Managerial Implications

- Design omnichannel strategies
  - Design cross channel worker incentives
  - Design events to increase store foot traffic and enhance customers' in-store experience
  - Design customer friendly and cross-channel return policies

- Examine possibility of low cost return centers if opening store is too costly

- Account complementary effect of stores on online sales to
  - Compute the viability of existing store
  - Conduct cost-benefit analysis of opening a new store

An Information Stock Model of Customer Behavior in Multichannel Customer Support Services

Anuj Kumar
Univ. of Florida

Kinshuk Jerath
Columbia Business School

Serguei Netessine
Wharton School

Multichannel Customer Support

- Customer support is the major determinant of satisfaction
  - Customer support accounts for 70% of B2C interactions.
  - Firms use a variety of channels to offer customer support

- Two types of channels
  - Assisted channels – E.g., telephone, email, SMS
  - Self-service channels – E.g., website, interactive voice response, web forums

- Telephone and Web are two major channels of customer support
  - Telephone accounts for 65% but Web is catching up fast
  - Web is an order of magnitude cheaper but its efficacy vis-à-vis telephone not known

Previous Work

- Survey-based studies on determinants of adoption of self-service channels (Meuter et al. 2000, Dhabolkar 2001)
- Substitution/augmentation across service channels (Xue et al. 2007, Campbell et al. 2010, Kumar and Telang 2010)
- Call center optimization (OM/OR angle)

Our Field Setting

- Health insurance service setting
  - Customers use insurance when they visit a doctor and a claim is filed
  - Customers call or visit the web to resolve a query regarding claim(s) and insurance plan
- Web provides:
  - Plan coverage and membership details
  - Track costs, claim status
  - Access personal health records
  - Information on health providers
  - Treatment options and general health information
- Telephone provides all of the above, and more

Key idea

- Customer use both telephone and web to obtain information
- Behavior driven by information need:
  - How does information need influence time of query and choice of channel?
  - What is the relative information value of web and telephone?

How to estimate the above from call center data?

Model – Basic Idea

- For every customer, given the times at which claims arrive, we need to model two processes:
  - Time of query arrival
  - Choice of channel at the time of query

- Every customer, at any point of time, has an information need I_i
  - I_i increases with claims
  - I_i decreases after a phone call or web visit (by different amounts)
  - I_i changes with seasons for all customers, such as flu season.
  - Query arrival and choice of channel depend on I_i

Model – Individual Level

- Components of *Information need*
  - Transactional information need
    - Every claim increases I_i by amount C
    - Every phone call decreases I_i by amount T
    - Every web visit decreases I_i by amount W

    I_i = (# claims) C – (# tel.) T – (# web) W
  - Seasonal information need I_m monthly(seasonal) information need constant for all customers

- At time t, query arrival is exponential with rate

  $\lambda_{io} \exp(\beta_\lambda (I_i + I_m))$ (λ_{io} is the baseline rate)

- For a query, probability of visiting the web is

$$p_{i0} \frac{\exp(\pi_T I_{ij} + \pi_S I_m)}{[1 - p_{i0} + p_{i0} \exp(\pi_T I_{ij} + \pi_S I_m)]}$$

(p_{io} is the baseline probability)

Model – Population level

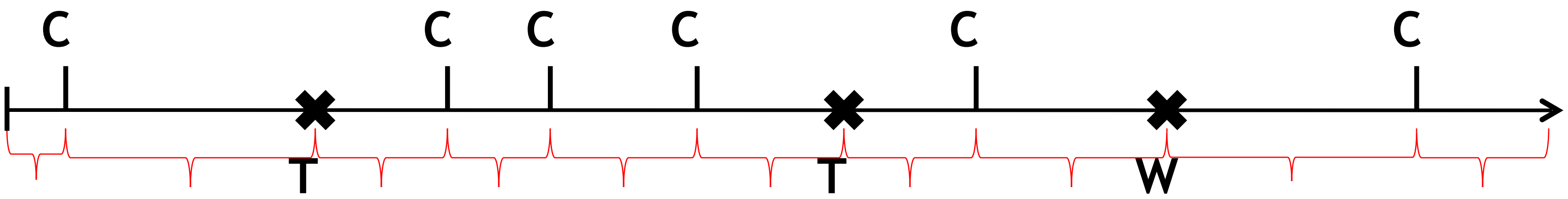
- Baseline query arrival rate: λ_{io} ~ gamma(γ, θ)

- Baseline web choice rate: p_{io} ~ beta(a, b)

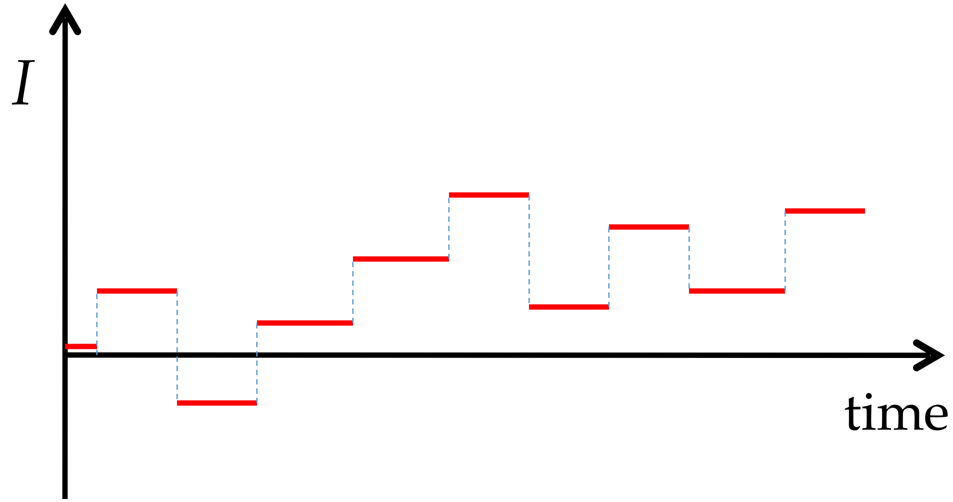
- For a claim, a larger customer liability (customer out of pocket expenses) may lead to a greater need for information

$$C_{ik} = C_o \exp(\beta_{CL} \log(1 + CustomerLiability_{ik}))$$

Illustration



$$I = C - T \; 2T - W$$



$$\lambda_0 \exp(\beta_\lambda I)$$

$$p_0 \frac{\exp(\beta_p I)}{1 + \exp(\beta_p I)}$$

Data

- Data from a large US health insurance firm with approx. 4 million customers

- Random sample of 2462 web registered customers from July 2005 to Dec 2007

- For each customer, data is available on
  - Date of each claim with Customer Liability of claim
  - Date of each query
  - Whether telephone or web was used for query

Estimation results

Baseline	γ	0.887
	θ	126.12
	a	0.492
	b	0.716
Augmentation by Information Gain	C_o	0.258
	W	0.835
	T	1.801
	π_T	-0.039
	π_S	0.303
Customer Liability	β_{CL}	-0.007 (not significant)

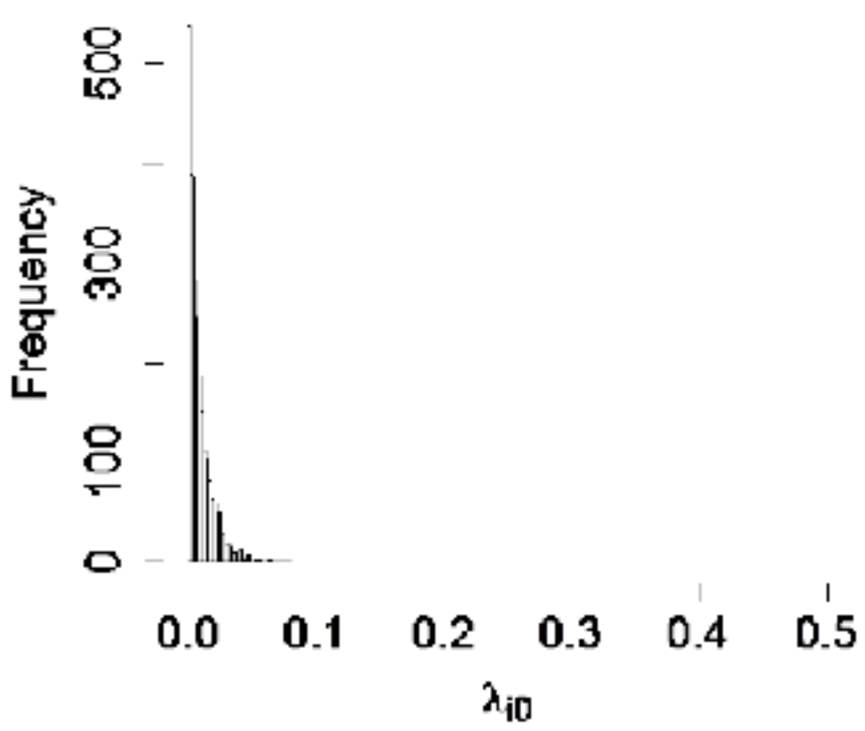
Inferences

- A telephone call , on average, provides more than double 2 times the information than that of a web visit (1.801/0.835=2.16)

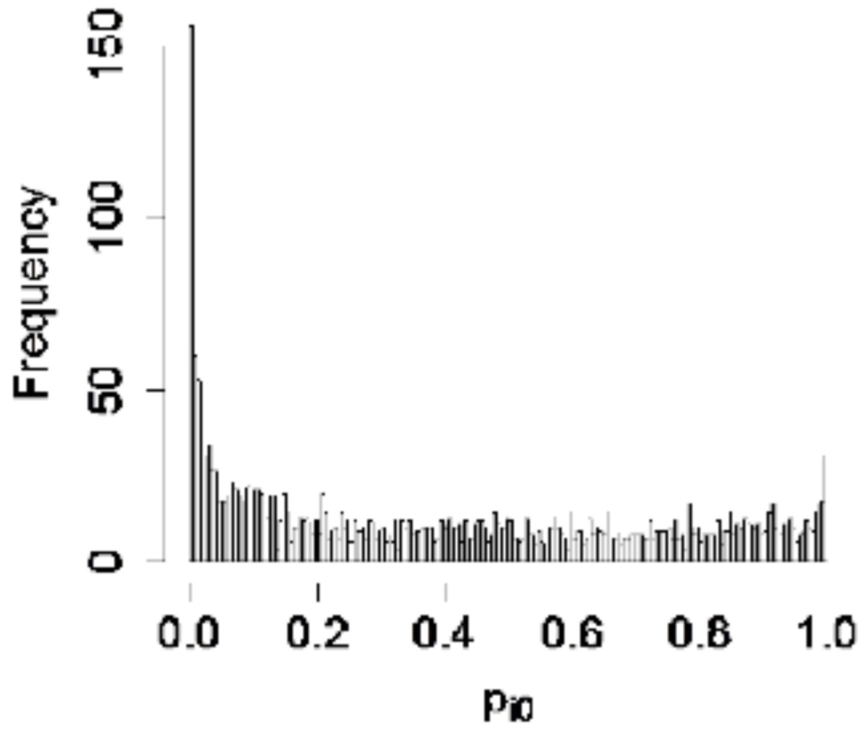
- A web visit costs $0.24 versus a telephone call costs $5.5. So information gain per dollar from web channel is 10 times higher than telephone

- Customers use telephone channel more for their transactional information needs (-ve π_T) but use web channel for their seasonal information needs (+ve π_S).

Baseline Processes



Distribution of baseline query arrival rates is skewed. Median 2.27 queries per year



Distribution of baseline probability of web choice is polarized. Median web choice probability =0.36

Aggregate Predictions

- Lower in-sample and out-of-sample aggregate prediction errors

	In-sample prediction error (%)			Out-of--sample prediction error (%)		
	Total	Telephone	Web	Total	Telephone	Web
Info. Stock Model	2.4	5.1	3.2	4.8	5.1	10.3
Benchmark Model	5.9	8.2	9.5	11.2	10.7	20.1

- More efficient capacity planning on the telephone and web channels with predictions from Info-stock model
  - Reduce CSR costs and call blockage /call abandonment

Individual-level Predictions

- Info-stock model tracks individual customer's information stock and thus should make superior individual-level predictions

Average values for 6 months (July-Dec'2007)	Benchmark model	Info-stock model
True Positive (Correctly predicted calling customers)	66.2	86.6
False Positive (Incorrectly predicted calling customers)	27.3	10.7
True Negative (Correctly predicted non-calling customers)	98.8	98.7
False Negative (Incorrectly predicted non-calling customers)	4.0	1.5

- Accurate individual-level predictions of calling customers can be used to make preemptive calls to them in non-peak time when CSR are free

Summary

- Probabilistic models incorporating customers' information needs result in
  - Significantly better individual-level and aggregate predictions
  - Insights in customers' channel choice and information value of different channels

- Information stock models can significantly reduce customer support costs and increase customer satisfaction

- Model can be easily extended to include more customer support channels
  - Extended to traditional marketing settings – E.g. relative efficacy of contacting customers via catalogs, personalized offers, targeted advertising, etc.

Thanks!