

Session Detail Information

[Add this session to your itinerary](#)

Cluster : Contributed Session

Session Information : Friday Jun 13, 15:30 - 17:00

Title: Predictive Modeling-Methods

Chair: Makoto Abe, Professor, The University of Tokyo, Graduate School of Economics, 7-3-1 Hongo, Bunkyo-ku, Tokyo 113-0033, Japan, abe@e.u-tokyo.ac.jp

Abstract Details

Title: Semi-supervised Learning for Response Modeling

Presenting Author: Sungzoon Cho, Professor, Department of Industrial Engineering, Seoul National University, San 56-1, Shillim-dong, Kwanak-gu, Seoul, Korea, Republic of, zoon@snu.ac.kr

Co-Author: Seong-seob Hwang, Department of Industrial Engineering, Seoul National University, San 56-1, Shillim-dong, Kwanak-gu, Seoul, Korea, Republic of, hss9414@snu.ac.kr

Hyoung-joo Lee, Dr., Robotics Group, University of Oxford, Parks Road, Oxford, United Kingdom, imhjlee@gmail.com

Douglas MacLachlan, Professor & Chair, Foster School of Business Dept of Mktg & IB, University of Washington, Box 353200, Seattle WA 98195, United States of America, macl@u.washington.edu

Hyunjung Shin, Professor, Department of Industrial and Systems Engineering, Ajou University, San 5, Wonchon-dong, Youngtong-gu, Suwon, Korea, Republic of, shin@ajou.ac.kr

Abstract: Response modeling is concerned with targeting customers who are likely to purchase a promoted product or service based on customers' data. A well developed response model can increase profitability while reducing marketing costs. Modeling is usually carried out in a supervised learning framework where only those customer data labeled from the past campaign are used. A response model would predict labels of other customers who are not labeled. One disadvantage of this approach is that only a small percentage of customer data is labeled, thus available for modeling. A semi-supervised learning framework has recently been proposed to improve classification performance by exploiting unlabeled as well as labeled data in modeling based on an intuitive assumption: similar attributes lead to similar labels. In order to address this issue, we propose to use semi-supervised learning approaches for response modeling where a large number of unlabeled data are available. In particular, the multiple graph algorithm fits well to response modeling where customer data come from qualitatively different sources. Through experiments on the CoLL Challenge 2000 dataset, it is shown that the algorithm results in a better model than conventional supervised algorithms in terms of a variety of performance measures. Semi-supervised learning is a viable option and merits further investigation.

Title: Contextual Determinants of the Predictive Accuracy of Binary Prediction Models

Presenting Author: Bas Donkers, School of Economics, Erasmus University Rotterdam, PO Box 1738, Rotterdam -, Netherlands, donkers@few.eur.nl

Co-Author: Aur?ie Lemmens, Assistant Professor, School of Economics, Erasmus University Rotterdam, P.O. Box 1738, Rotterdam ZH, Netherlands, lemmens@few.eur.nl

Peter C. Verhoef, Professor of Marketing, University of Groningen, P.O. Box 800, Groningen GR NL-9700 AV, Netherlands, P.C.Verhoef@rug.nl

Abstract: Previous research has investigated what models perform well in specific prediction tasks, including customer churn, CLV, etc. Recently, Lemmens et al. (2006) proposed bagging and in particular boosting as attractive alternatives to the frequently used logit model to predict customer defection in the US telecommunication industry. They advocate the advantages of these aggregation methods in presence of large datasets that contain a high number of predictors. However, further research is needed to investigate whether the high predictive power of these approaches can generalize to other settings. This paper therefore asks the question: What conditions affect the predictive performance of the various approaches to model binary choices? and What model should managers prefer under specific conditions? We compare four approaches, bagging, boosting, logit and latent class logit models. To investigate the determinants of predictive accuracy, we generate many datasets with varying characteristics, such as, among others, sample size, asymmetry of the distribution of 0/1, inclusion of irrelevant variables, nonlinearity of the relationship. The predictive performance is then investigated through a validation sample, using as performance criteria the hit-rate, the gini coefficient and the top-decile lift. We find that bagging does not live up to expectations, but that the other three approaches all have their merits. The non-parametric nature of bagging and boosting, relative to the logit models, proves beneficial for nonlinear relationships, while it hurts performance the more variables are included. We illustrate the findings from the simulation study using a number of real-life marketing datasets that differ in the dimensions identified as influential in the simulation study.

Title: Optimizing the Value of Incentives for New Customers

Presenting Author: Jeanette Heiligenthal, University of Frankfurt, Mertonstra? 17, Frankfurt 60054, Germany, heiligen@wiwi.uni-frankfurt.de

Co-Author: Bernd Skiera, Professor Dr., University of Frankfurt, Faculty of Business and Economics, Mertonst. 17-25, Frankfurt HE, Germany, skiera@wiwi.uni-frankfurt.de

Abstract: Firms often use very appealing incentives to attract new customers. Newspapers, for example, use reduced prices for the first year of subscription, sometimes even in combination with additional gifts. Thereby, firms need to determine the amount of incentives which they would like to provide. Very small incentives lead only to a limited number of (most likely) loyal customers, whereas high incentives lead to a high number of customers. Yet, those customers are likely to suffer from "adverse selection" because they tend to be rather non-loyal as they are primarily attracted by the incentive and not by the product itself. Hence, the value of the incentive has two effects: it determines the number of customers as well as the average customer lifetime value, which means that an optimal value for the incentive exists. Despite the importance of this problem, research on that problem is rare. Only Cao/Gruca (2005) deal with adverse selection in the credit approval process and present an approach for selecting customers for a cross-selling campaign based on response and approval likelihoods. Yet, they do not focus on the optimal value of possible incentives. Therefore, we address in our paper the problem of how to determine the optimal value of the incentive for new customers. We develop a model and use an empirical data-set from a major European Bank to illustrate the application of the model and to calculate a possible increase in long-term profitability. Based on those results we derive managerial implications for customer acquisition activities under adverse selection.

Title: Counting your Customers? One by One: A Hierarchical Bayes Extension to the Pareto/NBD Model

Presenting Author: Makoto Abe, Professor, The University of Tokyo, Graduate School of Economics, 7-3-1 Hongo, Bunkyo-ku, Tokyo 113-0033, Japan, abe@e.u-tokyo.ac.jp

Abstract: This research extends a Pareto/NBD model of customer-base analysis using a hierarchical Bayesian (HB) framework to suit today's customized marketing. The proposed HB model presumes three tried and tested assumptions of Pareto/NBD models: (1) a Poisson purchase process, (2) a memoryless dropout process (i.e., constant hazard rate), and (3) heterogeneity across customers, while relaxing the independence assumption of the purchase and dropout rates and incorporating customer characteristics as covariates. The model also provides useful output for CRM, such as a customer-specific lifetime and survival rate, as by-products of the MCMC estimation. Using three different types of databases --- music CD for e-commerce, FSP data for a department store and a music CD chain, the HB model is compared against the benchmark Pareto/NBD model. The study demonstrates that recency-frequency data, in conjunction with customer behavior and characteristics, can provide important insights into direct marketing issues, such as the demographic profile of best customers and whether long-life customers spend more.

Semi-Supervised Learning for Response Modeling

Hyung-joo Lee University of Oxford
Hyunjung Shin Ajou University
Seong-seob Hwang Seoul National University
Sungzoon Cho* Seoul National University
Douglas MacLachlan University of Washington

INFORMS Marketing Science Conference
13 June 2008

Response Modeling

Objective

Identifying a subset of likely responders to an offer

In the subset, an accurate response model include

- Many responders (Profit opportunities)
- Few non-responders (Marketing cost)

Empirical approaches

Statistical models

- Logistic regression, Discriminant analysis, etc.

Machine learning

- CBR, DT, K-NN, MLP, SVM, etc.

1

Supervised Learning

Supervised Learning for Response Modeling

Labels obtained from a preliminary campaign

Training a model based on the labeled customer data

Predicting the labels of the unlabeled (test) customer

Problems of Supervised Learning

Labeled data come only from actual campaigns

Only a small part of data are labeled

Unable to utilize the unlabeled data

2

Semi-supervised Learning

Motivation

To make use of unlabeled data

“Cluster” or “Smoothness” assumption

- Data with similar attributes lead to similar labels

Useful when ...

- Labeled data are difficult to collect
- Unlabeled data are readily available or relatively easy to collect

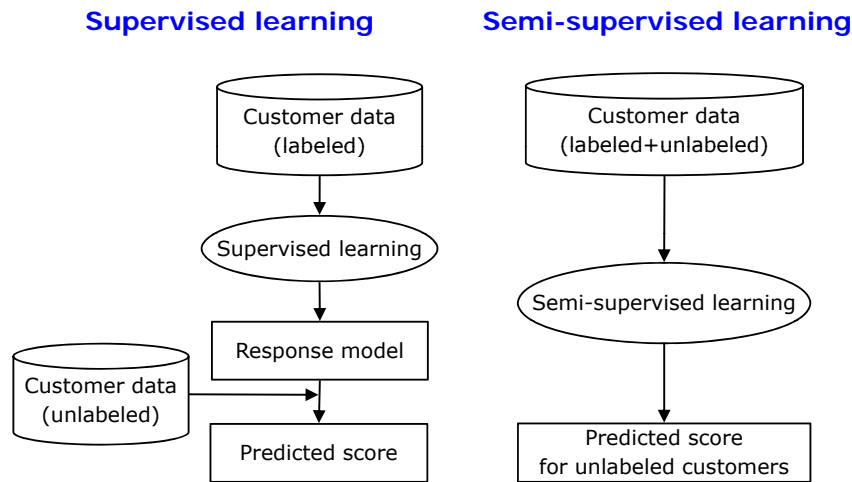
Semi-supervised Learning for Response Modeling

Predicting the labels of the unlabeled (test) customer

By considering the relationship between customers

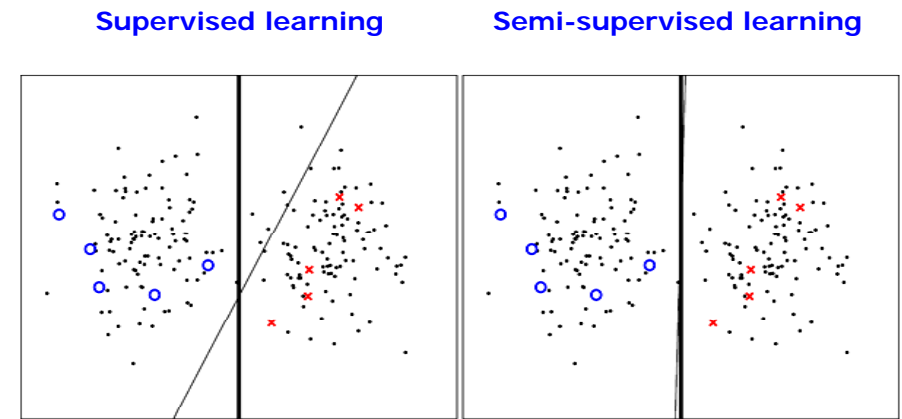
3

Semi-supervised Learning



4

Semi-supervised Learning



5

In This Presentation

Semi-supervised Learning for Response Modeling

To exploit the information of unlabeled customer data

Graph-based semi-supervised learning

- Single graph algorithm
- Multiple graph algorithm

Case Study: CoLL2000 Challenge Data Set

Supervised learning algorithms

- Logistic regression
- K-nearest neighbors (K-NN)
- Support vector machine (SVM)

Semi-supervised learning algorithms

- Single graph algorithm
- Multiple graph algorithm

6

Graph-based Learning

Graph Representation

Each customer: Each vertex

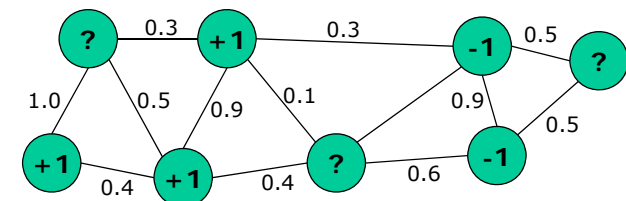
A similarity between customers: A weight of an edge

Labels: +1 / -1 / ? (unlabeled)

Vector → Graph Representation

RBF kernel for K-NN customers $w_{ij} = \exp\left(-\frac{\|\mathbf{x}_i - \mathbf{x}_j\|^2}{2\sigma^2}\right)$

Similarity matrix "W"



7

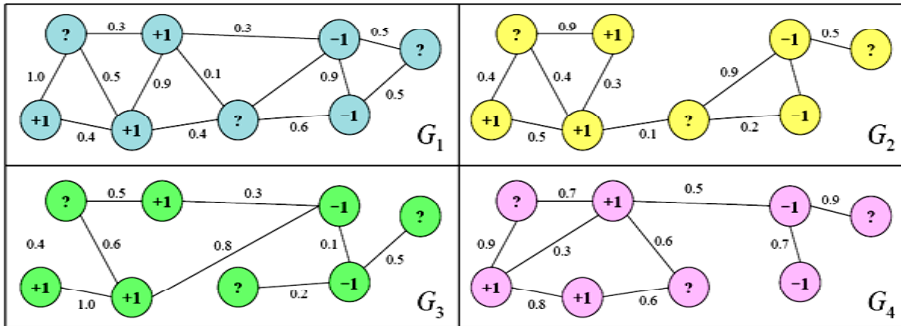
Graph-based Learning

Multiple Graph Learning

Multiple sources of data → Multiple graphs

Ex) Demographic, RFM, customer care, billing, etc.

Learning by integrating multiple graphs



8

Graph-based Learning

Outline of Graph-based Learning

Graph representation

- Transforming vector data into the graph representation
- Labeling vertices (+1 / -1 / ?)

Predicting the labels of unlabeled vertices

- By exploiting the graph structure of both labeled and unlabeled nodes
- Convex optimization

Single and multiple graph algorithm

- Single graph algorithm
- Multiple graph algorithm (multiple graphs available)

9

Single Graph Algorithm

$$\min \underbrace{\sum_i^{n_L} (\hat{y}_i - y_i)^2}_{\text{Loss function}} + \underbrace{C \sum_{i,j=1}^n w_{ij} (\hat{y}_i - \hat{y}_j)^2}_{\text{Smoothness}}$$

Trade-off coefficient

$$\Rightarrow \min_{\hat{\mathbf{y}}} (\hat{\mathbf{y}} - \mathbf{y})^\top (\hat{\mathbf{y}} - \mathbf{y}) + C \hat{\mathbf{y}}^\top \mathbf{L} \hat{\mathbf{y}}$$

L: graph Laplacian matrix

$$\Rightarrow \hat{\mathbf{y}} = (\mathbf{I} + C\mathbf{L})^{-1} \mathbf{y}$$

10

Multiple Graph Algorithm

$$\text{Primal} \quad \min_{\hat{\mathbf{y}}, \gamma} (\hat{\mathbf{y}} - \mathbf{y})^\top (\hat{\mathbf{y}} - \mathbf{y}) + C\gamma,$$

$$\text{subject to } \hat{\mathbf{y}}^\top \mathbf{L}_m \hat{\mathbf{y}} \leq \gamma, \quad m = 1, \dots, M$$

$$\text{Dual} \quad \min_{\alpha} d(\alpha) \equiv \mathbf{y}^\top (\mathbf{I} + \sum_{m=1}^M \alpha_m \mathbf{L}_m)^{-1} \mathbf{y}$$

$$\text{subject to } \sum_{m=1}^M \alpha_m = C$$

$$\text{Solution } \hat{\mathbf{y}} = (\mathbf{I} + \sum_{m=1}^M \alpha_m \mathbf{L}_m)^{-1} \mathbf{y}$$

Dual solution:
a weight of an individual graph

11

Case Study: CoLL2000 Challenge

Data Set

An insurance company case

Response modeling for caravan insurance policies

Training / Test: 5,822 / 4,000

86 variables: demographic, frequency, monetary

Preprocessing

Binary encoding for categorical variables

Graph representation

- Single graph based on all variables
- 3 graphs based on each source: demographic, frequency, monetary

30 random splits of training/test sets

12

Case Study: CoLL2000 Challenge

Response Models

Supervised learning (parameters)

- Logistic regression (none)
- K-NN (the number of neighbors, K)
- SVM (the kernel width, σ & the trade-off coefficient, C)

Semi-supervised learning (parameters)

- Single graph algorithm (K, σ, C)
- Multiple graph algorithm (K, σ, C)

Parameter selection

- The number of actual respondents among top 800 prospects

13

Case Study: CoLL2000 Challenge

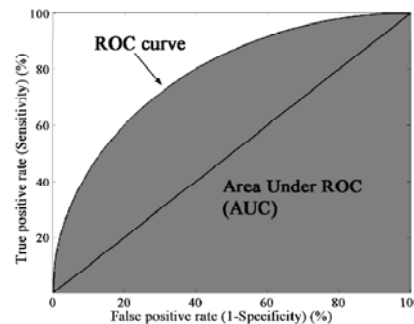
Performance Evaluation Criteria

The number of actual respondents among top 800 prospects

Lift chart

ROC chart

Area under ROC



14

Case Study: CoLL2000 Challenge

Top 800 Prospects

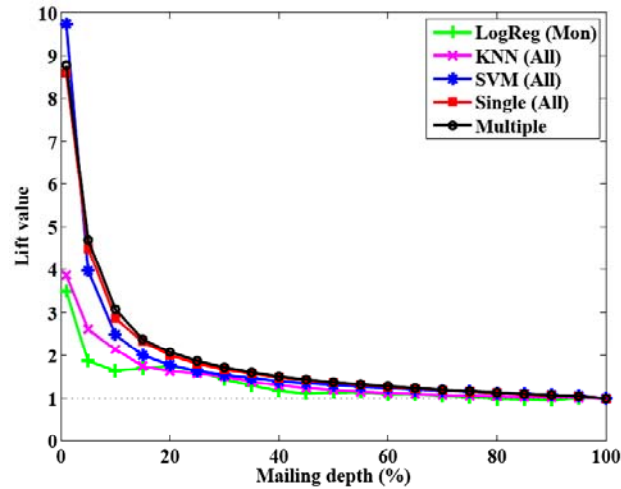
Data source	All	Demo	Freq	Mon	Ensemble
LogReg	52.80	66.43	46.70	75.53	65.37
K-NN	75.00	41.65	61.49	47.71	63.07
SVM	76.50 ³	61.37	44.63	65.67	69.20
Single	84.97 ²	73.57	52.17	59.43	-
Multiple	-	-	-	-	88.27 ¹

The best with $\alpha=0.10$

15

Case Study: CoIL2000 Challenge

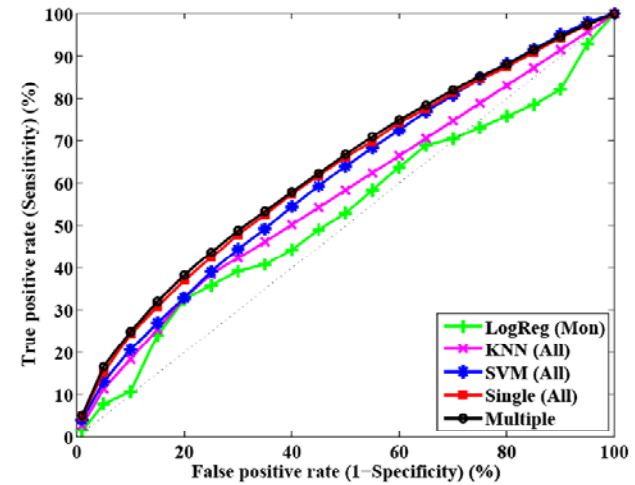
Lift Chart



16

Case Study: CoIL2000 Challenge

ROC Chart



17

Case Study: CoIL2000 Challenge

Area under ROC

Data source	All	Demo	Freq	Mon	Ensemble
LogReg	49.08	56.90	52.67	54.63	53.76
K-NN	57.00	47.94	49.63	53.17	53.52
SVM	60.61 ³	57.75	57.23	50.67	59.12
Single	62.29 ²	58.34	48.98	33.86	-
Multiple	-	-	-	-	63.02 ¹

The best with $\alpha=0.10$

18

Conclusion

Semi-supervised Learning for Response Modeling

To exploit information of the unlabeled customers

Graph-based learning

- Graph representation of data
- Convex optimization
- Single graph algorithm
- Multiple graph algorithm to deal with various sources of information

Case study on CoIL Challenge 2000 data set

- 3 sources \rightarrow 3 graphs
- Comparison with supervised learning algorithm
- Semi-supervised learning outperform supervised learning

Semi-supervised learning should be considered as a viable option for response modeling where a large number of unlabeled data are available

19

Future Research Directions

Computational Complexity

Missing Not At Random (MNAR)

Voluntary Response

Other Semi-supervised Learning Algorithms

Class Imbalance

Feature Selection

THE UNIVERSITY OF BRITISH COLUMBIA | Sauder School of Business

The Institute for Operations Research and the Management Sciences

[Home](#) [Doctoral Consortium](#) [Accommodations](#) [Sauder School of Business](#) [About Vancouver](#) [Map & Contact](#) [Past Conference Hosts](#)
[Conference Overview](#)[Conference Registration](#)[Call for Papers](#)[Abstract Submission](#)[Schedule at-a-glance](#)[Conference Program](#)[Conference Venue](#)[Sponsors](#)

Conference Overview

The Sauder School of Business of The University of British Columbia will host the 2008 INFORMS Marketing Science Conference to be held **from Thursday, June 12 to Saturday, June 14, 2008** at the Sheraton Wall Centre Hotel in Vancouver, British Columbia, Canada. This event brings academics, practitioners and policymakers from around the world together in one of the world's most beautiful cities to present, hear, and discuss new research in a broad range of marketing science topics.

Conference Organizing Committee is composed of the Marketing faculty of the Sauder School of Business:

[Chuck Weinberg](#) (Conference Chair)
[Darren Dahl](#) (co-chair)
[Dan Putler](#) (co-chair)
[Jack Chen](#)
[Srabana Dasgupta](#)

[Tirtha Dhar](#)
[Dale Griffin](#)
[Joey Hoegg](#)
[Tim Silk](#)
[Juliet Zhu](#)

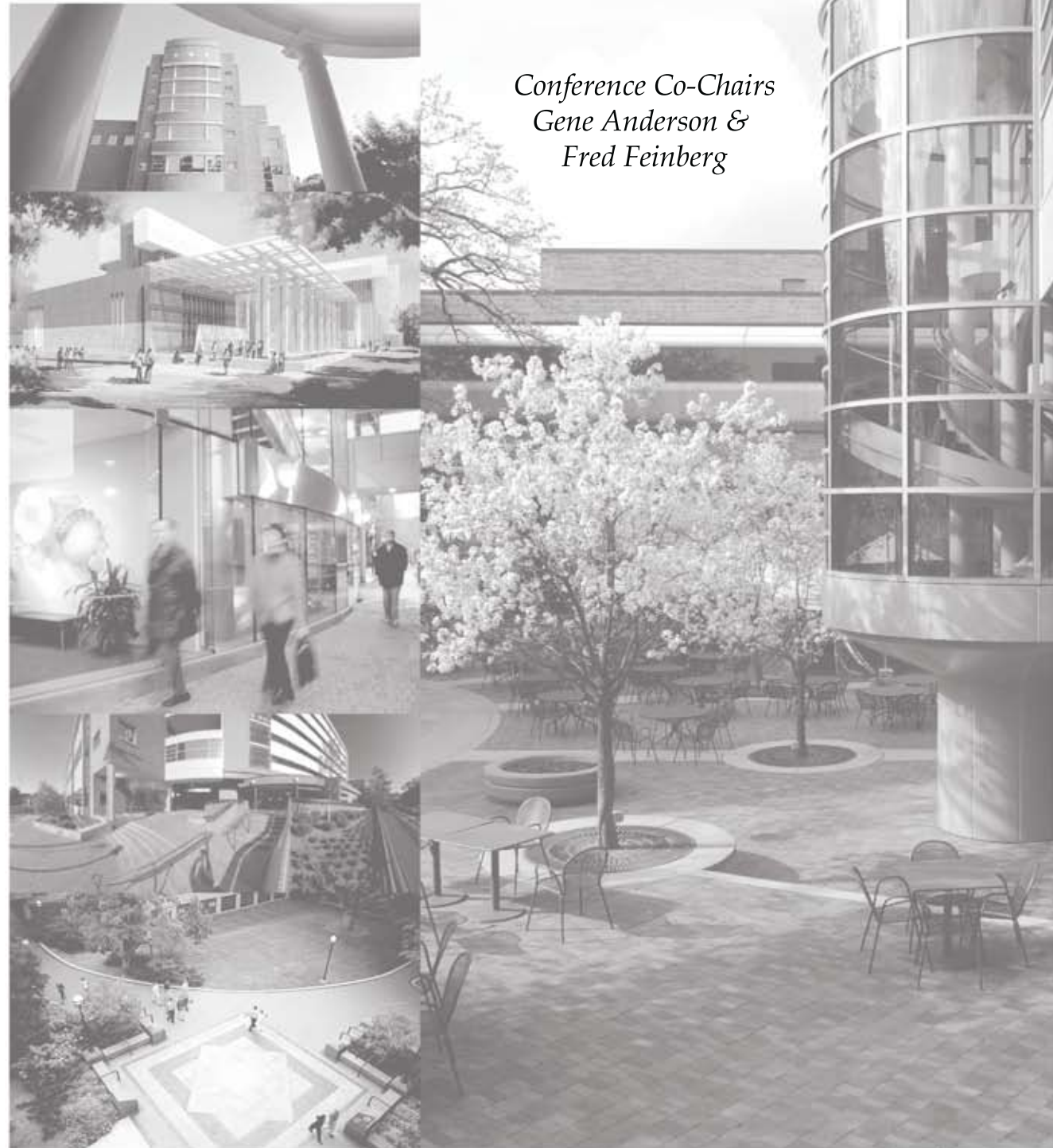
INFORMS MARKETING SCIENCE CONFERENCE

June 4 - 6, 2009

Join us in beautiful Ann Arbor, MI, USA

*The Stephen M. Ross School of Business
at the University of Michigan*

*Conference Co-Chairs
Gene Anderson &
Fred Feinberg*



INFORMS // Marketing Science // Conference 2010 // Cologne // Germany // June 16-19 //



// The Faculty of Management, Economics and Social Sciences at the University of Cologne is proud to host the 2010 INFORMS Society for Marketing Science Conference.

// Join us in the heart of Germany in Cologne from June 16-19 2010.

// Conference website
www.marketingscience2010.uni-koeln.de

// Conference Chair: Werner Reinartz
// Co-chairs: Karen Gedenk, Franziska Voelckner

2008 INFORMS Marketing Science Conference

Friday, June 13th, 2008 3:30-5:00PM (FD)

<p>FD09 – Junior Ballroom B</p> <p>Predictive Modeling-Methods</p> <p>Chair: Makoto Abe</p> <p>Semi-supervised Learning for Response Modeling <i>Sungzoon Cho, Hyunjung Shin, Seong-seob Hwang, Douglas MacLachlan, Hyoung-joo Lee</i></p> <p>Contextual Determinants of the Predictive Accuracy of Binary Prediction Models <i>Bas Donkers, Aurélie Lemmens, Peter C. Verhoef</i></p> <p>Optimizing the Value of Incentives for New Customers <i>Jeanette Heiligenthal, Bernd Skiera</i></p> <p>Counting your Customers' One by One: A Hierarchical Bayes Extension to the Pareto/NBD Model <i>Makoto Abe</i></p>	<p>FD10 – Junior Ballroom A</p> <p>Time Series Market Analysis</p> <p>Chair: Hernan Bruno</p> <p>Structural Dynamics of Competition and its Effects on Long-term Customer Profitability <i>Jens Keller, Sven Reinecke</i></p> <p>Dynamic Customer Management on Duration Time Structure <i>Geonha Kim, Dae Ryun Chang</i></p> <p>The Impact of Pricing on Customer Profitability: Evidence From an Industrial Market <i>Hernan Bruno, Bruce Hardie, Shantanu Dutta</i></p>	<p>FD11 – Port McNeill</p> <p>eCommerce-Consumer Choice and Learning</p> <p>Chair: Wenzel Drechsler</p> <p>The Effect of Internet Adoption on Catalog Customers' Buying Behavior <i>Junzhao Ma, Eric Anderson, Karsten Hansen</i></p> <p>The Impact of Internet Adoption on Customer Purchasing Behavior in Direct Marketing <i>Xiaojing Dong, Kirthi Kalyanam, Pradeep Chintagunta</i></p> <p>Does Opening a Physical Store Change Customer Behavior? <i>Yantao Wang, Eric Anderson, Karsten Hansen</i></p> <p>The Impact of Visualizing a Product's Price History at Price Comparison Sites on Consumer Decisions <i>Wenzel Drechsler, Martin Natter</i></p>	<p>FD12 – Parksville</p> <p>Price Promotions and Price Discrimination</p> <p>Chair: Amit Pazgal</p> <p>Market Segmentation for Pricing Strategy <i>Yoshiyuki Okuse</i></p> <p>Endogenous vs. Exogenous Wholesale Prices During Instant Rebates <i>Simon Sigue</i></p> <p>Endogenous and Exogenous Determinants of Inter-Category Price Sensitivity- A Spatial Approach <i>Joy Joseph, Aman Nanda</i></p> <p>Innovative Pricing Strategies Under Strategic Consumer Behavior <i>Amit Pazgal, Yossi Aviv</i></p>
<p>FD13 – Orca</p> <p>Questions on Brand Identity, Personality, and Extension</p> <p>Chair: Robert Kreuzbauer</p> <p>The Interchangeability of Brand Extensions <i>Joseph Chang, Bob Wu</i></p> <p>Designing Coolness: Brand Building in the Global Fashion Industry <i>Rajesh Chandy, Om Narasimhan, Paola Cillo, Jaideep Prabhu</i></p> <p>Brand Personality: Transference and Preference <i>Sharon Hodge, Kathryn Olinger</i></p> <p>Basic Social Motives of Brand Identity Signaling <i>Robert Kreuzbauer, Chi-yue Chiu, Vivian Vignoles</i></p>	<p>FD14 – Finback</p> <p>Advertising and Market Competition</p> <p>Chair: Nawel Amrouche</p> <p>Advertising and Market Structure: A Digital Video Recorder Field Experiment <i>Carl Mela, Bart Bronnenberg, Jean Pierre Dube</i></p> <p>Copycat Advertising and Strategic Theme Release <i>Chun (Martin) Qiu, Demetrios Vakratsas</i></p> <p>Search Advertising <i>Sridhar Moorthy, Avi Goldfarb</i></p> <p>Feedback Stackelberg Equilibrium Strategies When the Private Labels Competes with the National Brand <i>Nawel Amrouche, Guiomar Martín-Herrán, Georges Zaccour</i></p>	<p>FD15 – Granville</p> <p>Strategy and Consumer Behavior</p> <p>Chair: Wilfred Amaldoss</p> <p>Unbundling Music: Can Selling Individual Songs Reduce Competitive Intensity Among Music Producers? <i>Sherif Nasser, Nicholas Economides</i></p> <p>Turf Wars: Brand Extension in Markets with Preference Based Segmentation <i>Yogesh V. Joshi, David Reibstein, John Zhang</i></p> <p>Software Piracy in the Presence of Open Source Software <i>Rajiv Sinha, T.S. Raghu, Fernando Machado</i></p> <p>Strategic Implications of Reference Groups: An Experimental Investigation <i>Wilfred Amaldoss, Sanjay Jain</i></p>	<p>FD16 – Galiano</p> <p>Retail Store Environment</p> <p>Chair: Julien Schmitt</p> <p>Shelf Layout Effects for Sustainable Products <i>Erjen van Nierop, Erica van Herpen, Laurens Sloot</i></p> <p>Extreme Makeover: Financial and Perceptual Effects of a Remodeled Servicescape Over Time <i>Bram Foubert, Elisabeth Brügger, Dwayne Gremler</i></p> <p>Investigation of the Impact of Store Layout on Category Sales and Store Profitability <i>Yu Ma, Minakshi Trivedi, Dinesh Gauri</i></p> <p>Associations Between Purchases and In-store Behavior: An Extension of the Market Basket Analysis <i>Julien Schmitt, Ganael Bascoul</i></p>