Personalizing Retail Promotions through a DAG-based Representation of Customer Preferences

Srikanth Jagabathula
Associate Professor
NYU Stern School of Business, New York

Joint with:
Gustavo Vulcano
School of Business, Torquato di Tella University, Argentina
Dmitry Mitrofanov
NYU Stern School of Business
Collaborators:

**Gustavo Vulcano**
Professor of OM  
School of Business  
Torquato di Tella University  
Argentina

**Dmitry Mitrofanov**
PhD student in OM  
NYU Stern School of Business
Personalized promotions: an ongoing trend

- drive up sales
- increase both visits & basket size
- reduce competition
- stronger cust. relationship
- price discrimination

65% appreciate personalized prices
52% prefer weekly promotions

Forrester Consulting Study "Indiscriminate Promotions Cost Retailers" [2018]
Retailers want to move away from “mass promotions” to customized and targeted promotions

Goal: CRAFT PROMOTION STRATEGIES USING HISTORICAL TRANSACTION DATA

Is the customer going to buy (@ full price) this item anyway?

- **52%** of promotions go to customers who would pay full price
- **37%** of customers are neutral or negative to such promotions

Which brand(s), if any, can we induce a switch to through promotions?

- Need to consider customer’s: brand loyalty, promotion sensitivity, brand sensitivity

How to optimize promotions, not independently, but jointly?

- Need to consider cannibalization and cross-product elasticities

---

Forrester Consulting Study "Indiscriminate Promotions Cost Retailers" [2018]

Can save ~$60M
Our focus: personalized demand predictions using panel data

Panel data  

Sales transactions tagged by customer id

Challenges

- very few observations per customer
- product unavailability [stock-outs]
- switching [price/quantity promotion]
- latent consideration sets
Our focus: personalized demand predictions using panel data

Panel data: sales transactions tagged by customer id

Challenges:
- very few observations per customer
- product unavailability [stock-outs]
- switching [price/quantity promotion]
- latent consideration sets

Our approach:
use partial orders to efficiently extract preference info
Our focus:
personalized demand predictions using panel data

Panel data  sales transactions tagged by customer id

Challenges
- very few observations per customer
- product unavailability [stock-outs]
- switching [price/quantity promotion]
- latent consideration sets

Our approach
choice model
both utility- and rank-based
Our focus: personalized demand predictions using panel data

Panel data  sales transactions tagged by customer id

Challenges

‣ very few observations per customer
‣ product unavailability [stock-outs]
‣ switching [price/quantity promotion]
‣ latent consideration sets

Our approach

promoted product copies
**Our focus:** personalized demand predictions using panel data

| Panel data | sales transactions tagged by customer id |

**Challenges**

- very few observations per customer
- product unavailability [stock-outs]
- switching [price/quantity promotion]
- latent consideration sets

**Our approach**

data-driven imputation
Our focus: personalized demand predictions using panel data

Panel data: sales transactions tagged by customer id

Challenges:
- very few observations per customer
- product unavailability [stock-outs]
- switching [price/quantity promotion]
- latent consideration sets

Contributions:
- methodology for personalized promotions
- performance gains on IRI dataset
Related work: Rich work spanning various areas

Operations:
Estimation: [Farias, Jagabathula, Shah ’12], [Haensel, Koole ’11], [van Ryzin, Vulcano ’13], [Jagabathula, Rusmevichientong ’13], …
Optimization: [Mahajan, van Ryzin ’01], [Farias, Jagabathula, Shah ’13], [Papers by Rusmevichientong, Topaloglu, and Gallego], …
Airline RM: [Zhang, Cooper ’06], [Chaneton, Vulcano ’11], [Kunnumkal ’12], …

Marketing:
[Erdem, Imai, Keane ’03], [Dekimpe, Hanssens, Silva-Rissio ’99], [Foekens, Leeftang, Wittink ’99], [Aiwadi, Neslin ’98], [Mela, Gupta, Lehmann ’97], [Grover, Srinivasan ’92], [Gupta 88], …

Focus is on
- decision-making
- estimation from aggregated transactions

our focus: panel data & individual predictions

we complement by enriching inferences through partial orders

methodological connections
- extension to promoted items
- explain cycles in preferences
- new probability bounds
- promotion optimization

Jagabathula & Vulcano (Mgmt. Sci., 2017)
1. Model
assumptions, data for estimation

2. Inference framework
construction of DAGs, de-cycling, ML estimation

3. Numerical results
results on real-world panel data from IRI dataset

4. Summary/Conclusions
takeaway messages
Personalizing Retail Promotions through a DAG-based Representation of Customer Preferences

1. Model assumptions, data for estimation

2. Inference framework construction of DAGs, de-cycling, ML estimation

3. Numerical results results on real-world panel data from IRI dataset

4. Summary/Conclusions takeaway messages
Panel data: sales transactions tagged with customer ids

- **$n$** products
- **$m$** customers
  - mutually substitutable
  - make repeated purchases

**Observations:** for each customer and time period
- offer and promotion sets
- purchased product

<table>
<thead>
<tr>
<th>Time</th>
<th>Offer Set $S_t$</th>
<th>Promoted Set $P_t$</th>
<th>Purchase $i_t$</th>
<th>Products</th>
</tr>
</thead>
<tbody>
<tr>
<td>$t = 1$</td>
<td>{1,2,3,4,6,7}</td>
<td>{3,6,7}</td>
<td>3</td>
<td>{1,2,3,4,6,7,7}</td>
</tr>
<tr>
<td>$t = 2$</td>
<td>{1,2,3,4,5,6,7}</td>
<td>{6,7}</td>
<td>3</td>
<td>{1,2,3,4,6,7}</td>
</tr>
<tr>
<td>$t = 3$</td>
<td>{1,2,5,6,7}</td>
<td>{}</td>
<td>2</td>
<td>{1,2,6,7}</td>
</tr>
</tbody>
</table>
Weekly store sales and consumer panel data

31 product categories
11 years
2001 - 2011

We analyzed year 2007 panel data

27 product categories
83K user-category combinations
1.2M transactions across 52 weeks
<table>
<thead>
<tr>
<th>Category name</th>
<th># Vendors</th>
<th># Customers retained</th>
<th>Avg. # trans.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Beer</td>
<td>105</td>
<td>1154</td>
<td>7.11</td>
</tr>
<tr>
<td>Carbonated beverages</td>
<td>57</td>
<td>4387</td>
<td>17.63</td>
</tr>
<tr>
<td>Cigarettes</td>
<td>18</td>
<td>307</td>
<td>10.39</td>
</tr>
<tr>
<td>Coffee</td>
<td>73</td>
<td>2255</td>
<td>5.59</td>
</tr>
<tr>
<td>Cold cereal</td>
<td>45</td>
<td>3998</td>
<td>10.94</td>
</tr>
<tr>
<td>Deodorant</td>
<td>36</td>
<td>653</td>
<td>3.47</td>
</tr>
<tr>
<td>Facial tissue</td>
<td>13</td>
<td>2063</td>
<td>4.96</td>
</tr>
<tr>
<td>Frozen dinners/Entrees</td>
<td>97</td>
<td>3288</td>
<td>13.46</td>
</tr>
<tr>
<td>Frozen pizza</td>
<td>47</td>
<td>2846</td>
<td>7.83</td>
</tr>
<tr>
<td>Household cleaners</td>
<td>84</td>
<td>1699</td>
<td>4.14</td>
</tr>
<tr>
<td>Hot dogs</td>
<td>44</td>
<td>2187</td>
<td>3.82</td>
</tr>
<tr>
<td>Laundry detergent</td>
<td>24</td>
<td>2181</td>
<td>4.04</td>
</tr>
<tr>
<td>Margarine/Butter</td>
<td>19</td>
<td>2750</td>
<td>5.65</td>
</tr>
<tr>
<td>Mayonnaise</td>
<td>19</td>
<td>2386</td>
<td>3.28</td>
</tr>
<tr>
<td>Milk</td>
<td>37</td>
<td>4652</td>
<td>14.9</td>
</tr>
<tr>
<td>Mustard</td>
<td>60</td>
<td>2515</td>
<td>3.66</td>
</tr>
<tr>
<td>Paper towels</td>
<td>12</td>
<td>2051</td>
<td>5.2</td>
</tr>
<tr>
<td>Peanut butter</td>
<td>25</td>
<td>1923</td>
<td>3.89</td>
</tr>
<tr>
<td>Salt snacks</td>
<td>113</td>
<td>4446</td>
<td>15.09</td>
</tr>
<tr>
<td>Shampoo</td>
<td>81</td>
<td>738</td>
<td>3.73</td>
</tr>
<tr>
<td>Soup</td>
<td>123</td>
<td>4322</td>
<td>12.02</td>
</tr>
<tr>
<td>Spaghetti/Italian sauce</td>
<td>74</td>
<td>2698</td>
<td>5.46</td>
</tr>
<tr>
<td>Sugar substitutes</td>
<td>17</td>
<td>308</td>
<td>3.3</td>
</tr>
<tr>
<td>Toilet tissue</td>
<td>13</td>
<td>2817</td>
<td>5.1</td>
</tr>
<tr>
<td>Toothbrushes</td>
<td>52</td>
<td>499</td>
<td>3.06</td>
</tr>
<tr>
<td>Toothpaste</td>
<td>38</td>
<td>1186</td>
<td>3.58</td>
</tr>
</tbody>
</table>
| Yogurt                  | 32        | 3491                 | 19.81         

- aggregate items by vendors
- training: first 26 weeks
- retain customers ≥ 2 sales
Training data summary

- aggregate items by vendors
- training: first 26 weeks
- retain customers \( \geq 2 \) sales

After pre-processing

64K user-category combinations
600K transactions across 52 weeks

Average across 27 categories

50 vendors
2.3K customers
9.3 transactions per customer
Choice model: decomposes preferences into “strong” and “weak” prefs.

“STRONG” PREFS. Partial order: DAG (Directed Acyclic Graph)

- COLGATE
- CREST
- ORAL B

Offer set choice
Choice model: decomposes preferences into “strong” and “weak”

“STRONG” PREFS. Partial order: DAG (Directed Acyclic Graph)
Choice model: decomposes preferences into “strong” and “weak”

“STRONG” PREFS. Partial order: DAG (Directed Acyclic Graph)

“WEAK” PREFS. Distribution over rankings samples preference list consistent w/ DAG
Model benefits: efficiently extracts preference information from data

Infer preference relations with fewer samples

- **RUM model**
  - Rank lists differ across periods
  - Prefers 1 or 2 each period
  - Many samples needed to estimate

- **Strong preferences**
  - Rank lists consistent with partial order
  - Prefers 1 over 2 every period
  - Fewer samples needed to estimate

- Rank list same across periods
  - Only few samples needed to estimate
Consideration set: customers may not consider all the products on offer

Customers may choose from smaller consideration set

Avoids inferring spurious preference relations
Personalizing Retail Promotions through a DAG-based Representation of Customer Preferences

1. Model
   assumptions, data for estimation

2. Inference framework
   construction of DAGs, de-cycling, ML estimation

3. Numerical results
   results on real-world panel data from IRI dataset

4. Summary/Conclusions
   takeaway messages
DAG construction:
infer candidate edges from transactions

\[ t = 1; \text{offer set} = \{1, 2, 3, 4\}; \text{promotion set} = {} \]
DAG construction:
infer candidate edges from transactions

candidate edges

edge (2, 3) may have been sampled

product 4 may not have been considered
DAG construction:
infer candidate edges from transactions
DAG construction:
infer candidate edges from transactions

$t = 1; offer set = \{1, 2, 3, 4\}; promotion set = \{\}$

$t = 2; offer set = \{2, 5, 7\}; promotion set = \{\}$
DAG construction:
infer candidate edges from transactions

$t = 2$; offer set = \{2, 5, 7\}; promotion set = \{

$t = 3$; offer set = \{2, 5, 7\}; promotion set = \{

\[
\begin{align*}
p_{5,2} &= 2 \\
p_{5,7} &= 2
\end{align*}
\]
DAG construction:
infer candidate edges from transactions

$t = 3; \text{offer set} = \{2, 5, 7\}; \text{promotion set} = \{}$

$t = 4; \text{offer set} = \{4, 5\}; \text{promotion set} = \{}$
DAG construction:
infer candidate edges from transactions

\[ w_{5,2} = 2 \]
\[ w_{5,7} = 2 \]
De-cycling procedure removes spurious preference edges.

Spurious edges may be sampled from rank dist. from products not considered. Structural assumptions lead to fewest sampled edges/largest DAG. The largest possible consideration sets are found while being consistent with the transitive relationships.

Find largest weighted sub-graph that is a DAG.
To deal with promoted items, we maintain two copies of each item—promoted and nonpromoted.

$t = 1; offer\ set = \{1, 2, 3\}; promotion\ set = \{3\}$

purchased
<table>
<thead>
<tr>
<th>Category name</th>
<th># Vendors</th>
<th># Customers retained</th>
<th>% w/ Cycles</th>
<th>Size w/o cycles</th>
<th>Size w/ Cycles</th>
</tr>
</thead>
<tbody>
<tr>
<td>Beer</td>
<td>105</td>
<td>1154</td>
<td>60%</td>
<td>182.42</td>
<td>257.30</td>
</tr>
<tr>
<td>Carbonated beverages</td>
<td>57</td>
<td>4387</td>
<td>85%</td>
<td>83.61</td>
<td>130.70</td>
</tr>
<tr>
<td>Cigarettes</td>
<td>18</td>
<td>307</td>
<td>24%</td>
<td>28.62</td>
<td>36.85</td>
</tr>
<tr>
<td>Coffee</td>
<td>73</td>
<td>2255</td>
<td>51%</td>
<td>114.75</td>
<td>154.88</td>
</tr>
<tr>
<td>Cold cereal</td>
<td>45</td>
<td>3998</td>
<td>83%</td>
<td>76.91</td>
<td>132.91</td>
</tr>
<tr>
<td>Deodorant</td>
<td>36</td>
<td>653</td>
<td>47%</td>
<td>59.29</td>
<td>80.95</td>
</tr>
<tr>
<td>Facial tissue</td>
<td>13</td>
<td>2063</td>
<td>44%</td>
<td>19.92</td>
<td>26.00</td>
</tr>
<tr>
<td>Frozen dinners/Entrees</td>
<td>97</td>
<td>3288</td>
<td>71%</td>
<td>156.70</td>
<td>282.66</td>
</tr>
<tr>
<td>Frozen pizza</td>
<td>47</td>
<td>2946</td>
<td>59%</td>
<td>74.28</td>
<td>113.35</td>
</tr>
<tr>
<td>Household cleaners</td>
<td>84</td>
<td>1699</td>
<td>77%</td>
<td>140.72</td>
<td>199.94</td>
</tr>
<tr>
<td>Hot dogs</td>
<td>44</td>
<td>2187</td>
<td>46%</td>
<td>76.49</td>
<td>99.76</td>
</tr>
<tr>
<td>Laundry detergent</td>
<td>24</td>
<td>2181</td>
<td>37%</td>
<td>42.58</td>
<td>58.75</td>
</tr>
<tr>
<td>Margarine/Butter</td>
<td>19</td>
<td>2750</td>
<td>49%</td>
<td>39.39</td>
<td>58.00</td>
</tr>
<tr>
<td>Mayonnaise</td>
<td>19</td>
<td>2386</td>
<td>22%</td>
<td>30.70</td>
<td>37.30</td>
</tr>
<tr>
<td>Milk</td>
<td>37</td>
<td>4652</td>
<td>64%</td>
<td>69.38</td>
<td>102.10</td>
</tr>
<tr>
<td>Mustard</td>
<td>60</td>
<td>2515</td>
<td>64%</td>
<td>93.50</td>
<td>116.49</td>
</tr>
<tr>
<td>Paper towels</td>
<td>12</td>
<td>2051</td>
<td>52%</td>
<td>25.41</td>
<td>37.30</td>
</tr>
<tr>
<td>Peanut butter</td>
<td>25</td>
<td>1923</td>
<td>36%</td>
<td>37.59</td>
<td>47.25</td>
</tr>
<tr>
<td>Salt snacks</td>
<td>113</td>
<td>4446</td>
<td>86%</td>
<td>162.17</td>
<td>272.61</td>
</tr>
<tr>
<td>Shampoo</td>
<td>81</td>
<td>738</td>
<td>52%</td>
<td>113.97</td>
<td>147.62</td>
</tr>
<tr>
<td>Soup</td>
<td>123</td>
<td>4322</td>
<td>77%</td>
<td>186.19</td>
<td>291.87</td>
</tr>
<tr>
<td>Spaghetti/Italian sauce</td>
<td>74</td>
<td>2698</td>
<td>49%</td>
<td>108.28</td>
<td>149.98</td>
</tr>
<tr>
<td>Sugar substitutes</td>
<td>17</td>
<td>308</td>
<td>16%</td>
<td>24.09</td>
<td>28.22</td>
</tr>
<tr>
<td>Toilet tissue</td>
<td>13</td>
<td>2817</td>
<td>45%</td>
<td>27.75</td>
<td>41.76</td>
</tr>
<tr>
<td>Toothbrushes</td>
<td>52</td>
<td>499</td>
<td>48%</td>
<td>77.68</td>
<td>99.60</td>
</tr>
<tr>
<td>Toothpaste</td>
<td>38</td>
<td>1186</td>
<td>40%</td>
<td>57.14</td>
<td>74.26</td>
</tr>
<tr>
<td>Yogurt</td>
<td>32</td>
<td>3491</td>
<td>61%</td>
<td>49.91</td>
<td>77.22</td>
</tr>
</tbody>
</table>

% of customers with cycles and densities of DAGs
<table>
<thead>
<tr>
<th>Category name</th>
<th># Vendors</th>
<th># Customers retained</th>
<th>% w/ Cycles</th>
<th>Size w/o cycles</th>
<th>Size w/ Cycles</th>
</tr>
</thead>
<tbody>
<tr>
<td>Beer</td>
<td>105</td>
<td>1154</td>
<td>60%</td>
<td>182.42</td>
<td>257.30</td>
</tr>
<tr>
<td>Carbonated beverages</td>
<td>57</td>
<td>4387</td>
<td>85%</td>
<td>83.61</td>
<td>130.70</td>
</tr>
<tr>
<td>Cigarettes</td>
<td>18</td>
<td>307</td>
<td>24%</td>
<td>28.62</td>
<td>36.85</td>
</tr>
<tr>
<td>Coffee</td>
<td>73</td>
<td>2255</td>
<td>51%</td>
<td>114.75</td>
<td>154.88</td>
</tr>
<tr>
<td><strong>Cold cereal</strong></td>
<td><strong>45</strong></td>
<td><strong>3998</strong></td>
<td><strong>83%</strong></td>
<td><strong>76.91</strong></td>
<td><strong>132.91</strong></td>
</tr>
<tr>
<td><strong>Deodorant</strong></td>
<td><strong>36</strong></td>
<td><strong>653</strong></td>
<td><strong>47%</strong></td>
<td><strong>59.29</strong></td>
<td><strong>80.95</strong></td>
</tr>
<tr>
<td>Facial tissue</td>
<td>13</td>
<td>2063</td>
<td>44%</td>
<td>19.92</td>
<td>26.00</td>
</tr>
<tr>
<td>Frozen dinners/Entrees</td>
<td>97</td>
<td>3288</td>
<td>71%</td>
<td>156.70</td>
<td>282.66</td>
</tr>
<tr>
<td>Frozen pizza</td>
<td>47</td>
<td>2946</td>
<td>59%</td>
<td>74.28</td>
<td>113.35</td>
</tr>
<tr>
<td>Household cleaners</td>
<td>84</td>
<td>1699</td>
<td>77%</td>
<td>140.72</td>
<td>199.94</td>
</tr>
<tr>
<td>Hot dogs</td>
<td>44</td>
<td>2187</td>
<td>46%</td>
<td>76.49</td>
<td>99.76</td>
</tr>
<tr>
<td>Laundry detergent</td>
<td>24</td>
<td>2181</td>
<td>37%</td>
<td>42.58</td>
<td>58.75</td>
</tr>
<tr>
<td>Margarine/Butter</td>
<td>19</td>
<td>2750</td>
<td>49%</td>
<td>39.39</td>
<td>58.00</td>
</tr>
<tr>
<td>Mayonnaise</td>
<td>19</td>
<td>2386</td>
<td>22%</td>
<td>30.70</td>
<td>37.30</td>
</tr>
<tr>
<td>Milk</td>
<td>37</td>
<td>4652</td>
<td>64%</td>
<td>69.38</td>
<td>102.10</td>
</tr>
<tr>
<td>Mustard</td>
<td>60</td>
<td>2515</td>
<td>64%</td>
<td>93.50</td>
<td>116.49</td>
</tr>
<tr>
<td>Paper towels</td>
<td>13</td>
<td>2051</td>
<td>52%</td>
<td>25.41</td>
<td>37.30</td>
</tr>
</tbody>
</table>

% of customers with cycles and densities of DAGs

- **54%**
- customers with cycles

avg. # of edges in DAGs

- **73** in cust.
- **150** in cust.
- w/o cycles
- w/ cycles

avg. # of trans. per cust.

- **9.3**

avg. offer set size

- **16**
mostly purchases Crest & Colgate
if Colgate is promoted, will definitely buy
never purchases “Natures Gate” even if promoted
brand loyal to “Philip Morris”, then “Reynolds Light”
might switch to “Howard”, “Reynolds Light”, or “Commonwealth Light” if promoted and “Philip Morris” is stocked out
MNL model has closed-form expression for forest of trees with one root

\[
\text{Likelihood of DAG} \quad \prod_{i=1}^{n} \frac{v_i}{v_i + \sum_{j \in T_i} v_j}
\]

subtree of node \(i\)

\[
\text{prob of tree} = (\text{prob of subtree 2}) \times (\text{prob of subtree 3}) \times \left( \frac{v_1}{v_1 + \cdots + v_{11}} \right)
\]

\[
\text{prob of subtree 2} = (\text{prob of subtree 4}) \times (\text{prob of subtree 5}) \times \left( \frac{v_2}{v_2 + v_4 + v_5 + v_8 + v_9} \right)
\]

\[
\vdots
\]
MNL model has closed-form expression for forest of trees with one root

Theorem [informal]:
The likelihood expression above is
  ‣ lower bound approx. in general
  ‣ exact when DAG forest of directed trees, each with a unique root

Theorem [informal]:
We derive upper and lower bounds for posterior probability predictions
1. Model
   assumptions, data for estimation

2. Inference framework
   construction of DAGs, de-cycling, ML estimation

3. Numerical results
   results on real-world panel data from IRI dataset

4. Summary/Conclusions
   takeaway messages
Benchmark models compared: LC-MNL and RPL

**k-latent class MNL (LC-MNL)**

- Sample class membership; follow MNL for that class
- \# of parameters = \( k \ n \)
- EM-based regularized max likelihood estimation
- Best performance up to \( k = 10 \) reported

**Random Parameters Logit (RPL)**

- Sample MNL parameters \( \sim \) multivariate normal
- \# of parameters = \( 2 \ n \)
- Max simulated likelihood estimation (MSLE)
- Computationally intensive
Experiments conducted: one step-ahead prediction

\( U = \text{user set} \quad N = \text{product set} \quad T = \# \text{ of discrete time periods} \)

For any \( t = 1, 2, 3, \ldots, T \)

Given

\[ U_t, N, T \]

everything until time period \( t \)
(offer sets, promoted items, purchases of users in \( U \))

\( S_{t+1} = \text{offer set in period } t+1 \)
\( P_{t+1} = \text{promoted items in period } t+1 \)
\( U_{t+1} = \text{users purchasing in period } t+1 \)

Prediction

\( f_u(i, t+1) = 1 \text{ if } i \text{ has highest choice probability for } u \text{ in period } t+1 \)

for all \( u \in U_{t+1}, i \in S_{t+1} \)

\[ X^2 \text{ score } = \frac{1}{|U||N|} \sum_{u \in U, i \in N} \frac{(n_{ui} - \hat{n}_{ui})^2}{0.5 + \hat{n}_{ui}}, \quad \hat{n}_{ui} = \sum_{t \in T} f_u(i, t) \]

# of purchases of i by u

similar to \( \chi^2 \text{ score } \frac{(O - E)^2}{E} \)

lower is better

\[ \text{miss rate } = \frac{1}{|U||T|} \sum_{u \in U, t \in T} I[f_u(a_{j,u,t}) \neq 1] \quad \# \text{ obs. purchase in } t \text{ of } u \]
Key takeaways:

- Even single-class MNL outperforms RPL for most categories
- RPL has more parameters. Time: our method ~ 10 secs, RPL ~ 67 mins
- Gains higher for cust. w/o cycles because of strong prefs.
- De-cycling extends coverage to all customers
- Single-class: heterogeneity through DAGs
- Multi-class: additional heterogeneity through classes
Promotion optimization upon customer arrival to store: offer set fixed but promoted items can be changed

We formulate this problem as an MILP

$$\max_y \sum_{a_j \in S} \left[ r_j \cdot (1 - y_j) + (r_j - d_j) \cdot y_j \right] \times \mathbb{I}[a_j \text{ purchased}]$$

- $r_j$: revenue at full-price
- $(r_j - d_j)$: discounted revenue on promotion
- $\mathbb{I}[a_j \text{ purchased}]$: indicator that item $a_j$ is purchased

We formulate this problem as an MILP
24% revenue gain from personalizing promotions

HUGE revenue opportunity

impr w./o. mass = 23.93%
impr w. mass = 16.61%

MAE = 6.34%

impr w./o. mass = 23.93%
impr w. mass = 16.61%

observed revenue on hold-out offer sets

revenue predicted by our model on hold-out offer sets

accurate revenue predictions from our method
Largest revenue opportunity: for less frequently purchased categories
Customized Individual Promotions: Model, Optimization, and Prediction

1. Model assumptions, data for estimation
2. Inference framework construction of DAGs, de-cycling, ML estimation
3. Numerical results results on real-world panel data from IRI dataset
4. Summary/Conclusions takeaway messages
Summary and key findings

**KEY CONTRIBUTION**
Methodology to design personalized promotions from panel data through DAGs

**MAIN TAKEAWAYS**
- DAGs provide rich representation of preferences
- Better predictive accuracy: more than 15% better predictive accuracy compared to state-of-the-art benchmarks
- HUGE revenue opportunity: up to 51% revenue gain w/ personalized promotions