Heterogeneous Demand Modeling

UCSD MGT 100 Week 5

Kenneth C. Wilbur and Dan Yavorsky

Segmentation case study: Quidel

- Leading B2B manufacturer of home pregnancy tests
- Tests were quick and reliable
- Wanted to enter the B2C HPT market
- Market research found 2 segments of equal size;
  what were they?

	Segment 1	Segment 2
Segment Name		
Product Name		
Positioning		
Image on Package		
Package Size		
Location in Store		
Product Line Extensions		
Price		

How Africa's top handset maker designs phone cameras calibrated for darker skin tones

or Ethiopia's Amharic. But one feature that has distinguished its devices is their camera-centric nature and their ability to calibrate exposures for darker skin tones.

As its key focus market, Transsion has thousands of employees across Africa, working in production lines in its Ethiopian factory and as in-design and user interface personnel in Kenya and Nigeria. After conducting an in-depth analysis of consumers' photo habits and needs, the company found photo quality was important to not just younger consumers but increasingly wider age demographics. Phone cameras, especially front camera exposure, was the first feature customers inspected when considering buying a new mobile phone.

"We discovered ways to optimize photos, such as improving users' eyes, nose, skin color, and quality, which helps our users take a clearer, more natural, and more beautiful photo," says Robin Wang, the general manager of Transsion's hardware center.

Transsion's suite of brands (Tecno, Infinix, and Itel) has a **commanding 48.2% share of the African smartphone market,** three times that of its closest competitor Samsung (16%).

Digital Default Frequencies

- Classic study: 95% of MS Word users maintained original default options
- Classic study: 81-90% of users don't use Ctrl+F
- Classic study: Randomizing top 2 search results only changed click rates from 42%/8% to 34%/12%
- 2021 data: Safari has 90% share on iPhone, Chrome (74%) and Samsung Internet (15%) have 89% share on Android

    - Chrome always preinstalled on Android, Samsung Internet preinstalled on 58% of Android devices

- Why? Behavioral research:

    - Defaults are good enough
    - Changing defaults is hard
    - Changing defaults is uncertain/ambiguous
    - User assumes product designer knows best--sometimes correctly
    - Popularity implies utility
    - Possible fear of exclusion or norm deviation

- Implies importance of understanding customer needs (aka market research) prior to initial product offerings

Het. Demand Models

1. Discrete heterogeneity by segment
2. Continuous heterogeneity by customer attributes
3. Individual-level demand parameters

```
- We'll code 1 & 2
- 3 is often best but needs advanced techniques --> graduate study
```

MNL Demand

- Recall our MNL market share function $s_{jt} = \frac{e^{x_{jt}\beta - \alpha p_{jt}}}{\sum_{k=1}^{J} e^{x_{kt}\beta - \alpha p_{kt}}}$

```
  - Recall that the model can predict how *any* change in x_{jt} or p_{jt}
would affect *all* phones' market shares
  -  What is \alpha? What is \beta?
```

- What are the model's main limitations?

```
  1. Assumes all customers have the same preferences
  2. Assumes all customers have same price sensitivity
  3. IIA: Predictions become unreliable when choice sets change
  4. Requires exogenous price variation to estimate \alpha (all demand
models)
  5. Assumes iid \epsilon distribution: Convenient but unrealistic
  Modeling heterogeneity can alleviate 1-3 and enable better predictions
```

Het. Demand Models : Intuition

1. MNL estimates quality; Het MNL estimates quality & fit

```
    Recall vertical vs. horizontal product differentiation
```

2. Better "counterfactual predictions" for strategic variables that enter the model and predict sales

```
    - Pricing: price discrimination, two-part tariffs, fees, targeted
coupons
    - Advertising: Ad targeting, frequency, media, channels
    - Product: Targeted attributes, line extensions, brand extensions
    - Distribution: Partner selection, intensity/shelfspace, in-store
environment
    - M&A: Oft used in antitrust merger reviews
```

Biggest risks? Overfitting ; Misuse

```
    - Who has heard of cross-validation?
```

1. Discrete heterogeneity by segment

- Assume each customer $i = 1, ..., N$ is in exactly 1 of $l = 1, ..., L$ segments with sizes N_l and $N = \sum_{l=1}^{L} N_l$

```
- We will use 3 kmeans segments based on 6 usage variables
- We take usage variables as best available proxies for customer needs
```

- Assume preferences are uniform within segments & vary between segments

```
- Consistent with the definition of segments
```

- Replace $u_{ijt} = x_{jt}\beta - \alpha p_{jt} + \epsilon_{ijt}$ with $u_{ijt} = x_{jt}\beta_l - \alpha_l p_{jt} + \epsilon_{ijt}$
- That implies $s_{ljt} = \frac{e^{x_{jt}\beta_l - \alpha_l p_{jt}}}{\sum_{k=1}^{J} e^{x_{kt}\beta_l - \alpha_l p_{kt}}}$ and $s_{jt} = \sum_{l=1}^{L} N_l s_{ljt}$
- Alternatively, it is also possible to estimate segment memberships

```
  - Pro: don't have to define the segment memberships ex ante
  - Cons: noisy, demanding of the data; may change w time; may neglect available theory; possible numerical
problems. Need a lot of data to do this well
```

2. Continuous heterogeneity by customer attributes

- Let $w_{it} \sim F(w_{it})$ be observed customer attributes that drive demand, e.g. usage
- w_{it} is often a vector of customer attributes including an intercept
- Assume $\beta = \delta w_{it}$ and $\alpha = w_{it}\gamma$ :: δ & γ conformable matrices
- Then $u_{ijt} = x_{jt}\delta w_{it} - w_{it}\gamma p_{jt} + \epsilon_{ijt}$ and

$$s_{jt} = \int \frac{e^{x_{jt}\delta w_{it} - w_{it}\gamma p_{jt}}}{\sum_{k=1}^{J} e^{x_{kt}\delta w_{it} - w_{it}\gamma p_{kt}}} dF(w_{it}) \approx \frac{1}{N_t} \sum_i \frac{e^{x_{jt}\delta w_{it} - w_{it}\gamma p_{jt}}}{\sum_{k=1}^{J} e^{x_{kt}\delta w_{it} - w_{it}\gamma p_{kt}}}$$

```
- We usually approximate this integral with a Riemann sum
```

- What goes into w_{it}? What if $dim(x)$ and/or $dim(w)$ is large?

3. Individual demand parameters

- Assume $(\alpha_i, \beta_i) \sim F(\Theta)$

```
- Includes the Hierarchical Bayesian Logit
```

- Then $s_{jt} = \int \frac{e^{x_{jt}\alpha_i - \beta_i p_{jt}}}{\sum_{k=1}^{J} e^{x_{jt}\alpha_i - \beta_i p_{jt}}} dF(\Theta)$
- Typically, we assume $F(\Theta)$ is multivariate normal, for convenience, and estimate Θ
  - We usually have to approximate the integral, often use Bayesian techniques (MSBA/PhD)
  - Or, we can estimate F but that is very very data intensive
  - In theory, we can estimate all (α_i, β_i) pairs without $\sim F(\Theta)$ assumption, but requires numerous observations & sufficient variation for each i. Most data intensive

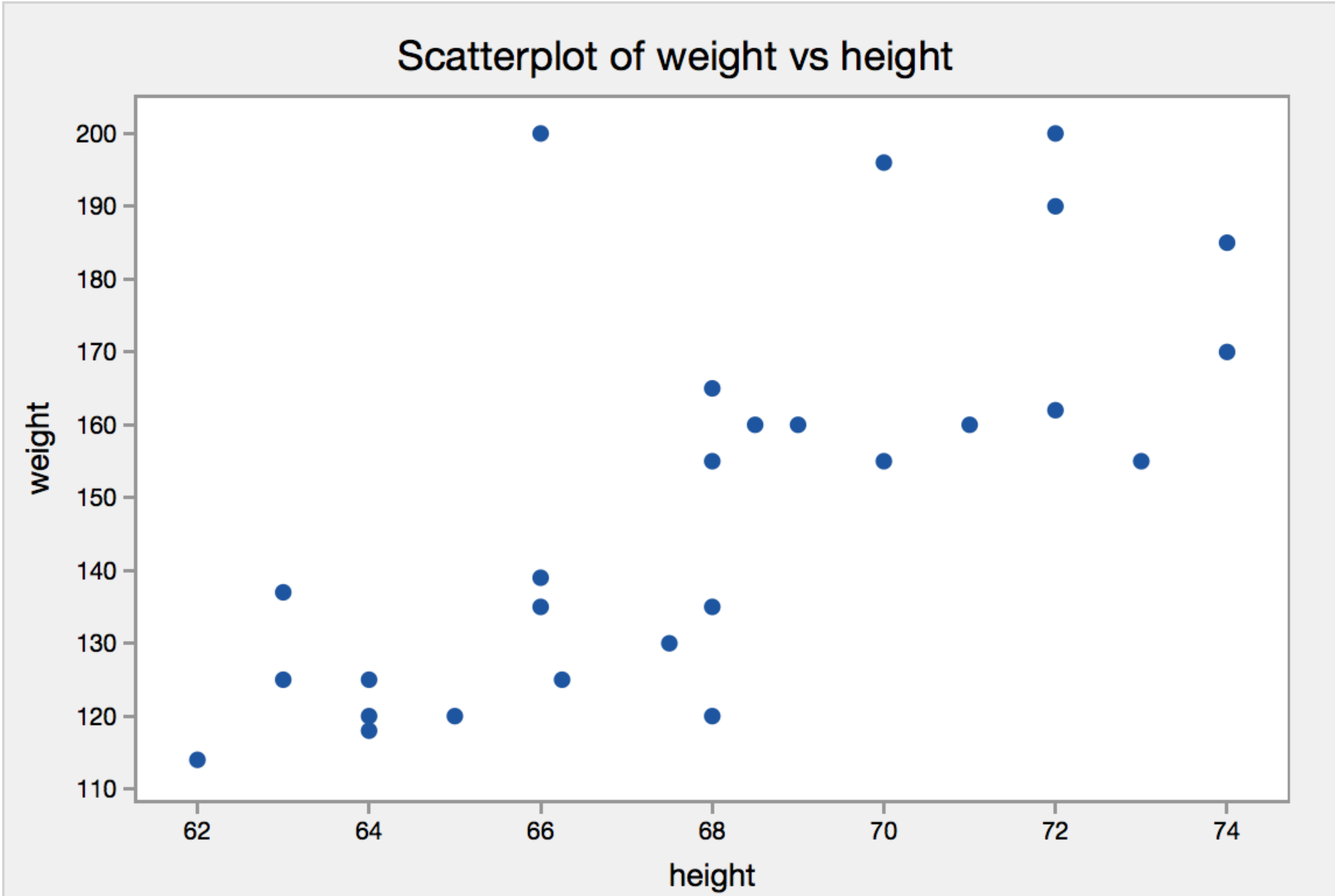
How to choose?

Humans choose the model. How do you know if you specified the best model?

```
        - "All models are wrong. Some models are useful" (Box)
        - “Truth is too complicated to allow anything but approximations.”
(von Neumann)
        - "The map is not the territory" (Box)
        - "Scientists generally agree that no theory is 100% correct. Thus,
the real test of knowledge is not truth, but utility" (Hariri)
        - No model is ever "correct," No assumption is ever "true" (why not?)
```

- Model selection: A Judgment Problem

```
        - How do you choose among plausible specifications?
        - Involves both model selection--which f() in y=f(x)--and covariate
   selection
        - Use modeling purpose and constraints as model selection criterion
        - What are our demand modeling objectives?
```


Scatterplot of weight vs height
weight
200
190
180
170
160
150
140
130
120
110
62
64
66
68
70
72
74
height

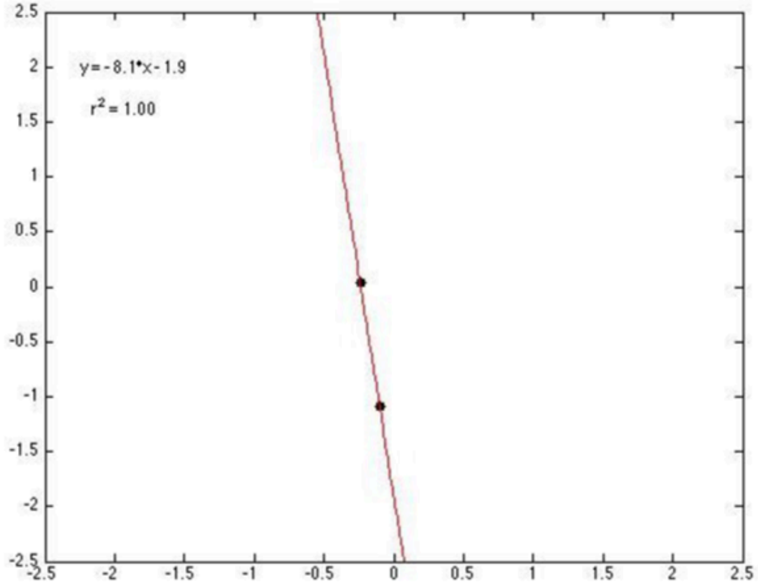
Model specification

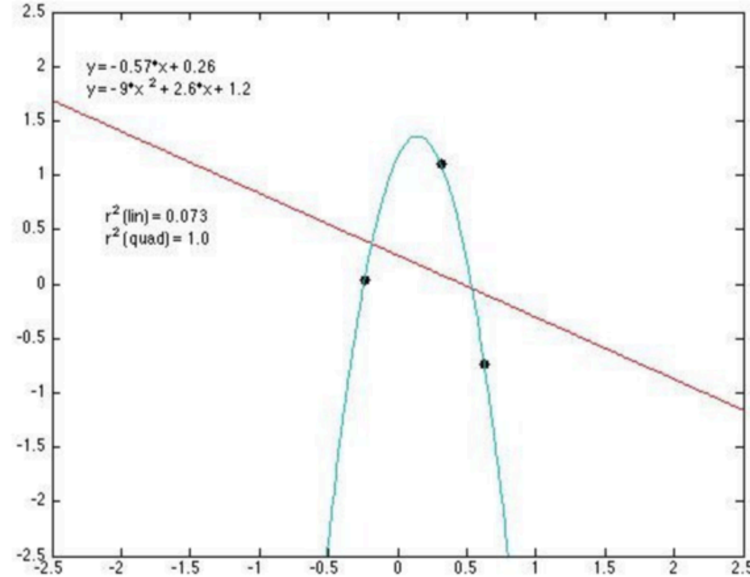
- Bias-variance tradeoff

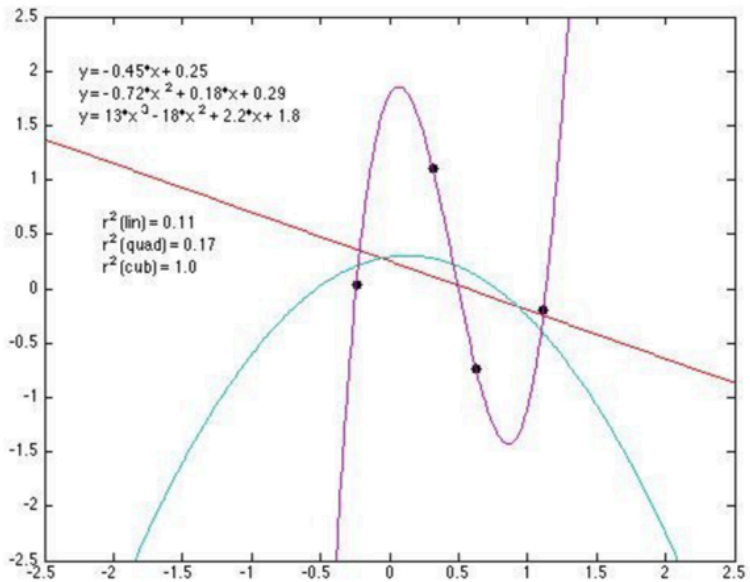
```
- Adding predictors always increases model fit
- Yet parsimony often improves predictions
```

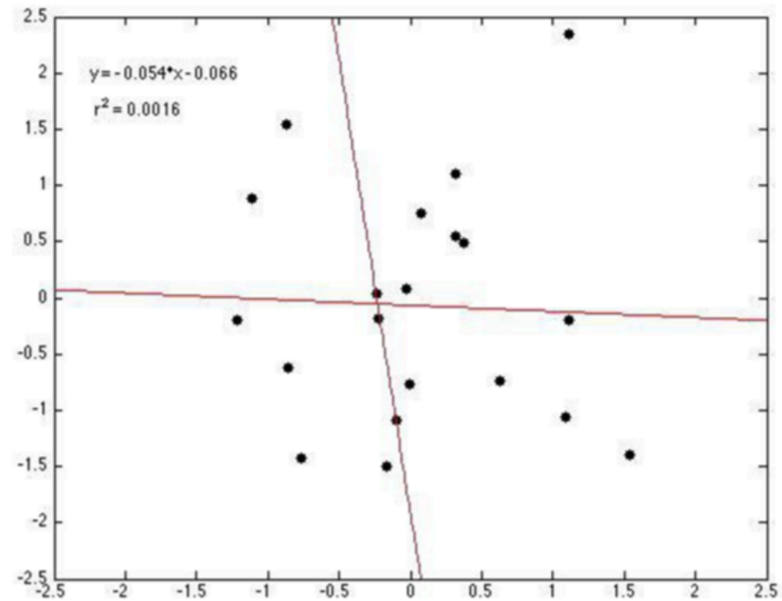
- Many criteria drive model selection

```
- Modeling objectives
- Theoretical properties
- Model flexibility
- Precedents & prior beliefs
- In-sample fit
- Prediction quality
- Computational properties
```


y = - 8.1*x - 1.9
r^2 = 1.00


y = - 0.57*x + 0.26
y = - 9*x ^{2} + 2.6*x + 1.2
r^2 (lin) = 0.073
r^2 (quad) = 1.0


y = - 0.45*x + 0.25
y = - 0.72*x ^{2} + 0.18*x + 0.29
y = 13*x ^{3} - 18*x ^{2} + 2.2*x + 1.8
r^2 (lin) = 0.11
r^2 (quad) = 0.17
r^2 (cub) = 1.0


y = - 0.054*x - 0.066
r^2 = 0.0016

How to evaluate overfitting?

- Retrodiction = “RETROspective preDICTION”

```
      - Knowing what happened enables you to evaluate prediction quality
      - We can compare different models and different specifications on
retrodictive accuracy
```

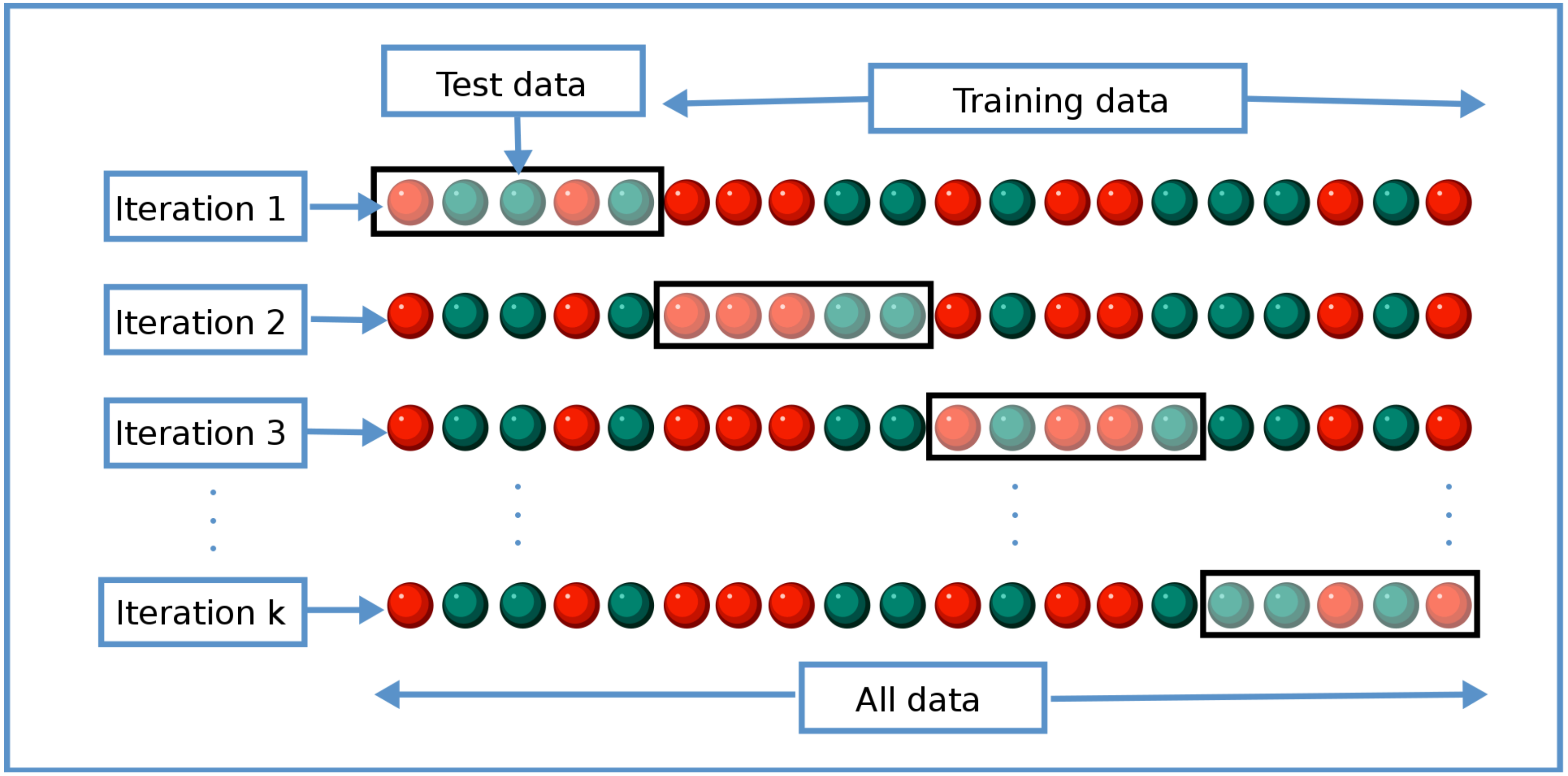
- We can even train a model to maximize retrodiction quality (“Cross-validation”)

```
      - Most helpful when the model's purpose is prediction
      - More approaches: Choose intentionally simple models
      - Penalize the model for uninformative parameters: Lasso, Ridge,
Elastic Net, etc.
```

Cross-validation

- General approach to evaluate retrodiction performance and overfitting risk among a set of competing models $m = 1, \ldots, M$. Algorithm:

1. Randomly divide the data into K distinct folds
2. Hold out fold k, use remaining $K - 1$ folds to estimate model m, then predict outcomes in fold k; store prediction errors
3. Repeat 2 for each k
4. Repeat 2&3 for every model m
5. Retain model m with minimal prediction errors, usually MAPE or MSPE



- You estimate every model K times (K=4 in the graphic)
- Each estimation uses a different (K-1)/K proportion of the data
- We evaluate each model's retrodiction quality K times, then average them
- When K=N, we call that "leave-one-out" cross-validation
- Important: cross-validation is just one tool in the box. Not the only

Ex-post evaluations

- Can a model be robust to major changes in the data-generating process?
- Non-random holdouts are strong tests, but can only be retrospective

informs
http://pubsonline.informs.org/journal/mnsc/

MANAGEMENT SCIENCE
Articles in Advance, pp. 1–21
ISSN 0025-1909 (print), ISSN 1526-5501 (online)

Dynamic Quality Ladder Model Predictions in Nonrandom Holdout Samples

Linli Xu,[a] Jorge M. Silva-Risso,[b] Kenneth C. Wilbur[c]

[a] Carlson School of Management, University of Minnesota, Minneapolis, Minnesota 55455; [b] A. Gary Anderson Graduate School of Management, University of California, Riverside, Riverside, California 95251; [c] Rady School of Management, University of California, San Diego, La Jolla, California 92093

Contact: linlixu@umn.edu, http://orcid.org/0000-0001-6832-0981 (LX); jorge.silva-risso@ucr.edu (JMS-R); kennethcwilbur@gmail.com (KCW)

Received: April 12, 2015
Revised: October 6, 2015; December 7, 2016
Accepted: January 19, 2017
Published Online in Articles in Advance: July 28, 2017

https://doi.org/10.1287/mnsc.2017.2780

Abstract. In light of recent calls for further validation of structural models, this paper evaluates the popular dynamic quality ladder (DQL) model using a nonrandom holdout approach. The model is used to predict data following a regime shift—that is, a change in the environment that produced the estimation data. The prediction performance is evaluated relative to a benchmark vector autoregression (VAR) model across three automotive categories and multiple prediction horizons. Whereas the VAR model performs better in all scenarios in the compact car category, the DQL model tends to perform better on multiple-year horizons in both the midsize car and full-size pickup categories. A supplementary data analysis suggests that DQL model performance in the nonrandom holdout prediction task is better in categories that are more affected by the regime shift, helping to validate the usefulness of the dynamic structural model for making predictions after policy changes.

History: Accepted by Matthew Shum, marketing.

Keywords: automobiles • product quality • dynamic oligopoly competition • product innovation • nonrandom holdout validation

1. Introduction

Structural dynamic oligopoly models have enabled re- … oligopoly literature, in particular, has not emphasized empirical model validation.[2]

Page 2 / 22

Het demand: Misuse Risks

- Customer attributes should reflect differences in customer needs
- Customer data should be high quality (GIGO, Errors-in-variables biases)
- Use needs to consider qualitative factors {effectiveness, legality, morality, privacy, conspicuousness, equity, reactance}
  - Guiding principle (not a rule):
    Using data to legally, genuinely serve customers' interests is usually OK
  - Using private data against customer interest can harm some consumers, break laws, incur liability. One lawsuit can kill a start-up
  - Major US laws: COPPA, GLBA, HIPAA, patchwork of state laws
- Adding heterogeneity to a demand model does not resolve price endogeneity. Still need exogenous price variation

Some evidence

- How does demand model performance depend on specification and training data?

Research question

- Suppose we
  1. Train demand model M to predict mayonnaise sales …
  2. … using information set X …
  3. … & choose targeted discounts for each consumer to maximize firm profits

```
- Essentially 3rd-degree price discrimination
```

- Separately, using different data, we nonparametrically estimate how each individual responds to price discounts

```
- This gives us ground-truth to assess each household's response to price discount
- But, the nonparametric estimate can't give counterfactual predictions; we need M for that
```

- How do targeted coupon profits depend on M and X?

```
      - We use model M and data X to predict profits of offering targeted price discounts to particular
households
      - We use ground-truth to calculate household response, then calculate profits across all households
      - We'll also compare to no-discount and always-discount strategies
```

- Optimal Price Targeting (MkSc 2022)

Lil bit of theory

- For any price discount < contribution margin, giving a targeted discount to…
  - … our own brand-loyal customer directly reduces profit
  - … a marginal customer may increase profit
  - … another brand's loyal customer does not change profit
- So the demand model's challenge is to distinguish marginal customers from loyal customers

```
- This research disregards the `post-promotion dip' for simplicity
```

Information sets X

1. Base Demographics:
   Income, HHsize, Retired, Unemployed, SingleMom
2. Extra Demographics: Age, HighSchool, College, WhiteCollar, #Kids, Married, #Dogs, #Cats, Renter, #TVs
3. Purchase History: BrandPurchaseShares, BrandPurchaseCounts, DiscountShare, FeatureShare, DisplayShare, #BrandsPurchased, TotalSpending

Demand Models M

1. Bayesian Logit models (3)

```
      - Based on utility maximization in which consumers compare utility
and price of each available product
      - Includes Hierarchical and Pooled versions
```

2. Multinomial Logit Regressions (2)

```
      - Estimated via Lasso and Elastic Net to reduce overfitting
```

3. Neural Network (2)

```
      - Including single-layer and deep NN
```

4. KNN: Nearest-Neighbor Algorithm (1)

5. Random Forests (2)

```
      - Including standard RF for bagging and XGBoost for boosting
```

How do we answer the question?

1. Economic criteria:
   What profit does each M-X combination imply?

```
 - Depends on counterfactual predictions: What if we had selected different customers to receive coupons?
 - Quantifies prediction quality in profit terms
```

2. Statistical criteria:
   How well does each M-X fit its training data?

```
 - Generally, what the models are generally trained to maximize
```

Economic and statistical criteria can be very different

```
 - Doing well on one does not imply doing well on the other
 - Which one do we care more about in customer analytics?
```

Panel (I): Average Profits (Per 100 Customers)

	Base Demos	Base Demos Extra Demos	Base Demos Extra Demos Purch. Hist.
Bayesian Hierarchical Logit			
– normal heterogeneity	7.53 (0.47)	7.48 (0.44)	
– mixtures of normals heterogeneity	7.38 (0.46)	7.29 (0.43)	
Bayesian Pooled Logit	5.74 (0.42)	5.77 (0.43)	7.14 (0.40)
Lasso	5.77 (0.41)	6.53 (0.43)	7.27 (0.41)
Elastic Net	5.50 (0.40)	6.57 (0.43)	7.27 (0.41)
Neural Network	5.87 (0.41)	6.35 (0.35)	7.41 (0.44)
Deep Neural Network	5.28 (0.40)	6.47 (0.38)	7.26 (0.42)
KNN	4.47 (0.26)	6.00 (0.30)	7.07 (0.41)
Random Forest	5.24 (0.30)	5.60 (0.27)	6.33 (0.35)
XGBoost	5.42 (0.44)	5.50 (0.42)	7.17 (0.43)
Blanket Coupon	6.45 (0.32)		
No Coupon	5.50 (0.42)		

Panel (II): Out-of-Sample Hit Probabilities			
Bayesian Hierarchical Logit			
– normal heterogeneity	0.929	0.929	
– mixtures of normals heterogeneity	0.929	0.929	
Bayesian Pooled Logit	0.927	0.928	0.929
Lasso	0.927	0.928	0.930
Elastic Net	0.927	0.927	0.930
Neural Network	0.928	0.930	0.931
Deep Neural Network	0.925	0.925	0.926
KNN	0.929	0.927	0.930
Random Forest	0.941	0.933	0.931
XGBoost	0.930	0.931	0.932

Table 4: **Targeting Policy Profits and Model Fit.** Panel (I) reports the average customer-level profits (scaled by 100) from each targeted pricing policy. Bootstrapped standard errors are reported in parentheses. Panel (II) reports the out-of-sample hit probabilities from each model.

Panel (II): Predictors of Pricing Policies

	Dependent Variable: Discount Dummy							
			Elastic		Deep		Random	
Model	Logit	Lasso	Net	NNet	NNet	KNN	Forest	XGBoost
Chain B Dummy	-0.015	-0.019	-0.017	-0.072***	-0.045*	-0.043*	0.01	0.059*
	(0.022)	(0.020)	(0.020)	(0.020)	(0.020)	(0.020)	(0.023)	(0.023)
Income (in $10,000s)	-0.015***	-0.001	-0.001	0.005	-0.001	0.002	-0.006	0.001
	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)
Family Size	-0.019	-0.016	-0.017	-0.013	-0.017	-0.017	-0.009	-0.022*
	(0.010)	(0.009)	(0.009)	(0.009)	(0.009)	(0.009)	(0.011)	(0.011)
Retired	-0.154***	-0.001	0.024	-0.034	-0.108*	-0.092*	-0.073	0.027
	(0.047)	(0.043)	(0.043)	(0.044)	(0.043)	(0.043)	(0.049)	(0.050)
Unemployed	-0.016	-0.01	-0.007	-0.008	-0.006	-0.019	0.016	0.002
	(0.027)	(0.025)	(0.024)	(0.025)	(0.025)	(0.024)	(0.028)	(0.028)
Single Mother	0.061	-0.008	-0.012	0.009	-0.035	-0.014	-0.025	-0.035
	(0.041)	(0.038)	(0.037)	(0.038)	(0.037)	(0.037)	(0.043)	(0.043)
Hellmann's Choice Share	-0.242***	-0.037	-0.04	-0.100*	0.003	0.001	-0.182***	-0.013
	(0.041)	(0.039)	(0.038)	(0.039)	(0.038)	(0.038)	(0.044)	(0.044)
Hellmann's Choice Count	0.052***	0.007	0.003	-0.031***	0.004	0.001	-0.01	0.028**
	(0.009)	(0.008)	(0.008)	(0.008)	(0.008)	(0.008)	(0.009)	(0.010)
Discount Share	0.783***	0.972***	0.991***	0.999***	0.926***	0.955***	0.812***	0.799***
	(0.031)	(0.029)	(0.029)	(0.030)	(0.029)	(0.029)	(0.033)	(0.034)
# Brands Purchased	-0.184***	-0.015	-0.024	-0.041	0.070**	0.037	0.001	0.088**
	(0.029)	(0.027)	(0.027)	(0.027)	(0.027)	(0.027)	(0.031)	(0.031)
Total Spending	-0.013***	-0.005**	-0.004*	0.002	-0.006***	-0.005**	-0.005**	-0.001
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
Observations	1,162	1,162	1,162	1,162	1,162	1,162	1,162	1,162

Table 5: **Policy Comparisons.** Panel (I) reports discount frequencies and the frequency of agreement between pricing policies. Panel (II) reports a linear probability regression of discount incidence on a selected set of customer characteristics. The unit of observation is a customer/chain combination. Significance codes: *p<0.05; **p<0.01; ***p<0.001.

Takeaways

- To predict behavior, use past behavior
- Economic theory can help demand models to perform well with limited behavioral data
- ML model performance depends critically on data quality & abundance. Counterfactual predictions do not always outperform economic models
- Statistical performance $\neq$ economic performance

Conjoint Analysis

- Generates stated-preference data to estimate heterogeneous demand model, to enable counterfactual predictions and optimal product designs
- Probably the most popular quant marketing framework: >10k studies/year (Sawtooth 2008)

Choosing Product Attributes

- Until now, we studied existing product attributes
  - What about choosing new product attribute levels?
  - Or what about introducing new products?
- Enter conjoint analysis: Attributes are **Con**sidered **joint**ly
  Survey and model to estimate attribute utilities

```
- Autos, phones, hardware, durables
- Travel, hospitality, entertainment
- Professional services, transportation
- Consumer package goods
```

- Combines well with cost data to select optimal attributes

Conjoint analysis implementation

1. Identify K product attributes and levels/values x_k : These constitute points in your attribute space

```
- Screen size: 5.5", 6", 6.5", 7"
- Memory: 8 GB, 16 GB, 32 GB, 64 GB, 128 GB
- Price: $199, $399, $599, $799, $999
```

2. Recruit consumer participants to make choices

```
- Choose a representative sample of your target market
- Offer 8-15 choices among 3-5 hypothetical attribute bundles
```

3. Sample from product space, record consumer choices
4. Specify model, i.e. $U_j = \sum_k x_{jk}\beta_k - \alpha_k p_j + \epsilon_j$ and $P_j = \frac{\sum_k x_{jk}\beta_k - \alpha_k p_j}{\sum_l \sum_k x_{lk}\beta_k - \alpha_k p_l}$

```
- Beware: p is price , P is choice probability or market share
```

5. Calibrate choice model to estimate attribute utilities
6. Combine estimated model with cost data to choose product locations and predict outcomes

Sample choice task

Study Attributes/Levels

Brand
Ford
Chevrolet
Dodge

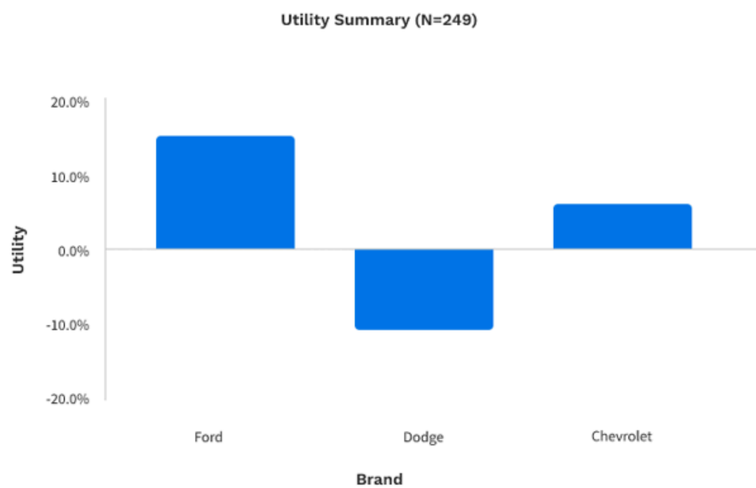
Type
Truck
SUV

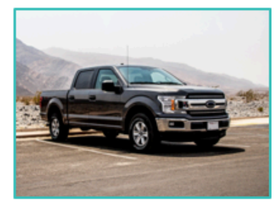
Engine
V6, 3.6 Liter
V8, 3.6 Liter
V8, 4.2 Liter

Price
$36,999
$42,999
$45,999
$49,999

If you were in the market for your next vehicle and these were the only alternatives, which would you choose?

Brand:	Ford	Chevy	Ford	Dodge
Type:	SUV	Truck	Truck	Truck
Engine:	V6, 3.6 liter	V8, 5.7 liter	V8, 4.2 liter	V8, 4.8 liter
Price:	$36,599	$42,999	$42,999	$45,999
	Selected	Select	Select	Select





Case study: UberPOOL

- In 2013, Uber hypothesized

```
- some riders would wait and walk for lower price
- some riders would trade pre-trip predictability for lower price
- shared ridership could ↓ average price and ↑ quantity
- more efficient use of drivers, cars, roads, fuel
```

Type of Trip	Illustration	Cost of Trip	Rider Price
Solo Trip (UberX)	A → B	\$12	\$12
Imperfect POOL	A1 → A2 → B	\$12	\$8 + \$4
The Perfect POOL	A → B	\$12	\$12/3 = \$4

Figure 2. Unit Economics of uberPOOL in 3 Theoretical Scenarios. 'Cost of trip' consists of the driver payout for providing the service. 'Rider price' is the price that the rider will have to pay. Numbers across the 3 theoretical scenarios are for illustrative purposes only.

Business case was clear! But ...

- Shared rides were new for Uber

```
      - Rider/driver matching algo could reflect various tradeoffs
      - POOL reduces routing and timing predictability
```

- Uber had little experience with price-sensitive segments

```
      - What price tradeoffs would incentivize new behaviors?
      - How much would POOL expand Uber usage vs cannibalize other
services?
```

- Coordination costs were unknown

```
      - "I will never take POOL when I need to be somewhere at a specific
time"
      - Would riders wait at designated pickup points?
      - How would comunicating costs upfront affect rider behavior?
```

- Uber used market research to design UberPOOL

Approach

1. 23 in-home diverse interviews in Chicago and DC

   - Interviewed {prospective, new, exp.} riders to (1) map rider's regular travel, (2) explore decision factors and criteria, (3) a ride-along for context
   - Findings identified 6 attributes for testing

2. Online Maximum Differentiation Survey

   - Selected participants based on city, Uber experience & product; N=3k, 22min

Maxdiff results

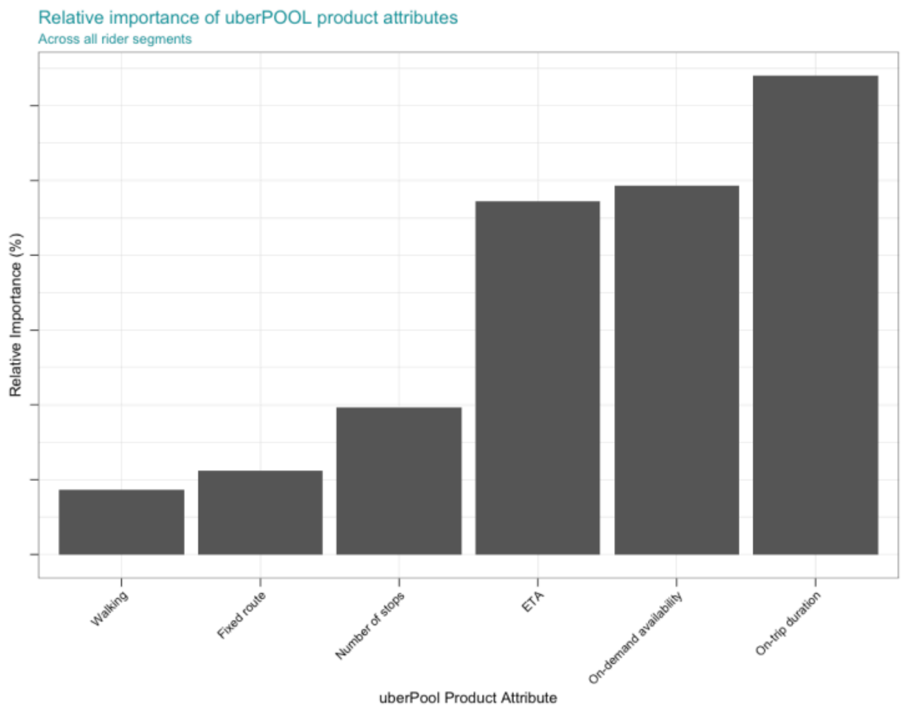


Figure 7. This plot shows the relative difference in the importance of feature utilities, as derived from the maxdiff analysis. Exact percentages have been abstracted to protect business insights, but all the features in the plot add to 100%.

Conjoint Attribute Space

Table 1. Conjoint Features and Levels - Research Design Matrix

Feature	Level 1	Level 2	Level 3	Level 4	Level 5
Estimated time of arrival	Request now and wait 5 mins	Request now and wait 10 mins	Book 15 mins ahead	Book 30 mins ahead	
Walking	No walking	Walk 1-block	Walk 2-3 blocks		
Trip length multiplier	1x	1.1x	1.2x	1.3x	1.4x
Trip variance multiplier	1.1x	1.2x	1.3x		
Discount multiplier	Very low	Low	Medium	High	Very high

Conjoint Sample Question

Look over the packages below and select the option that you would be most likely to use to commute to work.

	Option 1	Option 2	Option 3
Time to pickup	• Request now and **wait 10 mins**	• Book **15 mins** ahead	• Take my existing commute option
Walking to pickup or dropoff point	• Walk 2-3 blocks	• Walk 2-3 blocks	
Time on trip	• 22 to 26 mins	• 28 to 31 mins	
Trip cost	• $7.95	• $9.94	

<<< >>>

Figure 8. Example conjoint question.

Conjoint Model

The model was estimated using the bayesm package in the R programming language (Rossi & Allenby 2009).[15] The model can be described by the following specification:

$$y_i \sim Multinomial(Pr(X_i, \beta_i))$$
$$i = 1, \ldots, n\ units$$
$$\beta_i \sim \Delta' z_i + u_i$$
$$u_i \sim N(0, V_\beta)$$

Where the probability of choosing product y for respondent i is distributed multinomially as a function of covariates X_i and β_i. The part-worth estimates (β_i) are distributed logit parameters over respondent units with mean $\Delta' z_{,}$, with $\Delta' z_{.i}$ being a matrix containing mean-centered control variables for each respondent, with errors (u_i) that are normally distributed with variance V_β (Rossi & Allenby 2009). The posterior distribution of β_i is used to determine the overall utility of each product feature and is the main quantity of interest of the analysis. Researchers used the model's log-likelihood as a measure of goodness of fit, which converged successfully after 100,000 simulations.[16,17]

Control variables used in the analysis include historic Uber usage data, such as the home city of the respondent, rider tenure in days since signing up for an account, lifetime billings, as well as survey-based variables such as the time it took a respondent to complete the survey and demographic information. No other behavioral features were used in the analysis, per Uber's policy of respecting the privacy of user data (Privacy Policy - Uber).

Conjoint Findings

Product Redesign

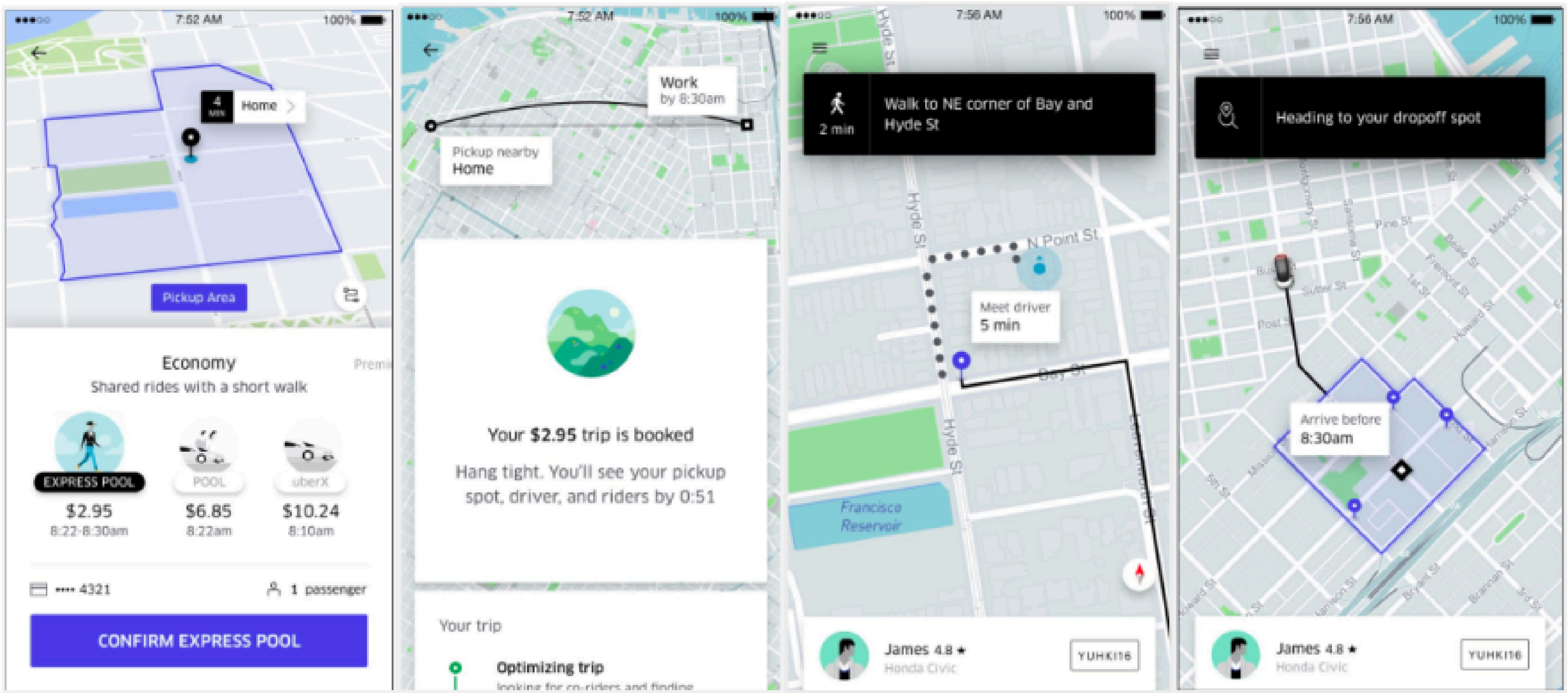


Figure 11. End-to-end Designs of Express POOL - Uber's new shared ride experience with walking and waiting. Designs as of Nov 2017.

Business Results

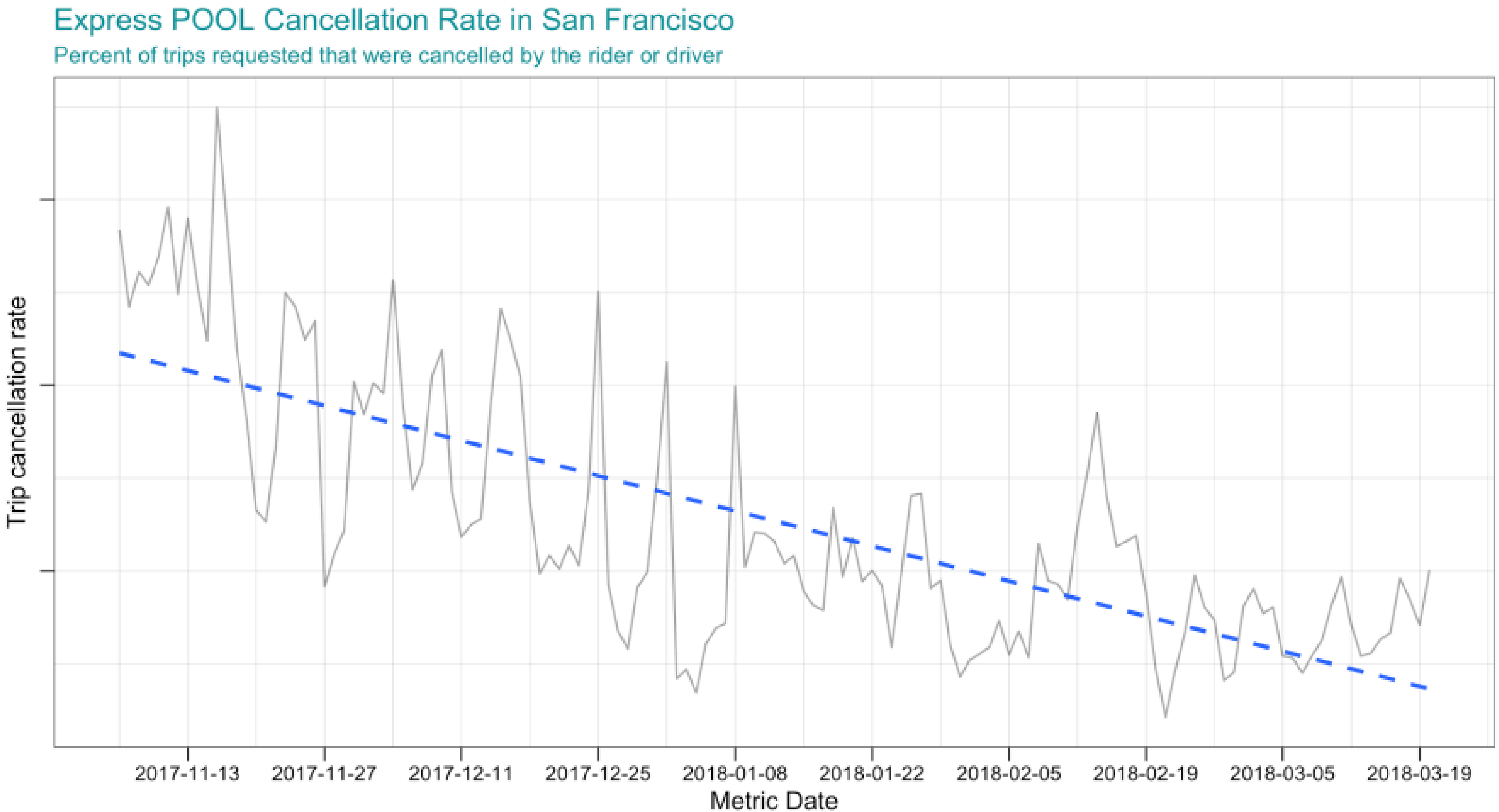
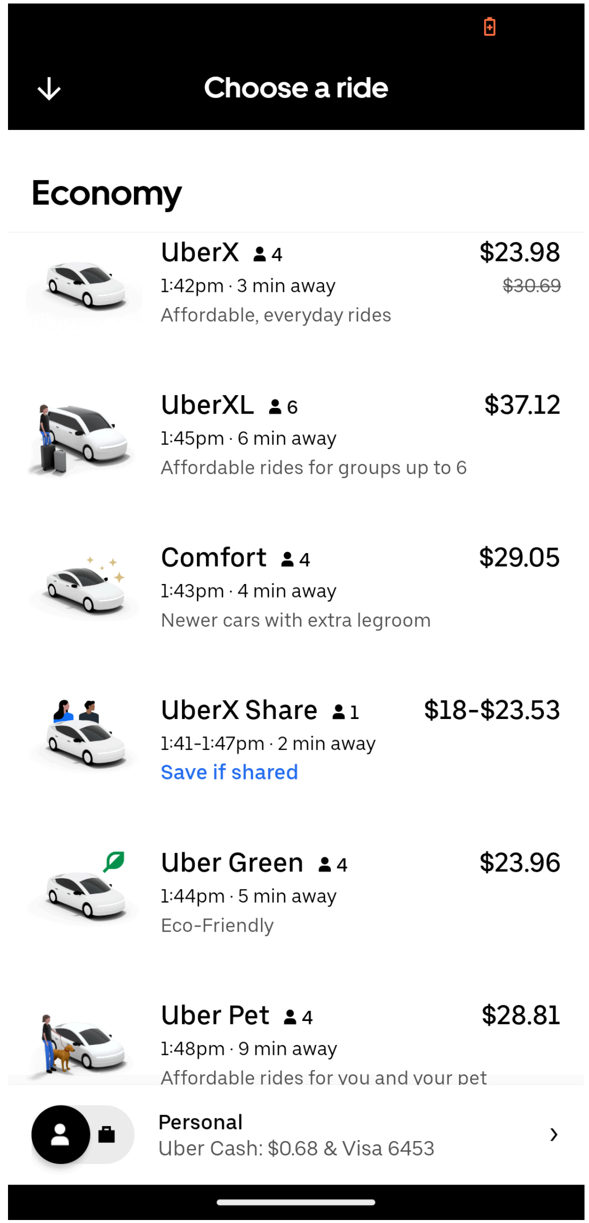


Figure 12. Express POOL Cancellation Rate

More Products

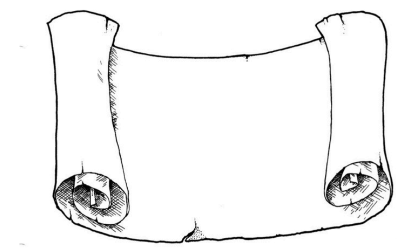


Conjoint: Limitations, Workarounds

Wrapping up

Class script

- Add heterogeneity to MNL model
- Individual-level heterogeneity via price-minutes interaction
- Segment-level heterogeneity via segment-attribute interactions
- Both



Recap

- Heterogeneous demand models enable personalized and segment-specific policy experiments
- Demand models can incorporate discrete, continuous and/or individual-level heterogeneity structures
- Heterogeneous demand models fit better, but beware overfitting and misuse
- Conjoint analysis uses stated-preference data to map markets and predict profits of product locations in attribute space

Going further

- Train (2009), Chapters 7-12
- Reconciling modern machine learning practice and the bias-variance trade-off
- MGT 108R to design & run conjoint analyses
- Conjoint literature is huge. Good entry points: Chapman 2015, Ben-Akiva et al 2019, Green 2022, Allenby et al. 2019