AC297r: IACS Capstone Project

Fall 2020, in collaboration with Wayfair

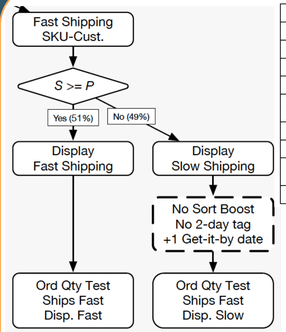
Impact of Fast Shipping on Sales

Yixing Guan, Paul-Emile Landrin, Jordan Turley, Lin Zhu

Overview and Objectives

1) **Estimate the average boost in sales** on different scopes
2) **Measure the marginal effect of sales boost**: how the sales boost changes as the portion of products offered with fast shipping increases.
3) **Identify key product characteristics** that will make the product sales **more sensitive to fast shipping.**

Experimental Design



Item	S-value	Fast Shipping	Attr
Couch X	25	False	...
Couch X	25	True	...
Couch X	26	False	...
Couch X	26	True	...
⋮	⋮	⋮	⋮
Couch X	75	True	...
Chair Z	25	False	...
⋮	⋮	⋮	⋮
Chair Z	75	True	...

Attributes: 1) Impression count
2) Average sort rank
3) Product display page click count
4) Order count
5) Quantities ordered
6) Other generic product attributes…

Treatment Effect Definition

- $Y_i(t_i)$: observed outcome for the chosen treatment → number of orders (numord), the quantity ordered (qtyord), etc.
- t_i: the treatment → fast shipping flag (0 or 1)
- x_i : covariates used for heterogeneity (i.e., the variable that interacts with the treatment) → **S-value**
- w_i : other parameters that might affect our response variables → generic product attributes like previous sales, rating, price, etc.

We assume the outcome **Y** is linear in the treatment vector, and we assume the following structure for our target quantity **τ** (heterogeneous treatment effect):

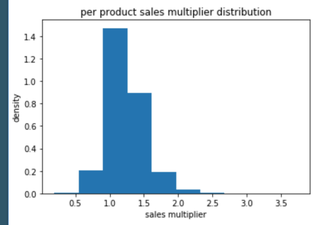
$$Y = H(X,W)T + g(X,W,\epsilon)$$
$$T = f(X,W,\eta)$$
$$\tau(t1,t0,x) = E[H(X,W) \mid X = x](t_1 - t_0)$$

Here **η** and **ϵ** are noise factor added, and **H, g, f** are functions.

Task 1: average boost in sales

Aggregation: total number of orders

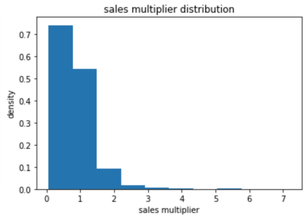
25% more sales



Issues:
Not factoring in the increase product exposure due to the sort-boost effect of fast-shipping.

Conversion rate: nb. of orders / nb. of impressions

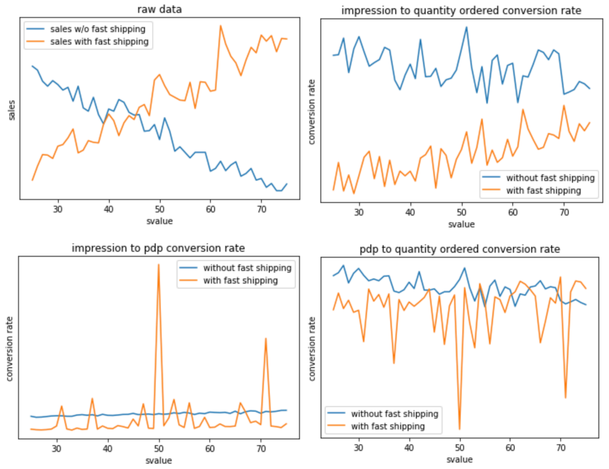
46% more sales



Issues:
Only 30% products receive sales boost under this framework. Estimate only applies to these products.

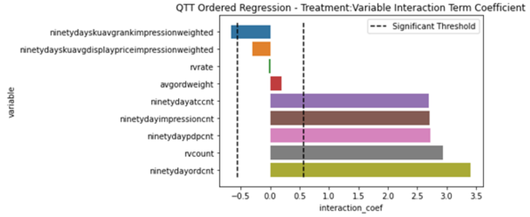
Task 2: impact of S-value

How various quantities of interest change as S-value changes:



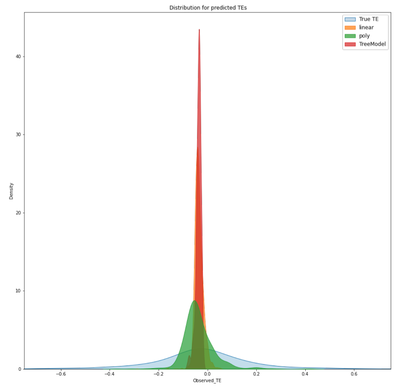
Task 3: important product features

Treatment Effect: absolute increase in number of orders
Baseline Model -- Linear Regression



Treatment Effect: increase in conversion rate
Model Performance Comparison

Variable	Model	MSE Score	95% Confidence Interval Accuracy
$TE(log(\frac{N_{atc}}{N_{pdp}}))$	Causal Forest	0.097	15.8%
$TE(log(\frac{N_{atc}}{N_{impression}}))$	Causal Forest	0.087	17.2%
$TE(log(\frac{N_{atc}}{N_{impression}}))$	LinearDRL with Polynomial Feature	0.06	42%



Coefficient Results

	point_estimate	stderr	zstat	pvalue	ci_lower	ci_upper
ninetydayskuavgdisplaypriceimpressionweighted	-0.004	0.018	-0.2	0.841	-0.033	0.026
ninetydayskuavgrankimpressionweighted	0.014	0.056	0.254	0.8	-0.077	0.105
rvrate	-0.02	0.019	-1.043	0.297	-0.052	0.012
avgordweight	0.006	0.016	0.4	0.689	-0.02	0.033
ninetydayordcnt	-0.023	0.017	-1.324	0.186	-0.061	0.006
ninetydayskuavgdisplaypriceimpressionweighted ninetydayskuavgrankimpressionweighted	-0.014	0.03	-0.462	0.644	-0.063	0.035
ninetydayskuavgdisplaypriceimpressionweighted rvrate	0.01	0.022	0.432	0.666	-0.027	0.046
ninetydayskuavgdisplaypriceimpressionweighted avgordweight	-0.004	0.006	-0.642	0.521	-0.013	0.006
ninetydayskuavgdisplaypriceimpressionweighted ninetydayordcnt	0.004	0.008	0.56	0.576	-0.008	0.017
ninetydayskuavgrankimpressionweighted rvrate	0.022	0.034	0.635	0.526	-0.034	0.078
ninetydayskuavgrankimpressionweighted avgordweight	-0.013	0.028	-0.461	0.645	-0.059	0.033
ninetydayskuavgrankimpressionweighted ninetydayordcnt	-0.029	0.026	-1.096	0.273	-0.071	0.014
rvrate avgordweight	-0.002	0.015	-0.11	0.913	-0.027	0.023
rvrate ninetydayordcnt	-0.004	0.01	-0.413	0.68	-0.021	0.013
avgordweight ninetydayordcnt	-0.001	0.006	-0.187	0.852	-0.012	0.009

Key findings: For popular products, adding fast shipping label won't have a significant contribution in boosting conversion rate.

Conclusion

Aggregation-based approach:

1. **25%** increase in sales.

2. We can expect that order number and quantities sold scale **linearly** as s-value increases.

Conversion rate-based approach:

1. **46%** more sales for products whose conversion rates does not degrade with fast-shipping.

2. The effect of fast-shipping has a much larger impact on the **conversion rate from product display page clicks to quantities ordered**, than the conversion rate from the impression to product display page clicks. The conversion rate from impression to quantities ordered with fast-shipping increases as s-value increases.

3. For popular products, adding fast shipping label won't have a significant contribution in boosting conversion rate.

Acknowledgements

We would like to thank Nathaniel Burbank (Wayfair), Nick Stern (Harvard), and Chris Tanner (Harvard) for their support and guidance throughout this work.

Reference

1. Chipman, H. et al. "BART: Bayesian Additive Regression Trees." *The Annals of Applied Statistics* 4 (2010): 266-298.
2. Wagner, S. and S. Athey. "Estimation and Inference of Heterogeneous Treatment Effects using Random Forests." *Journal of the American Statistical Association* 113 (2015): 1228 - 1242.
3. EconML: Python SDK, developed by the ALICE team at MSR New England
4. GRF: Generalized Random Forest, a pluggable package for forest-based statistical estimation and inference.