Product Return Policy Change

Laura Chamberlain, Laura Klipp,
Rahul Panicker, Nishita Sule, Yue Yu

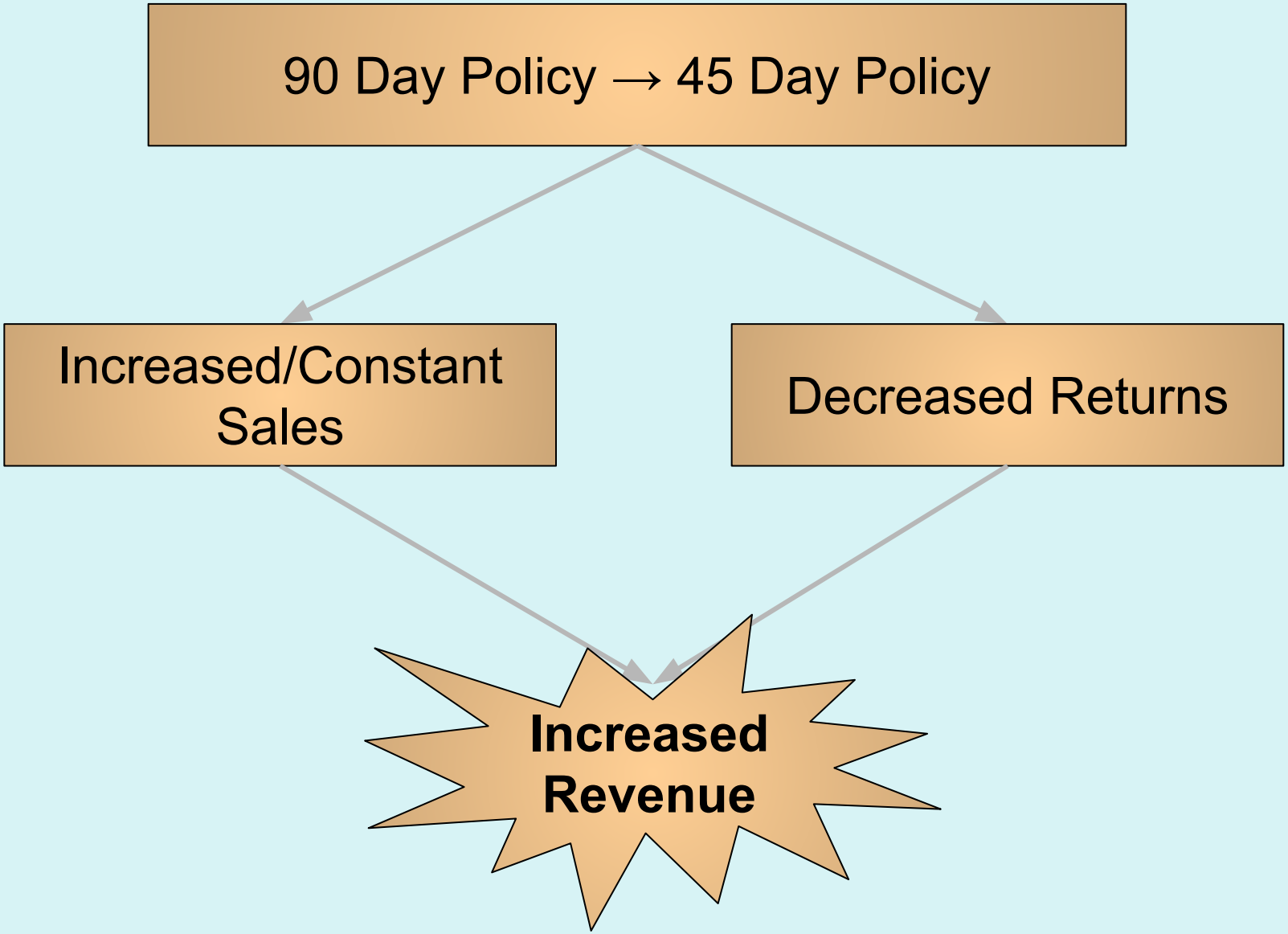
Agenda

1. Overview and Background
2. Research Questions
3. Models
4. Analysis
5. Summary

Overview

- On October 1st, 2013 our client changed their return policy from 90 days to 45 days.
- We are assessing the effect of the return policy change on the company's overall sales and returns by analyzing customer transaction data across our brands and comparing it to a sister brand whose policy did not change.

Desired Impact From Policy Change



Questions of Interest

What is the impact of decreasing the return policy from 90 days to 45 days on:

1. Online channel sales & returns
2. Physical store sales & returns
3. Online channel sales & returns by product category
4. Physical store sales & returns by product category

Dependent Variables

Sales Value

Sales Quantity

Return Value

Return Quantity

Independent Variable

Impact of Policy Change before and after October 1st, 2013.

Models Used

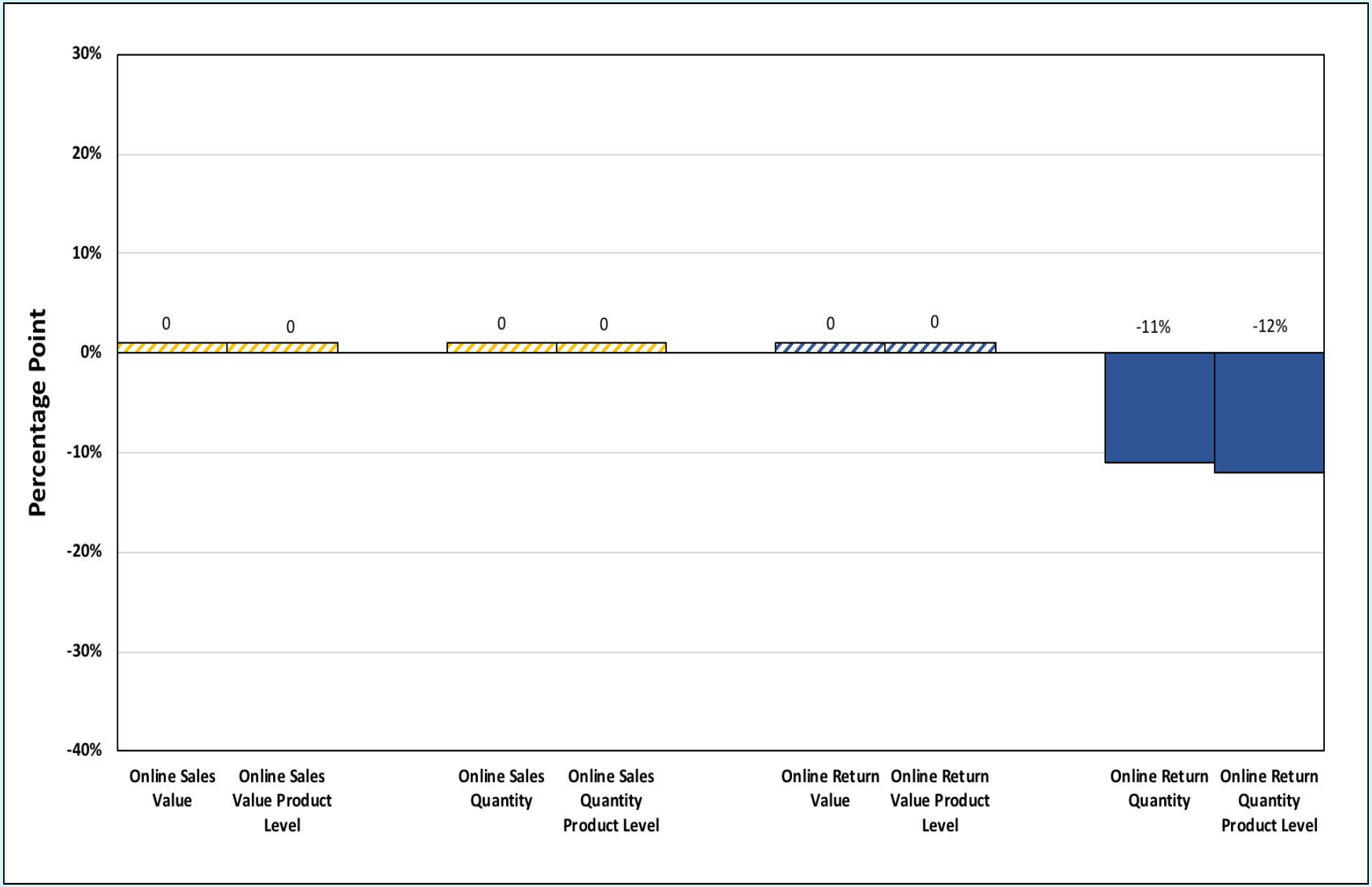
Models for overall sales and returns:

- **Sales Q/V** ~ Policy*Time + Policy + Time + Controls
- **Return Q/V** ~ Policy*Time + Policy + Time + Sales Q/V+ Controls
- **Profit** ~ Sales - Returns - Return Costs - Procurement Costs

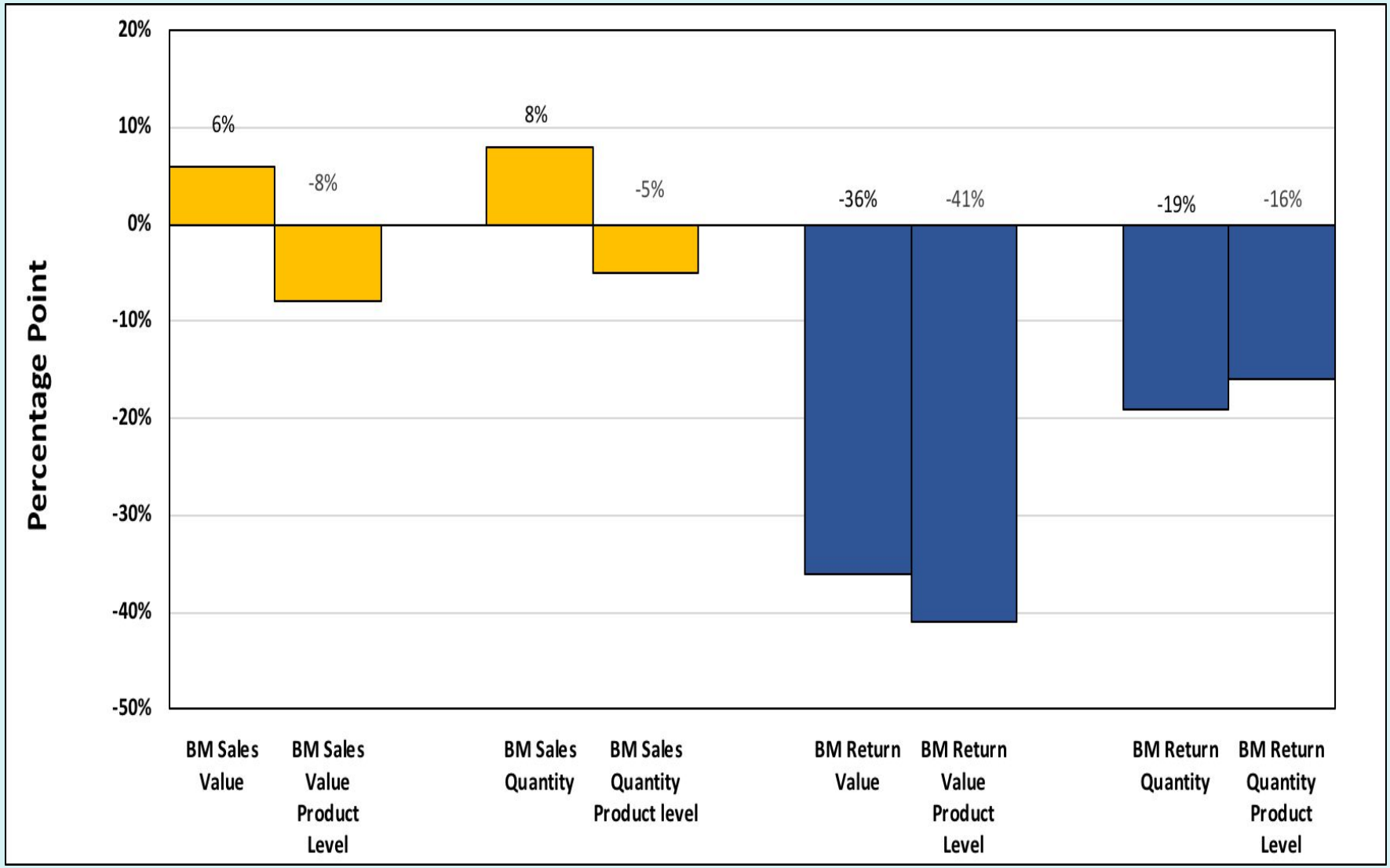
Models for sales and returns by product category (PC):

- **Sales Q/V** ~ Policy*Time*PC+ Policy + Time + PC + Controls
- **Return Q/V** ~ Policy*Time*PC + Policy + Time + PC + Sales Q/V +Controls
- **Profit** ~ Policy*Time + Policy + Time + Controls

Impact on Online Channel Sales and Returns



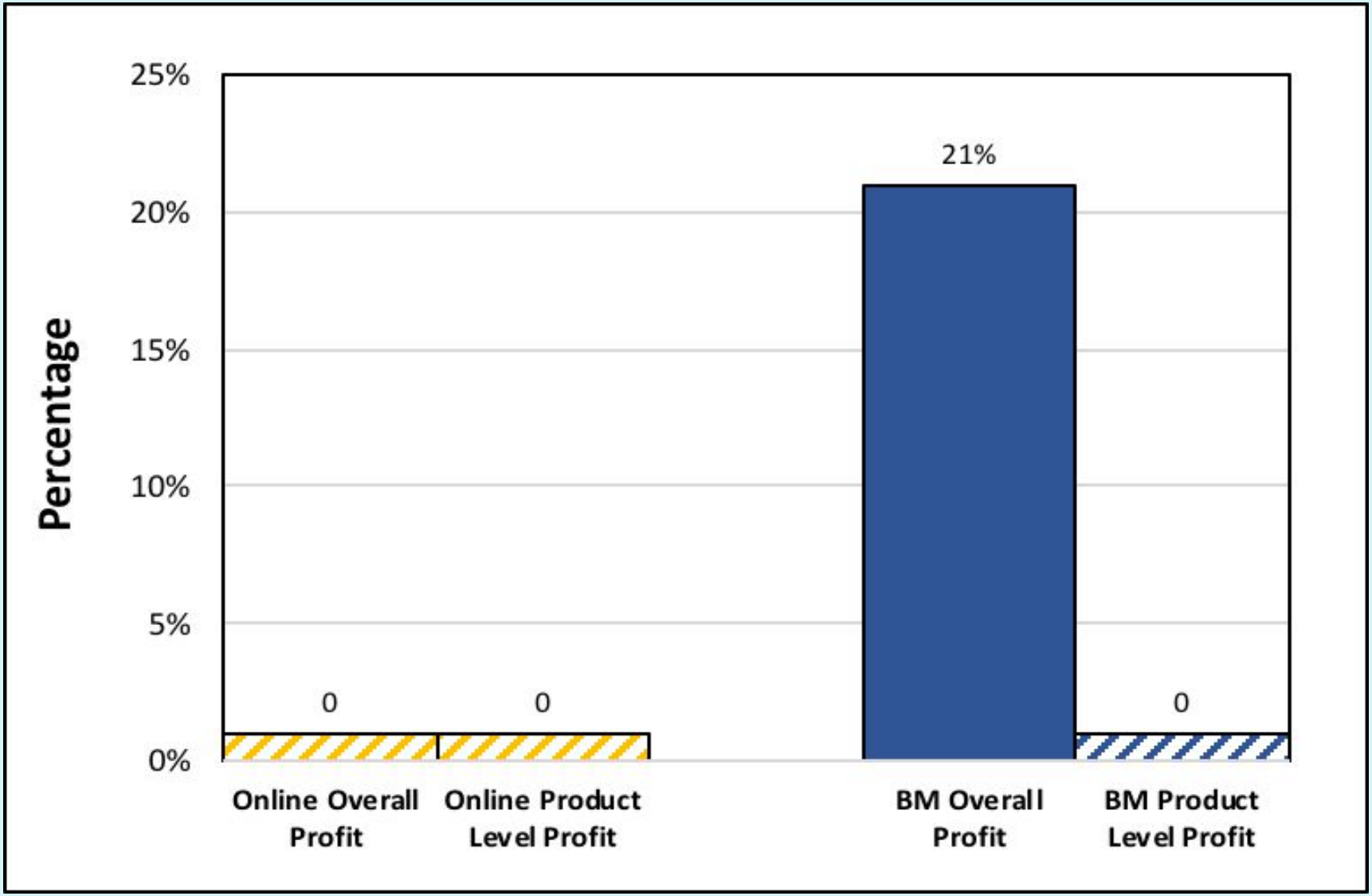
Impact on Physical Store Sales and Returns



Not Significant

Significant

Impact on Physical Store Sales and Returns



Not Significant

Significant

Results by Product Category

Product Category	Online Sales Value	Online Return Value	Online Overall Profit Impact	BM Sales Value	BM Return Value	BM Overall Profit Impact
1 - Bridal					-92%	24%
2 - Gold Wedding Bands					-71%	
3 - Solitaires						
4 - Diamond Fashion					-34%	
5 - Semi Precious		-87%		13%	-48%	18%
6 - Mens				-25%	-6%	
7 - Gold Earrings	-89%		-92%	-60%	-34%	-53%
8 - In House Special Event						
9 - Beads				-29%		
10 - Piercings/Close Out				-15%	-83%	
11 - Diamond Solitaires						
12 - Gold Chain/Jewelry				-31%	-22%	-23%
13 - Watches				46%	-101%	37%
14 - Pre-Owned				-40%	-31%	-33%
15 - Specialized Jewelry						
16 - Estate						
17 - Events						
18 - Trade Ins						
19 - Repair/Warranty						
20 - Diamond Wedding Bands						
21 - Sterling Silver					-38%	9%

Key
HIGH IMPACT (>|66%|)
MED IMPACT (|33%-66%|)
LOW IMPACT (<|33%|)
POSITIVE INCREASE
INSUFFICIENT DATA
INSIGNIFICANT RESULT

Summary

- Online stores did not experience a significant effect in overall sales and returns due to the policy change.
- At physical stores, the policy change resulted in decreased overall sales and returns. This led to an increase in overall profit.
- Product Category level analysis offered more granularity in analysis.

Product Category Level Summary

Categories with Increased Profit due to Policy Change

- Bridal (BM)
- Semi-Precious (BM)
- Watches (BM)
- Sterling Silver (BM)

Keep 45 Day Return Policy

Categories with Decreased Profit due to Policy Change

- Gold Earrings (OL/BM)
- Gold Chain/Jewelry (BM)
- Pre-Owned (BM)

Revert to 90 Day Return Policy

Recommendations

Enact a product category specific return policy.

By reverting the return policy for the product categories experiencing a decrease in profits, overall profitability for physical stores may increase.

Limitations

- Insufficient data for certain product categories (Eg: Product Category 3 (Solitaires) only had transaction after the policy change).
- Difference in differences modeling would be a more effective way to address purchases over time.
- Limited number of store data for physical stores.

Questions?

APPENDIX

Calculating Profit

Profit *= Revenue - Return Cost % (Return $) - (1-Unit Margin)(Sales $)*

Brick and Mortar Data

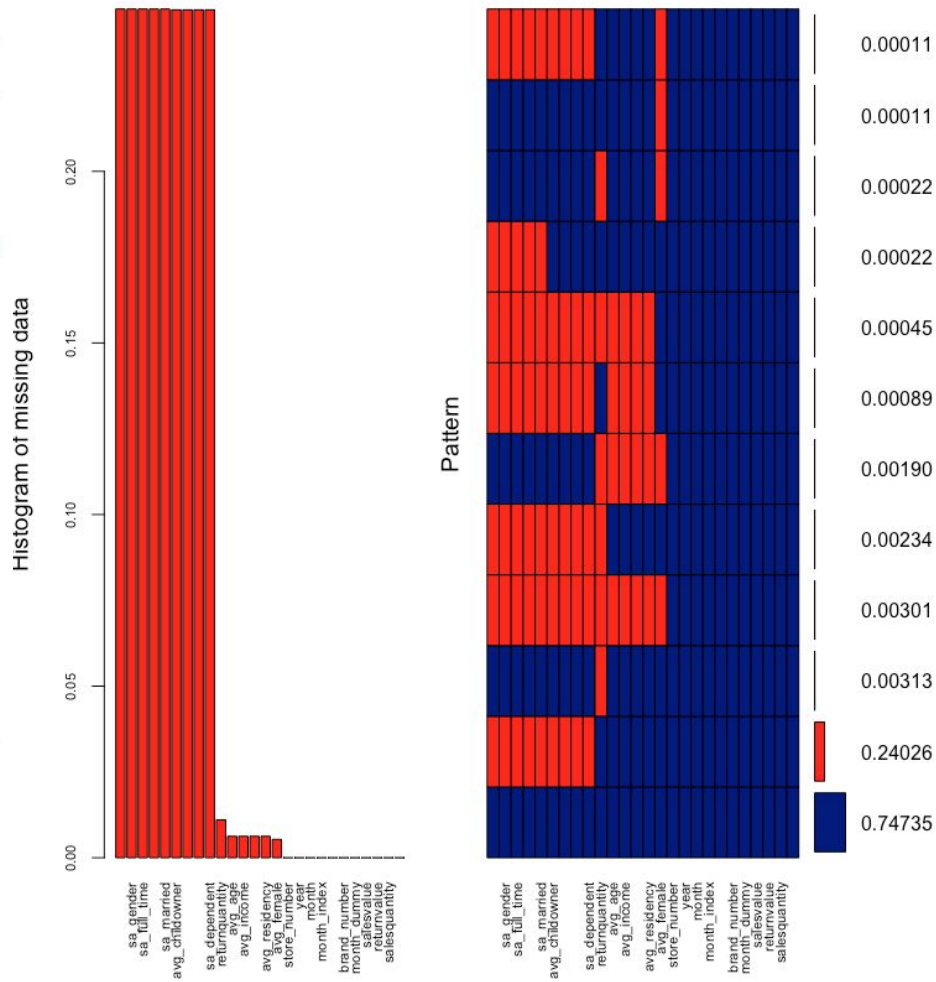
Missing Data

Brick and Mortar Sales and Returns

Descriptive Statistics

Statistic	N	Mean	St. Dev.	Min	Pctl(25)	Median	Pctl(75)	Max
store_number	8,957	1,765.8	1,426.6	15	403	1,469	2,753	4,975
year	8,957	2,013.2	0.4	2,013	2,013	2,013	2,013	2,014
month_index	8,957	50.4	3.4	45	47	50	53	56
brand_number	8,957	4.5	2.5	2	2	3	8	8
month_dummy	8,957	6.5	3.4	1	4	7	9	12
salesvalue	8,957	77,197.3	84,957.7	23.0	11,434.8	58,451.2	105,961.7	1,040,980.0
returnvalue	8,957	11,569.0	13,959.7	0.0	1,338.4	7,011.2	16,432.7	142,150.0
salesquantity	8,957	290.8	226.3	1	154	228	351	2,823
returnquantity	8,957	26.8	22.0	0	13	21	33	225
avg_female	8,858	0.5	0.2	0.0	0.4	0.5	0.6	1.0
avg_age	8,909	4.8	1.3	0.0	4.2	4.8	5.5	13.0
avg_income	8,901	5.2	1.0	1.0	4.7	5.2	5.8	9.0
avg_homeowner	8,901	0.7	0.2	0.0	0.6	0.7	0.8	1.0
avg_residency	8,901	7.2	1.8	0.0	6.2	7.1	8.1	15.0
avg_childowner	8,901	0.4	0.1	0.0	0.4	0.4	0.5	1.0
store_average_price	6,744	365.2	207.8	18.9	267.0	392.7	501.0	1,280.7
store_number_of_skus	6,744	223.4	170.2	13.0	132.0	178.0	245.0	1,775.0
sa_gender	6,742	0.8	0.2	0.0	0.7	0.9	1.0	1.0
sa_full_time	6,742	0.6	0.2	0.0	0.5	0.6	0.8	1.0
sa_avg_years_of_exp	6,742	3.8	2.5	0.0	2.0	3.2	5.2	17.3
sa_married	6,742	0.3	0.2	0.0	0.2	0.3	0.5	1.0
sa_avg_rate_of_pay	6,742	9.8	1.4	6.2	8.9	9.7	10.6	16.7
sa_dependent	6,744	0.2	0.2	0.0	0.1	0.2	0.3	0.8
sales_volume_group	6,744	2.7	1.3	1.0	2.0	3.0	3.0	6.0

While only **9%** of overall data is missing, if we delete the rows of data with missing values then **24.7%** of the data set would be missing and would skew the results.



9% of data missing

Descriptive Statistics B&M (After Data Imputation)

Descriptive Statistics Physical Stores

Statistic	N	Mean	St. Dev.	Min	Pctl(25)	Median	Pctl(75)	Max
salesquantity	8,957	290.8	226.3	1	154	228	351	2,823
salesvalue	8,957	77,197.3	84,957.7	23.0	11,434.8	58,451.2	105,961.7	1,040,980.0
returnquantity	8,957	26.8	22.0	0	13	21	33	225
returnvalue	8,957	11,569.0	13,959.7	0.0	1,338.4	7,011.2	16,432.7	142,150.0
avg_female	8,957	0.5	0.2	0.0	0.4	0.5	0.6	1.0
avg_age	8,957	4.8	1.3	0.0	4.2	4.8	5.5	13.0
avg_income	8,957	5.2	1.0	1.0	4.7	5.2	5.9	9.0
avg_homeowner	8,957	0.7	0.2	0.0	0.6	0.7	0.8	1.0
avg_residency	8,957	7.2	1.8	0.0	6.2	7.1	8.1	15.0
avg_childowner	8,957	0.4	0.1	0.0	0.4	0.4	0.5	1.0
timedummy	8,957	0.5	0.5	0	0	0	1	1
policydummy	8,957	0.7	0.5	0	0	1	1	1
sa_gender	8,957	0.8	0.2	0	0.8	0.9	1	1
sa_full_time	8,957	0.6	0.2	0.0	0.5	0.6	0.8	1.0
sa_avg_years_of_exp	8,957	3.7	2.5	0	1.9	3.1	5	17
sa_married	8,957	0.3	0.2	0.0	0.2	0.3	0.4	1.0
sa_avg_rate_of_pay	8,957	9.8	1.4	6.2	8.8	9.6	10.5	16.7
sa_dependent	8,957	0.2	0.2	0.0	0.1	0.2	0.3	0.8
sales_volume_group	8,957	2.4	1.3	1	1	2	3	6

VIF Scores for B&M

Before

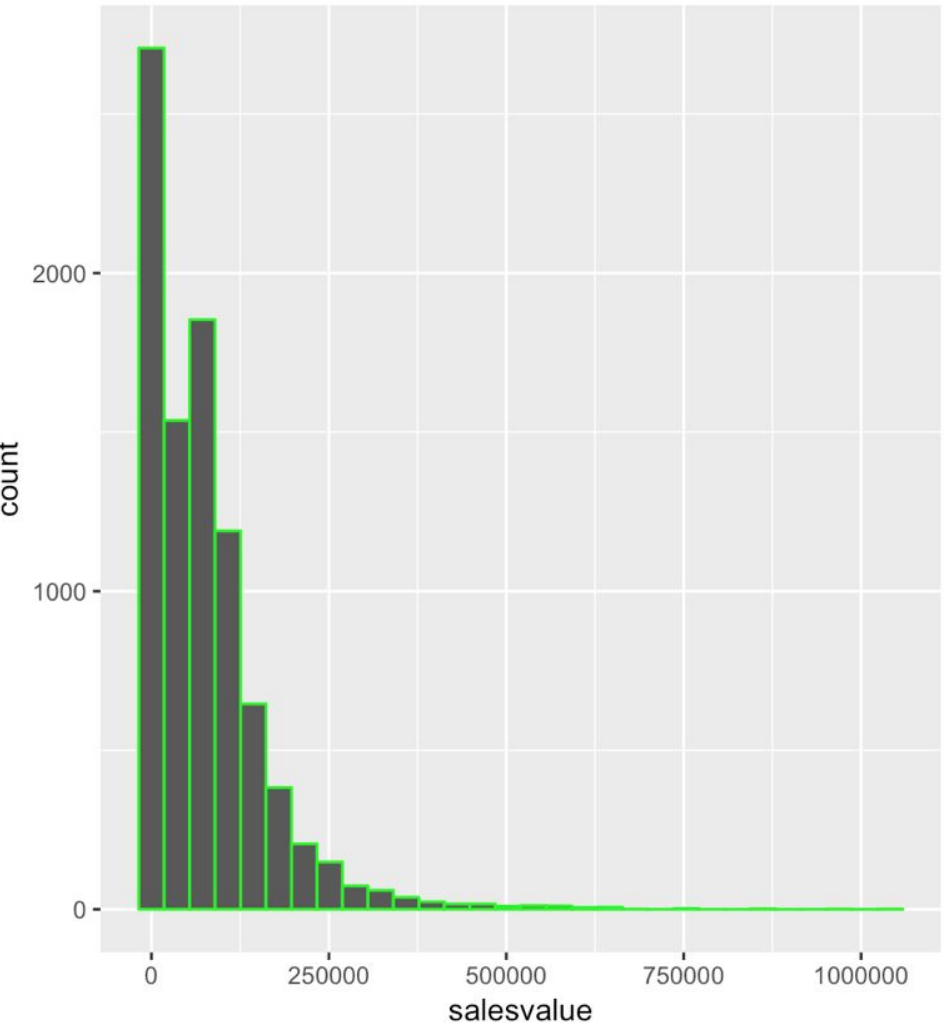
	Variables	VIF
1	store_number	46.811607
2	month_index	4.474937
3	brand_number	156.062905
4	month_dummy	1.267736
5	returnvalue	3.935411
6	returnquantity	4.421749
7	salesquantity	15.373558
8	avg_female	2.207364
9	avg_age	1.340710
10	avg_income	1.674063
11	avg_homeowner	1.819354
12	avg_residency	1.213383
13	avg_childowner	1.138577
14	timedummy	4.423444
15	policydummy	264.711024
16	store_average_price	6.775187
17	store_number_of_skus	18.703343
18	sa_gender	1.297011
19	sa_full_time	1.272080
20	sa_avg_years_of_exp	1.791858
21	sa_married	1.370944
22	sa_avg_rate_of_pay	1.318728
23	sa_dependent	1.961741
24	sales_volume_group	3.832801

After

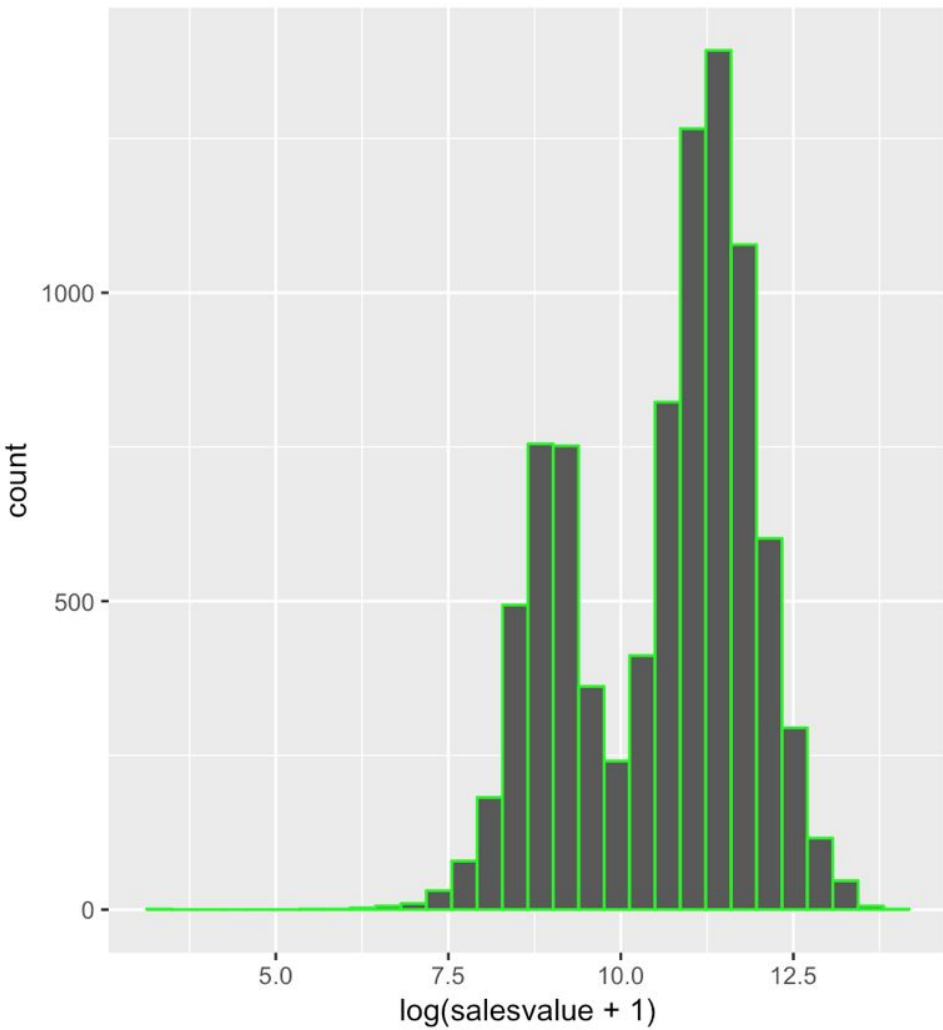
	Variables	VIF
1	month_dummy	1.183765
2	avg_female	1.236070
3	avg_age	1.136504
4	product_category	1.026955
5	avg_income	1.274406
6	avg_homeowner	1.429033
7	avg_residency	1.159441
8	avg_childowner	1.079102
9	timedummy	1.048268
10	policydummy	1.860914
11	sa_gender	1.267209
12	sa_full_time	1.263541
13	sa_avg_years_of_exp	1.725656
14	sa_married	1.363175
15	sa_avg_rate_of_pay	1.149384
16	sa_dependent	1.853232
17	store_number_of_skus	1.255372

Log Transforming Sales Value

Before



After



Output for B&M Sales Value

Regression Results

	Dependent variable:
	log(salesvalue + 1)
	BM Sales Value OLS
policydummy	2.11***
	(0.01)
timedummy	-0.07***
	(0.01)
avg_female	-0.04
	(0.02)
avg_age	-0.01***
	(0.003)
avg_income	0.01***
	(0.004)
avg_homeowner	-0.03
	(0.02)
avg_residency	-0.01***
	(0.002)
avg_childowner	0.06*
	(0.02)
sa_full_time	-0.02
	(0.02)
log_sa_avg_years_of_exp	0.10***
	(0.01)
log_sa_avg_rate_of_pay	0.47***
	(0.02)
log_store_number_of_skus	1.01***
	(0.01)
policydummy:timedummy	0.06***
	(0.01)
Constant	2.83***
	(0.06)
Observations	8,957
R2	0.95
Adjusted R2	0.95
Residual Std. Error	0.28
F Statistic	13,755.47***
Note:	*p<0.05; **p<0.01; ***p<0.001

studentized Breusch-Pagan test

data: bm_sv_linear_1
BP = 245.27, df = 13, p-value < 2.2e-16

Significant, meaning we have heteroskedasticity

Output for B&M Sales Quantity

```
Regression Results
==========================================================
                              Dependent variable:
                          ----------------------------
                                salesquantity
                                    IRRs
----------------------------------------------------------
policydummy                        0.4796***
                                   (0.0059)

timedummy                          0.9173***
                                   (0.0053)

avg_female                         1.0647***
                                   (0.0106)

avg_age                            1.0043***
                                   (0.0013)

avg_income                         1.0057**
                                   (0.0018)

avg_homeowner                       1.0123
                                   (0.0120)

avg_residency                      1.0037***
                                   (0.0009)

avg_childowner                      0.9934
                                   (0.0109)

sa_full_time                       0.9701**
                                   (0.0098)

log_sa_avg_years_of_exp            1.0106***
                                   (0.0032)

log_sa_avg_rate_of_pay             0.9475***
                                   (0.0121)

log_store_number_of_skus           2.9334***
                                   (0.0028)

policydummy:timedummy              1.0797***
                                   (0.0065)

Constant                           1.5910***
                                   (0.0311)

----------------------------------------------------------
Observations                          8,957
Log Likelihood                   -44,140.2400
theta                         63.8370*** (1.1940)
Akaike Inf. Crit.                 88,308.4800
==========================================================
Note:                        *p<0.05; **p<0.01; ***p<0.001
```

```
        studentized Breusch-Pagan test

data:  bm_sq_negbin_1
BP = 740.08, df = 13, p-value < 2.2e-16
```

Significant, meaning we have heteroskedasticity

```
Likelihood ratio test

Model 1: salesquantity ~ policydummy + timedummy + policydummy * timedummy +
    avg_female + avg_age + avg_income + avg_homeowner + avg_residency +
    avg_childowner + sa_full_time + sa_avg_years_of_exp + sa_avg_rate_of_pay +
    sales_volume_group
Model 2: salesquantity ~ 1
  #Df LogLik  Df  Chisq Pr(>Chisq)
1  15 -57062
2   2 -58212 -13 2299.9  < 2.2e-16 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

Negative Binomial model fits

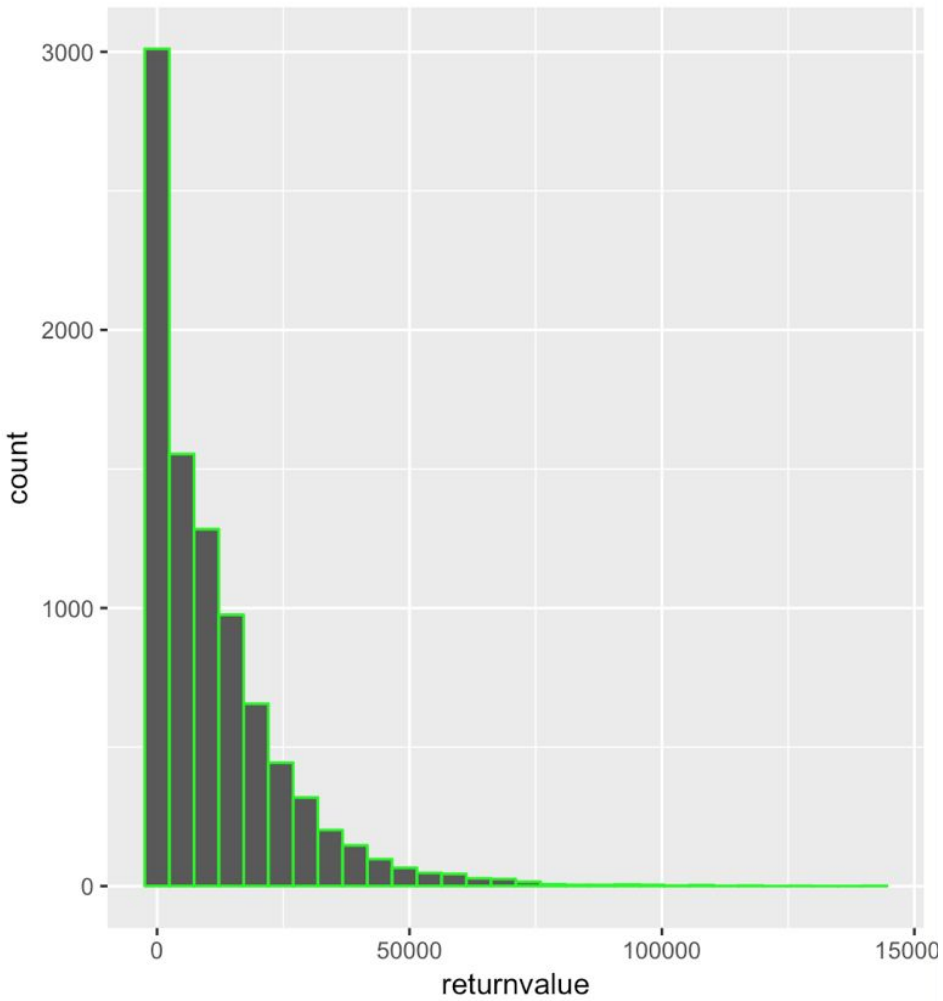
```
Likelihood ratio test

Model 1: salesquantity ~ policydummy + timedummy + policydummy * timedummy +
    avg_female + avg_age + avg_income + avg_homeowner + avg_residency +
    avg_childowner + sa_gender + sa_full_time + sa_avg_years_of_exp +
    sa_married + sa_avg_rate_of_pay + sa_dependent + sales_volume_group
Model 2: salesquantity ~ policydummy + timedummy + policydummy * timedummy +
    avg_female + avg_age + avg_income + avg_homeowner + avg_residency +
    avg_childowner + sa_full_time + sa_avg_years_of_exp + sa_avg_rate_of_pay +
    sales_volume_group
  #Df  LogLik Df  Chisq Pr(>Chisq)
1  17 -510365
2  15  -57062 -2 906605  < 2.2e-16 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

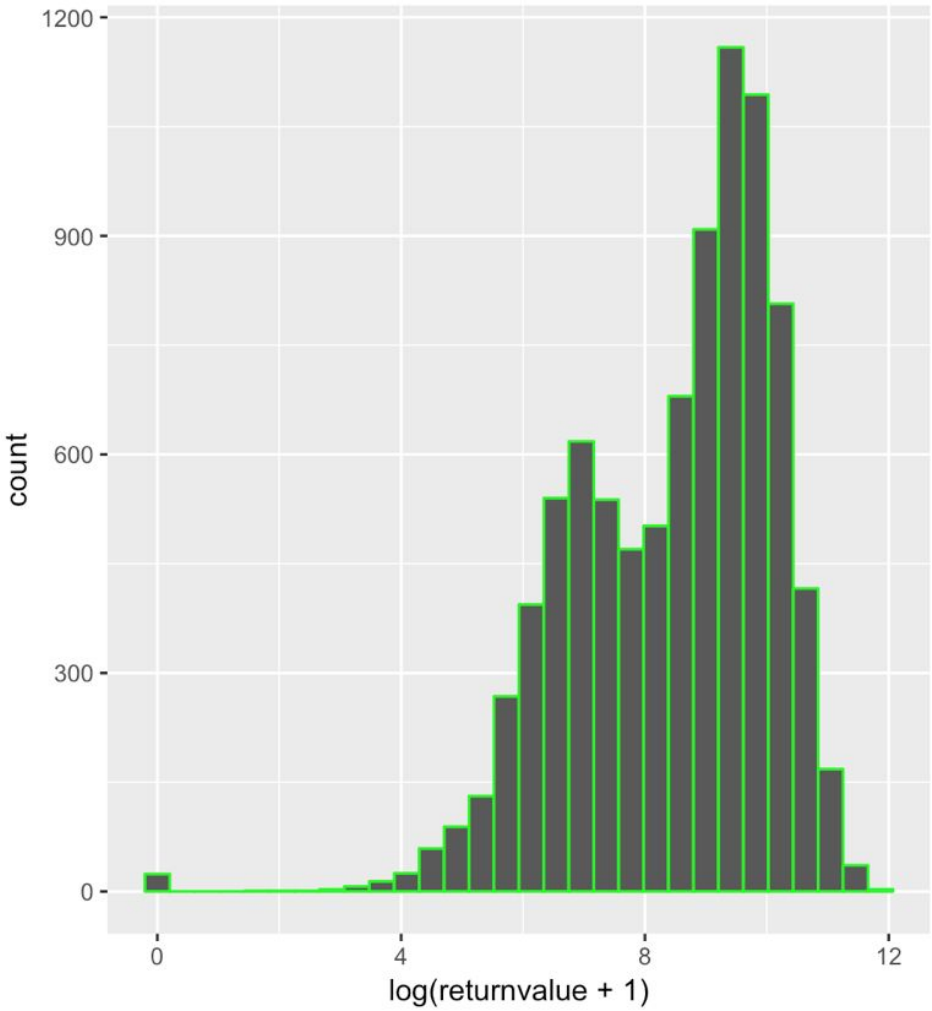
Choosing Neg. Binomial over Poisson model

Log Transforming Return Value

Before



After



Output for B&M Return Value

Std. Err. Regression Results

	Dependent variable:	
	log(returnvalue + 1)	
	Normal SE (1)	HW-Robust SE (2)
policydummy	-0.02	-0.02
	(0.04)	(0.05)
timedummy	0.12***	0.12***
	(0.03)	(0.03)
log(salesvalue + 1)	1.23***	1.23***
	(0.01)	(0.02)
avg_female	-0.02	-0.02
	(0.05)	(0.06)
avg_age	0.01	0.01
	(0.01)	(0.01)
avg_income	-0.01	-0.01
	(0.01)	(0.01)
avg_homeowner	0.09	0.09
	(0.06)	(0.07)
avg_residency	-0.03***	-0.03***
	(0.004)	(0.01)
avg_childowner	0.12*	0.12
	(0.05)	(0.06)
sa_full_time	-0.02	-0.02
	(0.04)	(0.05)
sa_avg_years_of_exp	0.0002	0.0002
	(0.003)	(0.003)
sa_avg_rate_of_pay	0.03***	0.03***
	(0.01)	(0.01)

```
studentized Breusch-Pagan test

data:  bm_rv_linear_1
BP = 392.74, df = 14, p-value < 2.2e-16
```

Significant, meaning we have heteroskedasticity

sales_volume_group	-0.05***	-0.05***
	(0.01)	(0.01)
policydummy:timedummy	-0.36***	-0.36***
	(0.03)	(0.03)
Constant	-4.62***	-4.62***
	(0.13)	(0.20)
Observations	8,957	8,957
R2	0.84	0.84
Adjusted R2	0.84	0.84
Residual Std. Error	0.65	0.65
F Statistic	3,346.75***	3,346.75***
Note:	*p<0.05; **p<0.01; ***p<0.001	

Output for B&M Return Quantity

Std. Err. Regression Results

	Dependent variable:	
	returnquantity	
	IRR Normal SE (1)	IRR HW-Robust SE (2)
policydummy	1.16***	1.16***
	(0.02)	(0.02)
timedummy	1.09***	1.09***
	(0.02)	(0.02)
salesquantity	1.00***	1.00***
	(0.0000)	(0.0000)
avg_female	1.11**	1.11**
	(0.04)	(0.04)
avg_age	1.01	1.01
	(0.004)	(0.005)
avg_income	1.02**	1.02**
	(0.01)	(0.01)
avg_homeowner	0.92*	0.92*
	(0.04)	(0.05)
avg_residency	0.99*	0.99*
	(0.003)	(0.004)
avg_childowner	1.15***	1.15***
	(0.04)	(0.04)
sa_full_time	1.02	1.02
	(0.03)	(0.03)
sa_avg_years_of_exp	1.00	1.00
	(0.002)	(0.002)
sa_avg_rate_of_pay	1.03***	1.03***
	(0.004)	(0.004)
sales_volume_group	1.10***	1.10***
	(0.01)	(0.01)
policydummy:timedummy	0.71***	0.71***
	(0.02)	(0.03)
Constant	6.03***	6.03***
	(0.06)	(0.06)
Observations	8,957	8,957
Log Likelihood	-33,587.19	-33,587.19
theta	5.64*** (0.11)	5.64*** (0.11)
Akaike Inf. Crit.	67,204.38	67,204.38

Note: *p<0.05; **p<0.01; ***p<0.001

```
studentized Breusch-Pagan test

data:  bm_rq_negbin_1
BP = 1210, df = 14, p-value < 2.2e-16
```

Significant, meaning we have heteroskedasticity

```
Likelihood ratio test

Model 1: returnquantity ~ policydummy + timedummy + policydummy * timedummy +
    salesquantity + avg_female + avg_age + avg_income + avg_homeowner +
    avg_residency + avg_childowner + sa_full_time + sa_avg_years_of_exp +
    sa_avg_rate_of_pay + sales_volume_group
Model 2: salesquantity ~ 1
  #Df LogLik  Df Chisq Pr(>Chisq)
1  16 -33586
2   2 -58212 -14 49252  < 2.2e-16 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

Negative Binomial model fits

```
Likelihood ratio test

Model 1: returnquantity ~ policydummy + timedummy + policydummy * timedummy +
    salesquantity + avg_female + avg_age + avg_income + avg_homeowner +
    avg_residency + avg_childowner + sa_full_time + sa_avg_years_of_exp +
    sa_avg_rate_of_pay + sales_volume_group
Model 2: returnquantity ~ policydummy + timedummy + policydummy * timedummy +
    salesquantity + avg_female + avg_age + avg_income + avg_homeowner +
    avg_residency + avg_childowner + sa_full_time + sa_avg_years_of_exp +
    sa_avg_rate_of_pay + sales_volume_group
  #Df LogLik Df Chisq Pr(>Chisq)
1  15 -48821
2  16 -33586  1 30469  < 2.2e-16 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

Choosing Neg. Binomial over Poisson model

Output for Overall B&M Profit

	Dependent variable:	
	log(Profit + 1)	
	Normal SE (1)	HW-Robust SE (2)
policydummy	1.92***	1.92***
	(0.02)	(0.02)
timedummy	-0.12***	-0.12***
	(0.01)	(0.01)
avg_female	-0.05	-0.05
	(0.03)	(0.04)
avg_age	-0.01**	-0.01*
	(0.004)	(0.004)
avg_income	0.02***	0.02***
	(0.01)	(0.01)
avg_homeowner	-0.03	-0.03
	(0.03)	(0.04)
avg_residency	-0.01**	-0.01**
	(0.002)	(0.003)
avg_childowner	0.003	0.003
	(0.03)	(0.04)
sa_full_time	-0.02	-0.02
	(0.03)	(0.03)
sa_avg_years_of_exp	-0.02***	-0.02***
	(0.01)	(0.005)
log(sa_avg_years_of_exp + 1)	0.19***	0.19***
	(0.03)	(0.03)
log(sa_avg_rate_of_pay)	0.27***	0.27***
	(0.03)	(0.04)
log(store_number_of_skus)	1.02***	1.02***
	(0.01)	(0.01)
policydummy:timedummy	0.21***	0.21***
	(0.02)	(0.02)
Constant	2.12***	2.12***
	(0.08)	(0.10)

Brick and Mortar Data at Product Level

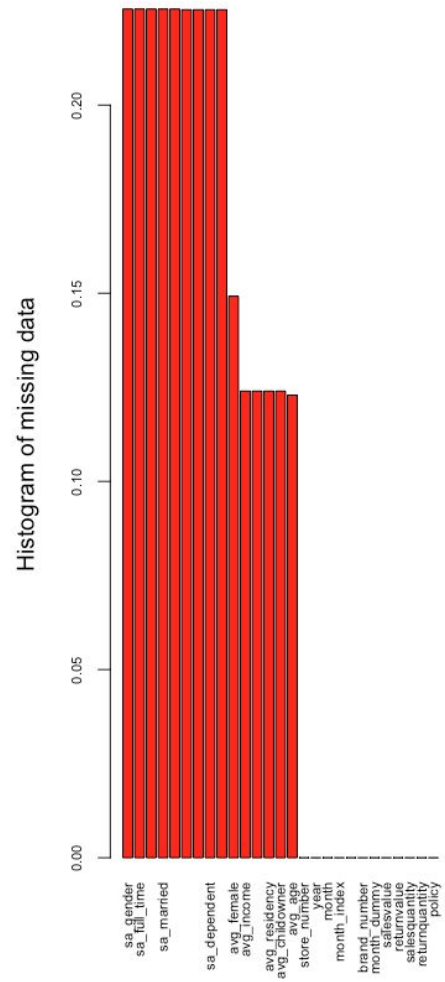
Missing Data

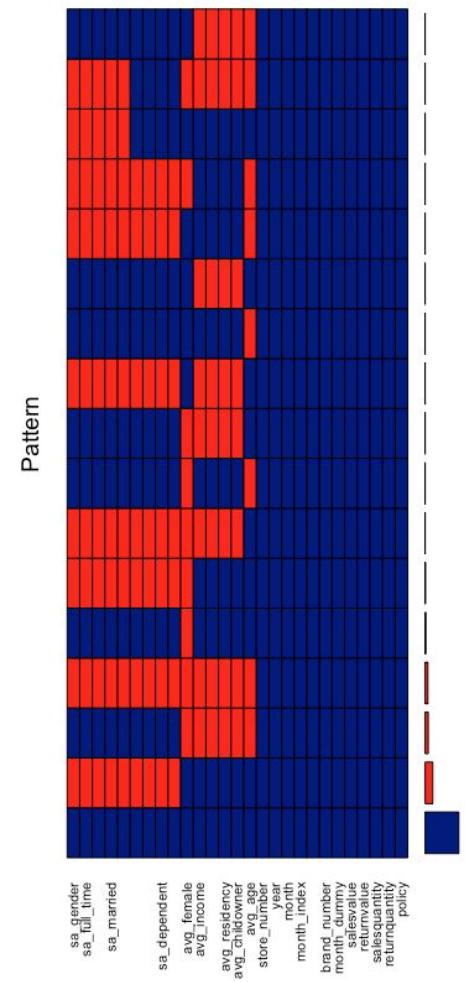
Brick and Mortar Sales and Returns by Product Category

Descriptive Statistics

Statistic	N	Mean	St. Dev.	Min	Pctl(25)	Median	Pctl(75)	Max
store_number	112,243	1,877.8	1,405.1	15	1,003	1,534	2,780	4,975
year	112,243	2,013.2	0.4	2,013	2,013	2,013	2,013	2,014
month_index	112,243	50.3	3.4	45	47	50	53	56
product_category	112,243	9.6	6.1	1	4	9	13	21
brand_number	112,243	4.2	2.4	2	2	3	5	8
month_dummy	112,243	6.6	3.4	1	4	7	9	12
salesvalue	112,243	5,868.7	10,698.9	0.0	560.3	2,378.7	6,161.1	209,777.8
returnvalue	112,243	886.0	2,398.0	0	0	123.5	679.0	72,684
salesquantity	112,243	22.4	37.2	1	5	11	22	627
returnquantity	112,243	2.1	3.1	0	0	1	3	130
avg_female	95,495	0.5	0.3	0.0	0.3	0.5	0.7	1.0
avg_age	98,446	4.8	2.3	0.0	3.5	4.8	6.0	13.0
avg_income	98,325	5.2	1.4	1.0	4.4	5.2	6.0	9.0
avg_homeowner	98,325	0.7	0.3	0.0	0.5	0.7	0.9	1.0
avg_residency	98,325	7.1	3.0	0.0	5.2	7.0	8.7	15.0
avg_childowner	98,325	0.4	0.3	0.0	0.2	0.4	0.5	1.0
store_average_price	86,950	380.8	198.0	18.9	294.1	401.5	506.7	1,280.7
store_number_of_skus	86,950	230.7	174.1	13.0	136.0	183.0	253.0	1,775.0
sa_gender	86,930	0.8	0.2	0.0	0.7	0.9	1.0	1.0
sa_full_time	86,930	0.6	0.2	0.0	0.6	0.7	0.8	1.0
sa_avg_years_of_exp	86,930	3.9	2.5	0.0	2.0	3.3	5.3	17.3
sa_married	86,930	0.3	0.2	0.0	0.2	0.3	0.5	1.0
sa_avg_rate_of_pay	86,930	9.9	1.4	6.2	8.9	9.7	10.6	16.7
sa_dependent	86,950	0.2	0.2	0.0	0.1	0.2	0.3	0.8
sales_volume_group	86,950	2.7	1.3	1.0	2.0	3.0	3.0	6.0

While only **11%** of overall data is missing, if we delete the rows of data with missing values then **22.6%** of the data set would be missing and would skew the results.





11% of data missing

Descriptive Statistics B&M by Product Category (After Data Imputation)

Descriptive Statistics

Statistic	N	Mean	St. Dev.	Min	Pctl(25)	Median	Pctl(75)	Max
X	112,243	56,122.0	32,401.9	1	28,061.5	56,122	84,182.5	112,243
store_number	112,243	1,877.8	1,405.1	15	1,003	1,534	2,780	4,975
year	112,243	2,013.2	0.4	2,013	2,013	2,013	2,013	2,014
month_index	112,243	50.3	3.4	45	47	50	53	56
brand_number	112,243	4.2	2.4	2	2	3	5	8
month_dummy	112,243	6.6	3.4	1	4	7	9	12
salesvalue	112,243	5,868.7	10,698.9	0.0	560.3	2,378.7	6,161.1	209,777.8
returnvalue	112,243	886.0	2,398.0	0	0	123.5	679.0	72,684
salesquantity	112,243	22.4	37.2	1	5	11	22	627
returnquantity	112,243	2.1	3.1	0	0	1	3	130
avg_female	112,243	0.5	0.3	0	0.3	0.5	0.8	1
avg_age	112,243	4.7	2.4	0.0	3.4	4.8	6.0	13.0
avg_income	112,243	5.2	1.5	1	4.3	5.2	6	9
avg_homeowner	112,243	0.7	0.3	0.0	0.5	0.7	0.9	1.0
avg_residency	112,243	7.1	3.2	0.0	5.1	7.0	8.8	15.0
avg_childowner	112,243	0.4	0.3	0.0	0.2	0.4	0.5	1.0
store_average_price	112,243	348.0	214.7	18.9	155.1	381.2	492.2	1,280.7
store_number_of_skus	112,243	224.7	169.3	13	133	179	246	1,775
sa_gender	112,243	0.8	0.2	0	0.7	0.9	1	1
sa_full_time	112,243	0.6	0.2	0.0	0.5	0.6	0.8	1.0
sa_avg_years_of_exp	112,243	3.8	2.5	0.0	1.9	3.1	5.1	17.3
sa_married	112,243	0.3	0.2	0	0.2	0.3	0.5	1
sa_avg_rate_of_pay	112,243	9.8	1.4	6.2	8.9	9.6	10.5	16.7
sa_dependent	112,243	0.2	0.2	0.0	0.1	0.2	0.3	0.8
sales_volume_group	112,243	2.6	1.3	1	2	2	3	6
log_salesvalue	112,243	7.5	1.8	0.0	6.3	7.8	8.7	12.3
policydummy	112,243	0.8	0.4	0	1	1	1	1
timedummy	112,243	0.5	0.5	0	0	0	1	1

Output for B&M Sales Value at Prod. Level

Std. Err. Regression Results

	Dependent variable:	
	log(salesvalue + 1)	
	Normal SE (1)	HW-Robust SE (2)
policydummy	1.88***	1.88***
	(0.01)	(0.02)
timedummy	-0.02	-0.02
	(0.01)	(0.02)
FactorPC2	-2.02***	-2.02***
	(0.02)	(0.02)
FactorPC3	-0.43***	-0.43***
	(0.02)	(0.02)
FactorPC4	-0.99***	-0.99***
	(0.02)	(0.02)
FactorPC5	-1.81***	-1.81***
	(0.02)	(0.02)
FactorPC6	-2.14***	-2.14***
	(0.02)	(0.02)
FactorPC7	-3.29***	-3.29***
	(0.02)	(0.03)
FactorPC8	-1.68***	-1.68***
	(0.03)	(0.03)
FactorPC9	-3.04***	-3.04***
	(0.02)	(0.02)
FactorPC10	0.02	0.02
	(0.02)	(0.02)
FactorPC11	-1.29***	-1.29***
	(0.02)	(0.02)
FactorPC12	-1.46***	-1.46***
	(0.02)	(0.02)
FactorPC13	-1.92***	-1.92***
	(0.02)	(0.02)
FactorPC14	-2.95***	-2.95***
	(0.02)	(0.02)
FactorPC15	-2.53***	-2.53***
	(0.03)	(0.03)
FactorPC16	-2.12***	-2.12***
	(0.03)	(0.02)
FactorPC17	-1.42***	-1.42***
	(0.03)	(0.03)
FactorPC18	-5.60***	-5.60***
	(1.05)	(0.02)
FactorPC19	-6.61***	-6.61***
	(0.47)	(0.14)
FactorPC20	-1.09***	-1.09***
	(0.02)	(0.01)
FactorPC21	-1.87***	-1.87***
	(0.02)	(0.01)
avg_female	-0.13***	-0.13***
	(0.01)	(0.01)
avg_age	-0.005***	-0.005**
	(0.001)	(0.002)
avg_income	0.01***	0.01***
	(0.002)	(0.003)
avg_homeowner	-0.02	-0.02
	(0.01)	(0.02)
avg_residency	-0.01***	-0.01***
	(0.001)	(0.001)
avg_childowner	0.05***	0.05***
	(0.01)	(0.02)
sa_full_time	0.04*	0.04*
	(0.02)	(0.02)
log(sa_avg_years_of_exp + 1)	0.08***	0.08***
	(0.01)	(0.01)
log(sa_avg_rate_of_pay)	0.34***	0.34***
	(0.02)	(0.02)
log(store_number_of_skus)	0.87***	0.87***
	(0.01)	(0.01)
policydummy:timedummy	-0.08***	-0.08***
	(0.01)	(0.02)
Constant	2.32***	2.32***
	(0.06)	(0.06)
Observations	112,243	112,243
R2	0.67	0.67
Adjusted R2	0.67	0.67
Residual Std. Error	1.05	1.05
F Statistic	6,769.21***	6,769.21***
Note:	*p<0.05; **p<0.01; ***p<0.001	

Output for B&M Sales Quantity at Prod. Level

Regression Results

	Dependent variable:	
	salesquantity	
	Normal SE IRR (1)	HW-Robust SE IRR (2)
policydummy	1.00	1.00
	(0.01)	(0.01)
timedummy	0.98*	0.98*
	(0.01)	(0.01)
month_dummy	1.01***	1.01***
	(0.001)	(0.001)
FactorPC2	0.99	0.99
	(0.01)	(0.01)
FactorPC3	1.25***	1.25***
	(0.01)	(0.01)
FactorPC4	1.59***	1.59***
	(0.01)	(0.01)
FactorPC5	2.30***	2.30***
	(0.01)	(0.01)
FactorPC6	0.52***	0.52***
	(0.01)	(0.01)
FactorPC7	1.29***	1.29***
	(0.01)	(0.01)
FactorPC8	0.35***	0.35***
	(0.02)	(0.02)
FactorPC9	2.40***	2.40***
	(0.01)	(0.01)
FactorPC10	9.51***	9.51***
	(0.02)	(0.02)
FactorPC11	0.89***	0.89***
	(0.01)	(0.01)
FactorPC12	1.85***	1.85***
	(0.01)	(0.01)
FactorPC13	1.05***	1.05***
	(0.01)	(0.01)
FactorPC14	8.31***	8.31***
	(0.01)	(0.02)
FactorPC15	0.47***	0.47***
	(0.02)	(0.02)
FactorPC16	2.09***	2.09***
	(0.02)	(0.02)
FactorPC17	0.26***	0.26***
	(0.02)	(0.02)
FactorPC18	0.04**	0.04**
	(1.23)	(0.01)
FactorPC19	0.16***	0.16***
	(0.43)	(0.29)
FactorPC20	1.04**	1.04**
	(0.01)	(0.01)
FactorPC21	1.90***	1.90***
	(0.01)	(0.01)
avg_female	0.95***	0.95***
	(0.01)	(0.01)
avg_age	1.00***	1.00***
	(0.001)	(0.001)
avg_income	1.00	1.00
	(0.002)	(0.002)
avg_homeowner	0.98	0.98
	(0.01)	(0.01)
avg_residency	1.00	1.00
	(0.001)	(0.001)
avg_childowner	1.05***	1.05***
	(0.01)	(0.01)
sa_full_time	1.01	1.01
	(0.01)	(0.01)
log(sa_avg_years_of_exp + 1)	1.03***	1.03***
	(0.005)	(0.004)
log(sa_avg_rate_of_pay)	1.07***	1.07***
	(0.02)	(0.02)
log(store_number_of_skus)	2.43***	2.43***
	(0.005)	(0.004)
policydummy:timedummy	0.95***	0.95***
	(0.01)	(0.01)
Constant	0.08***	0.08***
	(0.05)	(0.04)
Observations	112,243	112,243
Log Likelihood	-410,273.20	-410,273.20
theta	1.98*** (0.01)	1.98*** (0.01)
Akaike Inf. Crit.	820,616.30	820,616.30
Note:	*p<0.05; **p<0.01; ***p<0.001	

Output for B&M Return Value at Prod. Level

Std. Err. Regression Results

	Dependent variable:	
	log(returnvalue + 1)	
	Normal SE (1)	HW-Robust SE (2)
policydummy	-1.59***	-1.59***
	(0.03)	(0.03)
timedummy	0.10***	0.10***
	(0.03)	(0.02)
FactorPC2	1.29***	1.29***
	(0.04)	(0.05)
FactorPC3	-0.59***	-0.59***
	(0.04)	(0.06)
FactorPC4	0.75***	0.75***
	(0.04)	(0.04)
FactorPC5	0.76***	0.76***
	(0.04)	(0.05)
FactorPC6	0.37***	0.37***
	(0.04)	(0.05)
FactorPC7	0.79***	0.79***
	(0.05)	(0.05)
FactorPC8	-0.82***	-0.82***
	(0.06)	(0.08)
FactorPC9	0.87***	0.87***
	(0.05)	(0.05)
FactorPC10	-1.40***	-1.40***
	(0.05)	(0.05)
FactorPC11	-0.48***	-0.48***
	(0.04)	(0.05)
FactorPC12	1.01***	1.01***
	(0.04)	(0.04)
FactorPC13	0.43***	0.43***
	(0.04)	(0.05)
FactorPC14	0.07	0.07
	(0.05)	(0.05)
FactorPC15	0.88***	0.88***
	(0.06)	(0.06)
FactorPC16	1.51***	1.51***
	(0.06)	(0.05)
FactorPC17	-1.94***	-1.94***
	(0.06)	(0.08)
FactorPC18	-0.09	-0.09
	(2.44)	(0.06)
FactorPC19	2.31*	2.31***
	(1.09)	(0.39)
FactorPC20	0.01	0.01
	(0.04)	(0.05)
FactorPC21	0.78***	0.78***
	(0.04)	(0.05)
log(salesvalue + 1)	1.28***	1.28***
	(0.01)	(0.01)
avg_female	-0.03	-0.03
	(0.03)	(0.03)
avg_age	-0.01*	-0.01*
	(0.003)	(0.003)
avg_income	0.02***	0.02***
	(0.01)	(0.01)
avg_homeowner	-0.07*	-0.07*
	(0.03)	(0.03)
avg_residency	-0.004	-0.004
	(0.002)	(0.002)
avg_childowner	0.07**	0.07**
	(0.03)	(0.03)
sa_full_time	-0.11*	-0.11*
	(0.05)	(0.05)
log(sa_avg_years_of_exp + 1)	0.04*	0.04*
	(0.02)	(0.02)
log(sa_avg_rate_of_pay)	0.57***	0.57***
	(0.06)	(0.06)
log(store_number_of_skus)	0.47***	0.47***
	(0.01)	(0.01)
policydummy:timedummy	-0.41***	-0.41***
	(0.03)	(0.03)
Constant	-8.56***	-8.56***
	(0.14)	(0.15)
Observations	112,243	112,243
R2	0.40	0.40
Adjusted R2	0.40	0.40
Residual Std. Error	2.44	2.44
F Statistic	2,213.79***	2,213.79***
Note:	*p<0.05; **p<0.01; ***p<0.00	

Output for B&M Return Quantity at Prod. Level

Std. Err. Regression Results

	Dependent variable:	
	returnquantity	
	Normal SE (1)	HW-Robust SE (2)
policydummy	0.63***	0.63***
	(0.01)	(0.01)
timedummy	1.06***	1.06***
	(0.01)	(0.01)
FactorPC2	1.16***	1.16***
	(0.02)	(0.01)
FactorPC3	0.60***	0.60***
	(0.02)	(0.02)
FactorPC4	0.82***	0.82***
	(0.01)	(0.01)
FactorPC5	0.59***	0.59***
	(0.01)	(0.01)
FactorPC6	0.87***	0.87***
	(0.02)	(0.02)
FactorPC7	0.48***	0.48***
	(0.02)	(0.02)
FactorPC8	0.88***	0.88***
	(0.03)	(0.03)
FactorPC9	0.50***	0.50***
	(0.02)	(0.02)
FactorPC10	0.14***	0.14***
	(0.02)	(0.03)
FactorPC11	0.66***	0.66***
	(0.02)	(0.02)
FactorPC12	0.83***	0.83***
	(0.01)	(0.01)
FactorPC13	0.71***	0.71***
	(0.02)	(0.01)
FactorPC14	0.12***	0.12***
	(0.02)	(0.03)
FactorPC15	0.92***	0.92***
	(0.02)	(0.02)
FactorPC16	0.98	0.98
	(0.02)	(0.02)
FactorPC17	0.75***	0.75***
	(0.04)	(0.04)
FactorPC18	0.00	0.00
	(114,839.70)	(1.00)
FactorPC19	0.37	0.37
	(1.03)	(0.81)
FactorPC20	0.78***	0.78***
	(0.02)	(0.01)
FactorPC21	0.65***	0.65***
	(0.01)	(0.01)
log(salesquantity + 1)	2.62***	2.62***
	(0.005)	(0.01)
avg_female	1.00	1.00
	(0.01)	(0.01)
avg_age	1.00	1.00
	(0.001)	(0.001)
avg_income	1.01***	1.01***
	(0.002)	(0.002)
avg_homeowner	0.95***	0.95***
	(0.01)	(0.01)
avg_residency	1.00*	1.00*
	(0.001)	(0.001)
avg_childowner	1.04***	1.04***
	(0.01)	(0.01)
sa_full_time	0.95**	0.95**
	(0.02)	(0.02)
log(sa_avg_years_of_exp + 1)	1.01	1.01
	(0.01)	(0.01)
log(sa_avg_rate_of_pay)	1.29***	1.29***
	(0.02)	(0.02)
log(store_number_of_skus)	1.09***	1.09***
	(0.01)	(0.01)
policydummy:timedummy	0.84***	0.84***
	(0.01)	(0.01)
Constant	0.09***	0.09***
	(0.05)	(0.06)
Observations	112,243	112,243
Log Likelihood	-176,264.50	-176,264.50
theta	5.45*** (0.09)	5.45*** (0.09)
Akaike Inf. Crit.	352,599.00	352,599.00
Note:	*p<0.05; **p<0.01; ***p<0.001	

Output for B&M Profit at Prod. Level

Std. Err. Regression Results

	Dependent variable:	
	log(Profit + 1)	
	Normal SE (1)	HW-Robust SE (2)
policydummy	1.88***	1.88***
	(0.01)	(0.02)
timedummy	-0.06***	-0.06**
	(0.01)	(0.02)
FactorPC2	-2.10***	-2.10***
	(0.02)	(0.02)
FactorPC3	-0.41***	-0.41***
	(0.02)	(0.02)
FactorPC4	-0.95***	-0.95***
	(0.02)	(0.02)
FactorPC5	-1.59***	-1.59***
	(0.02)	(0.02)
FactorPC6	-2.05***	-2.05***
	(0.02)	(0.02)
FactorPC7	-3.07***	-3.07***
	(0.02)	(0.03)
FactorPC8	-1.54***	-1.54***
	(0.03)	(0.03)
FactorPC9	-2.77***	-2.77***
	(0.02)	(0.02)
FactorPC10	0.38***	0.38***
	(0.02)	(0.03)
FactorPC11	-1.16***	-1.16***
	(0.02)	(0.02)
FactorPC12	-1.53***	-1.53***
	(0.02)	(0.02)
FactorPC13	-1.79***	-1.79***
	(0.02)	(0.02)
FactorPC14	-2.62***	-2.62***
	(0.02)	(0.02)
FactorPC15	-2.57***	-2.57***
	(0.03)	(0.03)
FactorPC16	-2.45***	-2.45***
	(0.03)	(0.03)
FactorPC17	-1.20***	-1.20***
	(0.03)	(0.03)
FactorPC18	-5.18***	-5.18***
	(1.09)	(0.02)
FactorPC19	-6.26***	-6.26***
	(0.49)	(0.12)
FactorPC20	-1.03***	-1.03***
	(0.02)	(0.02)
FactorPC21	-1.84***	-1.84***
	(0.02)	(0.02)
avg_female	-0.14***	-0.14***
	(0.01)	(0.01)
avg_age	-0.005**	-0.005*
	(0.002)	(0.002)
avg_income	0.01***	0.01***
	(0.003)	(0.003)
avg_homeowner	-0.02	-0.02
	(0.02)	(0.02)
avg_residency	-0.01***	-0.01***
	(0.001)	(0.001)
avg_childowner	0.05***	0.05**
	(0.01)	(0.02)
sa_full_time	0.05*	0.05*
	(0.02)	(0.02)
sa_avg_years_of_exp	-0.004	-0.004
	(0.004)	(0.004)
log(sa_avg_years_of_exp + 1)	0.09***	0.09***
	(0.02)	(0.02)
log(sa_avg_rate_of_pay)	0.22***	0.22***
	(0.03)	(0.03)
log(store_number_of_skus)	0.84***	0.84***
	(0.01)	(0.01)
policydummy:timedummy	0.01	0.01
	(0.02)	(0.02)
Constant	1.56***	1.56***
	(0.07)	(0.07)
Observations	104,498	104,498
R2	0.64	0.64
Adjusted R2	0.64	0.64
Residual Std. Error	1.09	1.09
F Statistic	5,449.46***	5,449.46***
Note:	*p<0.05; **p<0.01; ***p<0.001	

Online Sales/Returns Data

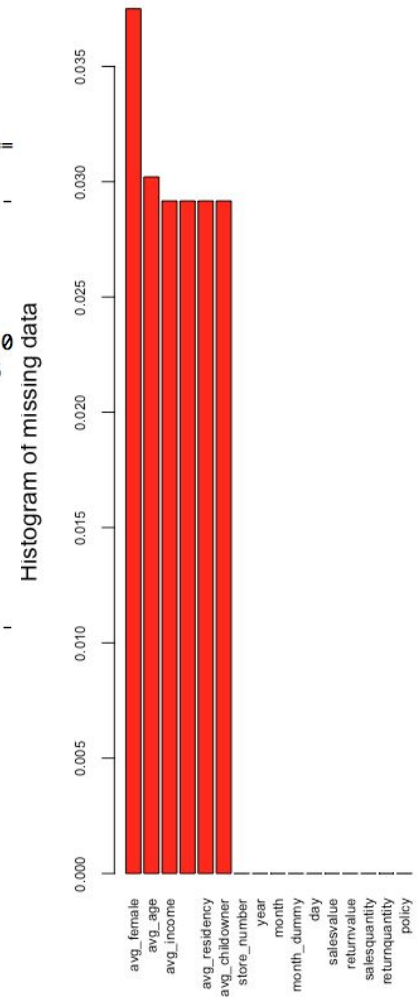
Missing Data

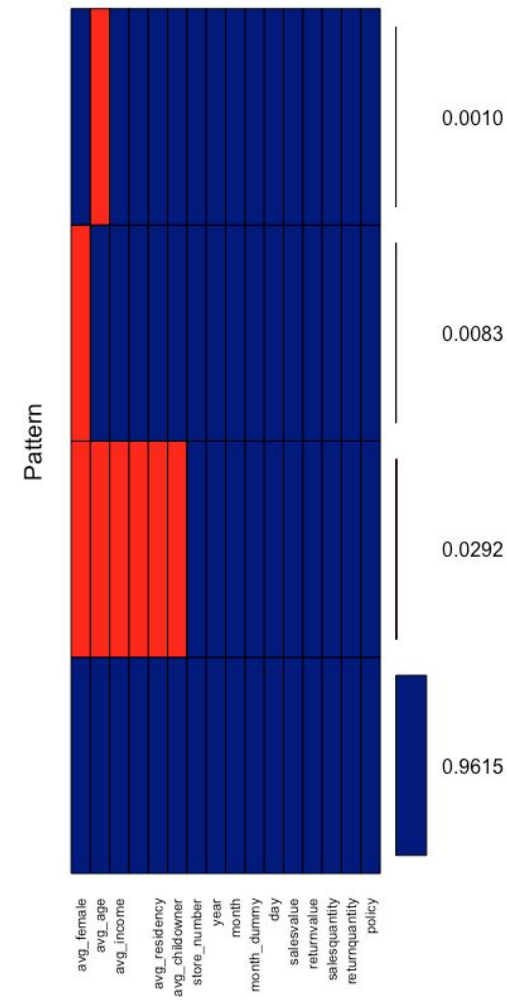
Online Sales and Returns

Descriptive Statistics

Statistic	N	Mean	St. Dev.	Min	Pctl(25)	Median	Pctl(75)	Max
store_number	960	5.8	3.3	2	2	6	10	10
year	960	2,013.2	0.4	2,013	2,013	2,013	2,013	2,014
month_dummy	960	6.6	3.4	1	4	7	10	12
day	960	179.0	104.7	1	88	178	265.2	365
salesvalue	960	103,682.1	219,624.5	0.0	4,138.5	12,333.3	138,976.5	2,133,233.0
returnvalue	960	14,748.5	28,608.2	0.0	310.1	1,549.6	18,883.4	254,989.6
salesquantity	960	638.9	1,333.9	1	59.5	135	778.8	13,074
returnquantity	960	56.4	109.6	0	4	10	71	1,033
avg_female	924	0.6	0.2	0.0	0.5	0.6	0.8	1.0
avg_age	931	4.4	1.3	0.0	3.8	4.4	5.0	13.0
avg_income	932	5.3	0.7	1.0	5.1	5.3	5.6	9.0
avg_homeowner	932	0.7	0.1	0.0	0.6	0.7	0.7	1.0
avg_residency	932	7.2	1.7	0.0	6.7	7.2	7.8	15.0
avg_childowner	932	0.4	0.1	0.0	0.4	0.4	0.5	1.0

Only **1%** of overall data is missing, if we delete the rows of data with missing values then **3.8%** of the data set would be missing.





1% of data missing

Descriptive Statistics Online Stores (After Data Imputation)

Descriptive Statistics Online

Statistic	N	Mean	St. Dev.	Min	Pctl(25)	Median	Pctl(75)	Max
salesvalue	960	103,682.1	219,624.5	0.0	4,138.5	12,333.3	138,976.5	2,133,233.0
salesquantity	960	638.9	1,333.9	1	59.5	135	778.8	13,074
returnvalue	960	14,748.5	28,608.2	0.0	310.1	1,549.6	18,883.4	254,989.6
returnquantity	960	56.4	109.6	0	4	10	71	1,033
avg_female	960	0.6	0.2	0.0	0.5	0.6	0.8	1.0
avg_age	960	4.4	1.5	0.0	3.8	4.4	5.0	13.0
avg_income	960	5.4	0.8	1.0	5.1	5.3	5.6	9.0
avg_homeowner	960	0.7	0.1	0.0	0.6	0.7	0.7	1.0
avg_residency	960	7.2	1.9	0.0	6.7	7.2	7.8	15.0
avg_childowner	960	0.4	0.2	0.0	0.4	0.4	0.5	1.0
timedummy	960	0.5	0.5	0	0	0	1	1
policydummy	960	0.7	0.5	0	0	1	1	1

VIF Scores for Online Sales

Before

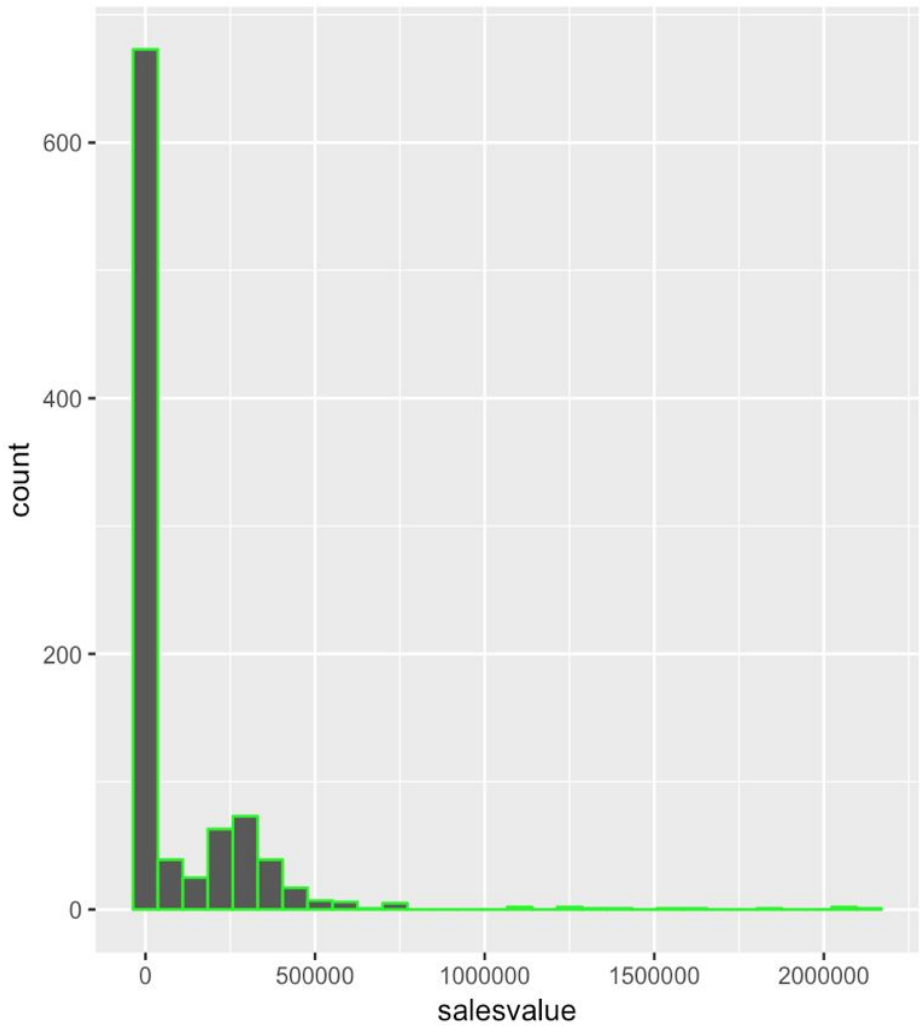
	Variables	VIF
1	store_number	6.379196
2	year	322.782908
3	month_dummy	151.676609
4	day	145.801956
5	returnvalue	23.875052
6	salesquantity	24.185957
7	returnquantity	58.095546
8	avg_female	1.987723
9	avg_age	1.217422
10	avg_income	1.399197
11	avg_homeowner	1.505152
12	avg_residency	1.077255
13	avg_childowner	1.089318
14	timedummy	4.426193
15	policydummy	4.973294

After

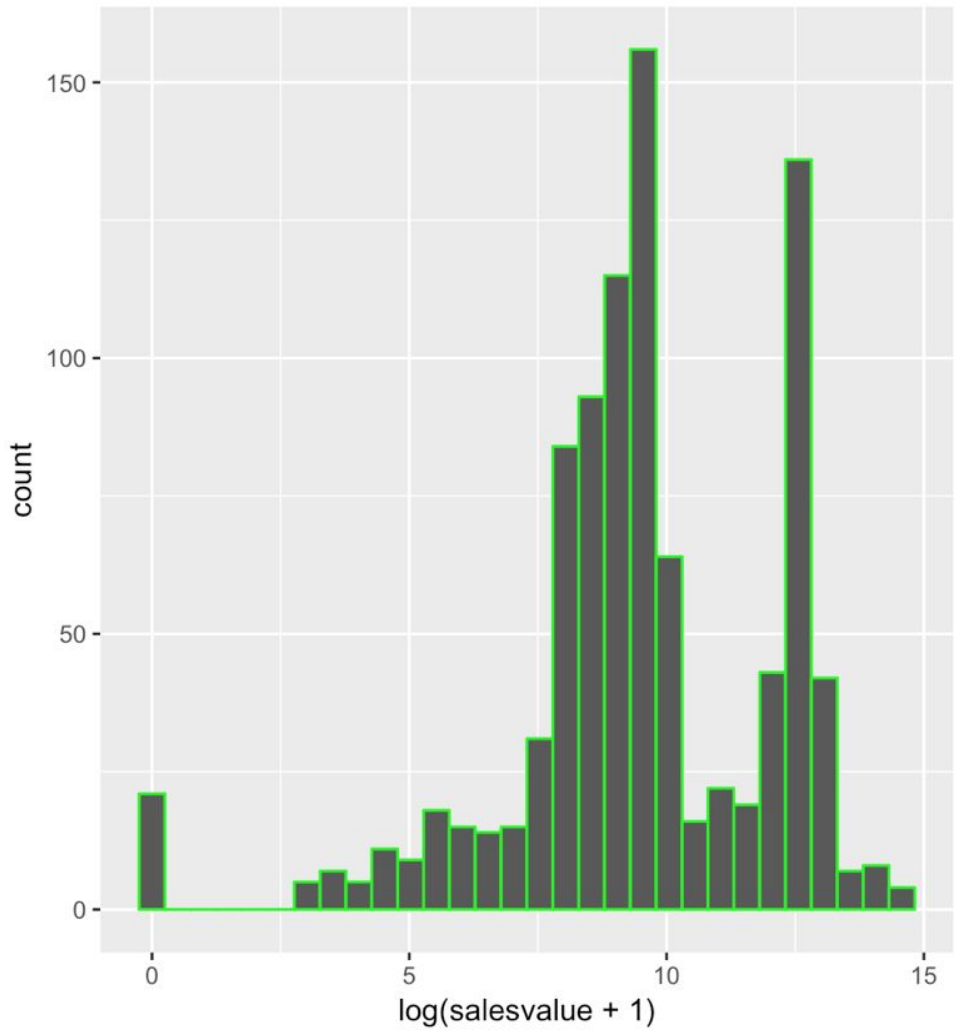
	Variables	VIF
1	salesquantity	1.102780
2	avg_female	1.945407
3	avg_age	1.165368
4	avg_income	1.354547
5	avg_homeowner	1.459239
6	avg_residency	1.068947
7	avg_childowner	1.086658
8	timedummy	1.049504
9	policydummy	1.873676

Log Transforming Sales Value

Before



After



Output for Online Sales Value

Std. Err. Regression Results

	Dependent variable:	
	log(salesvalue + 1)	
	Normal SE (1)	HW-Robust SE (2)
policydummy	1.60***	1.60***
	(0.27)	(0.32)
timedummy	0.17	0.17
	(0.27)	(0.24)
avg_female	-3.56***	-3.56***
	(0.50)	(0.95)
avg_age	0.05	0.05
	(0.05)	(0.09)
avg_income	0.16	0.16
	(0.11)	(0.19)
avg_homeowner	-1.51*	-1.51
	(0.61)	(1.22)
avg_residency	0.22***	0.22**
	(0.04)	(0.08)
avg_childowner	0.25	0.25
	(0.48)	(0.88)
policydummy:timedummy	-0.13	-0.13
	(0.32)	(0.30)
Constant	8.83***	8.83***
	(0.77)	(1.37)
Observations	960	960
R2	0.26	0.26
Adjusted R2	0.25	0.25
Residual Std. Error	2.28	2.28
F Statistic	36.88***	36.88***
Note:	*p<0.05; **p<0.01; ***p<0.001	

```
        studentized Breusch-Pagan test

data:  ol_sv_linear_1
BP = 44.071, df = 9, p-value = 1.37e-06
```

Significant, meaning we have heteroskedasticity

Output for Online Sales Quantity

Std. Err. Regression Results

	Dependent variable:	
	salesquantity	
	Normal SE (1)	HW-Robust SE (2)
policydummy	0.39***	0.39***
	(0.16)	(0.19)
timedummy	1.21	1.21
	(0.16)	(0.15)
avg_female	0.001***	0.001***
	(0.30)	(0.48)
avg_age	1.06	1.06
	(0.03)	(0.05)
avg_income	1.43***	1.43***
	(0.07)	(0.10)
avg_homeowner	0.02***	0.02***
	(0.36)	(0.55)
avg_residency	1.24***	1.24***
	(0.02)	(0.04)
avg_childowner	3.10***	3.10***
	(0.29)	(0.45)
policydummy:timedummy	1.06	1.06
	(0.19)	(0.18)
Constant	9,483.03***	9,483.03***
	(0.46)	(0.80)
Observations	960	960
Log Likelihood	-6,681.96	-6,681.96
theta	0.55*** (0.02)	0.55*** (0.02)
Akaike Inf. Crit.	13,383.93	13,383.93
Note:	*p<0.05; **p<0.01; ***p<0.001	

```
        Goldfeld-Quandt test

data:  ol_sq_negbin_1
GQ = 5.7137, df1 = 470, df2 = 470, p-value < 2.2e-16
alternative hypothesis: variance increases from segment 1 to 2
```

Significant, meaning we have heteroskedasticity

```
Likelihood ratio test

Model 1: salesquantity ~ policydummy + timedummy + timedummy * policydummy +
    avg_female + avg_age + avg_income + avg_homeowner + avg_residency +
    avg_childowner
Model 2: salesquantity ~ 1
  #Df  LogLik Df  Chisq Pr(>Chisq)
1  11 -6681.0
2   2 -6846.6 -9 331.37  < 2.2e-16 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

Negative Binomial model fits

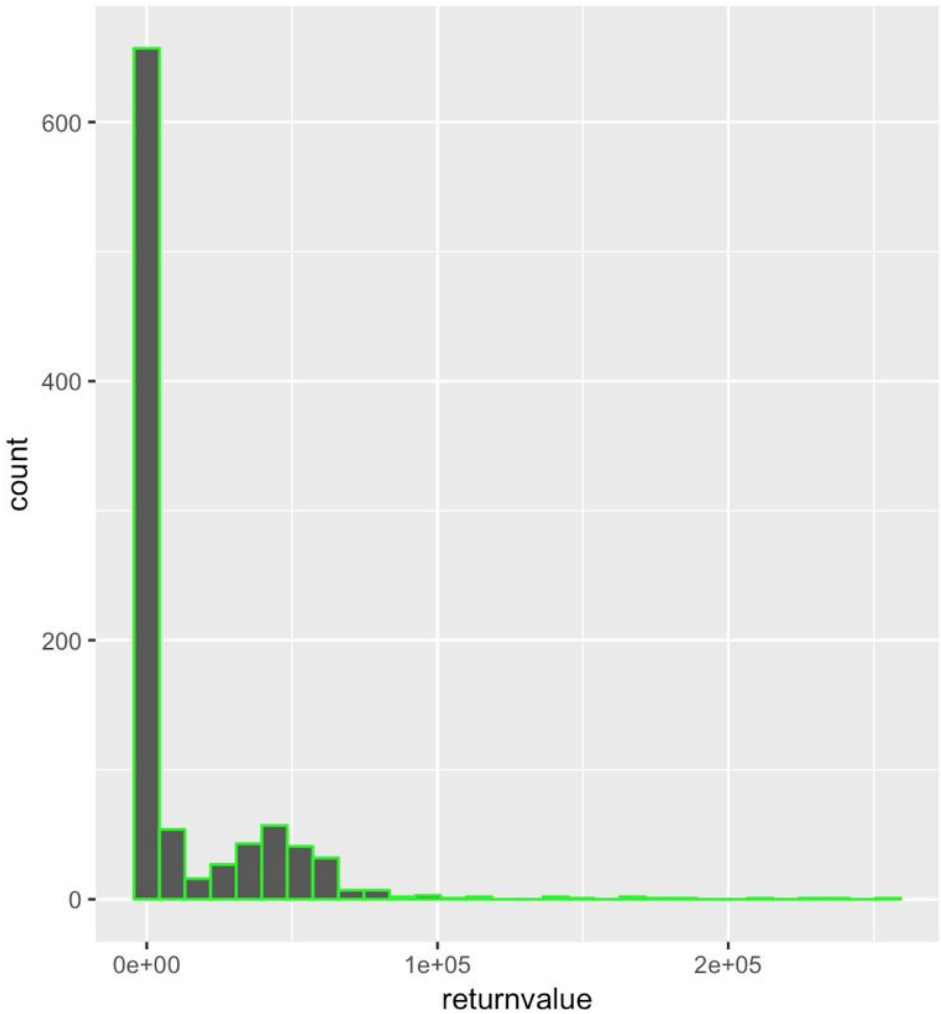
```
Likelihood ratio test

Model 1: salesquantity ~ policydummy + timedummy + policydummy * timedummy +
    avg_female + avg_age + avg_income + avg_homeowner + avg_residency +
    avg_childowner
Model 2: salesquantity ~ policydummy + timedummy + timedummy * policydummy +
    avg_female + avg_age + avg_income + avg_homeowner + avg_residency +
    avg_childowner
  #Df  LogLik Df   Chisq Pr(>Chisq)
1  10 -521705
2  11   -6681  1 1030048  < 2.2e-16 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

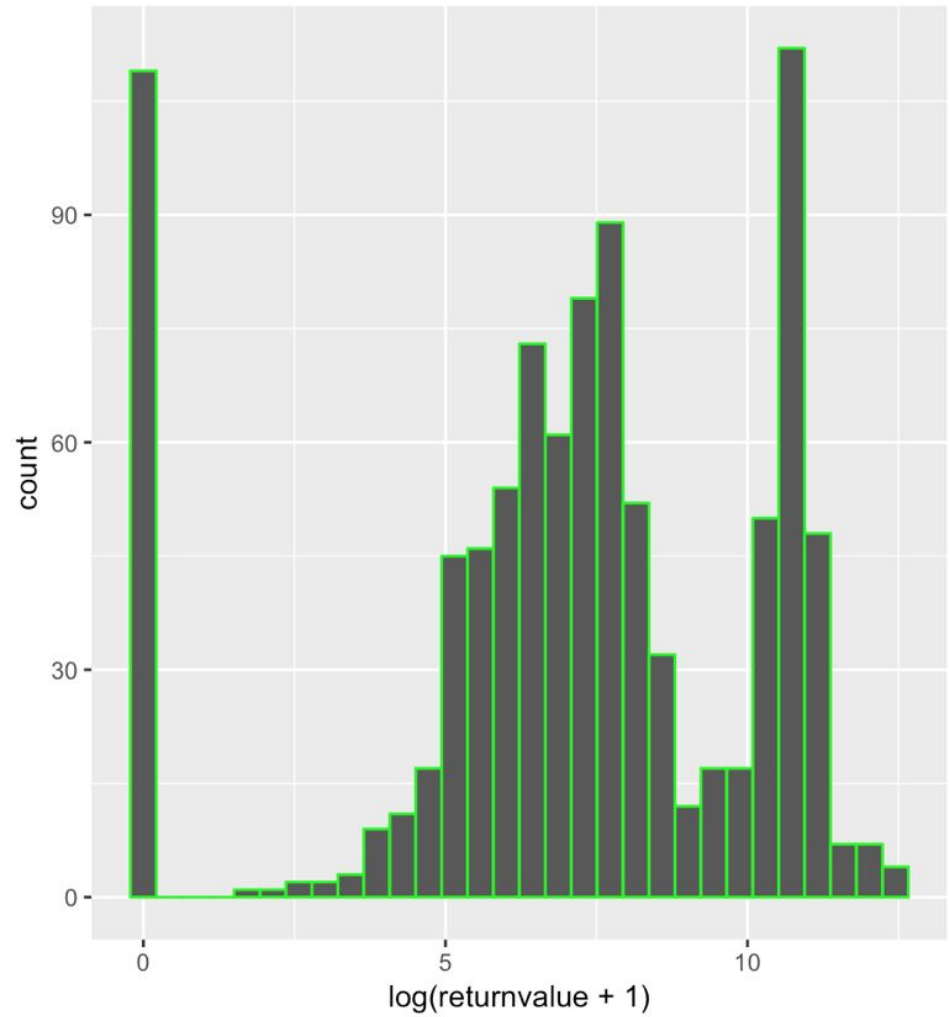
Choosing Neg. Binomial over Poisson model

Log Transforming Return Value

Before



After



Output for Online Return Value

Std. Err. Regression Results

	Dependent variable:	
	log(returnvalue + 1)	
	Normal SE (1)	HW-Robust SE (2)
policydummy	0.08	0.08
	(0.16)	(0.19)
timedummy	-0.08	-0.08
	(0.16)	(0.15)
log(salesvalue + 1)	1.11***	1.11***
	(0.02)	(0.03)
avg_female	0.07	0.07
	(0.31)	(0.48)
avg_age	0.02	0.02
	(0.03)	(0.05)
avg_income	0.03	0.03
	(0.07)	(0.10)
avg_homeowner	-0.38	-0.38
	(0.36)	(0.55)
avg_residency	-0.03	-0.03
	(0.02)	(0.04)
avg_childowner	0.26	0.26
	(0.28)	(0.45)
policydummy:timedummy	-0.11	-0.11
	(0.19)	(0.18)
Constant	-3.43***	-3.43***
	(0.49)	(0.80)
Observations	960	960
R2	0.83	0.83
Adjusted R2	0.82	0.82
Residual Std. Error	1.35	1.35
F Statistic	449.79***	449.79***
Note:	*p<0.05; **p<0.01; ***p<0.001	

```
        studentized Breusch-Pagan test

data:  ol_rv_linear_1
BP = 205.49, df = 10, p-value < 2.2e-16
```

Significant, meaning we have heteroskedasticity

Output for Online Return Quantity

Std. Err. Regression Results

	Dependent variable:	
	returnquantity	
	Normal SE (1)	HW-Robust SE (2)
policydummy	1.09	1.09
	(0.13)	(0.10)
timedummy	1.02	1.02
	(0.13)	(0.10)
salesquantity	1.00***	1.00***
	(0.0000)	(0.0000)
avg_female	0.25***	0.25***
	(0.24)	(0.19)
avg_age	0.98	0.98
	(0.03)	(0.02)
avg_income	1.07	1.07
	(0.05)	(0.04)
avg_homeowner	0.18***	0.18***
	(0.29)	(0.23)
avg_residency	1.12***	1.12***
	(0.02)	(0.02)
avg_childowner	1.19	1.19
	(0.23)	(0.19)
policydummy:timedummy	0.80	0.80
	(0.15)	(0.12)
Constant	21.57***	21.57***
	(0.37)	(0.29)
Observations	960	960
Log Likelihood	-3,982.88	-3,982.88
theta	0.96*** (0.05)	0.96*** (0.05)
Akaike Inf. Crit.	7,987.76	7,987.76
Note:	*p<0.05; **p<0.01; ***p<0.001	

```
studentized Breusch-Pagan test

data:  ol_rq_negbin_1
BP = 334.41, df = 10, p-value < 2.2e-16
```

Significant, meaning we have heteroskedasticity

```
Likelihood ratio test

Model 1: returnquantity ~ policydummy + timedummy + policydummy * timedummy +
    salesquantity + avg_female + avg_age + avg_income + avg_homeowner +
    avg_residency + avg_childowner
Model 2: salesquantity ~ 1
  #Df  LogLik  Df  Chisq Pr(>Chisq)
1  12 -3981.9
2   2 -6846.6 -10 5729.5  < 2.2e-16 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

Negative Binomial model fits

```
Likelihood ratio test

Model 1: returnquantity ~ policydummy + timedummy + policydummy * timedummy +
    salesquantity + avg_female + avg_age + avg_income + avg_homeowner +
    avg_residency + avg_childowner
Model 2: salesquantity ~ policydummy + timedummy + timedummy * policydummy +
    avg_female + avg_age + avg_income + avg_homeowner + avg_residency +
    avg_childowner
  #Df LogLik Df Chisq Pr(>Chisq)
1  11 -24952
2  11  -6681  0 36542  < 2.2e-16 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

Choosing Neg. Binomial over Poisson model

Output for Overall Online Profit

Std. Err. Regression Results

	Dependent variable:	
	log(Profit + 1)	
	Normal SE (1)	HW-Robust SE (2)
policydummy	1.51***	1.51***
	(0.26)	(0.31)
timedummy	0.14	0.14
	(0.26)	(0.23)
avg_female	-3.24***	-3.24***
	(0.50)	(0.92)
avg_age	0.002	0.002
	(0.05)	(0.09)
avg_income	0.24*	0.24
	(0.11)	(0.18)
avg_homeowner	-1.38*	-1.38
	(0.61)	(1.18)
avg_residency	0.22***	0.22**
	(0.04)	(0.07)
avg_childowner	0.24	0.24
	(0.47)	(0.84)
policydummy:timedummy	-0.09	-0.09
	(0.31)	(0.29)
Constant	7.37***	7.37***
	(0.75)	(1.28)
Observations	944	944
R2	0.25	0.25
Adjusted R2	0.24	0.24
Residual Std. Error	2.19	2.19
F Statistic	34.80***	34.80***
Note:	*p<0.05; **p<0.01; ***p<0.001	

Online Sales Data at Product Level

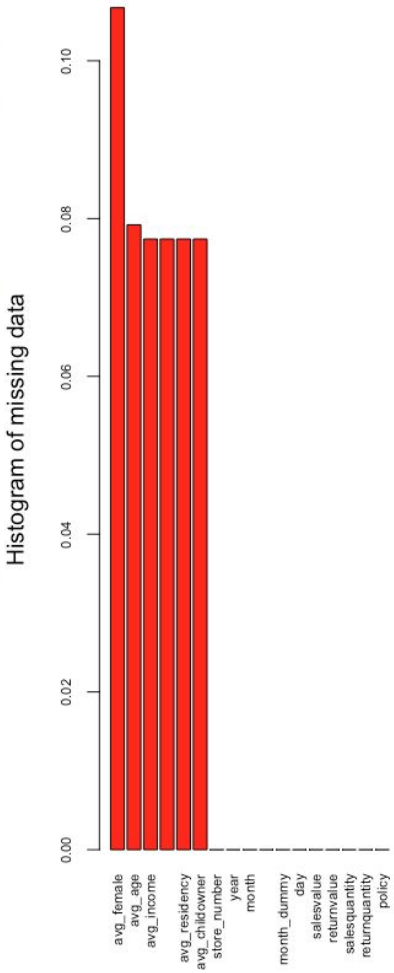
Missing Data

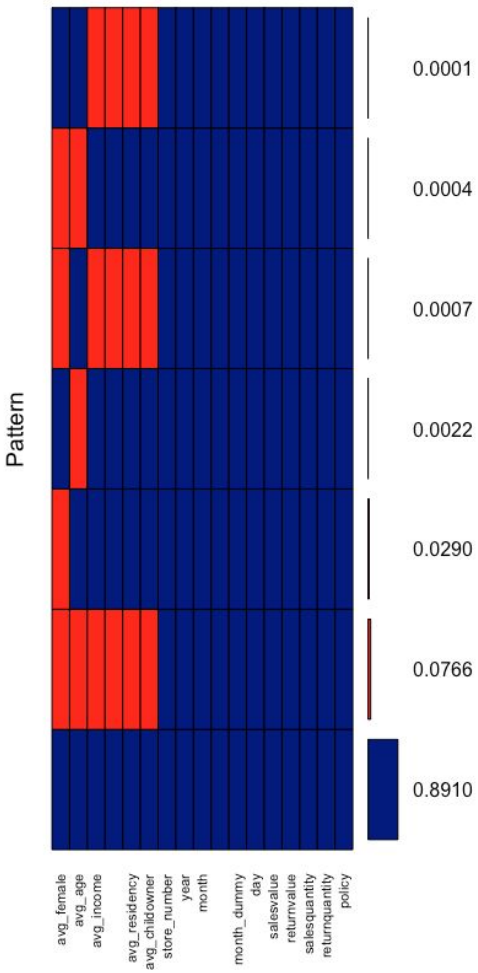
Online Sales and Returns by Product Category

Descriptive Statistics

Statistic	N	Mean	St. Dev.	Min	Pctl(25)	Median	Pctl(75)	Max
store_number	9,960	5.2	3.3	2	2	6	10	10
year	9,960	2,013.2	0.4	2,013	2,013	2,013	2,013	2,014
product_category	9,960	9.5	6.2	1	4	9	14	21
month_dummy	9,960	6.8	3.5	1	4	7	10	12
day	9,960	178.6	104.2	1	87	179	263	365
salesvalue	9,960	9,665.8	23,605.7	0.0	367.1	1,328.5	8,779.9	430,846.5
returnvalue	9,960	1,378.9	3,325.9	0.0	0.0	90.0	1,176.1	59,138.5
salesquantity	9,960	59.1	180.0	1	3	12	39	3,705
returnquantity	9,960	5.2	13.1	0	0	1	4	228
avg_female	8,897	0.6	0.3	0.0	0.4	0.6	0.8	1.0
avg_age	9,171	4.3	2.3	0.0	3.2	4.2	5.4	13.0
avg_income	9,189	5.2	1.4	1.0	4.7	5.3	6.0	9.0
avg_homeowner	9,189	0.7	0.3	0.0	0.5	0.7	0.8	1.0
avg_residency	9,189	7.0	3.1	0.0	5.5	7.0	8.2	15.0
avg_childowner	9,189	0.4	0.3	0.0	0.3	0.4	0.5	1.0

While only **3%** of overall data is missing, if we delete the rows of data with missing values then **10.7%** of the data set would be missing and might skew the results.





3% of data missing

Descriptive Statistics Online at Product Level (After Data Imputation)

Descriptive Statistics

Statistic	N	Mean	St. Dev.	Min	Pctl(25)	Median	Pctl(75)	Max
X	9,960	4,980.5	2,875.3	1	2,490.8	4,980.5	7,470.2	9,960
store_number	9,960	5.2	3.3	2	2	6	10	10
year	9,960	2,013.2	0.4	2,013	2,013	2,013	2,013	2,014
month_dummy	9,960	6.8	3.5	1	4	7	10	12
day	9,960	178.6	104.2	1	87	179	263	365
salesvalue	9,960	9,665.8	23,605.7	0.0	367.1	1,328.5	8,779.9	430,846.5
returnvalue	9,960	1,378.9	3,325.9	0.0	0.0	90.0	1,176.1	59,138.5
salesquantity	9,960	59.1	180.0	1	3	12	39	3,705
returnquantity	9,960	5.2	13.1	0	0	1	4	228
avg_female	9,960	0.6	0.3	0.0	0.4	0.6	1.0	1.0
avg_age	9,960	4.3	2.4	0.0	3.1	4.2	5.5	13.0
avg_income	9,960	5.2	1.4	1.0	4.7	5.3	6.0	9.0
avg_homeowner	9,960	0.7	0.3	0.0	0.5	0.7	0.9	1.0
avg_residency	9,960	7.0	3.3	0.0	5.4	7.0	8.3	15.0
avg_childowner	9,960	0.4	0.3	0.0	0.2	0.4	0.5	1.0
log_salesvalue	9,960	7.3	2.2	0.0	5.9	7.2	9.1	13.0
policydummy	9,960	0.7	0.4	0	0	1	1	1
timedummy	9,960	0.5	0.5	0	0	0	1	1

Output for Online Sales Value at Prod. Level

	Dependent variable:	
	log(salesvalue + 1)	
	Normal SE	HW-Robust SE
	(1)	(2)
policydummy	2.23***	2.23***
	(0.07)	(0.06)
timedummy	0.14	0.14*
	(0.07)	(0.06)
avg_female	-0.42***	-0.42***
	(0.06)	(0.07)
avg_age	0.01	0.01
	(0.01)	(0.01)
FactorPC2	-1.21***	-1.21***
	(0.10)	(0.10)
FactorPC3	-0.11	-0.11
	(0.11)	(0.11)
FactorPC4	-0.09	-0.09
	(0.09)	(0.09)
FactorPC5	-0.50***	-0.50***
	(0.09)	(0.09)
FactorPC6	-1.08***	-1.08***
	(0.10)	(0.09)
FactorPC7	-1.90***	-1.90***
	(0.11)	(0.12)
FactorPC8	-1.91***	-1.91***
	(0.17)	(0.10)
FactorPC9	-2.50***	-2.50***
	(0.11)	(0.12)
FactorPC10	-2.12***	-2.12***
	(0.22)	(0.15)
FactorPC11	-0.81***	-0.81***
	(0.11)	(0.11)
FactorPC12	-0.24*	-0.24*
	(0.09)	(0.10)
FactorPC13	-1.50***	-1.50***
	(0.11)	(0.10)
FactorPC14	-3.05***	-3.05***
	(0.12)	(0.11)
FactorPC15	-0.71***	-0.71***
	(0.14)	(0.11)
FactorPC16	-0.01	-0.01
	(0.14)	(0.12)
FactorPC17	-0.76***	-0.76***
	(0.14)	(0.14)
FactorPC20	-0.53***	-0.53***
	(0.11)	(0.11)
FactorPC21	-0.48***	-0.48***
	(0.09)	(0.09)
avg_income	0.09***	0.09***
	(0.01)	(0.02)
avg_homeowner	-0.21**	-0.21*
	(0.08)	(0.10)
avg_residency	0.003	0.003
	(0.01)	(0.01)
avg_childowner	-0.05	-0.05
	(0.07)	(0.08)
policydummy:timedummy	-0.02	-0.02
	(0.08)	(0.08)
Constant	6.30***	6.30***
	(0.12)	(0.12)
Observations	9,960	9,960
R2	0.33	0.33
Adjusted R2	0.33	0.33
Residual Std. Error	1.84	1.84
F Statistic	185.09***	185.09***
Note:	*p<0.05; **p<0.01; ***p<0.001	

Output for Online Sales Quantity at Prod. Level

Std. Err. Regression Results

	Dependent variable:	
	salesquantity	
	Normal SE (1)	HW-Robust SE (2)
policydummy	3.819***	3.819***
	(0.046)	(0.048)
timedummy	1.465***	1.465***
	(0.052)	(0.068)
FactorPC2	1.855***	1.855***
	(0.074)	(0.073)
FactorPC3	0.618***	0.618***
	(0.079)	(0.073)
FactorPC4	3.272***	3.272***
	(0.066)	(0.075)
FactorPC5	6.426***	6.426***
	(0.066)	(0.080)
FactorPC6	1.724***	1.724***
	(0.068)	(0.067)
FactorPC7	5.015***	5.015***
	(0.075)	(0.123)
FactorPC8	0.109***	0.109***
	(0.128)	(0.083)
FactorPC9	2.785***	2.785***
	(0.077)	(0.094)
FactorPC10	0.079***	0.079***
	(0.168)	(0.091)
FactorPC11	0.953	0.953
	(0.080)	(0.101)
FactorPC12	7.103***	7.103***
	(0.067)	(0.074)
FactorPC13	0.949	0.949
	(0.076)	(0.081)
FactorPC14	0.466***	0.466***
	(0.084)	(0.076)
FactorPC15	3.039***	3.039***
	(0.100)	(0.106)
FactorPC16	13.284***	13.284***
	(0.100)	(0.088)
FactorPC17	0.327***	0.327***
	(0.101)	(0.116)
FactorPC20	1.061	1.061
	(0.076)	(0.071)
FactorPC21	8.411***	8.411***
	(0.066)	(0.074)
avg_female	0.477***	0.477***
	(0.045)	(0.036)
avg_age	1.010	1.010
	(0.006)	(0.004)
avg_income	1.126***	1.126***
	(0.010)	(0.008)
avg_homeowner	0.760***	0.760***
	(0.054)	(0.039)
avg_residency	1.014**	1.014**
	(0.004)	(0.003)
avg_childowner	1.130**	1.130**
	(0.046)	(0.035)
policydummy:timedummy	0.996	0.996
	(0.060)	(0.077)
Constant	3.761***	3.761***
	(0.085)	(0.074)
Observations	9,960	9,960
Log Likelihood	-44,453.030	-44,453.030
theta	0.616*** (0.008)	0.616*** (0.008)
Akaike Inf. Crit.	88,962.050	88,962.050
Note:	*p<0.05; **p<0.01; ***p<0.001	

Output for Online Return Value at Prod. Level

Std. Err. Regression Results

	Dependent variable:	
	log(returnvalue + 1)	
	Normal SE (1)	HW-Robust SE (2)
policydummy	-0.13	-0.13
	(0.08)	(0.08)
timedummy	0.003	0.003
	(0.09)	(0.08)
log(salesvalue + 1)	1.23***	1.23***
	(0.01)	(0.01)
avg_female	0.21**	0.21*
	(0.08)	(0.09)
FactorPC2	0.86***	0.86***
	(0.12)	(0.13)
FactorPC3	-0.06	-0.06
	(0.13)	(0.17)
FactorPC4	0.19	0.19
	(0.11)	(0.12)
FactorPC5	0.46***	0.46***
	(0.11)	(0.12)
FactorPC6	0.38***	0.38**
	(0.11)	(0.12)
FactorPC7	0.67***	0.67***
	(0.13)	(0.13)
FactorPC8	0.10	0.10
	(0.20)	(0.27)
FactorPC9	0.73***	0.73***
	(0.13)	(0.14)
FactorPC10	-0.30	-0.30
	(0.26)	(0.34)
FactorPC11	0.004	0.004
	(0.13)	(0.15)
FactorPC12	0.72***	0.72***
	(0.11)	(0.12)
FactorPC13	0.54***	0.54***
	(0.13)	(0.14)
FactorPC14	1.07***	1.07***
	(0.14)	(0.16)
FactorPC15	0.06	0.06
	(0.17)	(0.17)
FactorPC16	0.15	0.15
	(0.17)	(0.16)
FactorPC17	-0.40*	-0.40
	(0.16)	(0.25)
FactorPC20	0.58***	0.58***
	(0.13)	(0.14)
FactorPC21	0.37***	0.37**
	(0.11)	(0.12)
avg_age	-0.06***	-0.06***
	(0.01)	(0.01)
avg_income	0.01	0.01
	(0.02)	(0.02)
avg_homeowner	-0.12	-0.12
	(0.09)	(0.10)
avg_residency	-0.004	-0.004
	(0.01)	(0.01)
avg_childowner	0.01	0.01
	(0.08)	(0.09)
policydummy:timedummy	-0.19	-0.19*
	(0.10)	(0.09)
Constant	-5.16***	-5.16***
	(0.16)	(0.19)
Observations	9,960	9,960
R2	0.60	0.60
Adjusted R2	0.60	0.60
Residual Std. Error	2.18	2.18
F Statistic	538.87***	538.87***
Note:	*p<0.05; **p<0.01; ***p<0.001	

Output for Online Return Quantity at Prod. Lev

Std. Err. Regression Results

	Dependent variable:	
	returnquantity	
	Normal SE (1)	HW-Robust SE (2)
policydummy	2.15***	2.15***
	(0.05)	(0.05)
timedummy	1.10	1.10
	(0.06)	(0.06)
salesquantity	1.01***	1.01***
	(0.0001)	(0.0001)
FactorPC2	1.22**	1.22**
	(0.07)	(0.06)
FactorPC3	0.74***	0.74***
	(0.08)	(0.07)
FactorPC4	0.94	0.94
	(0.07)	(0.06)
FactorPC5	0.60***	0.60***
	(0.07)	(0.06)
FactorPC6	0.87*	0.87*
	(0.07)	(0.06)
FactorPC7	0.74***	0.74***
	(0.08)	(0.08)
FactorPC8	0.27***	0.27***
	(0.15)	(0.12)
FactorPC9	0.38***	0.38***
	(0.09)	(0.10)
FactorPC10	0.20***	0.20***
	(0.21)	(0.19)
FactorPC11	0.64***	0.64***
	(0.08)	(0.07)
FactorPC12	0.81**	0.81**
	(0.07)	(0.06)
FactorPC13	0.62***	0.62***
	(0.08)	(0.07)
FactorPC14	0.34***	0.34***
	(0.10)	(0.09)
FactorPC15	0.61***	0.61***
	(0.12)	(0.13)
FactorPC16	0.91	0.91
	(0.11)	(0.09)
FactorPC17	0.30***	0.30***
	(0.12)	(0.09)
FactorPC20	1.22**	1.22**
	(0.08)	(0.06)
FactorPC21	0.71***	0.71***
	(0.07)	(0.06)
avg_female	0.94	0.94
	(0.05)	(0.04)
avg_age	0.97***	0.97***
	(0.01)	(0.005)
avg_income	1.08***	1.08***
	(0.01)	(0.01)
avg_homeowner	0.89*	0.89*
	(0.06)	(0.04)
avg_residency	1.00	1.00
	(0.005)	(0.004)
avg_childowner	0.97	0.97
	(0.05)	(0.04)
policydummy:timedummy	0.83**	0.83**
	(0.07)	(0.07)
Constant	1.02	1.02
	(0.09)	(0.08)
Observations	9,960	9,960
Log Likelihood	-21,017.70	-21,017.70
theta	0.75*** (0.02)	0.75*** (0.02)
Akaike Inf. Crit.	42,093.40	42,093.40
Note:	*p<0.05; **p<0.01; ***p<0.001	

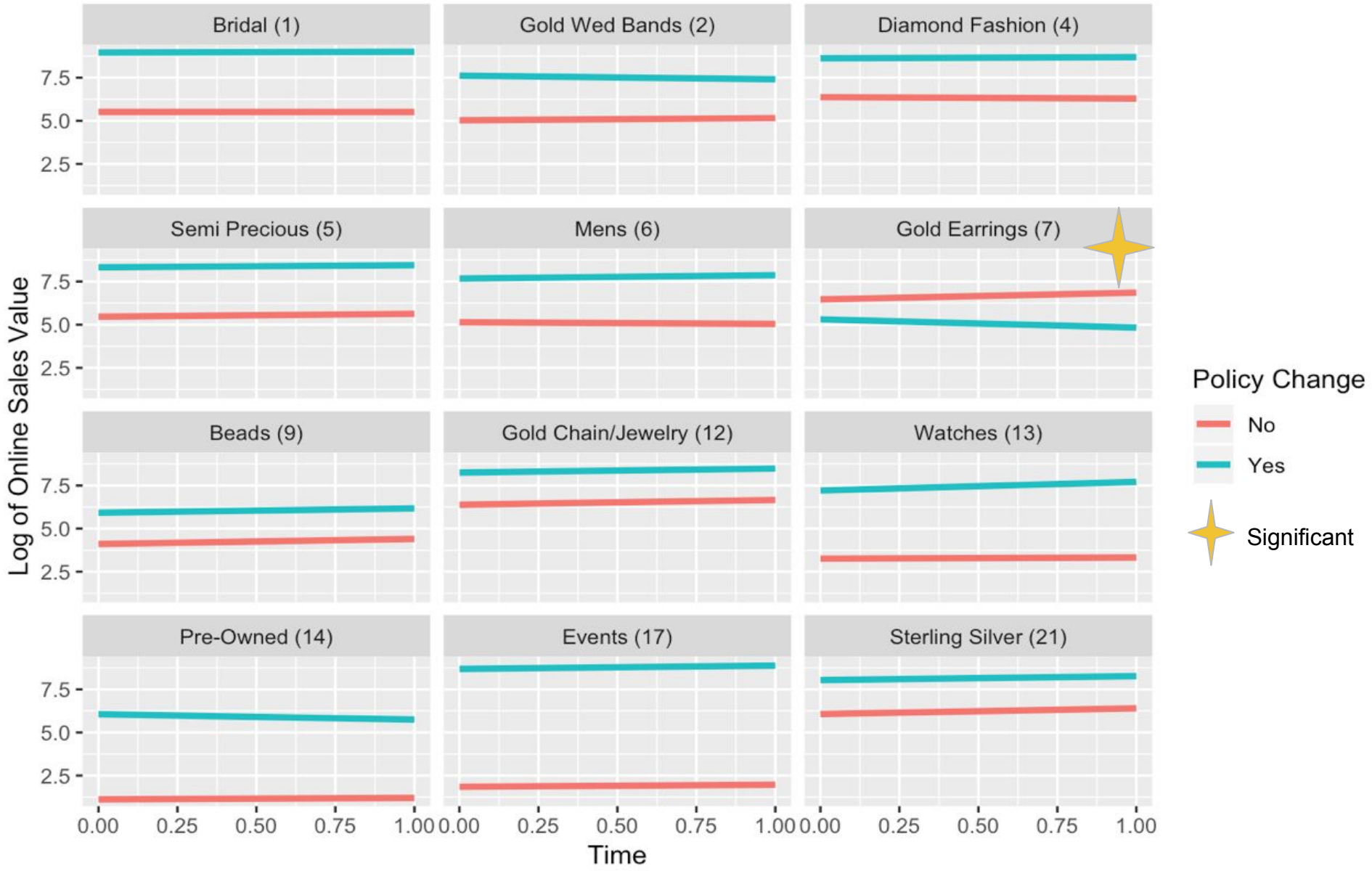
Output for Online Profit at Prod. Level

Std. Err. Regression Results

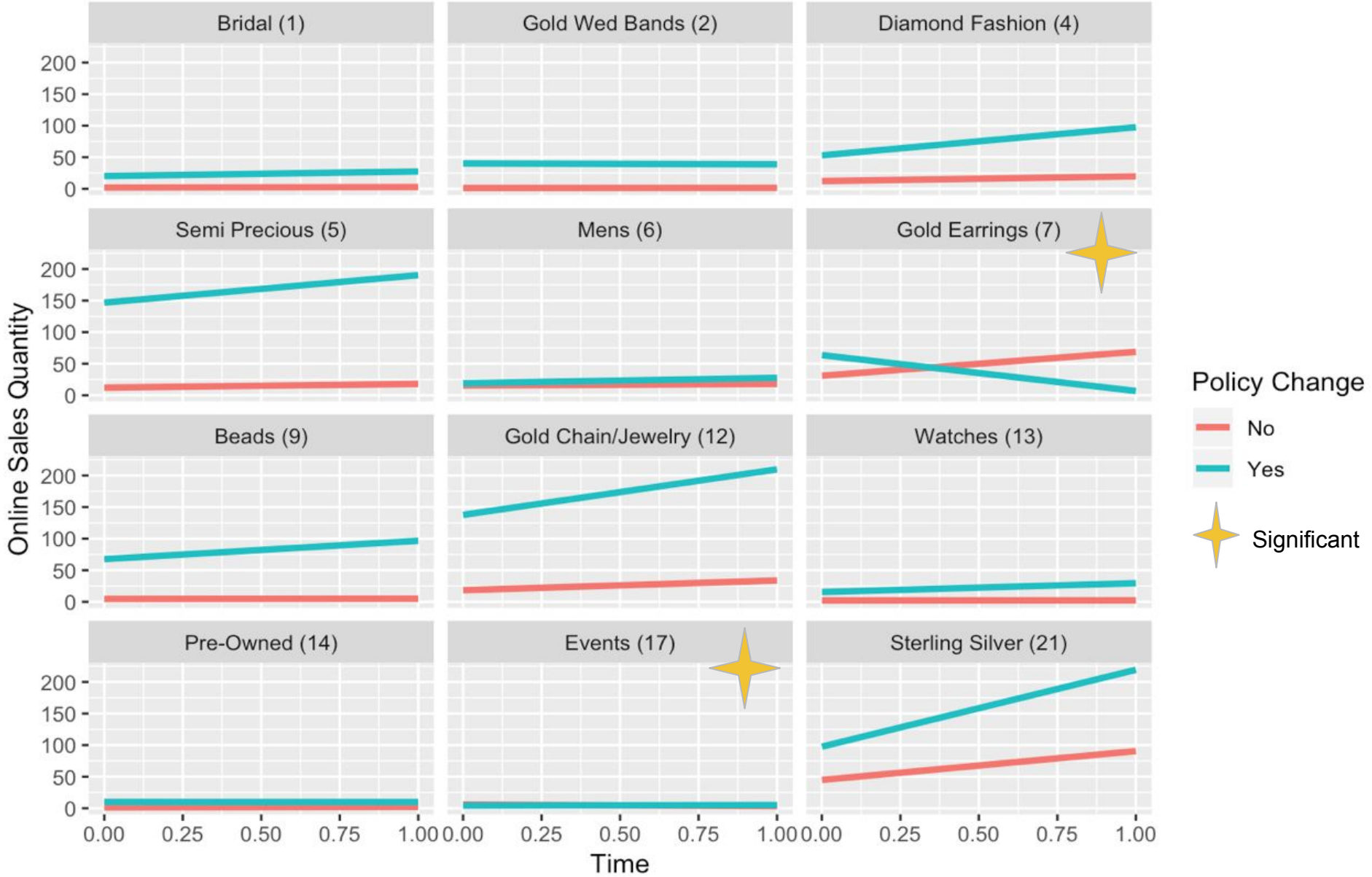
	Dependent variable:	
	log(Profit + 1)	
	Normal SE (1)	HW-Robust SE (2)
policydummy	2.10***	2.10***
	(0.06)	(0.06)
timedummy	0.14	0.14*
	(0.07)	(0.06)
avg_female	-0.39***	-0.39***
	(0.06)	(0.07)
avg_age	0.01	0.01
	(0.01)	(0.01)
FactorPC2	-1.26***	-1.26***
	(0.11)	(0.10)
FactorPC3	-0.10	-0.10
	(0.11)	(0.11)
FactorPC4	-0.12	-0.12
	(0.09)	(0.09)
FactorPC5	-0.45***	-0.45***
	(0.09)	(0.09)
FactorPC6	-1.08***	-1.08***
	(0.10)	(0.09)
FactorPC7	-1.76***	-1.76***
	(0.11)	(0.12)
FactorPC8	-1.89***	-1.89***
	(0.18)	(0.11)
FactorPC9	-2.32***	-2.32***
	(0.11)	(0.12)
FactorPC10	-2.09***	-2.09***
	(0.24)	(0.16)
FactorPC11	-0.74***	-0.74***
	(0.11)	(0.11)
FactorPC12	-0.27**	-0.27**
	(0.09)	(0.10)
FactorPC13	-1.38***	-1.38***
	(0.11)	(0.10)
FactorPC14	-3.04***	-3.04***
	(0.12)	(0.11)
FactorPC15	-0.70***	-0.70***
	(0.14)	(0.11)
FactorPC16	0.07	0.07
	(0.14)	(0.11)
FactorPC17	-0.75***	-0.75***
	(0.14)	(0.14)
FactorPC20	-0.65***	-0.65***
	(0.11)	(0.11)
FactorPC21	-0.41***	-0.41***
	(0.09)	(0.09)
avg_income	0.08***	0.08***
	(0.01)	(0.02)
avg_homeowner	-0.22**	-0.22*
	(0.08)	(0.09)
avg_residency	0.002	0.002
	(0.01)	(0.01)
avg_childowner	0.01	0.01
	(0.07)	(0.08)
policydummy:timedummy	0.03	0.03
	(0.08)	(0.08)
Constant	5.39***	5.39***
	(0.12)	(0.13)
Observations	9,371	9,371
R2	0.33	0.33
Adjusted R2	0.33	0.33
Residual Std. Error	1.77	1.77
F Statistic	169.48***	169.48***
Note:	*p<0.05; **p<0.01; ***p<0.001	

Sales and Return Analysis by Product Category

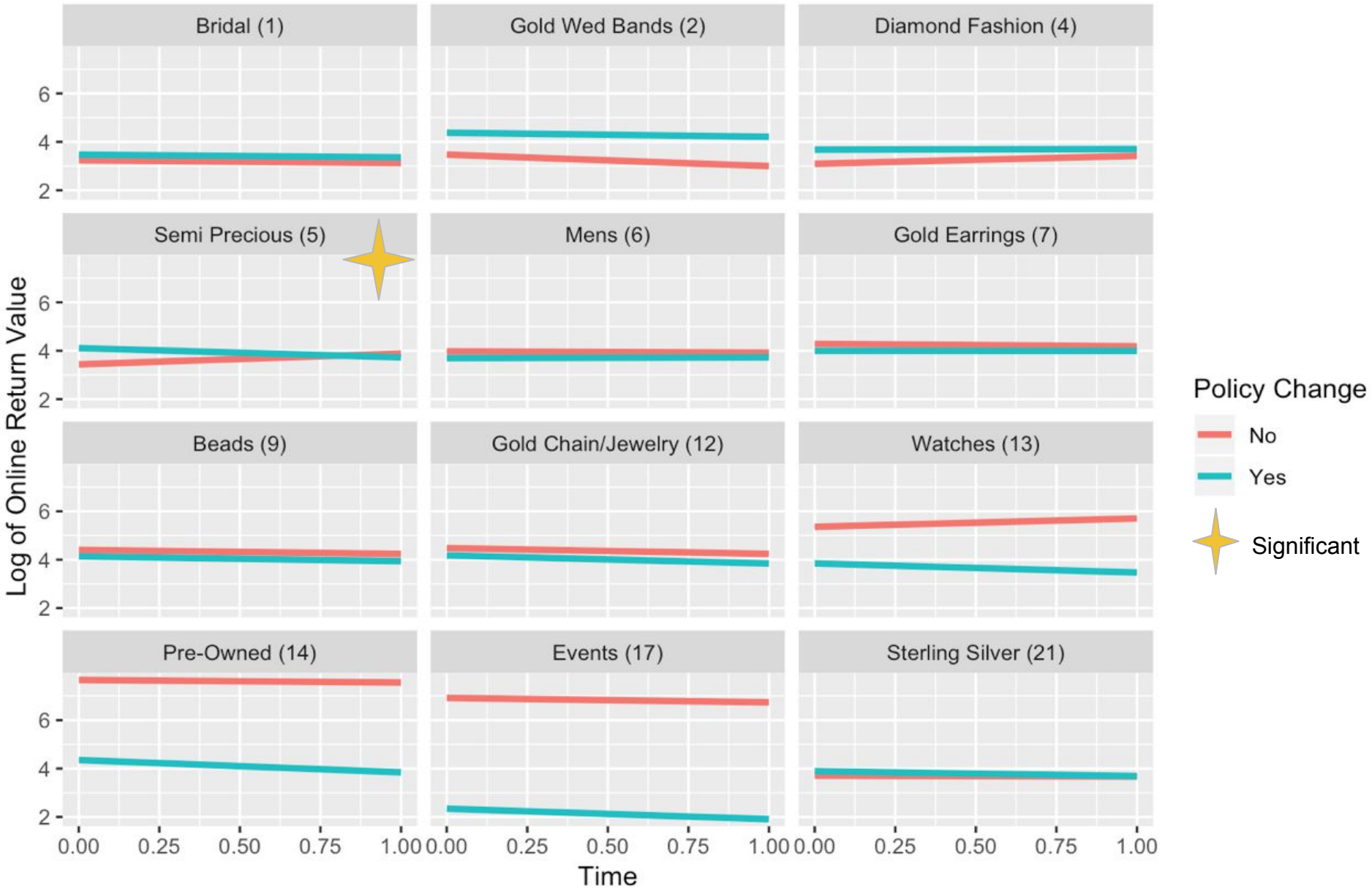
Online Sales Value



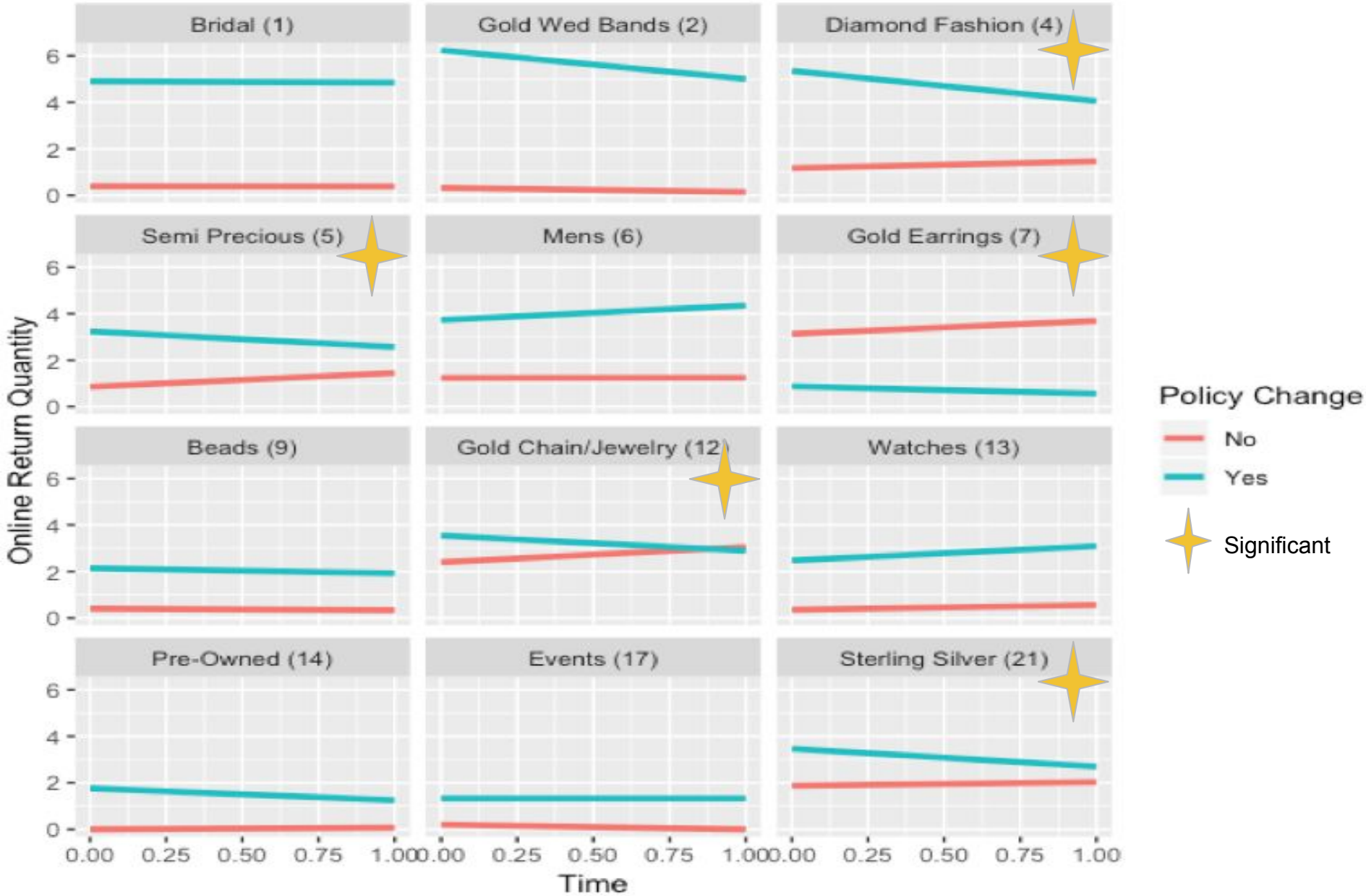
Online Sales Quantity



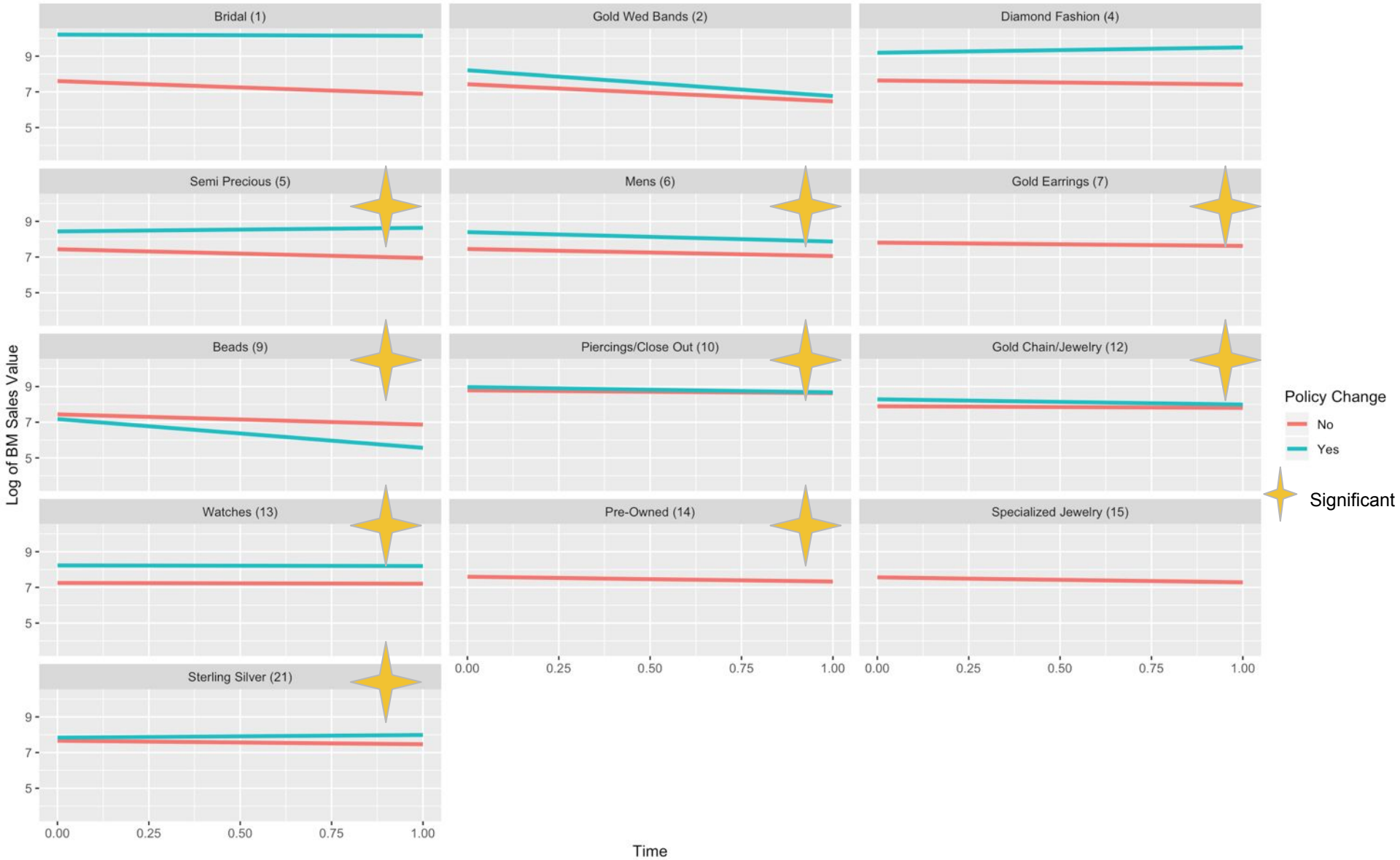
Online Return Value



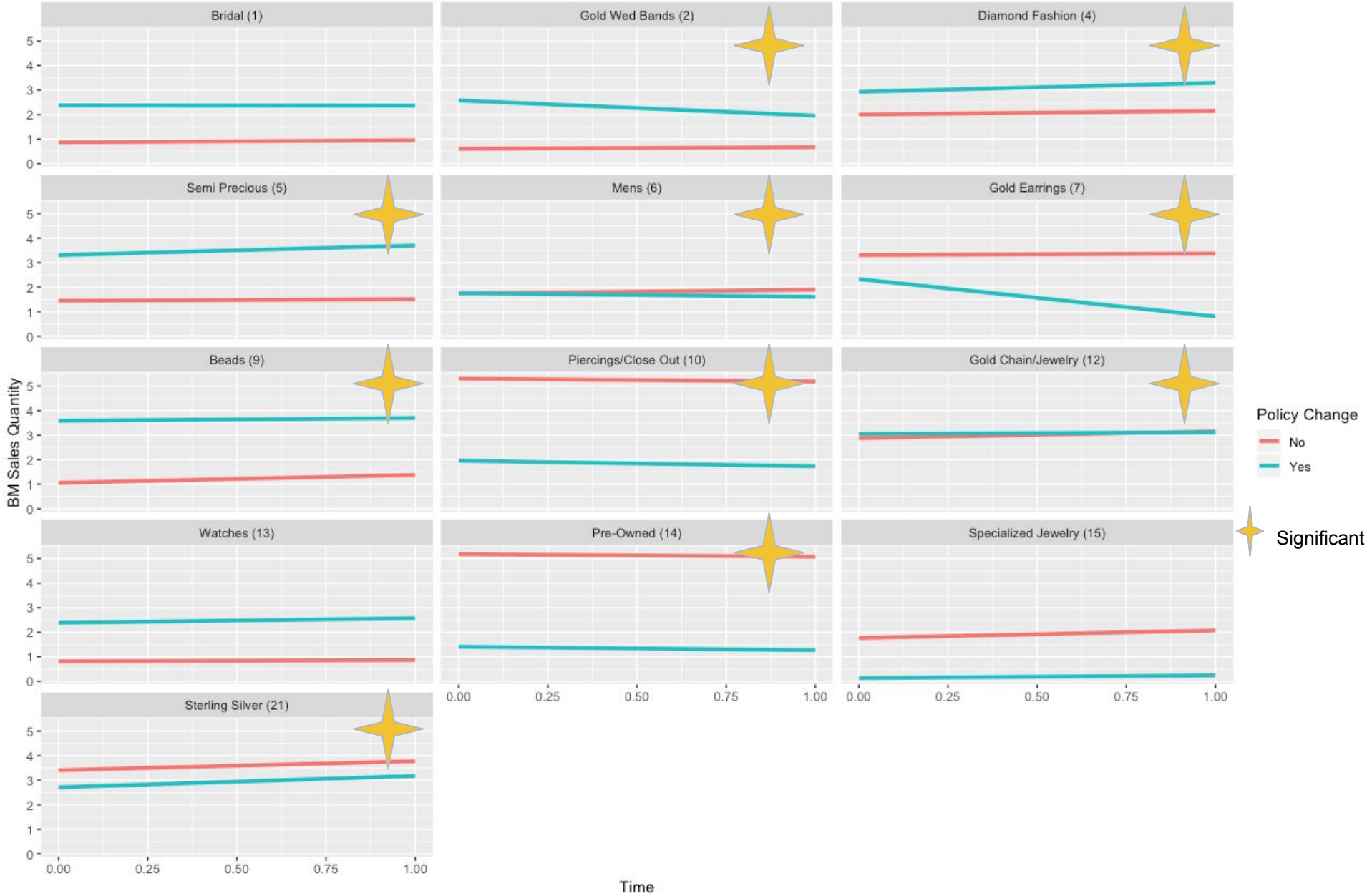
Online Return Quantity



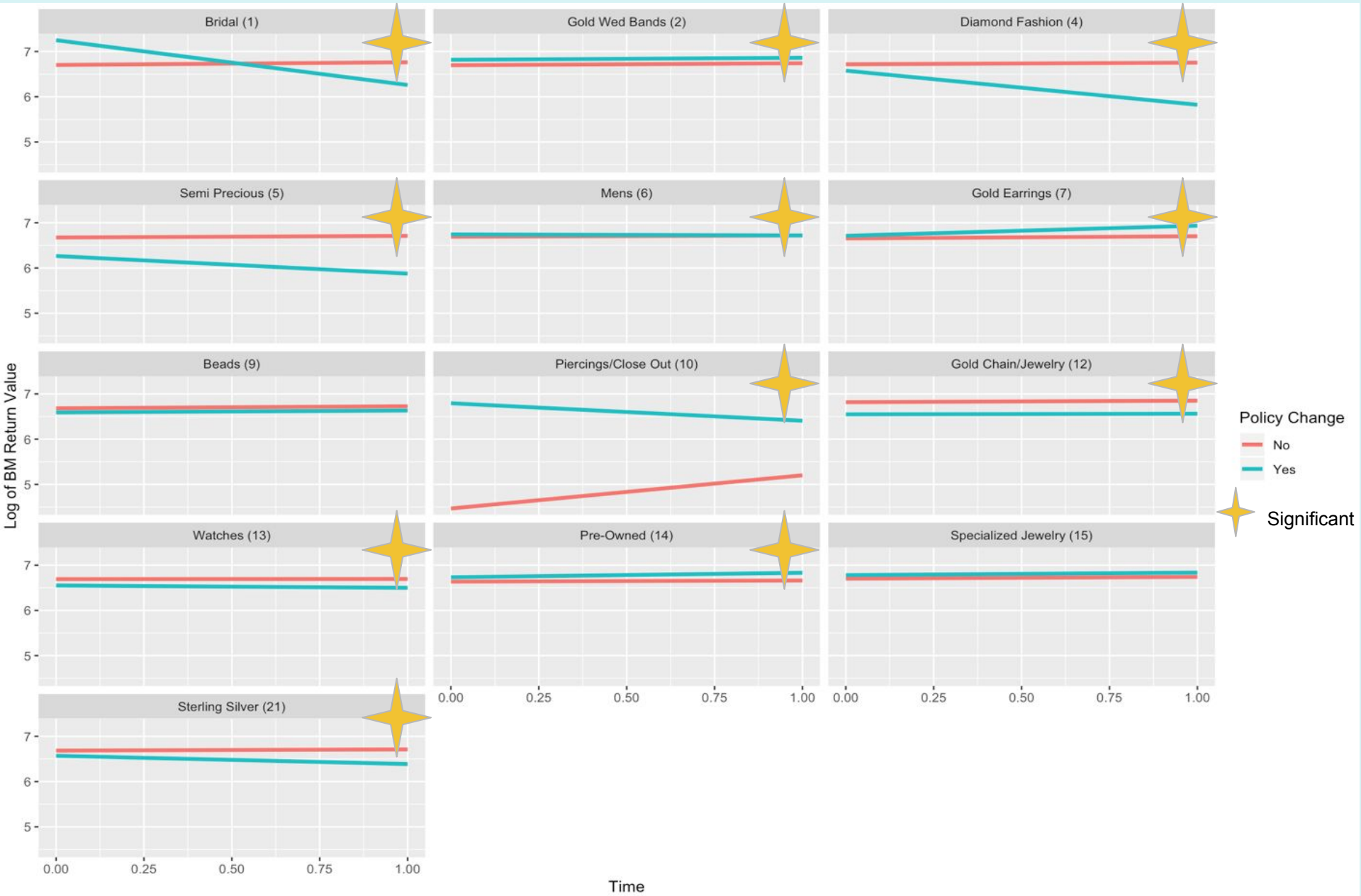
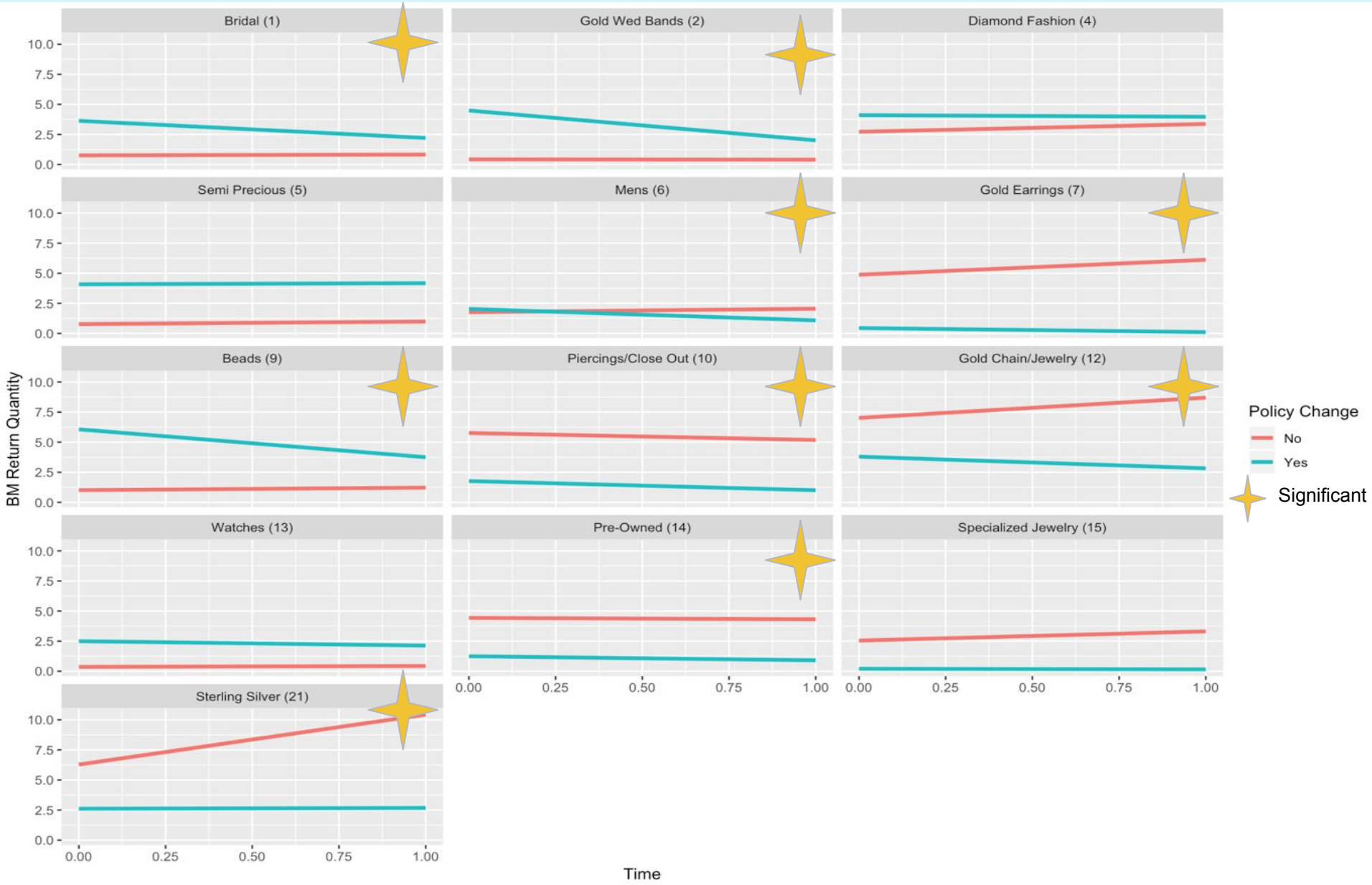
BM Sales Value



BM Sales Quantity



BM Return Value



BM Return Quantity