Econometrics with R
Business Analytics Project 3:

Impacts of Return Policy Change

Presenters: Amrita Sharma, Kyle Riener, Kevin Chiang, Tyler Young

Analytical Process

1. **Background**
2. **Research Questions**
3. **Key Variables**
4. **Questions 1 to 6**
   - Data Handling
   - Multicollinearity Test
   - Heteroskedasticity
   - Model discussion
5. **Conclusion**

Background

- Large jewelry retail company with several brands and categories
- Thousands of stores, (both physical and online) with millions of transaction data
- Data collected from **April 1st, 2013 to Mar 31st, 2014**
- Primary and secondary brands changed their return policy from **90 to 45 days** on **Oct 1st , 2013**
- Sister brand did NOT change its return policy which remained at **60 days**

Why Return Policy is Important?

In 2015 , retailers surveyed estimate that total annual returns in US reach $261 billion, or 8% of total retail sales, with $9.1 billion of retailers' annual returns expected to be fraudulent, or 3.5% of the industry's total returns.

Resource:National Retail Federation,Kathy Grannis Allen

Research Questions

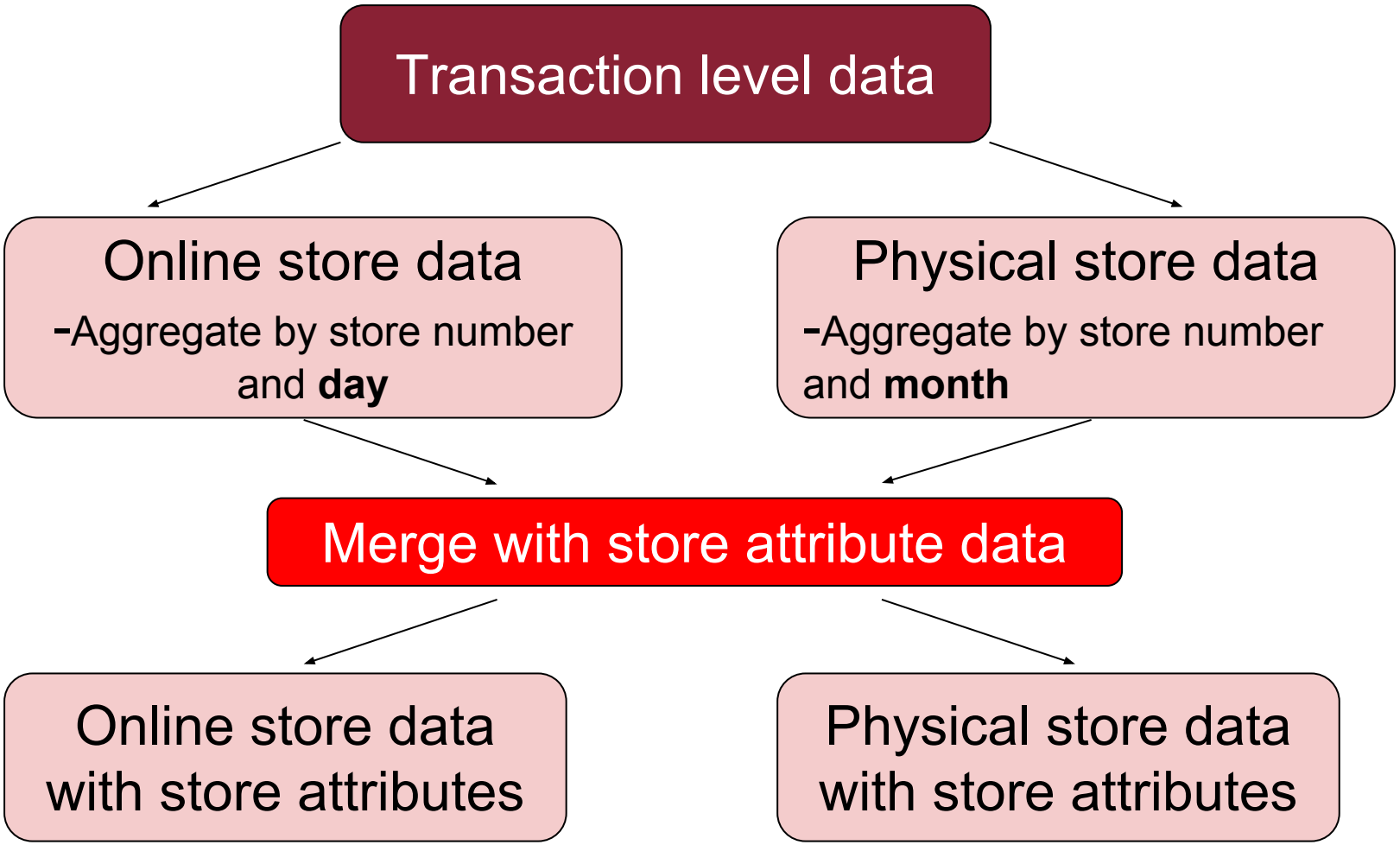
1. What is the impact of the policy change on online channel **sales**?
2. What is the impact of the policy change on physical store **sales**?
3. What is the impact of the policy change on online channel **returns**?
4. What is the impact of the policy change on physical store **returns**?
5. What is the impact of the policy change on **customers purchase behavior**?
6. What is the impact of the policy change on **customers return behavior**?

Key Variables Definition

Policy Change
0:transaction happened before the policy change(2013, Oct.1st)
1:transaction happened after the policy change

Study Group
0:the return policy remain 60 days(online store 10 and all sister brand for physical stores)
1:the return policy change from 90 days to 45 days

Data Handling for Question 1 to 4



Online Stores

Q1. What is the impact of the policy change on sales for online stores?

Dependent Variables

1. **Log(Net purchase amount+1)**

Key Independent Variable

1. **Policy change * Study Group**

Multicollinearity Check

Variables*	store_number_of_skus	study group	policy change
store_number_of_skus	1.00000000	0.4163970	0.09986499e
study group	0.41639704	1.00000000	-7.099896e-05
policy change	0.09986499	7.099896e-05	1.00000000

Variables*	VIF
store_number_of_skus	1.22
study group	1.21
policy change	1.01

*Variables for analyzing the impact of sales of online stores
*VIF of Monthly index is over 30.

OLS Model: Analyzing Impact on Sales in online Stores

log(net_purchase_amount+1)~studygrp*polchange+ store_number_of_skus

Variable	Coefficient	Std. Error	P-value
Intercept	7.084e+00	1.709e-01	< 2e-16
Study group	1.470e+00	2.098e-01	4.63e-12
Policy change	1.359e-01	2.418e-01	0.574
store_number_of_skus	2.683e-04	1.471e-05	< 2e-16
Study Group*Policy Change	**-4.139e-01**	**2.895e-01**	**0.153**

Heteroskedasticity

Test	P-value	Significant?
Goldfeld-Quandt	0.8896	No
Breusch-Pagan	0.5015	No

Conclusion: We find the standard error is not biased

Conclusion of Question 1

What is the impact of the policy change on sales at online stores?

We conclude that the return policy change **does not have significant effect** on sales at online stores.

Online Stores

Q3. What is the impact of the policy change on returns for online stores?

Dependent Variable

- Log(return amount+1)

Key Independent Variable

- Policy Change * Study Group

Multicollinearity Check

Variables*	store_number_of_skus	log net_purchase_amount	study group	policy change
store_number_of_skus	1.00000000	0.5949152	4.163970e-01	9.986499e-02
log.net_purchase_amount	0.5949152	1.00000000	4.321084e-01	2.160670e-02
study group	0.41639704	0.4321084	1.000000e+00	-7.099896e-05
policy change	0.09986499	0.0216067	-7.099896e-05	1.000000e+00

Variables*	VIF
store_number_of_skus	1.647549
log.net_purchase_amount	1.655293
study group	1.293618
policy change	1.013598

*Variables for analyzing the impact of return of online stores

OLS Model: Analyzing Impact on Return in online Stores

log(return_amount+1)~store_number_of_skus+log(net_purchase_amount+1)+studygrp*polchange

Variable	Coefficient	Std. Error	P-value
Intercept	-3.355e+00	1.878e-01	< 2e-16
store_number_of_skus	9.831e-06	1.122e-05	0.381
log.net_purchase_amount	1.097e+00	2.125e-02	< 2e-16
study group	4.337e-02	1.413e-01	0.759
policy change	-9.478e-02	1.588e-01	0.551
Study Group*Policy Change	**-1.055e-01**	**1.903e-01**	**0.580**

Heteroskedasticity

Test	P-value	Significant?
Goldfeld-Quandt	5.784e-06	Yes
Breusch-Pagan	< 2.2e-16	Yes

Conclusion:there is heteroskedasticity, so we recalculate model with robust standard errors

Impact on Online Store Returns

***R output with Robust Standard Errors**

Variable	Coefficient	Std. Error	P-value
Intercept	-3.3551e+00	3.3333e-01	<2e-16
store_number_of_skus	9.8312e-06	1.3868e-05	0.4786
log(net_purchase_amount + 1)	1.0969e+00	4.0598e-02	<2e-16
study group	4.3373e-02	1.4881e-01	0.7708
policy change	-9.4777e-02	1.4814e-01	0.5225
Study Group*Policy Change	**-1.0545e-01**	**1.8395e-01**	**0.5666**

Conclusion of Question 3

What is the impact of the policy change on returns at online stores?

We conclude that the return policy change **does not have significant effect** on returns at online stores.

Physical Stores (Brick & Mortar)

Q2.What is the impact of the policy change on sales for physical stores?

Dependent Variable

- Log(net purchase amount+1)

Key Independent Variable

- Policy change*Study Group

Multicollinearity Check

Variables*	VIF
Study Group	2.143
Policy Change	1.048
Store Number of SKUs	1.170
Ratio of Full Time Sales Associates	1.245
Gender Ratio of Sales Associates	1.215
Average Years of Exp of Sales Associates	1.700
Ratio of Married Sales Associates	1.294
Ratio of Sales Associates with a dependent	1.896
Sales Volume Group	2.266

*Variables for analyzing the impact on sales at physical stores

OLS Model: Analyzing Impact on Sales in Physical Stores

log(net_purchase_amount+1) ~ studygrp*polchange + store_number_of_skus + sa_full_time + sa_gender + sa_avg_years_of_exp + sa_married + sa_dependent + sales_volume_group

Variable	Coefficient	Std. Error	P-value
Intercept	8.490	3.145e-02	< 2e-16
Study Group	1.830	1.749e-02	< 2e-16
Policy Change	-0.07589	1.766e-02	1.76e-05
Store Number of SKUs	2.553e-03	2.559e-05	< 2e-16
Ratio of Full Time Sales Associates	0.07031	2.648e-02	0.007949
Gender Ratio of Sales Associates	-0.1285	2.565e-02	5.54e-07
Avg Years of Exp for Sales Associates	-4.060e-03	2.064e-03	0.049234
Ratio of Married Sales Associates	-0.09927	2.144e-02	3.74e-06
Ratio of Sales Associates with a dependent	0.1767	3.488e-02	4.15e-07
Sales Volume Group	0.2053	4.649e-03	< 2e-16
Study Group*Policy Change	**-0.06885**	**1.979e-02**	**0.000505**

Heteroskedasticity

Test	P-value	Significant?
Goldfeld-Quandt	0.9976	No
Breusch-Pagan	< 2.2e-16	Yes

Conclusion: Recalculate model with robust standard errors

Impact on Physical Store Sales

***R output with Robust Standard Errors**

Variable	Coefficient	Std. Error	P-value
Intercept	8.496	0.0339	< 2.2e-16
Study Group	1.830	0.0176	< 2.2e-16
Policy Change	-0.07589	0.0191	7.334e-05
Store Number of SKUs	2.553e-03	0.000045	< 2.2e-16
Ratio of Full Time Sales Associates	0.07031	0.02889	0.01498
Gender Ratio of Sales Associates	-0.1285	0.0263	1.070e-06
Avg Years of Exp for Sales Associates	-4.060e-03	0.00212	0.05631
Ratio of Married Sales Associates	-0.09927	0.0223	9.398e-06
Ratio of Sales Associates with a dependent	0.1767	0.0361	9.955e-07
Sales Volume Group	0.2053	0.00438	< 2.2e-16
Study Group*Policy Change	**-0.06885**	**0.0209**	**0.000993**

Conclusion of Question 2

What is the impact of the policy change on sales at physical stores?

We conclude that changing the return policy from 90 to 45 days **reduced sales by 6.89%** for physical stores.

Physical Stores (Brick & Mortar)

Q4. What is the impact of the policy change on returns for physical stores?

Dependent Variable

- Log(return amount+1)

Key Independent Variable

- Policy Change* Study Group

Multicollinearity Check

Variables*	VIF
Net Purchase Amount	1.830
Study Group	1.920
Policy Change	1.033
Average Years of Experience of Sales Associates	1.719
Ratio of Sales Associates with a dependent	1.786
Sales Associate Average Rate of Pay	1.103
Sales Volume Group	2.822
Gender Ratio of Sales Associates	1.197

*Variables for analyzing the impact on returns at physical stores

OLS Model: Analyzing Impact on Returns in Physical Stores

log(return_amount+1) ~ log(net_purchase_amount+1) + studygrp*polchange + sa_avg_years_of_exp + sa_dependent + sa_avg_rate_of_pay + sales_volume_group + sa_gender

Variable	Coefficient	Std. Error	P-value
Intercept	-4.690	0.1421	< 2e-16
log(Net Purchase Amount + 1)	1.202	0.0141	< 2e-16
Study Group	0.1059	0.0388	0.00636
Policy Change	0.1922	0.0322	2.56e-09
Average Years of Experience for Sales Associates	-0.0047	0.0038	0.2166
Ratio of Sales Associates with a dependent	-0.0115	0.0624	0.8527
Average Rate of Pay for Sales Associates	0.0362	0.0055	3.61e-11
Sales Volume Group	-0.0301	0.0095	0.00158
Gender Ratio of Sales Associates	0.1243	0.0467	0.00783
Study Group*Policy Change	**-.4251**	**0.0361**	**<2e-16**

Heteroskedasticity

Test	P-value	Significant?
Goldfeld-Quandt	1	No
Breusch-Pagan	< 2.2e-16	Yes

Conclusion: Recalculate model with robust standard errors

Impact on Physical Store Returns

***R output with Robust Standard Errors**

Variable	Coefficient	Std. Error	P-value
Intercept	-4.690	0.1779	< 2.2e-16
log(Net Purchase Amount + 1)	1.202	0.0183	< 2.2e-16
Study Group	0.1059	0.0438	0.0156
Policy Change	0.1922	0.0390	8.656e-07
Average Years of Experience for Sales Associates	-0.0047	0.0036	0.1839
Ratio of Sales Associates with a dependent	-0.0115	0.0589	0.8441
Average Rate of Pay for Sales Associates	0.0362	0.0057	2.471e-10
Sales Volume Group	-0.0301	0.0089	0.000819
Gender Ratio of Sales Associates	0.1243	0.0507	0.01428
Study Group*Policy Change	**-.4251**	**0.0419**	**< 2.2e-16**

Conclusion of Question 4

What is the impact of the policy change on returns at physical stores?

We conclude that changing the return policy from 90 to 45 days **reduced dollar amount of returns by 42.5%** for physical stores.

Customer Purchase Behavior

Q5. What is the impact of the policy change on customer purchase behavior?

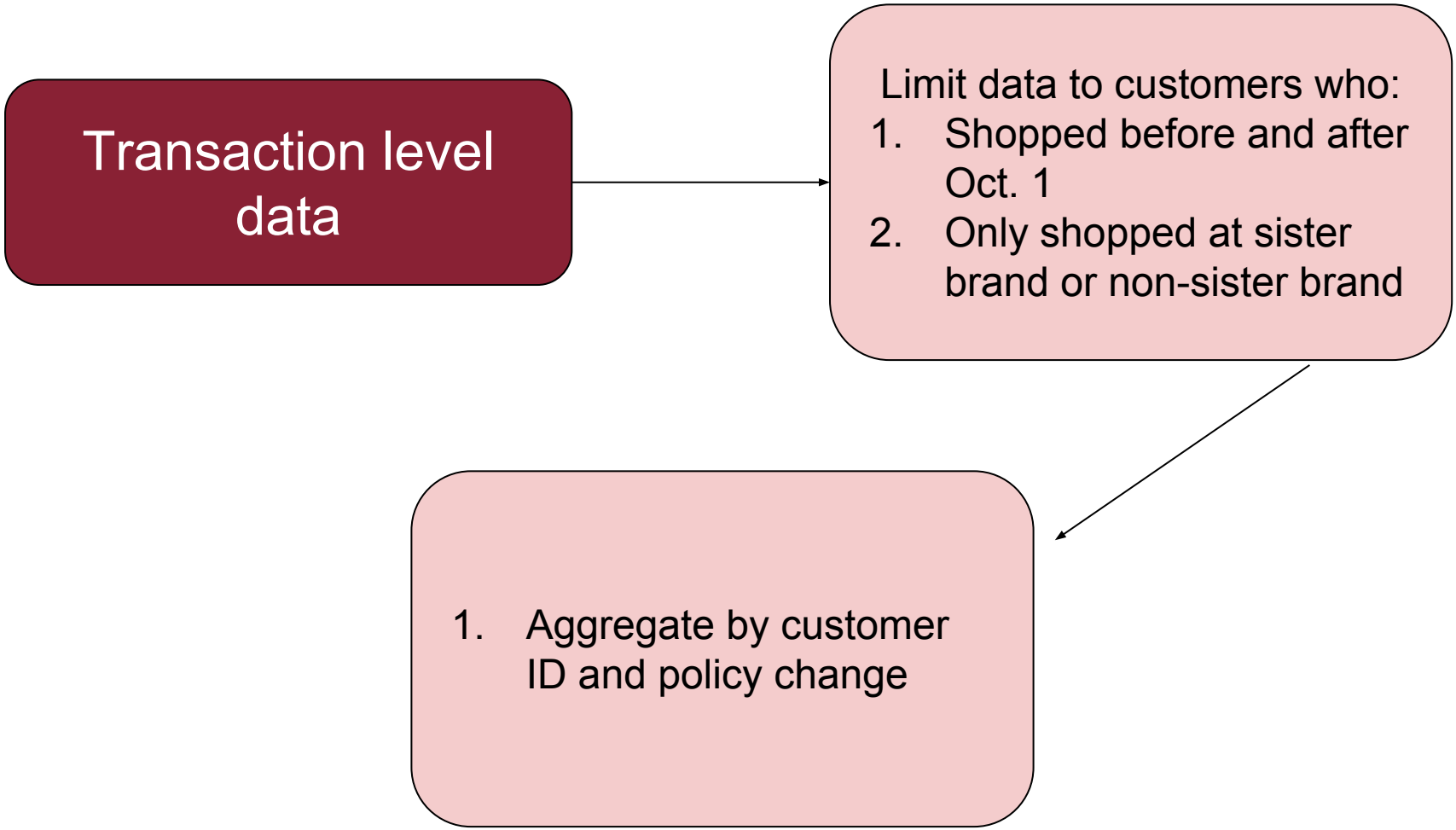
Dependent Variable

- Log(net purchase amount+1)

Key Independent Variable

- Policy change * Study Group

Data Handling for Question 5



Multicollinearity Check

Variables*	VIF
policy change	1.004419
study group	1.020659
est_income_code	1.044810
age_band	1.098334
length_of_residence	1.045146

OLS Model: Analyzing Impact on customer purchase behavior

log(1 + net_purchase_amount) ~ polchange * studygrp + factor(est_income_code) + factor(age_band) + factor(gender) + length_of_residence + factor(child)

Variable	Coefficients	Std. Error	P-Values
(Intercept)	3.8903739	0.040910021	0
custpurch$polchange	-0.034032	0.051811781	0.511285895
custpurch$studygrp	1.7217445	0.037267538	0
factor(custpurch$est_income_code)2	-0.0149796	0.021043877	0.47657464
factor(custpurch$est_income_code)3	0.0362372	0.016851539	0.031527352
factor(custpurch$est_income_code)4	0.0439496	0.01693625	0.009460163
factor(custpurch$est_income_code)5	0.0316464	0.017045509	0.063372777
factor(custpurch$est_income_code)6	0.0795119	0.014786869	7.58E-08
factor(custpurch$est_income_code)7	0.0987891	0.016566631	2.48E-09
factor(custpurch$est_income_code)8	0.1133201	0.019495445	6.16E-09
factor(custpurch$est_income_code)9	0.1437647	0.016923619	2.00E-17
factor(custpurch$age_band)1	0.4324707	0.03190455	7.93E-42
factor(custpurch$age_band)2	-0.3300869	0.179915527	0.066555995
factor(custpurch$age_band)3	-0.1663194	0.023530465	1.58E-12
factor(custpurch$age_band)4	-0.0463505	0.016555738	0.005116412
factor(custpurch$age_band)5	-0.0484128	0.015153209	0.001399225
factor(custpurch$age_band)6	-0.0634559	0.0160001	7.31E-05
factor(custpurch$age_band)7	-0.0560428	0.016957708	0.000950522
factor(custpurch$age_band)8	-0.0062232	0.016901545	0.712720319
factor(custpurch$age_band)9	0.0236926	0.01683442	0.159313751
factor(custpurch$age_band)10	0.0315928	0.017643371	0.073354742
factor(custpurch$age_band)11	0.0179519	0.020829697	0.388776068
factor(custpurch$age_band)12	-0.0775227	0.026303726	0.003207142
factor(custpurch$age_band)13	-0.0760427	0.026548703	0.004180407
factor(custpurch$gender)F	0.1135136	0.016727953	1.16E-11
factor(custpurch$gender)M	0.3438485	0.016309482	1.73E-98
factor(custpurch$gender)U	0.3259063	0.019722497	2.85E-61
custpurch$length_of_residence	-0.005724	0.000712675	9.69E-16
factor(custpurch$child)Y	-0.0752425	0.007917542	2.07E-21
custpurch$polchange:custpurch$studygrp	-0.1717836	0.052381122	0.00104023

Heteroskedasticity

Test	P-value	Significant?
Goldfeld-Quandt	0.7931	No
Breusch-Pagan	2.2e-16	Yes

Conclusion: Recalculate model with robust standard errors

Impact on purchase behaviour

*R output with Robust Standard Errors

	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept)	3.89037388	0.03222957	120.7082	< 2.2e-16
custpurch$polchange	-0.03403198	0.03928039	-0.8664	0.386280
custpurch$studygrp	1.72174449	0.02725965	63.1609	< 2.2e-16
factor(custpurch$est_income_code)2	-0.01497955	0.02106341	-0.7112	0.476983
factor(custpurch$est_income_code)3	0.03623717	0.01687371	2.1476	0.031751
factor(custpurch$est_income_code)4	0.04394962	0.01676648	2.6213	0.008761
factor(custpurch$est_income_code)5	0.03164642	0.01694623	1.8675	0.061840
factor(custpurch$est_income_code)6	0.07951192	0.01470107	5.4086	6.365e-08
factor(custpurch$est_income_code)7	0.09878912	0.01651615	5.9814	2.219e-09
factor(custpurch$est_income_code)8	0.11332007	0.01952488	5.8039	6.496e-09
factor(custpurch$est_income_code)9	0.14376474	0.01697238	8.4705	< 2.2e-16
factor(custpurch$age_band)1	0.43247071	0.03292188	13.1363	< 2.2e-16
factor(custpurch$age_band)2	-0.33008693	0.16445816	-2.0071	0.044739
factor(custpurch$age_band)3	-0.16631941	0.02247956	-7.3987	1.384e-13
factor(custpurch$age_band)4	-0.04635049	0.01652093	-2.8056	0.005024
factor(custpurch$age_band)5	-0.04841277	0.01512431	-3.2010	0.001370
factor(custpurch$age_band)6	-0.06345587	0.01585439	-4.0024	6.274e-05
factor(custpurch$age_band)7	-0.05604283	0.01705703	-3.2856	0.001018
factor(custpurch$age_band)8	-0.00622325	0.01706744	-0.3646	0.715390
factor(custpurch$age_band)9	0.02369263	0.01649924	1.4360	0.151010
factor(custpurch$age_band)10	0.03159280	0.01754647	1.8005	0.071781
factor(custpurch$age_band)11	0.01795191	0.02055983	0.8732	0.382580
factor(custpurch$age_band)12	-0.07752268	0.02677452	-2.8954	0.003788
factor(custpurch$age_band)13	-0.07604272	0.02582104	-2.9450	0.003230
factor(custpurch$gender)F	0.11351361	0.01644718	6.9017	5.163e-12
factor(custpurch$gender)M	0.34384852	0.01595490	21.5513	< 2.2e-16
factor(custpurch$gender)U	0.32590631	0.01970938	16.5356	< 2.2e-16
custpurch$length_of_residence	-0.00572405	0.00071327	-8.0251	1.024e-15
factor(custpurch$child)Y	-0.07524254	0.00790613	-9.5170	< 2.2e-16
custpurch$polchange:custpurch$studygrp	-0.17178357	0.04003533	-4.2908	1.782e-05

Conclusion of Question 5

What is the impact of the policy change on customer purchase behaviour?

We conclude that changing the return policy from 90 to 45 days **reduced** the customer net purchase amount by **17%**

Customer Return Behavior

Q6. What is the impact of the policy change on customer return behavior?

Dependent Variable

- Return

Key Independent Variable

- Policy Change*Study Group

Data Handling for Question 6

Transaction level data

Multicollinearity Check

Variables	VIF
Net Purchase Amount	1.012
Study Group	1.010
Policy Change	1.001
Estimated Income Code	1.021
Age Band	1.018

Probit Model

Variables	Coefficient	Std. Error	P-value
Intercept	-1.309	9.182e-03	< 2e-16
Policy Change	9.248e-02	1.140e-02	4.85e-16
Study Group	7.517e-02	8.920e-03	< 2e-16
Net Purchase Amount	1.302e-04	1.366e-06	< 2e-16
Estimated Income Code	2.567e-03	5.149e-04	6.21e-07
Age Band	-1.143e-02	2.879e-04	< 2e-16
Study Group*Policy Change	**-1.854e-01**	**1.166e-02**	**< 2e-16**

Goodness of fit p-value: 0
Accuracy: 90.83%

Heteroskedasticity

Test	P-value	Significant?
Goldfeld-Quandt	< 2.2e-16	Yes
Breusch-Pagan	< 2.2e-16	Yes

Conclusion: recalculate model with robust standard errors

Impact on return behaviour

***Marginal Effects with Robust Standard Errors**

Variable	dF/dx	Std. Error	P-value
Policy Change	1.6146e-02	1.9735e-03	2.799e-16
Study Group	1.2648e-02	1.4350e-03	< 2.2e-16
Net Purchase Amount	2.2897e-05	3.5515e-07	< 2.2e-16
Estimated Income Code	4.5130e-04	9.0740e-05	6.574e-07
Age Band	-2.0093e-03	5.0832e-05	< 2.2e-16
Study Group * Policy Change	**-3.2900e-02**	**2.0879e-03**	**< 2.2e-16**

Conclusion of Question 6

What is the impact of the policy change on customer return behaviour?

We conclude that the policy change **decreases** the probability of return by **3.29 percentage points**

Conclusion of the Study

Physical Stores

- Sales amount decreases by 7% and Return amount decreases by 42%

Online Stores

- Does not have influence on return and sales

Customer Purchase Behavior

- Net purchase amount decreased by 17%

Customer Return Behavior

- Probability of return decreased by 3.29 percentage points

Thank You!
All the Best for finals!