

# **DBA5101 Analytics in Managerial Economics**

## **Homework 2:**

# **Third-degree Price Discrimination by DiD Analysis**

Meritxell Camp Garcia (A0280366B)

October 27, 2023

1. Executive Summary and Recommendations	1
2. Overview and Treatment Group Identification	1
3. Price Elasticity of Demand	2
4. Pricing Strategy Recommendations	5
5. Limitations and Next Steps	7
6. Annex	8
6.1. Figures	8

### 1. Executive Summary and Recommendations

This analysis explores opportunities to increase revenue through an optimized third-degree price discrimination strategy, for a software firm serving education and business clients. We discover that the education sector is more price-sensitive than the business sector. By optimizing the firm's pricing strategy based on these sensitivities, we recommend a price of \$87 per license for education clients, down from \$100, and \$379.5 for business clients, up from \$250. Adopting this third-degree price discrimination could elevate annual revenue by 11% compared to the 2021 pricing setup and by 19% when contrasted with a unitary pricing model that doesn't differentiate between sectors. We also discovered that were the firm not to price discriminate, the whole education segment would remain unserved, showing how price discrimination can bring economic welfare.

However, it's essential to approach a 51.8% price hike for business clients with caution; real-world implications and client perceptions must be managed. Before executing such a considerable price increase, we recommend pursuing an additional RTC test for business clients to ensure the success and sustainability of the new pricing framework.

### 2. Overview and Treatment Group Identification

The firm we are consulting sells its software product to two primary client segments: educational institutions and businesses. Of their 250 distinct clients, a significant 66% are from the business sector.

Currently, the company employs third-degree price discrimination, charging educational clients \$100 and business clients \$250 per license. Given the nature of software products, we infer that the marginal costs of serving each client are negligible and consistent across both segments. Consequently, business clients present a higher profitability than their educational counterparts.

The firm executed an RTC test, implementing a 30% price increase for 47% of its educational clients and for 56% of its business clients from quarter 5 onwards. This means that from quarter 5 to quarter 8, 40 education clients paid \$130 per license, and 93 business clients paid \$325 per license. Treatment and control groups can be distinguished since the treatment group had a 30% price increase from quarter 5 onwards.

Our goal is to evaluate the impact of this test on quantity and optimize the company's pricing approach to maximize revenue.

### 3. Price Elasticity of Demand

We employed the Difference-in-Differences (DiD) approach to estimate the causal effect of the pricing intervention on the quantity sold. The essence of the DiD approach is to compare the changes in

outcomes over time between a population that is subject to an intervention (the treated group) and a population that is not (the control group).

Please note that we assume that in the absence of treatment, the treated and control groups would have followed the same trend over time and that there is no other shock that might have influenced the behavior of one of the two groups.

In our analysis, for each sector (education and business):

- Treatment Group: This group comprises clients who experienced the 30% price increase.
- **Control Group:** This group consists of clients who did not experience the price change.
- **Pre-Treatment Period:** This refers to the period before the price change (quarters 1 to 4).
- **Post-Treatment Period:** This is the period after the price change was implemented (quarters 5 to 8).

We utilize a Difference-in-Differences (DiD) estimation framework incorporating fixed effects and control variables, executed using PanelOLS, for each of the sectors. To ensure the robustness of our findings, multiple models were assessed. Central to our analysis is the interaction term *treatment&time*, an interaction term dummy variable assigned a value of 1 for clients in the treatment group during the post-treatment period. The two other dummy variables include *treatment* (which equals 1 if a client is in the treatment group) and *time\_dummy* (which equals 1 for quarters subsequent to quarter 4). We also incorporate *log\_revenue* as a control variable. This represents the natural logarithm transformation of a client's revenue and is introduced due to its substantial correlation with the logarithmically transformed license quantity (*log\_quantity*) bought by clients (0.88 and 0.95 for education and business respectively, See Figures 1 and 2). Notably, client revenue demonstrated an upward trajectory throughout the entire experimental span (from quarter 1 to quarter 8). Consequently, it's important to control for this revenue uptick to ensure we don't mistakenly attribute revenue growth effects on quantity to the pricing adjustment effects. A comprehensive breakdown of these models is detailed in the subsequent table (see Figures 3 to 11 for the models' summaries):

Education Sector Models (dependant variable: log\_quantity)

Significance levels are indicated by \*\*\* (p < 0.01), \*\* (p < 0.05) and \* (p < 0.1)

	Model Employed	Robustness Test 1	Robustness Test 2	Robustness Test 3
"treatment&time" coef	-0.405423***	-0.405423***	-0.405130***	-0.402683***
Constant	YES	YES	YES	YES
Dummies	NO	YES	YES	NO
Control (log_revenue)	YES	YES	YES	NO
Fixed Effects	YES	YES	NO	YES

R-squared (incl. Fixed effects)	0.87258	0.87258	0.85334	0.86085
Comments		Dummy variables were absorbed (collinear) in the regression and were dropped.		

#### Business Sector Models (dependant variable: log\_quantity)

Significance levels are indicated by \*\*\* (p < 0.01), \*\* (p < 0.05) and \* (p < 0.1)

	Model Employed	Robustness Test 1	Robustness Test 2	Robustness Test 3
"treatment&time" coef	-0.147473***	-0.147473***	-0.149929***	-0.162812***
Constant	YES	YES	YES	YES
Dummies	NO	YES	YES	NO
Control (log_revenue)	YES	YES	YES	NO
Fixed Effects	YES	YES	NO	YES
R-squared (incl. Fixed effects)	0.94404	0.94404	0.91830	0.94404
Comments		Dummy variables were absorbed (collinear) in the regression and were dropped.		

The coefficient of the interaction term *treatment&time* captures the differential change in *log\_quantity* for the treatment group, relative to the control group, from the pre-treatment to the post-treatment period. By regressing *log\_quantity* on this interaction term using Panel OLS, we obtained the coefficient of the interaction term, which provides the average treatment effect of the price increase. This coefficient essentially tells us the percentage change in quantity as a result of the price increase, holding all else constant.

We opted for the initial model that only retained the interaction term and control variable (*log\_revenue*) and excluded the *treatment* and *time\_dummy* variables. The rationale for this decision is because of the phenomenon of perfect multicollinearity. Specifically, these dummy variables were fully absorbed in Robustness Test 1, indicating they were redundant given the fixed effects of the model. By excluding them, we mitigate potential multicollinearity issues.

When excluding the control variable *log\_revenue* from the model in Robustness Test 3, there is a substantial shift in the coefficient of the interaction term, especially for the business sector. We opted to retain *log\_revenue* in the model due to this observable change, which suggests that it provides crucial information. Neglecting to include it could lead to omitted variable bias, potentially distorting the true effect of the interaction term.

In refining our model, we also employed the *cov\_type="clustered"* specification paired with *cluster\_entity=True* for our standard error calculations. This is because observations within the same entity (i.e., the same *client\_id*) over time might not be independent of each other. By clustering at the entity level, we account for potential correlations in the error terms within these entities. This approach corrects for any potential underestimation of standard errors due to intra-cluster correlation.

With the coefficient of the interaction term in hand, and knowing the magnitude of the price change, we can then compute the price elasticity of demand (PED) for both sectors, which measures the percentage change in quantity demanded for a one-percent change in price. The formula for PED is as follows:

```
PED = \% Change in Quantity \div % Change in Price PED education = -40.5\% \div 30.0\% = -1.3514 PED business = -14.7\% \div 30.0\% = -0.4916
```

Our analysis first reveals that the education sector exhibits greater price sensitivity than the business sector. The absolute magnitude of their respective price elasticities, 1.35 for the education sector and 0.49 for the business sector, suggests that the education sector responds more dramatically to price fluctuations. Conversely, the business sector has a softer reaction to similar changes, at these price points.

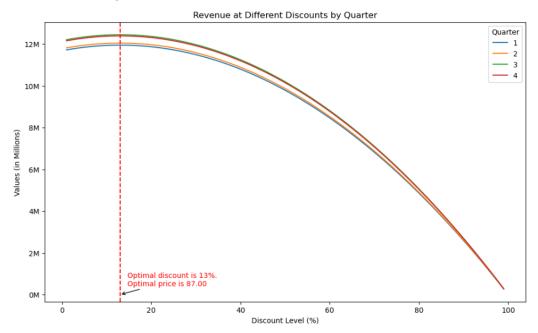
### 4. Pricing Strategy Recommendations

The optimal pricing strategy is setting a price of \$87 per license for education clients, and a price of \$379.5 per license for business clients. This strategy would have increased revenue by 11% vs the 2021 price setting annually and by 19% vs the optimal unitary price strategy (without price discrimination).

To determine the optimal price for each sector, we employ the concept of price elasticity of demand (PED) to project how the quantity demanded would have shifted during the pre-treatment period (2021) in response to hypothetical price changes. This projection is feasible because, by definition, PED stipulates that a 1% modification in price will result in a corresponding change in quantity equal to the PED%.

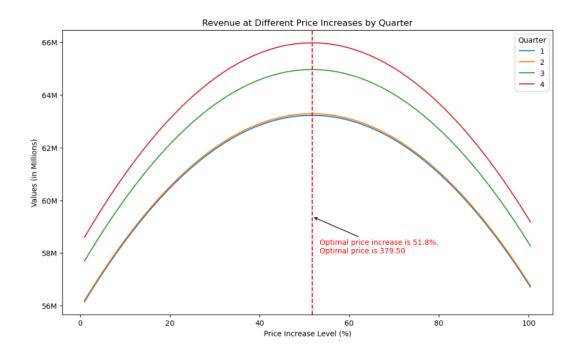
The price elasticity for education clients (-1.3514) suggests that the firm might gain from price reductions for education clients. A lowered price is likely to stimulate a higher demand quantity. To ascertain the most advantageous price point, we modeled the potential revenue of quarter 1 to quarter 4 of 2021 (pre-treatment) from the education sector for a spectrum of price discounts, ranging from 1% to 99% (the price should not be \$0), starting from the baseline price of \$100. This modeling produced a concave function, peaking at a 13% price reduction. This translates to an optimal price of \$87 for all quarters of 2021. This analysis suggests that, at this price point, the firm would have earned \$48,849,625 from the education sector, an increase of \$1,091,025 (+2%) compared to the actual revenue in 2021.

#### Simulation Results for Education:



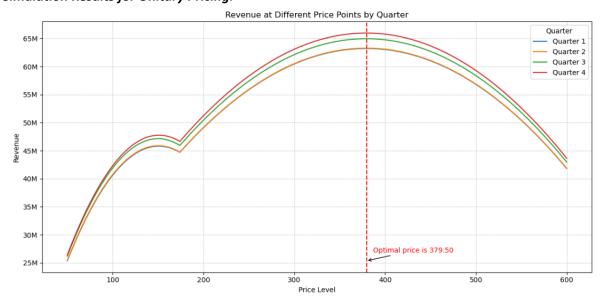
The price elasticity of business clients (-0.4916) suggests that the firm can increase revenue with a price increase. A higher price is likely to lead to a small quantity reduction, which could be compensated by a higher price point. To determine the optimal pricing strategy, we simulated potential revenue outcomes for the business sector across various price increments ranging from 1% to 100%, starting from the baseline price of \$250. This modeling produced a concave function, peaking at a 51.8% price increase. This means that the optimal price was \$379.5 for all quarters of 2021. This analysis suggests that, at this price point, the firm would have earned \$257,447,771 from the business sector, an increase of \$29.912.271 (+13%) compared to the actual revenue in 2021.

#### Simulation Results for Business:



By looking at the price sensitivity for each sector, we can find the optimal price if the firm were to employ a unitary pricing strategy. Just like before, we assessed how quantity and revenue in 2021 would have changed for unitary prices between \$50 and \$600. The demand from education clients gets to zero when the price exceeds \$174. Our findings show the best approach is to sell only to business clients for a price of \$379.5 for all quarters of 2021. This would have brought in \$257,447,771 — the same amount as if we only sold to business customers at the optimized price. However, this is 6% less than what the firm actually earned in 2021 without fine-tuning its pricing discrimination strategy.

#### Simulation Results for Unitary Pricing:





Our proposed third-degree price discrimination could boost revenue by 19% compared to a unitary price strategy. This additional revenue would be derived solely from sales to education clients. (See Figures 12 to 14) This situation illustrates the economic benefits of price discrimination. Without it, a significant market segment - the education market - would be left unserved.

### 5. Limitations and Next Steps

We must acknowledge the constraints of our analysis. Our study operates under the assumption that price elasticity remains consistent across the demand curve for both sectors. Importantly, the proposed price increase of 51.8% for the business sector extends beyond our previously tested range of a 30% hike. Additionally, we assume that no other factor than the price increase motivated the behavior of the firm's clients ( parallel trends assumption).

Furthermore, practical business considerations come into play. Firstly, this analysis does not take into consideration serving costs, which might be higher than the proposed \$87 per license for the education clients. Secondly, while educational clients will welcome a 13% discount, a 51.8% price surge may not sit well with business clients. Such a dramatic increase could pose challenges for account managers when justifying the new rates. Before rolling out this adjustment, it would be prudent to conduct another RTC test with business clients to assess their response and confirm the viability of the proposed plan.

### 6. Annex

## 6.1. Figures

Figure 1: Education population correlation matrix

	revenue	price	quantity	log_revenue	log_price	log_quantity
revenue	1.000000	-0.001425	0.959525	0.926099	-0.001425	0.827246
price	-0.001425	1.000000	-0.218544	-0.028252	1.000000	-0.290857
quantity	0.959525	-0.218544	1.000000	0.895326	-0.218544	0.892192
log_revenue	0.926099	-0.028252	0.895326	1.000000	-0.028252	0.884531
log_price	-0.001425	1.000000	-0.218544	-0.028252	1.000000	-0.290857
log_quantity	0.827246	-0.290857	0.892192	0.884531	-0.290857	1.000000

Figure 2: Business population correlation matrix

	revenue	price	quantity	log_revenue	log_price	log_quantity
revenue	1.000000	0.064731	0.977470	0.912603	0.064731	0.918431
price	0.064731	1.000000	-0.044573	0.029910	1.000000	-0.083651
quantity	0.977470	-0.044573	1.000000	0.896634	-0.044573	0.950724
log_revenue	0.912603	0.029910	0.896634	1.000000	0.029910	0.951667
log_price	0.064731	1.000000	-0.044573	0.029910	1.000000	-0.083651
log_quantity	0.918431	-0.083651	0.950724	0.951667	-0.083651	1.000000

Figure 3: Education - Model Employed test results

		std er				
treatment&time						
log_revenue						
		PanelOLS Est		-		
Dep. Variable:						0.2168
Estimator:		Pane10LS				0.8267
No. Observation				(Within):		0.8419
Date:		Oct 27 2023				0.8417
Time:		16:23:03	_	ihood		23.122
Cov. Estimator:		Clustered				
			F-statist:	ic:		81.083
Entities:		8	P-value			0.0000
Avg Obs:		85.000	Distribut:	ion:		F(2,586)
Min Obs:		85.000				
Max Obs:		85.000	F-statist:	ic (robust):		56.241
			P-value			0.0000
Time periods:		85	Distribut:	ion:		F(2,586)
Avg Obs:		8.0000				
Min Obs:		8.0000				
Max Obs:		8.0000				
		Paramete	er Estimate:	5		
		Std. Err.				
const						
treatment&time	-0.4054	0.0409	-9.9198	0.0000	-0.4857	-0.3252
log_revenue	0.8505	0.0817	10,415	0.0000	0.6901	1.0109

F-test for Poolability: 0.9828 P-value: 0.5277 Distribution: F(91,586)

Included effects: Entity, Time

### Figure 4: Education - Robustness Test 1 results

R2: 0.872583672	coeff   -0.405423	std err   0.04087	70   0.	000000		
		(collinear)			has been d	Ironned
		(collinear)				
		PanelOLS Esti				
Dep. Variable:	1	og_quantity	R-squared	i:		0.2168
Estimator:		Pane10LS	R-squared	(Between):		0.8267
No. Observation	s:	680	R-squared	(Within):		0.8419
Date:	Fri,	Oct 27 2023	R-squared	(Overall):		0.8417
Time:		10:55:40	Log-likel	ihood		23.122
Cov. Estimator:		Clustered				
			F-statist	ic:		81.083
Entities:		8	P-value			0.0000
Avg Obs:		85.000	Distribut	ion:		F(2,586)
Min Obs:		85.000				
Max Obs:		85.000	F-statist	ic (robust)	:	56.241
			P-value			0.0000
Time periods:		85	Distribut	ion:		F(2,586)
Avg Obs:		8.0000				
Min Obs:		8.0000				
Max Obs:		8.0000				
		Paramete	er Estimate	s		
		Std. Err.				
		0.4865				
treatment&time	-0.4054	0.0409	-9.9198	0.0000	-0.4857	-0.3252
log_revenue	0.8505	0.0817	10.415	0.0000	0.6901	1.0109

F-test for Poolability: 0.9828 P-value: 0.5277 Distribution: F(91,586)

Figure 5: Education - Robustness Test 2 results

R2: 0.853341314	3742858					
	coeff	std en	ror   p-	-value		
treatment&time	-0.405130	0.0381	27   0.	.000000		
log_revenue						
treatment	0.008275	0.00974	48   0.	396240		
time_dummy	-0.017636	0.0099	27   0.	.076101		
	I	PanelOLS Est:	imation Sur	nmary		
Dep. Variable:	10	og_quantity	R-squared	i:		0.8533
Estimator:		Pane10LS	R-squared	d (Between):		0.8573
No. Observation	s:	680	R-squared	d (Within):		0.8533
Date:	Fri, (	Oct 27 2023	R-squared	d (Overall):		0.8533
Time:		10:55:40	Log-like	lihood		-24.699
Cov. Estimator:		Clustered				
			F-statist	tic:		981.88
Entities:		8	P-value			0.0000
Avg Obs:		85.000	Distribut	tion:		F(4,675)
Min Obs:		85.000				
Max Obs:		85.000	F-statist	tic (robust)	:	147.55
			P-value			0.0000
Time periods:		85	Distribut	tion:		F(4,675)
Avg Obs:		8.0000				
Min Obs:		8.0000				
Max Obs:		8.0000				
			er Estimate			
	Parameter	Std. Err.	T-stat	P-value	Lower CI	Upper CI
const treatment&time		0.3175				
log_revenue						
treatment						
time_dummy						

Figure 6: Education - Robustness Test 3 results

	coeff	std er	ror   p-	value		
treatment&time	-0.402683	0.0430	55   0.	000000		
		PanelOLS Est	imation Sum	mary		
Dep. Variable:						0.1446
Estimator:		Pane10LS	R-squared	(Between):		0.7057
No. Observation	5:	680	R-squared	(Within):		0.0746
Date:	Fri,	Oct 27 2023	R-squared	(Overall):		0.0837
Time:		10:55:40	Log-likel	ihood		-6.8339
Cov. Estimator:		Clustered				
			F-statist	ic:		99.239
Entities:		8	P-value			0.0000
Avg Obs:		85.000	Distribut	ion:		F(1,587)
Min Obs:		85.000				
Max Obs:		85.000	F-statist	ic (robust):		87.435
			P-value			0.0000
Time periods:		85	Distribut	ion:		F(1,587)
Avg Obs:		8.0000				
Min Obs:		8.0000				
Max Obs:		8.0000				
		Paramet	er Estimate	S		
		Std. Err.				
const		0.0101				
treatment&time	-0.4027	0.0431	-9.3507	0.0000	-0.4873	-0.3181

F-test for Poolability: 35.984 P-value: 0.0000 Distribution: F(91,587)

Figure 7: Business - Model Employed test results

		std er		value		
treatment&time	-0.147473	0.0121	15   0.	000000		
log_revenue	0.782106	0.0489	37   0.	000000		
		PanelOLS Est:	imation Sum	nmary		
Dep. Variable:						0.2398
Estimator:		Pane10LS				0.7041
No. Observations				(Within):		0.8853
Date:	Fri,	Oct 27 2023				0.8851
Time:		16:23:24	Log-likel	lihood		787.37
Cov. Estimator:		Clustered				
			F-statist	ic:		180.72
Entities:		8	P-value			0.0000
Avg Obs:		165.00	Distribut	ion:		F(2,1146)
Min Obs:		165.00				
Max Obs:		165.00	F-statist	ic (robust)	:	332.18
			P-value			0.0000
Time periods:		165	Distribut	ion:		F(2,1146)
Avg Obs:		8.0000				
Min Obs:		8.0000				
Max Obs:		8.0000				
		Paramete	er Estimate	:s		
		Std. Err.	T-stat	P-value	Lower CI	Upper CI
		0.2961				
treatment&time						
log_revenue						

P-value: 0.0000 Distribution: F(171,1146)

Included effects: Entity, Time

### Figure 8: Business - Robustness Test 1 results

		std er				
treatment&time	-0.147473	0.0121	15   0.	000000		
log_revenue	0.782106	0.0489	37   0.	000000		
treatment	is absorbed	(collinear)	in the reg	ression and	has been	dropped.
time_dummy		(collinear)			has been	dropped.
		PanelOLS Est				
Dep. Variable:						0.2398
Estimator:		PanelOLS				
No. Observation				(Within):		
Date:		Oct 27 2023				0.8851
Time:		10:55:40	Log-likel	ihood		787.37
Cov. Estimator:		Clustered				
			F-statist			180.72
Entities:			P-value			0.0000
Avg Obs:			Distribut	ion:		F(2,1146)
Min Obs:		165.00				
Max Obs:				ic (robust)	:	
			P-value			0.0000
Time periods:			Distribut	ion:		F(2,1146)
Avg Obs:		8.0000				
Min Obs:		8.0000				
Max Obs:		8.0000				
		Paramet	er Estimate	s		
		Std. Err.				
const						
treatment&time	-0.1475	0.0121	-12.173	0.0000	-0.1712	-0.1237
log revenue						

F-test for Poolability: 3.0885 P-value: 0.0000 Distribution: F(171,1146)

Figure 9: Business - Robustness Test 2 results

R2: 0.918296622	1359756					
	coeff	std en	ror   p-	value		
treatment&time	-0.149929   0.01158		37   0.	000000		
log_revenue	0.656857	0.0173	38   0.	000000		
treatment	0.010581	0.0089	9 0.235683			
time_dummy	0.004284	0.01030	01   0.	1 0.677569		
	Pane:	10LS Est:	imation Sum	mary		
Dep. Variable:	log_q	uantity	R-squared	:		0.9183
Estimator:		ane10LS	R-squared	(Between):		0.8415
No. Observations	5:	1320	R-squared	(Within):		0.9184
Date:	Fri, Oct	27 2023	R-squared	(Overall):		0.9183
Time:	10	0:55:40	Log-likel	ihood		537.57
Cov. Estimator:	Cl	ustered				
			F-statist	ic:		3695.0
Entities:	8		P-value			0.0000
Avg Obs:	165.00		Distribution:			F(4,1315)
Min Obs:		165.00				
Max Obs:		165.00	F-statist	ic (robust):		766.36
			P-value			0.0000
Time periods:		165	Distribut	ion:		F(4,1315)
Avg Obs:		8.0000				
Min Obs:		8.0000				
Max Obs:		8.0000				
			er Estimate	-		
	Parameter Std					
const	3.1763	0.1041	30.503	0.0000	2.9720	3.3806
treatment&time	-0.1499	0.0116	-12.939	0.0000	-0.1727	-0.1272
log_revenue	0.6569	0.0173	37.886	0.0000	0.6228	0.6909
treatment	0.0106	0.0089	1.1864	0.2357	-0.0069	0.0281
time dummy	0.0043	0103	0 4159	0.6776	-0.0159	0.0245

Figure 10: Business - Robustness Test 3 results

	coeff	std er	ror I p-	value		
treatment&time						
		PanelOLS Est	imation Sum	marv		
Dep. Variable:	1	og_quantity	R-squared	i:		0.0698
Estimator:		Pane10LS	R-squared	(Between):		-3.4241
No. Observations	:	1320	R-squared	(Within):		0.0085
Date:	Fri,	Oct 27 2023	R-squared	(Overall):		0.0048
Time:		10:55:40	Log-likel	ihood		654.17
Cov. Estimator:		Clustered				
			F-statist	ic:		86.031
Entities:		8	P-value			0.0000
Avg Obs:		165.00	Distribution:			F(1,1147)
Min Obs:		165.00				
Max Obs:		165.00	F-statist	ic (robust):		104.98
			P-value			0.0000
Time periods:		165	Distribut	ion:		F(1,1147)
Avg Obs:		8.0000				
Min Obs:		8.0000				
Max Obs:		8.0000				
		Paramet	er Estimate	·s		
	Parameter	Std. Err.		P-value	Lower CI	Upper CI
const	7.1445	0.0045			7.1357	7.1533
treatment&time	-0.1639	0 0150	-10 246	0 0000	-0 1040	-0.1316

F-test for Poolability: 90.568 P-value: 0.0000 Distribution: F(171,1147)

Figure 11: Price Elasticity of Demand by Model

EDUCATION PARAMETERS		
Educ Test num	Coefficient	PED
Model employed	-0.405423	-1.351411
Robustness test 1	-0.405423	-1.351411
Robustness test 2	-0.405130	-1.350432
Robustness test 3	-0.402683	-1.342276
BUSINESS PARAMETERS		
Busi Test num	Coefficient	PED
Model employed	-0.147473	-0.491576
Robustness test 1	-0.147473	-0.491576
Robustness test 2	-0.149929	-0.499764
Robustness test 3	-0.162812	-0.542705

Figure 12: Revenue comparison by price strategy

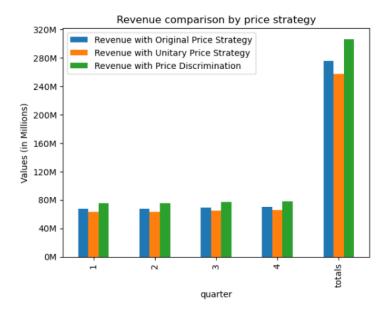


Figure 13: Potential increase in revenue by price strategy

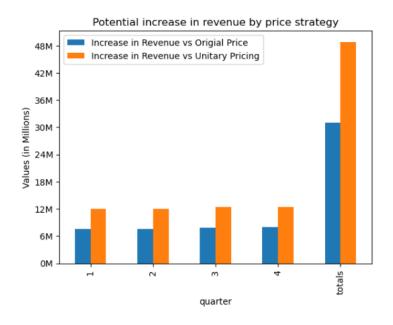


Figure 14: Potential % increase in revenue by price strategy

