The Badge Effect: Measuring the Impact of Price Signals on Airbnb Engagement

Team 4 - Mitun Adenuga, Pleng Witayaweerasak, Yitian Xu

Our Business Question

Does highlighting an Airbnb listing's favorable pricing through a badge increases user interest?




Our Hypothesis

Null Hypothesis (H_0):

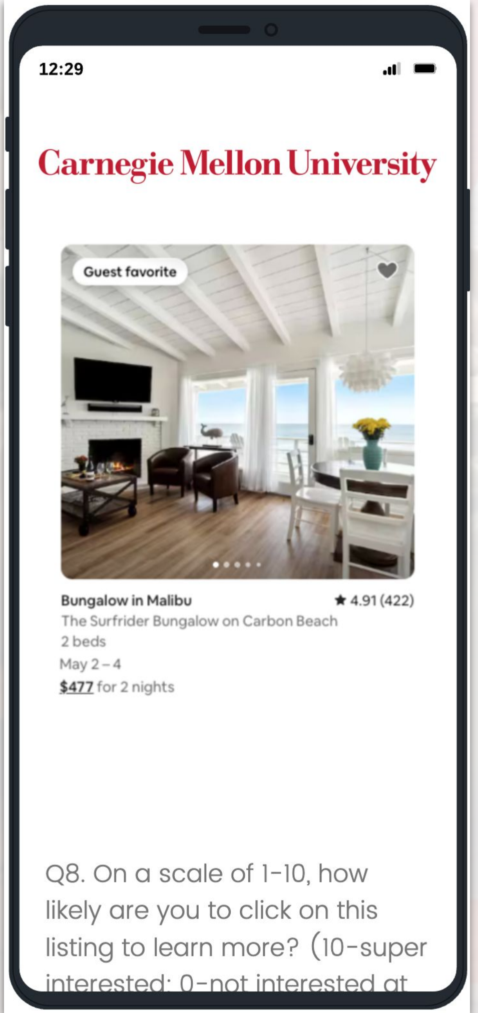
Displaying a “Good Price” badge has no effect on user interest in a listing.

Alternative Hypothesis (H_a):

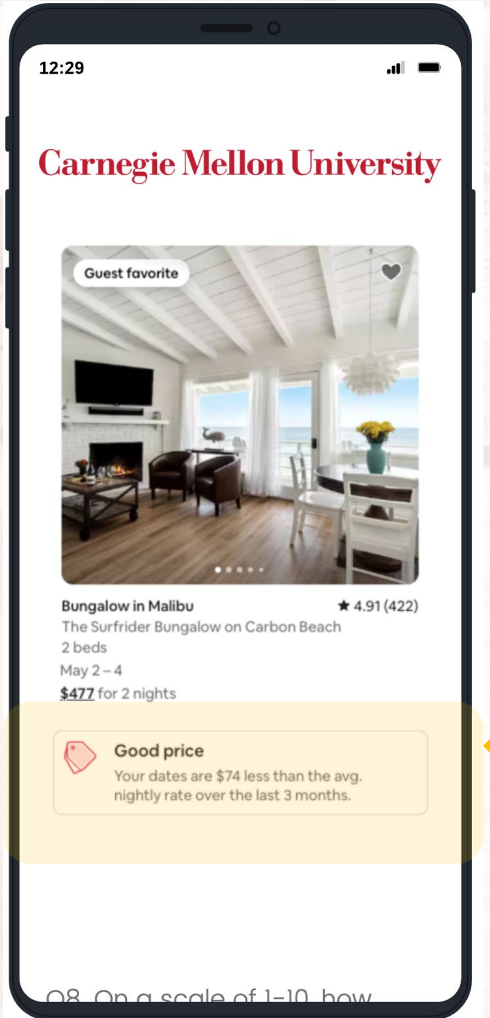
Displaying a “Good Price” badge increases user interest in a listing.

Survey Design

Control



Treatment



Good price tag

User Profile & Preferences Questions

- Age
- Gender
- Annual Household Income
- Vacations per year
- Nightly Budget
- Priorities When Booking

User Interest Questions

Click Likelihood

How likely are you to click on this listing to learn more?

Save Or Not

Based on your view of the listing, will you save it as your “favorite” to view it later?

Decision Time

Once you see the listing, how long do you need to decide and book this listing?

Rating Across 4 Dimensions

How do you like the above listing? Rate price, location, cover photo, and reviews respectively.

How We Analyzed Data

Data Cleaning

1. Exclude preview data
2. Merge control & treatment blocks
3. Convert textual columns to numeric
4. Merge with LLM synthetic data

Data Analysis

1. Balance check
2. User demographics distribution
3. Priority factors during booking
4. User engagement metrics
5. Significant correlations

Balance Checks

Balance check for: Age

	coef	std err	t	P>|t|	[0.025	0.975]
Intercept	30.5000	0.551	55.347	0.000	29.402	31.598
Treatment	-0.9444	0.801	-1.180	0.242	-2.540	0.651

Balance check for: Gender

	coef	std err	t	P>|t|	[0.025	0.975]
Intercept	0.7000	0.075	9.318	0.000	0.550	0.850
Treatment	-0.0611	0.109	-0.560	0.577	-0.279	0.156

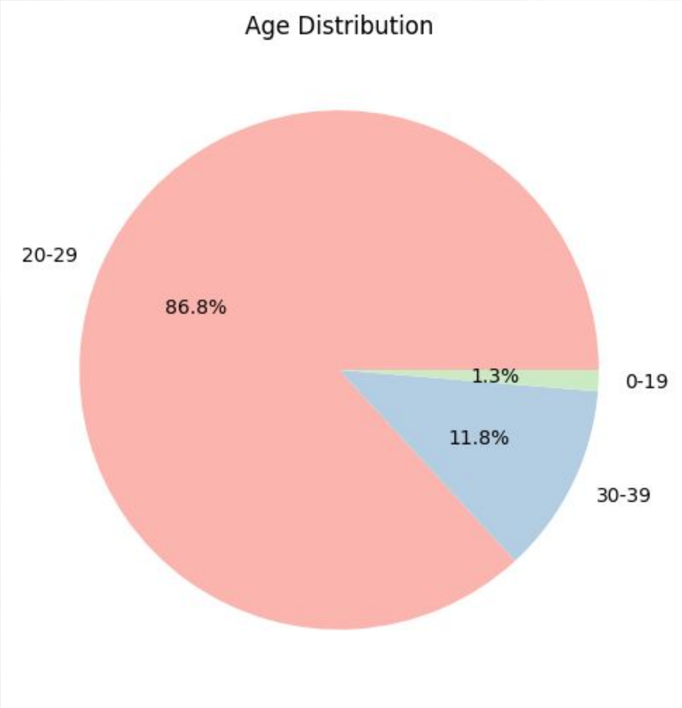
Balance check for: Income

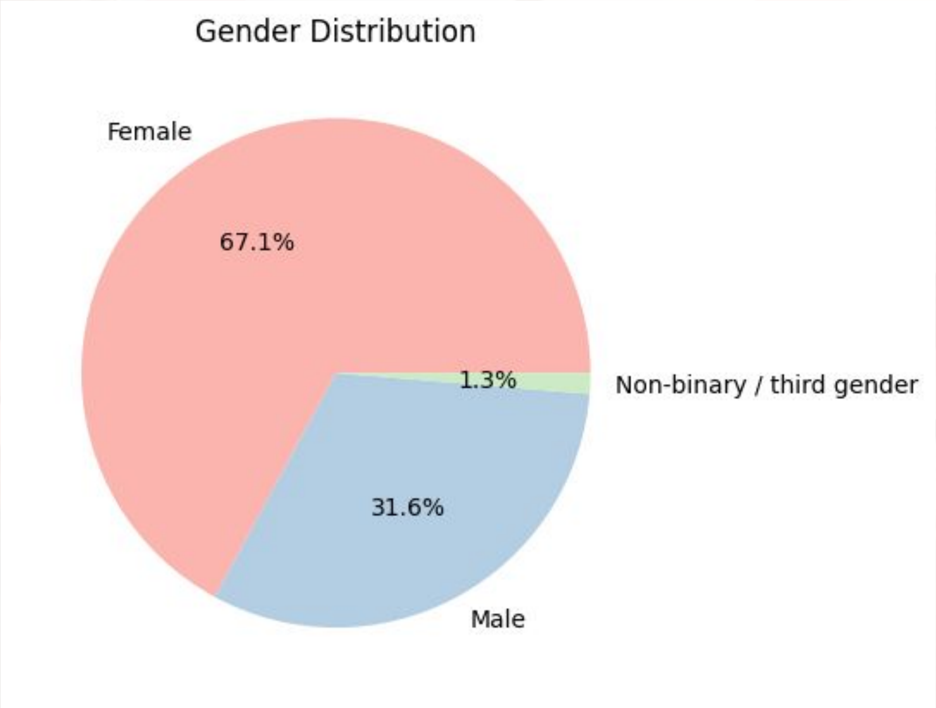
	coef	std err	t	P>|t|	[0.025	0.975]
Intercept	7.307e+04	1.04e+04	7.021	0.000	5.23e+04	9.38e+04
Treatment	9580.7421	1.51e+04	0.634	0.528	-2.06e+04	3.97e+04

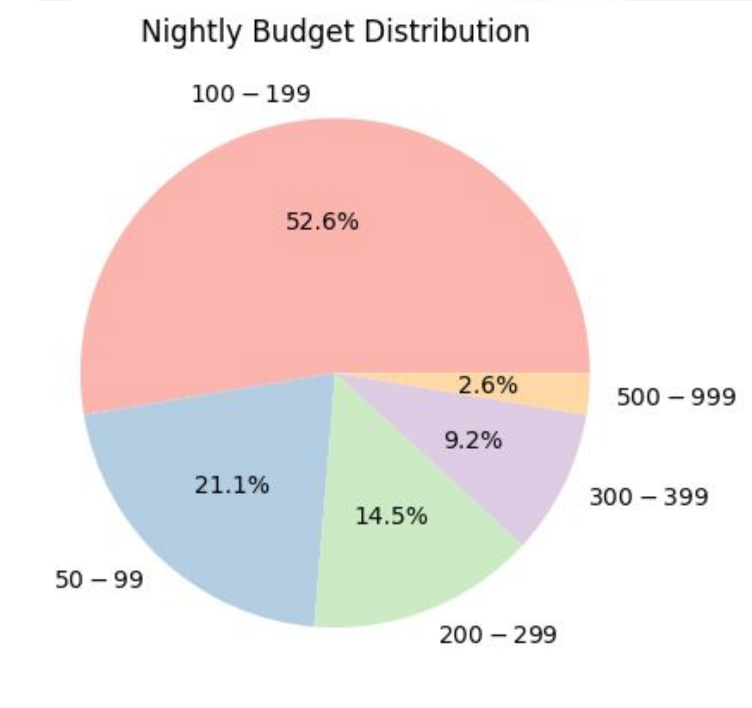
Balance check for: Nightly Budget

	coef	std err	t	P>|t|	[0.025	0.975]
Intercept	181.2500	19.431	9.328	0.000	142.534	219.966
Treatment	3.4722	28.232	0.123	0.902	-52.781	59.726

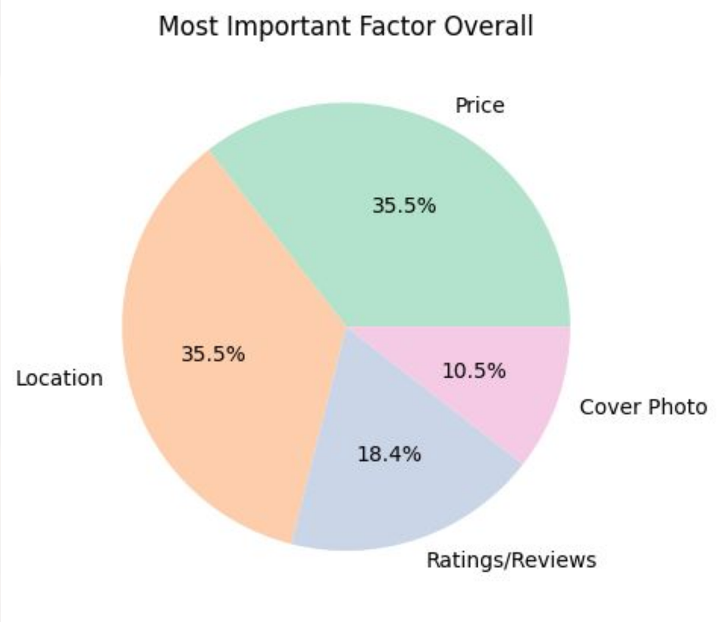
User Demographics

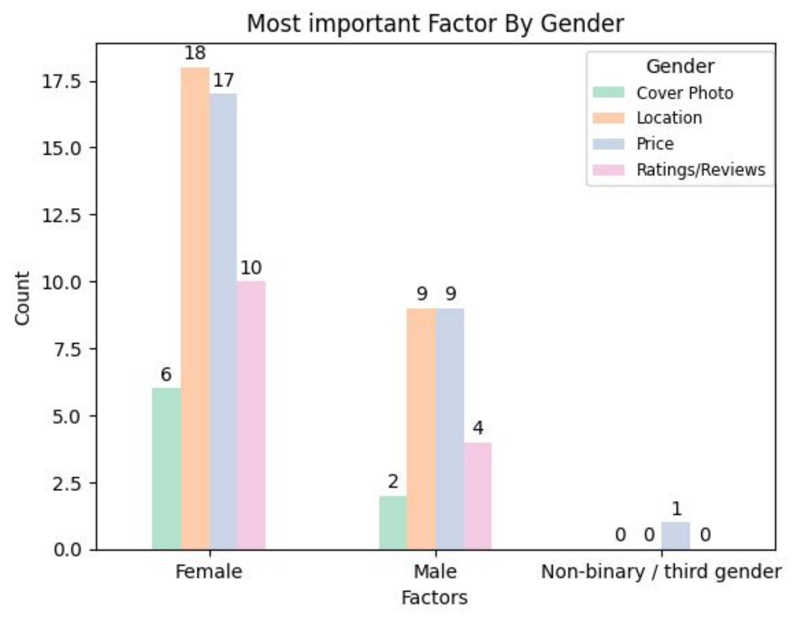


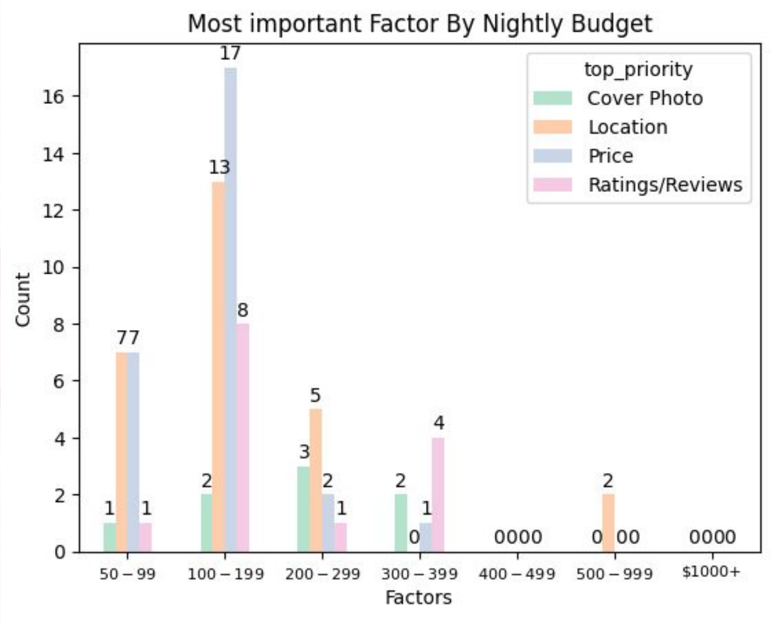




Top Booking Priorities: What Matters Most to Users







User Engagement & Decision Metrics

Metric	Mean	Min	Max
click likelihood	**6.58**	0	10
save or not	**0.68**	0	1
decision time	**95.20**	1	720

Human Responses: As Income Rises, So Does Nightly Spending

```
# price per night ~ income
smf.ols(formula='Q6_total ~ Q3_total', data = df4).fit().summary()
```

OLS Regression Results

Dep. Variable:	Q6_total	**R-squared:**	0.092
Model:	OLS	**Adj. R-squared:**	0.080
Method:	Least Squares	**F-statistic:**	7.507
Date:	Sat, 19 Apr 2025	**Prob (F-statistic):**	0.00770
Time:	18:15:07	**Log-Likelihood:**	-468.82
No. Observations:	76	**AIC:**	941.6
Df Residuals:	74	**BIC:**	946.3
Df Model:	1		
Covariance Type:	nonrobust		

	coef	**std err**	**t**	**P>|t|**	**[0.025**	**0.975]**
Intercept	139.0348	20.897	6.653	0.000	97.396	180.674
Q3_total	0.0006	0.000	2.740	0.008	0.000	0.001

Omnibus:	67.787	**Durbin-Watson:**	2.200
Prob(Omnibus):	0.000	**Jarque-Bera (JB):**	418.051
Skew:	2.741	**Prob(JB):**	1.66e-91
Kurtosis:	13.098	**Cond. No.**	1.58e+05

Human Responses: Treatment Effects

```
# Click Likelihood ~ Treatment * Nightly Budget
smf.ols(formula='Q8_total ~ Treatment * Q6_total', data = df4).fit().summary()
```

OLS Regression Results

Dep. Variable:	Q8_total	**R-squared:**	0.074
Model:	OLS	**Adj. R-squared:**	0.036
Method:	Least Squares	**F-statistic:**	1.929
Date:	Sat, 19 Apr 2025	**Prob (F-statistic):**	0.132
Time:	18:21:36	**Log-Likelihood:**	-171.15
No. Observations:	76	**AIC:**	350.3
Df Residuals:	72	**BIC:**	359.6
Df Model:	3		
Covariance Type:	nonrobust		

	coef	std err	t	P>|t|	[0.025	0.975]
Intercept	7.0905	0.669	10.606	0.000	5.758	8.423
Treatment	-1.4852	0.984	-1.509	0.136	-3.447	0.477
Q6_total	-0.0038	0.003	-1.246	0.217	-0.010	0.002
Treatment:Q6_total	0.0102	0.004	2.266	0.026	0.001	0.019

Omnibus:	3.362	**Durbin-Watson:**	1.938
Prob(Omnibus):	0.186	**Jarque-Bera (JB):**	3.292
Skew:	-0.495	**Prob(JB):**	0.193
Kurtosis:	2.755	**Cond. No.**	1.01e+03

```
# Nightly Budget Distribution
df4['Q6_total'].describe()
```

	Q6_total
count	76.000000
mean	182.894737
std	122.079928
min	75.000000
25%	150.000000
50%	150.000000
75%	250.000000
max	750.000000

LLM + Human responses also shows statistically significant relationship of income vs budget per night

```
# price per night ~ income
print(smf.ols(formula='Q6_total ~ Q3_total', data=df4).fit().summary())
```

```
                            OLS Regression Results
==============================================================================
Dep. Variable:               Q6_total   R-squared:                       0.087
Model:                            OLS   Adj. R-squared:                  0.086
Method:                 Least Squares   F-statistic:                     102.5
Date:                Sat, 19 Apr 2025   Prob (F-statistic):           4.40e-23
Time:                        23:54:00   Log-Likelihood:                -6704.6
No. Observations:                1076   AIC:                         1.341e+04
Df Residuals:                    1074   BIC:                         1.342e+04
Df Model:                           1
Covariance Type:            nonrobust
==============================================================================
                 coef    std err          t      P>|t|      [0.025      0.975]
------------------------------------------------------------------------------
Intercept    142.4450      5.690     25.034      0.000     131.280     153.610
Q3_total       0.0006   5.83e-05     10.127      0.000       0.000       0.001
==============================================================================
Omnibus:                      638.220   Durbin-Watson:                   1.987
Prob(Omnibus):                  0.000   Jarque-Bera (JB):             5154.508
Skew:                           2.700   Prob(JB):                         0.00
Kurtosis:                      12.263   Cond. No.                     1.48e+05
==============================================================================
```

LLM + Human : Treatment Effects on Decision Time

With Controls on Budget per Night

```
# Treatment effect on decision time
print(smf.ols(formula='Q11_total ~ Treatment + Q6_total',data=df4).fit().summary())
```

```
                            OLS Regression Results
==============================================================================
Dep. Variable:              Q11_total   R-squared:                       0.033
Model:                            OLS   Adj. R-squared:                  0.032
Method:                 Least Squares   F-statistic:                     18.53
Date:                Sun, 20 Apr 2025   Prob (F-statistic):           1.22e-08
Time:                        00:17:10   Log-Likelihood:                -6707.7
No. Observations:                1076   AIC:                         1.342e+04
Df Residuals:                    1073   BIC:                         1.344e+04
Df Model:                           2
Covariance Type:            nonrobust
==============================================================================
                 coef    std err          t      P>|t|      [0.025      0.975]
------------------------------------------------------------------------------
Intercept    113.8861      7.543     15.098      0.000      99.085     128.687
Treatment    -45.2054      7.549     -5.988      0.000     -60.019     -30.392
Q6_total       0.0287      0.029      0.981      0.327      -0.029       0.086
==============================================================================
Omnibus:                      785.090   Durbin-Watson:                   2.017
Prob(Omnibus):                  0.000   Jarque-Bera (JB):            10491.413
Skew:                           3.340   Prob(JB):                         0.00
Kurtosis:                      16.762   Cond. No.                         551.
==============================================================================

Notes:
[1] Standard Errors assume that the covariance matrix of the errors is correctly specified.
```

LLM + Human: Treatment Effects

Without Controls

```
# Simple OLS without controls
simple_model = smf.ols("Q8_total ~ Treatment", data=df4).fit()
print("=== SIMPLE REGRESSION: Treatment Effect Only ===")
print(simple_model.summary())
```

```
=== SIMPLE REGRESSION: Treatment Effect Only ===
                            OLS Regression Results
==============================================================================
Dep. Variable:               Q8_total   R-squared:                       0.009
Model:                            OLS   Adj. R-squared:                  0.008
Method:                 Least Squares   F-statistic:                     9.478
Date:                Sat, 19 Apr 2025   Prob (F-statistic):            0.00213
Time:                        21:23:48   Log-Likelihood:                -2472.1
No. Observations:                1076   AIC:                             4948.
Df Residuals:                    1074   BIC:                             4958.
Df Model:                           1
Covariance Type:            nonrobust
==============================================================================
                 coef    std err          t      P>|t|      [0.025      0.975]
------------------------------------------------------------------------------
Intercept      6.2439      0.101     62.082      0.000       6.047       6.441
Treatment      0.4533      0.147      3.079      0.002       0.164       0.742
==============================================================================
Omnibus:                       61.865   Durbin-Watson:                   2.087
Prob(Omnibus):                  0.000   Jarque-Bera (JB):               36.499
Skew:                          -0.306   Prob(JB):                     1.19e-08
Kurtosis:                       2.338   Cond. No.                         2.55
==============================================================================

Notes:
[1] Standard Errors assume that the covariance matrix of the errors is correctly specified.
```

With Demographics Controls

```
# OLS with demographic controls for age_group, gender, budger per night

demographic_model = smf.ols(
    "Q8_total ~ Treatment + C(Q1) + C(Q2) + Q6_total",
    data=df4
).fit()
print("\n=== REGRESSION WITH DEMOGRAPHIC CONTROLS ===")
print(demographic_model.summary())
```

```
=== REGRESSION WITH DEMOGRAPHIC CONTROLS ===
                            OLS Regression Results
==============================================================================
Dep. Variable:               Q8_total   R-squared:                       0.087
Model:                            OLS   Adj. R-squared:                  0.082
Method:                 Least Squares   F-statistic:                     16.98
Date:                Sat, 19 Apr 2025   Prob (F-statistic):           8.40e-19
Time:                        23:59:08   Log-Likelihood:                -2427.8
No. Observations:                1076   AIC:                             4870.
Df Residuals:                    1069   BIC:                             4905.
Df Model:                           6
Covariance Type:            nonrobust
===================================================================================================
                                      coef    std err          t      P>|t|      [0.025      0.975]
---------------------------------------------------------------------------------------------------
Intercept                           3.8318      0.650      5.897      0.000       2.557       5.107
C(Q1)[T.20-29]                      2.5324      0.636      3.982      0.000       1.285       3.780
C(Q1)[T.30-39]                      3.6107      0.667      5.410      0.000       2.301       4.920
C(Q2)[T.Male]                      -0.7460      0.154     -4.836      0.000      -1.049      -0.443
C(Q2)[T.Non-binary / third gender] -4.4785      0.599     -7.480      0.000      -5.653      -3.304
Treatment                           0.4879      0.147      3.316      0.001       0.199       0.777
Q6_total                            0.0005      0.001      0.858      0.391      -0.001       0.002
==============================================================================
Omnibus:                       41.639   Durbin-Watson:                   2.120
Prob(Omnibus):                  0.000   Jarque-Bera (JB):               44.574
Skew:                          -0.482   Prob(JB):                     2.09e-10
Kurtosis:                       2.744   Cond. No.                     3.54e+03
==============================================================================
```

LLM Only : Treatment Effects

Without Controls

```python
# Simple OLS without controls
simple_model = smf.ols("Q8_total ~ Treatment", data=df5).fit()
print("=== SIMPLE REGRESSION: Treatment Effect Only ===")
print(simple_model.summary())
```

```
=== SIMPLE REGRESSION: Treatment Effect Only ===
                            OLS Regression Results
==============================================================================
Dep. Variable:               Q8_total   R-squared:                       0.009
Model:                            OLS   Adj. R-squared:                  0.008
Method:                 Least Squares   F-statistic:                     9.008
Date:                Sun, 20 Apr 2025   Prob (F-statistic):            0.00275
Time:                        00:01:37   Log-Likelihood:                -2298.1
No. Observations:                1000   AIC:                             4600.
Df Residuals:                     998   BIC:                             4610.
Df Model:                           1
Covariance Type:            nonrobust
==============================================================================
                 coef    std err          t      P>|t|      [0.025      0.975]
------------------------------------------------------------------------------
Intercept      6.2322      0.104     59.726      0.000       6.027       6.437
Treatment      0.4588      0.153      3.001      0.003       0.159       0.759
==============================================================================
Omnibus:                       59.442   Durbin-Watson:                   2.086
Prob(Omnibus):                  0.000   Jarque-Bera (JB):               33.641
Skew:                          -0.296   Prob(JB):                     4.96e-08
Kurtosis:                       2.323   Cond. No.                         2.55
==============================================================================
```

With Demographics Controls

```python
# OLS with demographic controls for age_group, gender, budger per night

demographic_model = smf.ols(
    "Q8_total ~ Treatment + C(Q1) + C(Q2) + Q6_total",
    data=df5
).fit()
print("\n=== REGRESSION WITH DEMOGRAPHIC CONTROLS ===")
print(demographic_model.summary())
```

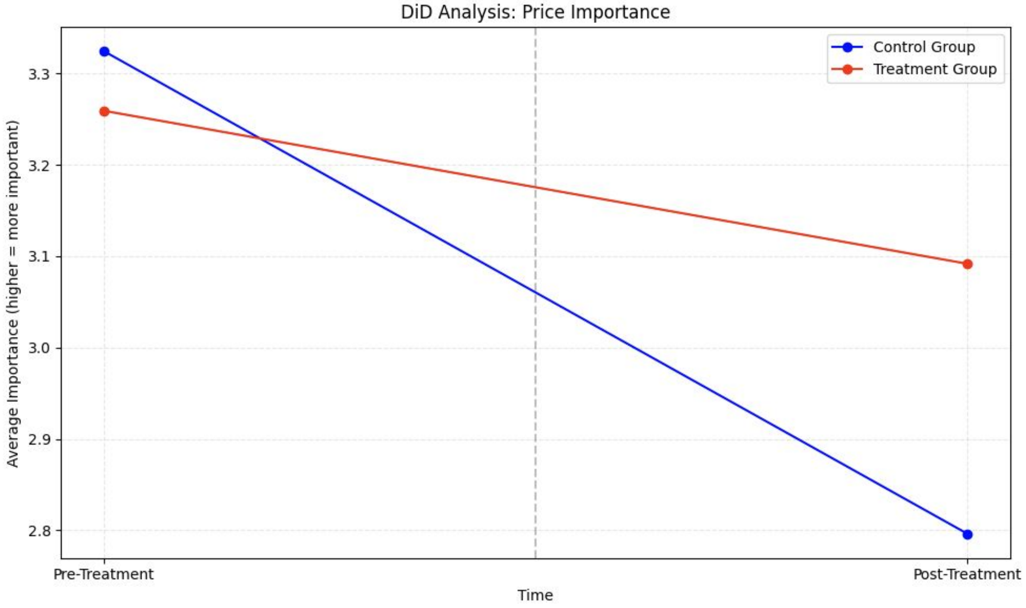
```
=== REGRESSION WITH DEMOGRAPHIC CONTROLS ===
                            OLS Regression Results
==============================================================================
Dep. Variable:               Q8_total   R-squared:                       0.088
Model:                            OLS   Adj. R-squared:                  0.082
Method:                 Least Squares   F-statistic:                     15.88
Date:                Sun, 20 Apr 2025   Prob (F-statistic):           1.75e-17
Time:                        00:01:47   Log-Likelihood:                -2256.8
No. Observations:                1000   AIC:                             4528.
Df Residuals:                     993   BIC:                             4562.
Df Model:                           6
Covariance Type:            nonrobust
===============================================================================================================
                                           coef    std err          t      P>|t|      [0.025      0.975]
---------------------------------------------------------------------------------------------------------------
Intercept                                3.8403      0.675      5.694      0.000       2.517       5.164
C(Q1)[T.20-29]                           2.5214      0.660      3.819      0.000       1.226       3.817
C(Q1)[T.30-39]                           3.5926      0.693      5.185      0.000       2.233       4.952
C(Q2)[T.Male]                           -0.7565      0.160     -4.726      0.000      -1.071      -0.442
C(Q2)[T.Non-binary / third gender]      -4.4671      0.617     -7.235      0.000      -5.679      -3.256
Treatment                                0.4959      0.153      3.244      0.001       0.196       0.796
Q6_total                                 0.0005      0.001      0.788      0.431      -0.001       0.002
==============================================================================
Omnibus:                       37.799   Durbin-Watson:                   2.124
Prob(Omnibus):                  0.000   Jarque-Bera (JB):               40.031
Skew:                          -0.471   Prob(JB):                     2.03e-09
Kurtosis:                       2.726   Cond. No.                     3.55e+03
==============================================================================
```

LLM + Human: DiD Treatment Effects

Difference in Difference of Price Ranking after Good Price Badge

```
=== DiD ANALYSIS: PRICE IMPORTANCE ===
Note: Values are converted so higher numbers = higher importance
==============================================================================
                 coef    std err          t      P>|t|      [0.025      0.975]
------------------------------------------------------------------------------
Intercept      2.7962      0.041     67.672      0.000       2.715       2.877
treatment      0.2955      0.060      4.884      0.000       0.177       0.414
post           0.5279      0.058      9.034      0.000       0.413       0.642
did           -0.3605      0.086     -4.214      0.000      -0.528      -0.193
==============================================================================

DiD Coefficient: -0.3605
P-value: 0.0000
SIGNIFICANT: Treatment DECREASED the importance of price (p < 0.05)
```



Data Insights

+0.45 point

Good Price badge **significantly increases likelihood of clicking the listing** without control for budget per night.
This demonstrates improvement in user interest.

+0.49 point

Good Price badge **significantly increases likelihood of clicking the listing** with control for budget per night.
This demonstrates improvement in user interest.

-45.21 hours

Good Price badge also **significantly decreases time to purchase** by more than 45.21 hours (1.9 days) from the baseline of 113 hours (4.7 days).
This demonstrates acceleration in time to purchase decision

Business Recommendations

Recommendation 1

Since Good Price Badge increased the likelihood of clicking on the listing, AirBnb should **bring the badge to the main feed**, not just within the listing

Recommendation 2

Badge effectiveness depends on the user's budget. Consider **segmenting by income in future analysis** or tailoring treatments for higher-budget users.

Recommendation 3

The Good Price badge is highly effective in accelerating purchase decisions. For users who tend to save listings for later or abandon them after browsing, highlighting the **Good Price can serve as a powerful nudge to encourage immediate booking.**

Limitations

1

Small Sample Size

Only 76 human responses are not representative of AirBnB users

2

Gender Imbalance

Much higher responses from females (67% vs other genders)

3

Age Concentration

86% of participants are aged 20–29, limiting representation of older age groups.

4

Isolated Listing View

Users viewed only a single listing instead of browsing in a scrollable feed, which may not reflect natural decision-making context.

Thank You

Please ask any question in the Q&A