Stage 1: Linear Competitor Price Decomposition
Hotel Pricing 2SRI Analysis - Corrected Implementation

Pricing Service Analysis

September 25, 2025

1 Statistical Significance Notation

Throughout this report, statistical significance is indicated using the following notation:

- * = p $<$ 0.05 (significant at 5% level)
- ** = p $<$ 0.01 (significant at 1% level)
- *** = p $<$ 0.001 (significant at 0.1% level)
- No asterisk = not statistically significant (p $\geq$ 0.05)

2 Methodology

2.1 Data Alignment Improvements

The corrected implementation addresses previous NaN value issues through:

- Proper date intersection across all datasets (focal and competitor prices)
- Complete observation filtering ensuring no missing values
- Uniform sample size (n=288) across all competitor models
- Clean temporal feature alignment

2.2 Corrected Temporal Feature Engineering

Multicollinearity issues resolved through streamlined feature set:

- Monthly seasonality: $\sin(\frac{2\pi \cdot month}{12})$, $\cos(\frac{2\pi \cdot month}{12})$
- Weekly patterns: $\sin(\frac{2\pi \cdot dayofweek}{7})$, $\cos(\frac{2\pi \cdot dayofweek}{7})$
- Business periods: Holiday season, summer travel, spring break indicators

Eliminated redundant features: day_of_week, is_weekend, and week_of_year-based calculations.

2.3 Stage 1 Regression Model

For each competitor i:

$$P_{i,t} = \alpha_i + \beta_{i,1} \sin_{month,t} + \beta_{i,2} \cos_{month,t} + \beta_{i,3} \sin_{day,t} + \beta_{i,4} \cos_{day,t} + \gamma_i X_t + \varepsilon_{i,t} \quad (1)$$

where X_t represents business period indicators and $\varepsilon_{i,t}$ are the residuals used as instruments in Stage 2.

3 Individual Competitor Model Results

3.1 Aqua Pacific Monarch (R^2=0.457, F=26.7)

$$\begin{aligned} P_{Aqua,t} =& 264.437^{***} + 23.400^{***} \sin_{month,t} - 6.581^{*} \cos_{month,t} \\ & - 3.548^{*} \sin_{day,t} - 2.666 \cos_{day,t} + 19.418^{***} Holiday_t \\ & + 20.841^{***} Summer_t - 2.491 Spring_t + \varepsilon_{Aqua,t} \end{aligned} \quad (2)$$

Key Findings:

- Strong monthly seasonality ($\sin_{month}$: t=8.23, p<0.001)
- Significant holiday premium ($19.42, t=4.16, p<0.001)
- Summer travel premium ($20.84, t=4.36, p<0.001)
- Weekly pattern effects ($\sin_{day}$: t=-2.11, p=0.036)
- Strong instrument validity (F=33.7 > 10)

3.2 Castle Kamaole Sands (R^2=0.553, F=49.5)

$$\begin{aligned} P_{Castle,t} =& 322.805^{***} + 31.986^{***} \sin_{month,t} + 49.132^{***} \cos_{month,t} \\ & + 0.409 \sin_{day,t} - 1.819 \cos_{day,t} - 35.971^{***} Holiday_t \\ & + 72.118^{***} Summer_t - 3.291 Spring_t + \varepsilon_{Castle,t} \end{aligned} \quad (3)$$

Key Findings:

- Strongest seasonal patterns (both $\sin_{month}$ and $\cos_{month}$ highly significant)
- Holiday discount (-$35.97, t=-8.37, p<0.001) - unique among competitors
- Highest summer premium ($72.12, t=16.40, p<0.001)
- Excellent instrument strength (F=49.5)
- No significant weekly patterns

3.3 Courtyard Marriott Airport (R^2=0.749, F=119.2)

$$\begin{aligned} P_{Courtyard,t} =& 433.954^{***} + 5.646 \sin_{month,t} - 41.752^{***} \cos_{month,t} \\ & - 0.992 \sin_{day,t} + 2.018 \cos_{day,t} - 15.613^{**} Holiday_t \\ & + 4.836 Summer_t + 29.994^{***} Spring_t + \varepsilon_{Courtyard,t} \end{aligned} \tag{4}$$

Key Findings:

- Highest model fit (R^2=0.749)
- Strong negative cosine monthly pattern (t=-11.47, p<0.001)
- Holiday discount (-$15.61, t=-2.89, p=0.004)
- Strong spring break premium ($29.99, t=6.04, p<0.001)
- Strongest instrument (F=119.2, highest F-statistic)
- Minimal summer travel effect (not significant)

3.4 Kohea Kai Resort (R^2=0.243, F=12.8)

$$\begin{aligned} P_{Kohea,t} =& 277.372^{***} + 3.602 \sin_{month,t} + 6.812^{*} \cos_{month,t} \\ & - 5.981^{***} \sin_{day,t} - 4.124^{*} \cos_{day,t} + 6.003 Holiday_t \\ & - 6.223 Summer_t + 8.072 Spring_t + \varepsilon_{Kohea,t} \end{aligned} \tag{5}$$

Key Findings:

- Only competitor with significant weekly patterns (both $\sin_{day}$ and $\cos_{day}$)
- Moderate monthly cosine effect (t=2.26, p=0.024)
- Lower seasonal variation compared to other competitors
- Adequate instrument strength (F=12.8 > 10)
- Spring break effect approaching significance (p=0.050)

3.5 Ohana Waikiki Malia (R^2=0.708, F=96.9)

$$\begin{aligned} P_{Ohana,t} =& 250.656^{***} + 7.116^{*} \sin_{month,t} - 32.396^{***} \cos_{month,t} \\ & - 3.886^{*} \sin_{day,t} - 3.222 \cos_{day,t} + 32.656^{***} Holiday_t \\ & + 33.011^{***} Summer_t - 16.920^{***} Spring_t + \varepsilon_{Ohana,t} \end{aligned} \tag{6}$$

Key Findings:

- High model fit (R^2=0.708)
- Strong negative monthly cosine pattern (t=-10.10, p<0.001)
- Largest holiday premium ($32.66, t=6.85, p<0.001)
- Spring break discount (-$16.92, t=-3.87, p<0.001) - unique pricing strategy
- Strong instrument (F=96.9)
- Significant weekly day pattern ($\sin_{day}$: t=-2.26, p=0.024)

4 Instrument Strength Validation

Competitor Hotel	R^2	F-Statistic	Strength	Sample Size
Aqua Pacific Monarch	0.457	33.7	Strong	288
Castle Kamaole Sands	0.553	49.5	Strong	288
Courtyard Marriott Airport	0.749	119.2	Strong	288
Kohea Kai Resort	0.243	12.8	Strong	288
Ohana Waikiki Malia	0.708	96.9	Strong	288

Table 1: Instrument Strength Assessment (Threshold: F > 10 for Strong Instruments)

All 5 competitors exceed the weak instrument threshold (F > 10), with 4 out of 5 showing F-statistics above 30, indicating very strong instruments for Stage 2 endogeneity correction. The uniform sample size of 288 ensures consistent estimation quality across all models.

5 Residual Quality Assessment

The generated residuals demonstrate proper statistical properties:

- Zero mean: All residual means $< 1\times10^{-6}$ (numerical zero)
- Standard deviations range from \$18.39 to \$23.16
- No systematic patterns in residual distributions
- Uniform sample sizes (n=288) across all competitors
- Ready for use as instruments in Stage 2 2SRI implementation

6 Temporal Pattern Analysis

6.1 Monthly Seasonality

- **Castle and Aqua**: Strong positive sine patterns (summer peaks)
- **Courtyard and Ohana**: Strong negative cosine patterns (winter/spring adjustments)
- **Kohea**: Moderate positive cosine pattern

6.2 Weekly Patterns

- **Kohea**: Strong significant daily patterns (both sine and cosine)
- **Aqua and Ohana**: Moderate weekly variation ($\sin_{day}$ significant)
- **Castle and Courtyard**: Minimal weekly variation

6.3 Business Periods

- **Holiday Season**: Mixed effects - premiums for Aqua and Ohana; discounts for Castle and Courtyard
- **Summer Travel**: Strong premiums for Castle ($72.12) and Ohana ($33.01); moderate for Aqua ($20.84)
- **Spring Break**: Significant premium for Courtyard ($29.99); discount for Ohana (-$16.92)

7 Performance Improvements

Compared to Previous Implementation:

- Average R^2 improved from 0.414 to 0.542 (31% increase)
- All competitors now have uniform sample size (288 vs previous 307-364 range)
- Strong instrument count: 5/5 vs previous 4/5
- F-statistics range improved: 12.8-119.2 vs previous 9.2-63.8
- Complete elimination of NaN values in output datasets

8 Stage 2 Readiness

Data Exports Generated:

- competitor_residuals_linear.csv: Clean instruments for Stage 2
- competitor_prices_actual.csv: Aligned endogenous variables
- temporal_features.csv: Consistent exogenous controls
- focal_base_rates.csv: Aligned dependent variable for Stage 2
- stage1_linear_metadata.json: Complete model diagnostics

Key Success Metrics:

- Strong instruments: 5/5 competitors (100%)
- Average R^2: 0.542 (54.2% temporal variation explained)
- F-statistic range: 12.8 to 119.2 (all well above weak instrument threshold)
- Uniform sample size: 288 observations per competitor
- Zero NaN values in all output datasets