Two-Stage Residual Inclusion (2SRI) Implementation
Hotel Pricing Competition Analysis

PricingService.ai Research Team

September 27, 2025

1 Data Summary

- **Sample Size:** 365 daily observations (2025-09-16 to 2026-09-14)
- **Focal Hotel:** Daily base rates (minimum room prices)
- **Competitors:** 5 hotels with complete price coverage
- **Missing Values:** Zero

2 Data Preprocessing

2.1 Key Steps

1. **Missing Value Treatment:** Forward fill + median imputation for competitor gaps
2. **Outlier Removal:** IQR-based filtering (3.0 multiplier)
3. **Cross-Hotel Date Filling:** Ensured complete coverage across all focal hotel dates
4. **Base Rate Extraction:** Daily minimum prices from focal hotel room types
5. **Normalization:** MAD-based cross-hotel price scaling

2.2 Final Dataset

- **Focal Hotel:** 365 daily base rates, range \$209–\$379
- **Competitors:** 365×5 complete price matrix
- **Data Completeness:** 100% (zero missing values)

3 Stage 1: Competitor Price Decomposition

3.1 Methodology

Linear regression models for each competitor using purely cyclical temporal features:

$$
\begin{aligned}
P_{c,t} = \alpha_c &+ \gamma_1 \sin\left(\frac{2\pi \cdot \text{month}_t}{12}\right) + \gamma_2 \cos\left(\frac{2\pi \cdot \text{month}_t}{12}\right) \\
&+ \gamma_3 \sin\left(\frac{2\pi \cdot \text{day}_t}{7}\right) + \gamma_4 \cos\left(\frac{2\pi \cdot \text{day}_t}{7}\right) \\
&+ \gamma_5 \sin\left(\frac{2\pi \cdot \text{week}_t}{52}\right) + \gamma_6 \cos\left(\frac{2\pi \cdot \text{week}_t}{52}\right) \\
&+ \gamma_7 \mathbb{I}_{\text{holiday},t} + \varepsilon_{c,t}
\end{aligned} \tag{1}
$$

3.2 Individual Competitor Models

3.2.1 Aqua Pacific Monarch

$$
\begin{aligned}
P_{\text{MONA},t} = &273.798 + 18.389 \sin(\text{month}) - 9.324 \cos(\text{month}) \\
&- 3.039 \sin(\text{day}) - 1.290 \cos(\text{day}) \\
&- 1.899 \sin(\text{week}) - 5.319 \cos(\text{week}) \\
&+ 20.735^{***} \mathbb{I}_{\text{holiday}} + \varepsilon_{\text{MONA},t}
\end{aligned} \tag{2}
$$

Model Statistics: $R^2 = 0.306$, $F = 22.5$, $n = 365$

3.2.2 Castle Kamaole Sands

$$
\begin{aligned}
P_{\text{SAND},t} = &338.811 + 9.561 \sin(\text{month}) + 17.695 \cos(\text{month}) \\
&+ 0.166 \sin(\text{day}) - 1.485 \cos(\text{day}) \\
&+ 6.104 \sin(\text{week}) + 1.039 \cos(\text{week}) \\
&- 15.287^{**} \mathbb{I}_{\text{holiday}} + \varepsilon_{\text{SAND},t}
\end{aligned} \tag{3}
$$

Model Statistics: $R^2 = 0.214$, $F = 13.9$, $n = 365$

3.2.3 Courtyard Marriott Airport

$$
\begin{aligned}
P_{\text{AIRP},t} = &441.189 - 67.915^{***} \sin(\text{month}) - 8.157 \cos(\text{month}) \\
&+ 1.987 \sin(\text{day}) + 1.934 \cos(\text{day}) \\
&+ 76.399^{***} \sin(\text{week}) - 24.562 \cos(\text{week}) \\
&+ 7.641 \mathbb{I}_{\text{holiday}} + \varepsilon_{\text{AIRP},t}
\end{aligned} \tag{4}
$$

Model Statistics: $R^2 = 0.391$, $F = 32.8$, $n = 365$

3.2.4 Kohea Kai Resort Maui

$$
\begin{aligned}
P_{\text{MAUI},t} = &278.654 - 15.626 \sin(\text{month}) + 3.972 \cos(\text{month}) \\
&- 4.490^{**} \sin(\text{day}) - 4.732^{**} \cos(\text{day}) \\
&+ 21.865^{*} \sin(\text{week}) + 11.811 \cos(\text{week}) \\
&+ 5.126 \mathbb{I}_{\text{holiday}} + \varepsilon_{\text{MAUI},t}
\end{aligned} \tag{5}
$$

Model Statistics: $R^2 = 0.254$, $F = 17.4$, $n = 365$

3.2.5 Ohana Waikiki Malia

$$\begin{aligned} P_{\text{MALI},t} = {} & 261.143 - 19.434 \sin(\text{month}) - 33.965^{*} \cos(\text{month}) \\ & - 4.058 \sin(\text{day}) - 4.174 \cos(\text{day}) \\ & + 7.458 \sin(\text{week}) - 15.729 \cos(\text{week}) \\ & + 57.976^{***} \mathbb{I}_{\text{holiday}} + \varepsilon_{\text{MALI},t} \end{aligned} \tag{6}$$

Model Statistics: $R^2 = 0.486$, $F = 48.2$, $n = 365$

3.3 Instrument Strength Validation

All F-statistics exceed the threshold of 10, confirming strong instruments:

- F-statistics range: 13.9 to 48.2
- All residuals have zero mean and appropriate variance
- Temporal features successfully capture seasonal patterns

4 Stage 2: Two-Stage Residual Inclusion (2SRI)

4.1 Methodology

The 2SRI model addresses endogeneity in competitive pricing:

$$P_{\text{focal},t} = \alpha + \sum_{c=1}^{5} \beta_c P_{c,t} + \sum_{c=1}^{5} \theta_c \hat{\varepsilon}_{c,t} + \sum_{j=1}^{7} \gamma_j X_{j,t} + u_t \tag{7}$$

Where:

- $P_{\text{focal},t}$ = focal hotel base rate
- $P_{c,t}$ = competitor c's price (endogenous)
- $\hat{\varepsilon}_{c,t}$ = Stage 1 residuals (instruments)
- $X_{j,t}$ = temporal controls
- β_c = direct competitive effects
- θ_c = endogeneity correction parameters

4.2 Complete 2SRI Estimation Results

$$
\begin{aligned}
P_{\text{focal},t} = & -72.289 + 0.563 \times P_{\text{Aqua Pacific Monarch}} \\
& - 0.862 \times P_{\text{Castle Kamaole Sands}} \\
& + 0.105 \times P_{\text{Courtyard Marriott Airport}} \\
& + 1.758^{**} \times P_{\text{Kohea Kai Resort Maui}} \\
& - 0.241 \times P_{\text{Ohana Waikiki Malia}} \\
& - 0.597 \times \hat{\varepsilon}_{\text{Aqua Pacific Monarch}} \\
& + 1.144^{*} \times \hat{\varepsilon}_{\text{Castle Kamaole Sands}} \\
& + 0.202 \times \hat{\varepsilon}_{\text{Courtyard Marriott Airport}} \\
& - 1.659^{**} \times \hat{\varepsilon}_{\text{Kohea Kai Resort Maui}} \\
& + 0.330 \times \hat{\varepsilon}_{\text{Ohana Waikiki Malia}} \\
& + 0.490 \times \sin(\text{month}) + 1.770 \times \cos(\text{month}) \\
& + 8.696^{**} \times \sin(\text{day}) + 4.175 \times \cos(\text{day}) \\
& + 0.851 \times \sin(\text{week}) + 2.578 \times \cos(\text{week}) \\
& + 2.691 \times \mathbb{I}_{\text{holiday}} + u_t
\end{aligned} \tag{8}
$$

4.3 Model Performance

- $R^2 = 0.512$, RMSE = \$23.91
- Sample Size: 365 observations
- Total Features: 17 (5 prices + 5 residuals + 7 temporal)

5 Key Results

5.1 Endogeneity Detection

Significant θ coefficients ($p < 0.05$) in 2 out of 5 competitors:

- **Castle Kamaole Sands:** $\theta = 1.144^{*}$ ($p = 0.025$)
- **Kohea Kai Resort Maui:** $\theta = -1.659^{**}$ ($p = 0.004$)

This confirms simultaneity bias was present and corrected through 2SRI.

5.2 Competitive Strategy Analysis

- **Total competitive effect:** 1.323 (complementary pricing strategy)
- **Dominant competitor:** Kohea Kai Resort Maui ($\beta = 1.758^{**}$)
- **Competitive dynamics:** 3 complementary, 2 competitive relationships

5.3 Statistical Significance

Significance levels: $^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$

Key significant relationships:

- **Kohea Kai Resort Maui:** Strong complementary pricing ($\beta = 1.758^{**}$)
- **Weekly patterns:** Significant day-of-week effects ($\gamma_{\sin(\text{day})} = 8.696^{**}$)
- **Endogeneity correction:** Validated in 40% of competitor relationships

5.4 Implementation Pipeline

1. **Preprocessing:** Complete missing value treatment and outlier removal
2. **Stage 1:** Individual competitor temporal decomposition (5 models)
3. **Stage 2:** Joint 2SRI estimation with endogeneity correction
4. **Validation:** F-statistics, significance testing, economic interpretation