A Comp Decision Engine for Luxury Hotel Service Recovery

A focused prototype and evaluation plan for manager-facing comp decisions

Grant McCurdy

Evidence boundary. The operating records, historical comp actions, costs, and outcomes in this prototype are synthetic. Official Santa Monica Proper sources inform the property-relevant recovery menu and guest-facing value anchors only. No Proper Hotels guest records, policies, inventory, margins, comp history, or performance data are used.

The business task

Luxury-hotel recovery decisions carry two risks at once. A response that is too small can lose the guest and damage the brand; a response that is poorly matched or unnecessarily expensive can erode rate and create inconsistent treatment. The question is not simply how to reduce comp spend:

> How can managers choose the right gesture for the guest and the situation while making cost, constraints, and reasoning visible?

The proposed decision product

Build a manager-facing Comp Decision Engine that places a **minimum recovery obligation** on the service failure itself, then ranks only gestures that are adequate, available, and within approval rules. Guest relationship value can justify additional generosity, but it cannot reduce what the failure warrants. A repeat-comp pattern triggers review; it does not erase the recovery obligation.

For each incident, the product returns one recommended gesture, guest-facing value, estimated cost range, delivery timing, alternatives, reasons, and approval path. The manager confirms operational availability, accepts or overrides the recommendation, and remains responsible for the guest interaction.

Test the synthetic prototype: Open the Decision Desk. Use synthetic scenarios only; do not enter actual guest or reservation information.

An illustrative recommendation

For a synthetic returning guest with a severity-4 room-readiness delay and high hotel responsibility, the prototype produces:

> **Recommended recovery:** late checkout + personal manager note
> **Modeled guest-facing value:** $100
> **Assumed internal-cost range:** $8-$45
> **Condition:** confirm live room availability before offering it

The reasoning is visible: the hotel owns a serious disruption, recovery can still happen during the stay, and late checkout meets the configured recovery floor at lower assumed cost than a direct refund. The manager can compare alternatives and override the result. The amount and cost range demonstrate the output contract; they are not estimates of property economics.

How the real model would be chosen

The current rules are a testable starting point. With actual property data, evaluate four approaches on the same future-dated cases:

1. **Manager judgment:** the operating baseline for choice, reason, speed, cost, and outcome.
2. **Explicit recovery rules:** a transparent standard tested for adequacy, consistency, feasibility, and escalation.
3. **Interpretable statistical model:** calibrated recovery and cost estimates tested over time and by segment.
4. **Boosted-tree benchmark:** a test of nonlinear improvement without weaker recovery safeguards.

Estimate satisfaction recovery, review risk, repeat-stay behavior, actual marginal cost, and operational feasibility separately. Compare calibration and decision quality on temporal holdouts, use paired resampling for differences between approaches, and inspect performance by severity, issue type, and guest segment. A model advances only if it preserves recovery and escalation safeguards while improving consistency, cost visibility, and manager usability.

Historical actions are not automatically causal because managers offered different gestures to different situations. Where comparable treatment overlap is weak, shadow observation or a controlled test is required before attributing an outcome to the recommendation.

A focused first step

Start with a 90-minute data and policy workshop. Define adequate recovery, approval rules, feasible gestures, marginal cost, outcome windows, and the systems that own each field. Then run the prototype invisibly for four weeks or 50 eligible cases, whichever is later, to test data capture, recommendation feasibility, and manager trust.

That first sample is for workflow discovery, not proof of impact. A powered controlled test should begin only after the data and operating process are credible.