

Forecasting Sparse Airline Demand for Revenue Management: Hierarchical Count Models and Fare-Sensitive Premium-Cabin Capacity Sharing

Irakli Gudavadze

`Irakli.gudavadze@utexas.edu`

The University of Texas at Austin

Research Article

Keywords: airline revenue management, sparse demand, Poisson–Gamma, empirical Bayes, premium cabin capacity sharing, cabin floating

Posted Date: October 16th, 2026

DOI: <https://doi.org/10.21203/rs.3.rs-11236942/v1>

License: © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Additional Declarations: No competing interests reported.

Forecasting Sparse Airline Demand for Revenue Management: Hierarchical Count Models and Fare-Sensitive Premium-Cabin Capacity Sharing

ABSTRACT

Low-volume airline demand is difficult to forecast because route-level estimates are unstable and can collapse to zero. We propose a hierarchical exposure-adjusted Poisson–Gamma (HEPG) model that partially pools demand across markets and yields a negative-binomial posterior predictive distribution. Synthetic experiments show lower negative log predictive scores than pooled and route-level Poisson benchmarks. We then connect the predictive distribution to premium-cabin capacity sharing (“cabin floating”), deriving a fare-sensitive seat-protection rule. In decision simulations, HEPG consistently reduces revenue regret relative to plug-in Poisson rules, with the benefit increasing as the premium-to-economy fare ratio rises.

Keywords: airline revenue management; sparse demand; Poisson–Gamma; empirical Bayes; premium cabin capacity sharing; cabin floating.

INTRODUCTION

Airline revenue management depends on forecasts at granular levels such as flight leg, origin–destination market, fare product, booking-horizon interval, or point-of-sale segment. When expected demand in one of these cells is well below a few requests per period, the forecasting problem changes qualitatively: zeros are common, the outcome is discrete, and route-level parameter estimates can be dominated by sampling noise. A short sequence of zeros may reflect genuinely low demand, but it does not prove that the underlying arrival rate is zero.

$$0, 0, 1, 0, 0, 0, 2, 0, 0, \dots$$

is not well represented by a symmetric continuous distribution. Zero is a valid and often probable outcome, the support is nonnegative and integer-valued, and the variance is strongly connected to the level of the process.

The issue is operational rather than merely statistical. Forecast distributions feed booking limits, seat protection, availability, pricing, and other revenue-management controls (Talluri and Van Ryzin 2004; Weatherford and Belobaba 2002). For thin demand, the probability of no request, the probability of at least one late premium request, and the upper tail of remaining demand can matter more than a single point forecast.

Poisson demand is not new to airline revenue management and is not claimed as the novelty of this study. JRPM has published airline allocation work with nonhomogeneous Poisson demand (Korkmaz, Kundakcioglu, and Sivrikaya 2022), while Meyer et al. (2025) model airline booking intensity with a Poisson process for price-sensitive demand. The present contribution is the combination of exposure-aware count modeling, network-level empirical-Bayes shrinkage, full predictive uncertainty, and a fare-sensitive capacity-sharing decision designed specifically for sparse premium demand.

1. a count likelihood with explicit exposure, so that different numbers of selling opportunities or scheduled departures can be compared on a common rate scale;
2. empirical-Bayes partial pooling across markets, which stabilizes sparse route estimates without forcing all routes to have the same demand;
3. a closed-form posterior predictive distribution that automatically becomes wider when route information is scarce; and
4. a diagnostic variance correction that relaxes the conditional Poisson assumption when the data display extra-Poisson variability; and
5. a decision layer for premium-cabin capacity sharing that converts the sparse premium-demand predictive distribution into a transparent seat-protection rule based on fare differences.

The framework is called the *hierarchical exposure-adjusted Poisson–Gamma* model, abbreviated HEPG. If market i has latent rate λ_i , period t has exposure e_{it} , and pooled calendar multiplier s_{it} , then

$$Y_{it} \mid \lambda_i \sim \text{Poisson}(e_{it}s_{it}\lambda_i), \quad \lambda_i \sim \text{Gamma}(a, b),$$

where the Gamma distribution is parameterized by shape a and rate b . Conditional on the pooled covariate effects, this hierarchy is conjugate. The posterior rate estimate is a weighted average of the route-only estimate and the network prior mean. For a future period, integrating over route uncertainty

yields a negative-binomial predictive distribution even though the conditional arrival mechanism is Poisson.

This structure is especially attractive in the low-demand regime because parameter uncertainty is a first-order effect. When a route has only a few effective exposure units, the forecast should be wider than a plug-in Poisson distribution built from a single estimated mean. HEPG provides that widening analytically rather than through an ad hoc safety factor.

The paper makes an additional practical distinction that is essential in airline data. A recorded zero is not always a realization of zero demand. A zero can also arise because the flight did not operate, inventory was closed, a fare product was unavailable, or observed bookings were censored by controls. The model developed here is intended for latent request counts, open and uncensored booking observations, or bookings that have been appropriately unconstrained. Airline unconstraining has a large literature precisely because accepted bookings need not equal underlying demand (Queenan et al., 2007; Weatherford & Pölt, 2002). Treating an availability-induced zero as a demand zero will bias any count model.

The paper first positions the method within airline forecasting, intermittent-demand modeling, and cabin-capacity sharing. It then develops the HEPG forecast, evaluates it in controlled sparse-demand simulations, derives a cabin-floating seat-protection rule, and measures the revenue regret produced by alternative forecasts. The final sections translate the method into an implementation workflow and identify the empirical validation still required before operational adoption.

RELATED LITERATURE

Airline demand forecasting and revenue management

Revenue management converts uncertain demand forecasts into inventory and pricing decisions. Forecast error can materially affect seat-allocation performance (Weatherford and Belobaba 2002), and unconstraining remains essential when accepted bookings are lower than latent demand because of availability controls (Weatherford and Pölt 2002; Queenan et al. 2007). More recently, Lee and Abdelghany (2026) describe airline demand as a “short-but-wide” problem: individual routes may have little stable history, but an airline observes many related routes. Their cross-series transfer-learning results motivate the same information-sharing principle used here, although HEPG targets low integer counts and provides an analytic predictive distribution rather than a neural forecasting architecture.

Distributional assumptions also matter. Swan (2002) discusses the shape of airline demand and spill, Korkmaz, Kundakcioglu, and Sivrikaya (2022) use nonhomogeneous Poisson demand in a single-leg allocation model, and Meyer et al. (2025) estimate Poisson booking intensities as part of a price-sensitive demand model. Other JRPM work connects forecasting inputs to joint pricing and allocation decisions (Jayaram et al. 2024) and to segmented continuous pricing (Long and Belobaba 2024). These studies reinforce the practical value of connecting a demand model to a revenue decision rather than evaluating prediction accuracy in isolation.

The closest operational antecedent is premium-cabin capacity sharing. Escovar-Álvarez and Belobaba (2022) evaluate accepting additional economy bookings when premium seats are expected to remain empty. Their PODS simulations show potential revenue gains but also show that sharing too early can displace later premium demand. Feng et al. (2024) study dynamic and flexible capacity management between business and economy compartments and include free upgrades as an extension. The

contribution here is complementary: instead of proposing a new cabin-sharing mechanism, we ask how sparse premium-demand uncertainty should change the number of seats released at a given fare gap.

Intermittent demand and count forecasting

Sparse airline booking series are related to intermittent-demand problems in inventory forecasting. Croston (1972) proposed separate smoothing of demand size and inter-arrival time. Subsequent work studied the bias and stochastic interpretation of Croston-type procedures (Shenstone & Hyndman, 2005; Syntetos & Boylan, 2005). For probabilistic planning, Snyder et al. (2012) compare Poisson, negative-binomial, and hurdle-type models and emphasize the value of modeling the full predictive distribution rather than only a point forecast.

The count literature also provides extensions for excess zeros and overdispersion. Zero-inflated count time-series models have been developed for settings where zeros occur more often than a standard Poisson model predicts (Maiti et al., 2015). Jiang (2020) separates zero and positive-demand mechanisms using a mixed zero-truncated Poisson approach. More recently, Aktekin et al. (2026) developed Bayesian zero-inflated Poisson–Gamma state-space models for sequential count forecasting. These models motivate a diagnostic principle used in this paper: begin with the simplest coherent count process, then relax it only when residual diagnostics show that Poisson variability is insufficient.

Hierarchical Poisson models

Hierarchical Poisson models and empirical-Bayes estimation are classical statistical tools for count data (Cameron & Trivedi, 2013; Tu & Piegorsch, 2003). Their relevance to thin airline demand is direct. A network contains many related but nonidentical markets. Complete pooling assumes all markets share one rate, while no pooling estimates a separate rate for every market. A hierarchical model lies between these extremes: sparse markets borrow more information from the network, and mature markets rely primarily on their own histories.

The statistical ingredients are established; the contribution is their assembly around a thin-market revenue-management use case. HEPG occupies a middle ground between complete pooling and route-by-route estimation. It borrows more information for new or sparse cells and less for mature cells, and it produces predictive survival probabilities that can be inserted directly into a Littlewood-style marginal seat-protection comparison (Littlewood 2005).

PROBLEM FORMULATION

Let $i = 1, \dots, I$ index forecasting cells. A cell can be a route, leg, fare product, point-of-sale segment, or another stable market definition. Let $t = 1, \dots, T$ index time periods. The latent variable of interest is

$$Y_{it} = \text{number of demand requests in cell } i \text{ during period } t.$$

The period should be chosen so that low demand is meaningful at the decision granularity. Examples include daily requests for a thin market, requests in a days-to-departure bucket, or bookings for a narrow fare product. We call a cell *thin* when its expected count per modeling interval is small enough that zeros and single-request outcomes have material probability. No universal numerical threshold is required, but expected counts below one are a natural regime in which discrete modeling becomes especially important.

Exposure

Let $e_{it} \geq 0$ denote known exposure. Depending on the application, exposure can represent the number of bookable departures, the duration of an observation window, or another quantity proportional to opportunity for requests. If every observation has the same opportunity, set $e_{it} = 1$.

Exposure is not a substitute for availability logic. If no service is offered, $e_{it} = 0$ and the record should not be interpreted as evidence of zero latent demand at a positive exposure level. Likewise, if a booking class is closed, the analyst must distinguish “no observed bookings” from “no customer requests.”

Calendar and demand multipliers

Let x_{it} contain known predictors such as day-of-week indicators, seasonal harmonics, holiday flags, booking-horizon effects, or market-level covariates. Define

$$s_{it}(\beta) = \exp(x_{it}^\top \beta).$$

The conditional mean is

$$\mu_{it} = e_{it} s_{it}(\beta) \lambda_i.$$

This formulation separates a market-specific baseline intensity λ_i from multiplicative effects shared across many markets. Richer specifications can include interactions or hierarchical coefficients, but the simple decomposition preserves interpretability and conjugacy in λ_i conditional on β .

HIERARCHICAL EXPOSURE-ADJUSTED POISSON–GAMMA MODEL

Conditional count model

The core observation model is

$$Y_{it} \mid \lambda_i, \beta \sim \text{Poisson}(e_{it} s_{it}(\beta) \lambda_i).$$

For genuinely low counts, the Poisson model has three attractive properties. It has the correct support, it assigns positive probability to zero, and its likelihood naturally weights information by exposure. Its main restriction is equidispersion conditional on the latent rate:

$$\text{Var}(Y_{it} \mid \lambda_i, \beta) = \text{E}(Y_{it} \mid \lambda_i, \beta).$$

The adaptive extension below relaxes this restriction.

Network prior and partial pooling

Market intensities are modeled as

$$\lambda_i \sim \text{Gamma}(a, b),$$

with shape a and rate b . Thus

$$\text{E}(\lambda_i) = \frac{a}{b}, \quad \text{Var}(\lambda_i) = \frac{a}{b^2}.$$

For fixed β , define the total count and effective exposure

$$Y_i = \sum_{t=1}^T Y_{it}, \quad S_i = \sum_{t=1}^T e_{it} s_{it}(\beta).$$

Conjugacy gives

$$\lambda_i \mid \mathcal{D} \sim \text{Gamma}(A_i, B_i), \quad A_i = a + Y_i, \quad B_i = b + S_i.$$

The posterior mean has an illuminating shrinkage form:

$$\hat{\lambda}_{i,\text{HEPG}} = \frac{a + Y_i}{b + S_i}.$$

Equivalently,

$$\hat{\lambda}_{i,\text{HEPG}} = w_i \frac{Y_i}{S_i} + (1 - w_i) \frac{a}{b}, \quad w_i = \frac{S_i}{b + S_i}.$$

The weight w_i depends only on effective exposure. A mature route with large S_i is barely shrunk. A sparse route with small S_i is pulled toward the network distribution. Most importantly, if $Y_i = 0$, HEPG does not force the future rate to zero. The forecast remains positive unless the learned network prior itself degenerates at zero.

The shrinkage figure below illustrates this behavior.

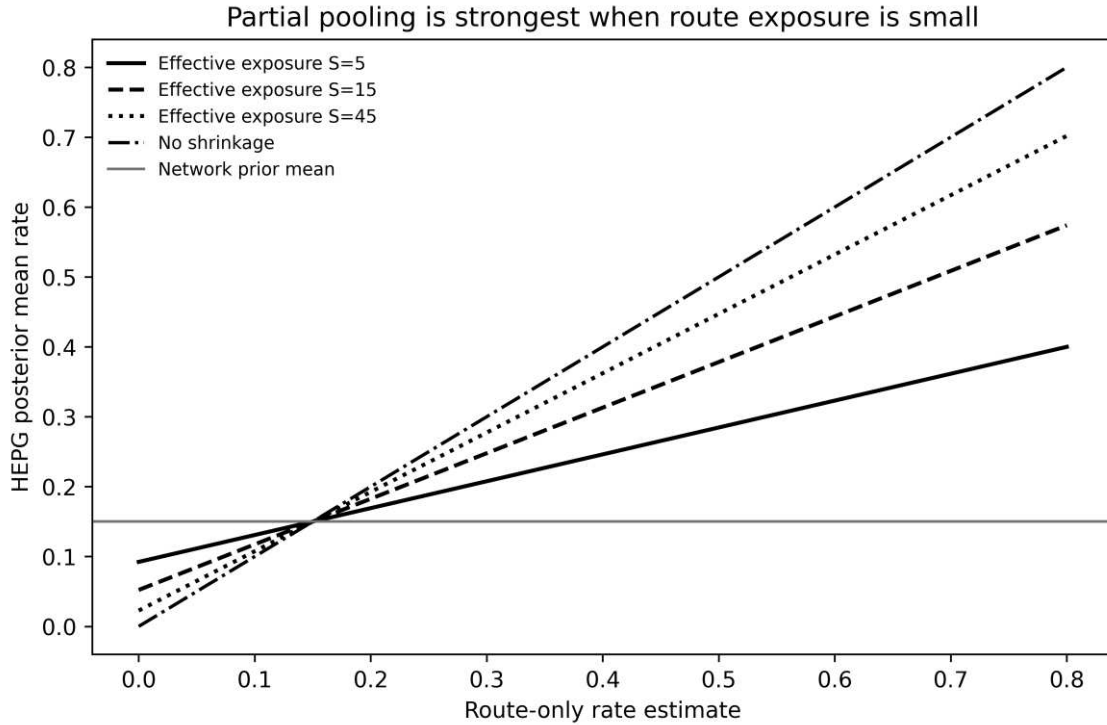


Figure 1. Partial pooling is strongest when effective route exposure is small and fades as route information accumulates.

Posterior predictive distribution

For a future period with

$$c_{i,f} = e_{i,f} s_{i,f}(\beta),$$

integrating the Poisson likelihood over the Gamma posterior gives

$$Y_{i,f} \mid \mathcal{D} \sim \text{NegBin}\left(A_i, p_i = \frac{B_i}{B_i + c_{i,f}}\right),$$

using the convention in which the negative-binomial random variable counts events before A_i successes.

The predictive mean and variance are

$$m_{i,f} = c_{i,f} \frac{A_i}{B_i},$$

and

$$v_{i,f} = m_{i,f} + \frac{m_{i,f}^2}{A_i}.$$

The second term in the variance is parameter uncertainty. When A_i is small, uncertainty about the sparse route rate materially widens the predictive distribution. As the route accumulates history, A_i grows and the predictive variance approaches the plug-in Poisson variance.

This is a central practical advantage of the hierarchy. It retains a Poisson arrival interpretation but does not pretend that an imprecisely estimated rate is known exactly.

Empirical-Bayes estimation

The hyperparameters (a, b) are learned from the network rather than fixed arbitrarily. Conditional on β , integrating out λ_i gives the route-level marginal likelihood

$$p(\mathbf{y}_i \mid a, b, \beta) = \frac{b^a \Gamma(a + Y_i)}{\Gamma(a)(b + S_i)^{a+Y_i}} \prod_{t=1}^T \frac{[e_{it} s_{it}(\beta)]^{Y_{it}}}{Y_{it}!}.$$

The empirical-Bayes estimator maximizes the product of the marginal likelihood above over markets.

Two implementation strategies are useful. First, one may jointly optimize (a, b, β) . Second, for operational simplicity, one may estimate pooled calendar effects first, compute S_i , and then estimate (a, b) by marginal likelihood. The simulation study uses the second strategy. In production, joint optimization or alternating updates are preferable when exposure and route intensity are correlated.

ADAPTIVE EXTRA-POISSON VARIANCE CORRECTION

A Poisson model can be too narrow when demand contains multiplicative shocks, group bookings, unobserved event effects, or other sources of heterogeneity. The usual negative-binomial variance function is

$$\text{Var}(Y \mid \mu) = \mu + \frac{\mu^2}{k},$$

where smaller k means stronger overdispersion.

We use the HEPG posterior mean as the stable route-intensity estimate and diagnose remaining extra-Poisson variation with the moment estimator

$$\hat{\theta} = \max\left\{0, \frac{\sum_{it} [(Y_{it} - \hat{\mu}_{it})^2 - \hat{\mu}_{it}]}{\sum_{it} \hat{\mu}_{it}^2}\right\}, \quad \hat{k} = \hat{\theta}^{-1}.$$

When $\hat{\theta}$ is essentially zero, $\hat{k}$ is treated as very large and the model reduces to HEPG.

For a variance-robust predictive approximation, suppose a multiplicative shock U has

$$U \sim \text{Gamma}(k, k), \quad E(U) = 1, \quad \text{Var}(U) = 1/k.$$

Combining U with posterior route uncertainty yields a predictive mean $m_{i,f}$ and an overdispersed variance. A negative-binomial distribution can be moment-matched to these two moments using effective size

$$r_{i,f} = \frac{A_i k}{k + A_i + 1}.$$

We call this predictive correction *adaptive HEPG* (A-HEPG). It should be interpreted as a robust moment-matched predictive approximation, not as an exact conjugate posterior for a negative-binomial observation model.

WHEN ZERO INFLATION OR HURDLE MODELS ARE PREFERABLE

The Poisson distribution naturally generates zeros. Therefore a large number of zeros is not, by itself, evidence that a zero-inflated model is needed. The appropriate diagnostic is whether the observed zero frequency materially exceeds the zero probability implied by a well-fitted count model after known no-service and unavailable-inventory periods have been removed.

If excess zeros remain, a two-process model may be more appropriate:

$$P(Y_{it} = 0) = \pi_{it} + (1 - \pi_{it})f_{\text{count}}(0),$$

where π_{it} is a structural-zero probability. A hurdle alternative models $P(Y_{it} > 0)$ separately and uses a zero-truncated count distribution for positive outcomes. These ideas are well developed in the intermittent-demand literature (Aktekin et al., 2026; Jiang, 2020; Maiti et al., 2015).

For airline applications, however, “structural zero” must be interpreted carefully. A canceled or nonexistent service should typically be represented as zero exposure, not as a stochastic structural-demand state. Likewise, a fare product that is closed by inventory control produces censored observation, not necessarily structural zero demand.

CENSORING AND UNCONSTRAINING

The core model assumes that Y_{it} represents latent requests or an observation that equals latent demand. When accepted bookings are constrained, the likelihood must reflect censoring or the data must first be unconstrained.

For a simple right-censoring threshold C_{it} , the contribution to the likelihood is

$$L_{it} = \begin{cases} P(Y_{it} = y_{it} \mid \mu_{it}), & \text{if uncensored,} \\ P(Y_{it} \geq C_{it} \mid \mu_{it}), & \text{if censored.} \end{cases}$$

The same principle applies to negative-binomial or hurdle extensions. In a production airline setting, demand preprocessing and availability reconstruction can be as important as the choice of forecast distribution. The simulation study below intentionally removes this complication to isolate the sparse-count estimation problem.

SIMULATION STUDY

Objectives

The simulation has three goals. First, it tests whether partial pooling improves probabilistic forecasts when true demand is sparse and heterogeneous. Second, it evaluates the specific failure mode of independent route estimation when short histories contain no requests. Third, it tests whether the adaptive variance correction is useful when the conditional Poisson assumption is violated.

All simulations are synthetic. No proprietary airline data are used.

Data-generating process

Each replication contains $I = 400$ markets. Baseline route intensities are drawn from

$$\lambda_i \sim \text{Gamma}\left(1.2, \frac{1.2}{\bar{\lambda}}\right),$$

so that $E(\lambda_i) = \bar{\lambda}$. The Poisson experiments use

$$\bar{\lambda} \in \{0.08, 0.15, 0.30\}$$

requests per exposure unit.

Daily exposure starts from a route-level frequency of one, two, or three units with probabilities 0.60, 0.30, and 0.10, respectively, plus small random day-to-day variation. Weekly seasonality follows

$$s_t \propto \exp[0.25\sin(2\pi t/7) - 0.10\cos(2\pi t/7)],$$

normalized to have weekly mean one.

For the correctly specified experiments,

$$Y_{it} \sim \text{Poisson}(e_{it}s_t\lambda_i).$$

Each replication uses 14 training days and 28 test days, deliberately creating a short-history environment.

The robustness experiments use negative-binomial observations with mean $e_{it}s_t\lambda_i$ and size

$$k \in \{5, 2, 0.8\},$$

with smaller k representing stronger extra-Poisson variation. These experiments use 28 training days and 28 test days.

Thirty independent replications are run for every scenario. Pooled weekly seasonality is estimated from the training sample in each replication.

Competing methods

Five forecasting methods are compared:

1. **Pooled Poisson**: one network-wide rate shared by all markets.
2. **Route Poisson MLE**: a separate maximum-likelihood rate for each route. For numerical scoring, zero estimates are clipped at a negligible positive value.
3. **Weak-prior Poisson–Gamma**: independent route models with a fixed Gamma(0.5,0.5) prior, illustrating generic smoothing without network-learned hyperparameters.
4. **HEPG**: the proposed empirical-Bayes Poisson–Gamma model.
5. **A-HEPG**: HEPG with the extra-Poisson variance correction described above.

Evaluation metrics

The primary metric is average negative log predictive probability,

$$\text{NLL} = -\frac{1}{N_{\text{test}}} \sum \log p(Y_{it}^{\text{test}} | \mathcal{D}),$$

a proper scoring rule for the full predictive distribution. Lower is better.

We also report mean absolute error (MAE), the empirical coverage of central 90% predictive intervals, and the Brier score for the zero event. Because very low-count distributions are discrete, nominal equal-tailed intervals are expected to be conservative; coverage should therefore be interpreted together with interval width and proper scores.

RESULTS

Sparse Poisson demand

The table below reports results under the correctly specified Poisson data-generating process.

Table 1. Synthetic sparse-Poisson forecasting results. Lower NLL, MAE, and Brier scores are better; coverage refers to central 90% predictive intervals.

Mean rate	Method	NLL	MAE	Coverage	Zero Brier
0.08	Pooled Poisson	0.381	0.215	0.987	0.095
	Route Poisson MLE	0.608	0.193	0.975	0.092
	Weak-prior PG	0.362	0.218	0.979	0.091
	HEPG	0.355	0.201	0.981	0.090
	A-HEPG	0.355	0.201	0.981	0.090
0.15	Pooled Poisson	0.585	0.362	0.973	0.146
	Route Poisson MLE	0.731	0.304	0.972	0.133
	Weak-prior PG	0.530	0.324	0.977	0.132
	HEPG	0.525	0.316	0.981	0.131
	A-HEPG	0.525	0.316	0.981	0.131
0.30	Pooled Poisson	0.909	0.597	0.955	0.207
	Route Poisson MLE	0.899	0.470	0.970	0.170
	Weak-prior PG	0.774	0.483	0.975	0.169

Mean rate	Method	NLL	MAE	Coverage	Zero Brier
	HEPG	0.770	0.484	0.978	0.169
	A-HEPG	0.770	0.484	0.978	0.169

At $\bar{\lambda} = 0.15$, HEPG reduces average NLL from 0.585 for pooled Poisson to 0.525, a reduction of about 10%. Relative to route-level Poisson MLE, the reduction is about 28%. At $\bar{\lambda} = 0.30$, the HEPG NLL is 0.770 versus 0.909 for pooled Poisson and 0.899 for route-level MLE.

The route-level MLE has slightly lower MAE in these experiments. This is an important limitation and not a contradiction. The route MLE is a sharp point estimate and, under absolute error, can occasionally outperform a posterior mean even while assigning implausibly small probability to unseen positive counts after a zero history. NLL penalizes such overconfidence. The principal HEPG benefit is therefore probabilistic rather than universal point-error dominance.

The weak-prior Poisson–Gamma model already removes the catastrophic zero-rate problem. HEPG improves modestly over that baseline because it learns the amount and center of shrinkage from the network instead of fixing the prior arbitrarily.

The following figure shows how the probabilistic advantage changes with the average rate.

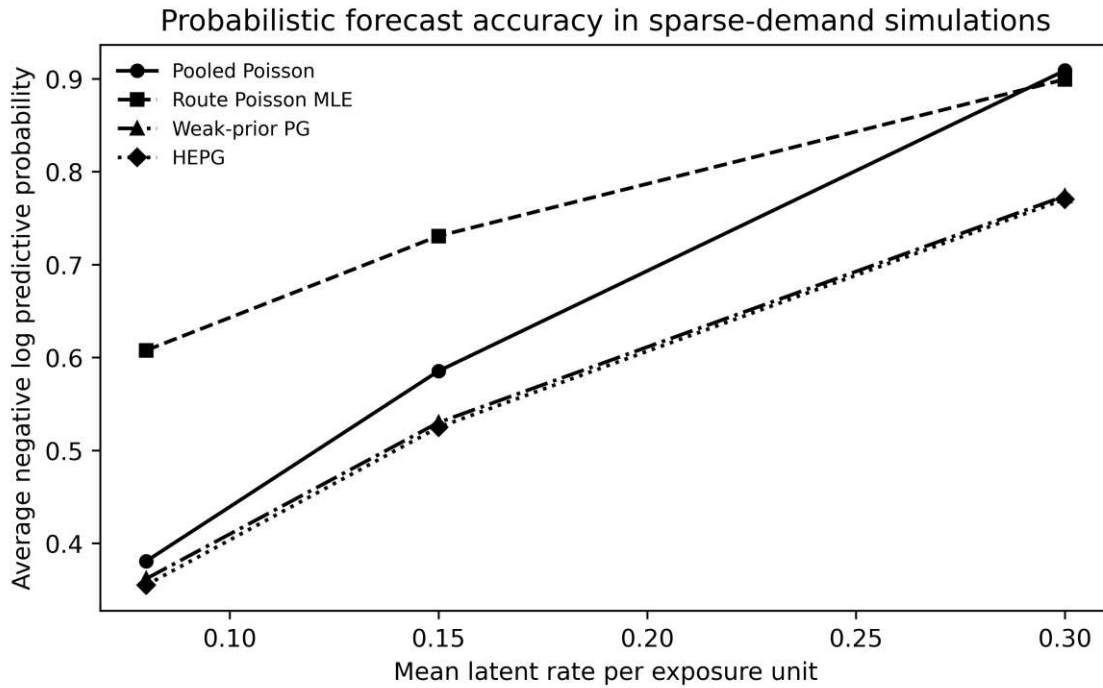


Figure 2. Average negative log predictive probability in sparse-demand simulations. Lower is better.

Robustness to overdispersion

The table below reports the negative-binomial robustness experiments.

Table 2. Synthetic overdispersed forecasting results at mean rate 0.15. Smaller true k indicates stronger overdispersion.

True k	Method	NLL	MAE	Coverage	Zero Brier
5.0	Pooled Poisson	0.590	0.366	0.972	0.144

True k	Method	NLL	MAE	Coverage	Zero Brier
2.0	Route Poisson MLE	0.586	0.311	0.973	0.129
	Weak-prior PG	0.524	0.321	0.978	0.128
	HEPG	0.521	0.318	0.978	0.128
	A-HEPG	0.520	0.318	0.979	0.128
	Pooled Poisson	0.597	0.369	0.969	0.140
0.8	Route Poisson MLE	0.598	0.318	0.970	0.127
	Weak-prior PG	0.532	0.329	0.975	0.127
	HEPG	0.529	0.325	0.975	0.126
	A-HEPG	0.524	0.325	0.978	0.126
	Pooled Poisson	0.624	0.381	0.964	0.134
0.4	Route Poisson MLE	0.627	0.337	0.963	0.125
	Weak-prior PG	0.554	0.348	0.968	0.125
	HEPG	0.552	0.342	0.968	0.124
	A-HEPG	0.529	0.342	0.974	0.121
	Pooled Poisson	0.624	0.381	0.964	0.134

When overdispersion is mild ($k = 5$), A-HEPG and HEPG are nearly identical. At $k = 2$, A-HEPG lowers NLL from 0.529 to 0.524. Under strong overdispersion ($k = 0.8$), the improvement is larger: NLL falls from 0.552 to 0.529. This pattern is desirable. The adaptive correction changes little when the Poisson approximation is adequate and matters more when the variance materially exceeds the mean.

Interpretation of interval coverage

The central 90% predictive intervals cover more than 90% of outcomes in these very sparse experiments. The conservatism is largely a consequence of discreteness: when most predictive mass lies on 0 and 1, an equal-tailed integer interval cannot achieve every nominal coverage level exactly. For operational use, proper distributional scores and decision-specific quantiles are more informative than demanding exact nominal coverage from a discrete interval.

APPLICATION: CABIN FLOATING AND PREMIUM-CABIN CAPACITY SHARING

A central application occurs when paid premium demand is weak while economy demand is strong. The airline can accept additional economy bookings against forecast-unused premium seats and later accommodate those passengers through complimentary upgrades. Practitioners often describe this as cabin floating; the academic literature generally calls it premium-cabin capacity sharing (Escovar-Álvarez and Belobaba 2022). The decision is economically asymmetric: an empty premium seat wastes capacity, but a seat released too aggressively may displace a late high-fare premium customer.

This is precisely where a distributional forecast is more useful than a mean alone. The optimal release depends on the probability that each additional protected premium seat will be required. HEPG supplies these survival probabilities while incorporating uncertainty about a sparse market's underlying rate.

A one-period seat-protection model

Let C_p denote the number of premium seats available for the remaining booking horizon. Let D_p denote remaining paid premium demand with predictive distribution obtained from HEPG or A-HEPG. Suppose economy demand is sufficiently high that every floated seat can be sold. Let p_E be the net revenue from

accepting one additional economy booking and p_P be the net revenue from one premium booking. If a complimentary upgrade carries incremental variable cost c_U , replace p_E below by $g_E = p_E - c_U$.

Choose $y \in \{0, \dots, C_P\}$ premium seats to protect for paid premium demand. Then

$$x = C_P - y$$

seats are floated to economy. Ignoring revenue from the fixed economy cabin because it is common to all choices, the expected relevant revenue is

$$V(y) = p_P E[\min(D_P, y)] + g_E(C_P - y).$$

The marginal value of protecting the y th premium seat rather than floating it is

$$V(y) - V(y - 1) = p_P P(D_P \geq y) - g_E.$$

Therefore the optimal protection level is

$$y^* = \max \left\{ y \leq C_P : P(D_P \geq y) \geq \frac{g_E}{p_P} \right\}, \quad x^* = C_P - y^*.$$

The cabin-floating rule above provides a direct answer to the price-difference question. A larger premium fare relative to the economy fare lowers the threshold g_E/p_P , causing more premium seats to be protected and fewer seats to be floated. When the price gap is small, the opportunity cost of using a premium seat for an economy passenger is also small, so the policy becomes more aggressive.

Under HEPG, D_P is negative binomial from the HEPG predictive distribution. Under a plug-in Poisson forecast, the same rule holds with a Poisson survival probability. The decision framework is therefore distribution-agnostic once a coherent predictive distribution is available.

Price-difference cases

For a transparent price-gap illustration, consider a synthetic flight with 12 premium seats available to the remaining booking horizon and abundant excess economy demand. Let the economy contribution be \$200 and assume no incremental upgrade cost. The HEPG predictive distribution has mean remaining paid premium demand of 4 and posterior shape $A=3$, which implies variance greater than the mean because the rate itself is uncertain. Table 3 applies the marginal protection rule across five premium fares.

$$4 + \frac{4^2}{3} = 9.33.$$

The wider-than-Poisson variance is intentional: with little route history, forecast uncertainty should influence protection.

The solution becomes more conservative as the premium/economy fare ratio rises.

Table 3. Illustrative cabin-floating decisions for different premium–economy fare gaps. Economy contribution is \$200, premium capacity is 12 seats, and predictive mean paid premium demand is 4.

Premium fare	Fare gap	Fare ratio	Protect premium	Float to economy	Expected gain vs. no sharing
\$220	\$20	1.10	1	11	\$1,532
\$300	\$100	1.50	2	10	\$1,326
\$400	\$200	2.00	3	9	\$1,154
\$600	\$400	3.00	5	7	\$937

Premium fare	Fare gap	Fare ratio	Protect premium	Float to economy	Expected gain vs. no sharing
\$1,000	\$800	5.00	6	6	\$688

When premium is only 10% above economy, the rule protects one premium seat and floats eleven. At a three-to-one fare ratio, five seats are protected; at five-to-one, six are protected. The increasing protection level is driven by the higher opportunity cost of displacing premium demand, not by a change in the demand forecast.

The final column is a synthetic contribution comparison within the 12-seat premium cabin only; it is not total flight revenue. It excludes the fixed economy cabin, loyalty effects, service costs other than the stated upgrade cost, reaccommodation, and long-run customer-experience effects.

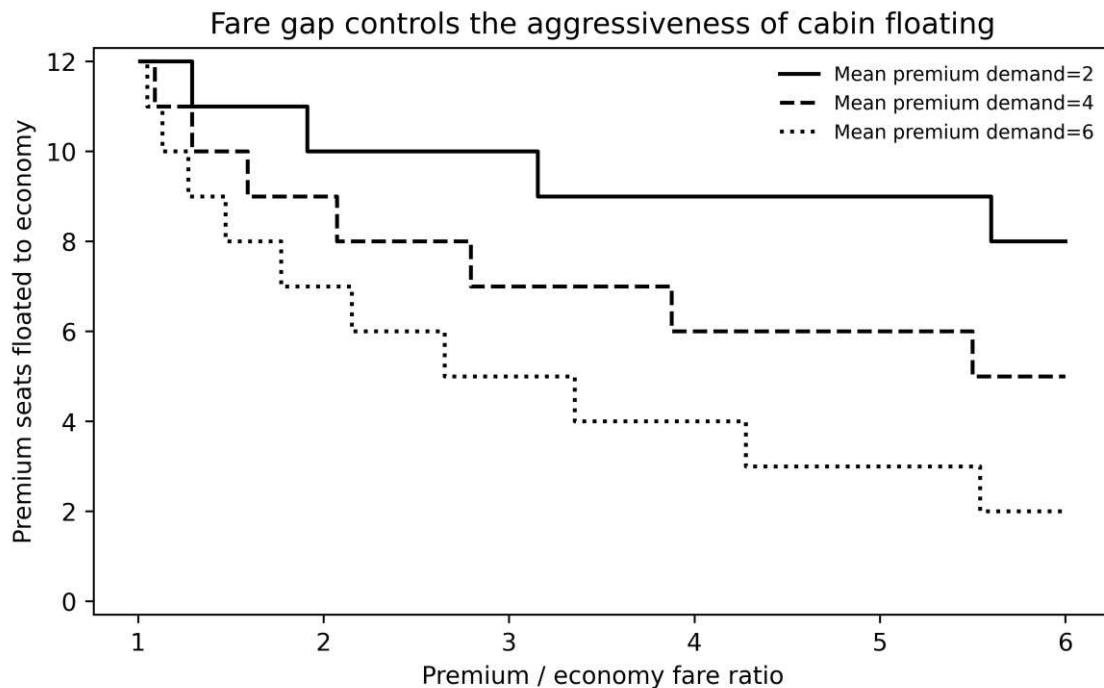


Figure 3. Optimal premium seats floated to economy as the premium/economy fare ratio changes. Higher paid-premium demand or a wider fare gap leads to more protection.

Decision simulation: from forecast uncertainty to revenue

To test whether the forecast differences matter economically, we simulate 100 independent replications of 600 premium-demand markets. Market intensities are drawn from a Gamma distribution with shape 1.2 and mean 0.4 requests per exposure unit. Each market receives between 2 and 14 training exposure units; remaining paid premium demand is then generated over 8 future exposure units. Premium capacity is 12 seats and economy contribution is \$200. The experiment uses abundant excess economy demand so that the only decision risk is displacing paid premium demand. This isolates the value of the premium-demand predictive distribution.

For each synthetic market, pooled Poisson, route-level Poisson MLE, a weak-prior Poisson–Gamma model, HEPG, and a mean-only HEPG rule choose a protection level. An oracle uses the true latent route rate. Relevant revenue is premium fare multiplied by paid premium customers accommodated in protected seats plus economy contribution from seats floated into economy. Revenue regret is the

difference between the oracle's relevant revenue and the revenue produced by each forecast-based rule. It is a decision-quality metric for the shared premium cabin, not a claim about total network revenue.

Table 4. Cabin-floating decision simulation. Values are averages across 100 replications of 600 synthetic markets; revenue and regret refer only to the 12-seat shared premium-cabin decision.

Fare ratio	HEPG protect / float	HEPG regret	Route-MLE regret	Pooled regret	Mean-only regret
1.1	0.7 / 11.3	\$9.0	\$23.7	\$43.6	\$146.9
1.5	1.9 / 10.1	\$50.2	\$73.3	\$141.6	\$101.0
2.0	2.7 / 9.3	\$88.6	\$117.2	\$240.9	\$97.8
3.0	3.7 / 8.3	\$144.7	\$199.9	\$385.6	\$158.8
5.0	4.8 / 7.2	\$212.6	\$353.8	\$594.4	\$370.1

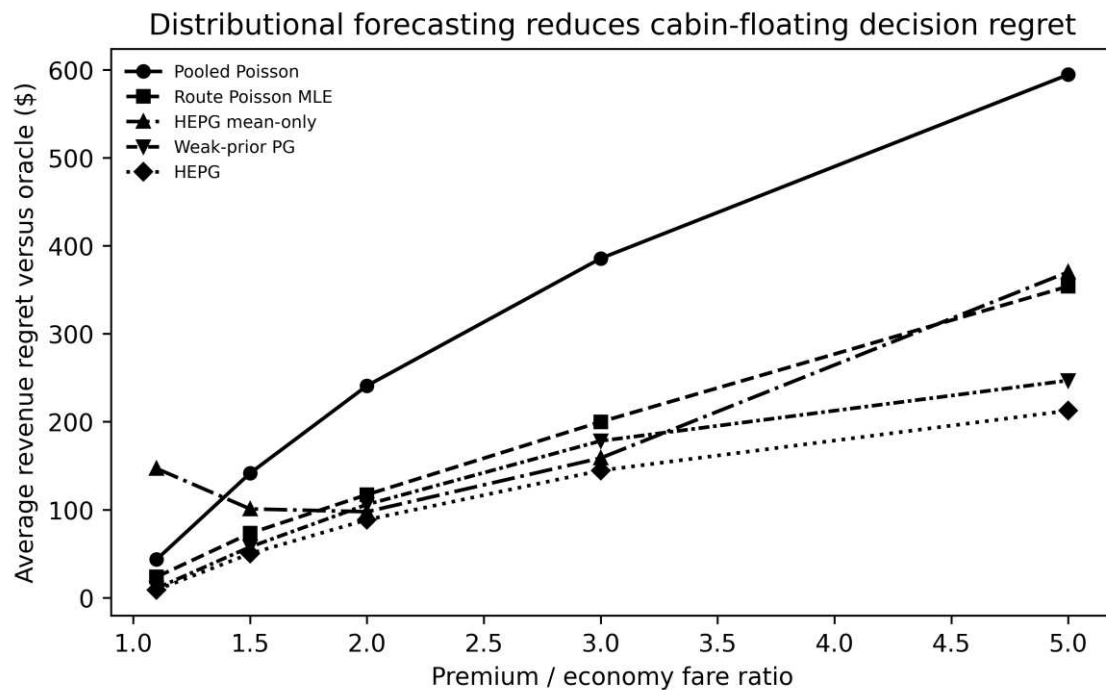


Figure 4. Distributional forecasting reduces cabin-floating revenue regret relative to pooled, route-level, and mean-only rules as the premium/economy fare ratio changes.

HEPG has the lowest regret among the implementable forecast rules at every tested fare ratio. At a 1.1 premium/economy ratio, HEPG regret averages \$9.0 compared with \$23.7 for route-level Poisson MLE and \$43.6 for pooled Poisson. At a five-to-one ratio, the corresponding regrets are \$212.6, \$353.8, and \$594.4. Thus, the dollar consequence of forecast misspecification grows with the fare gap. The mean-only HEPG rule is also inferior to the full-distribution rule in every scenario, illustrating why the upper tail of remaining premium demand matters for protection decisions.

These simulations do not establish a production revenue lift. They show a mechanism: when premium demand is sparse, a forecast that stabilizes route rates and preserves parameter uncertainty can make better capacity-sharing decisions than either complete pooling or a brittle route-only plug-in estimate. Real revenue impact must be measured with airline data and realistic economy spill, choice, cancellations, no-shows, and network displacement.

Uncertain excess economy demand

If excess economy demand is not guaranteed, let $D_{E,\text{excess}}$ denote demand above the physical economy capacity. The marginal value of the k th floated seat becomes

$$g_E P(D_{E,\text{excess}} \geq k).$$

The protection comparison is then

$$p_P P(D_P \geq y) \quad \text{versus} \quad g_E P(D_{E,\text{excess}} \geq k).$$

Thus the same sparse-count framework can forecast both sides of the trade: low premium demand and uncertain overflow economy demand. The simple rule in the cabin-floating rule above is recovered when excess economy demand is effectively certain.

Dynamic interpretation

The rule is most naturally reevaluated as departure approaches. Early in the booking horizon, uncertainty about high-value premium demand is larger and premium seats should generally be protected more conservatively. Late in the horizon, if the HEPG posterior continues to indicate weak premium demand while economy demand remains strong, more capacity can be released. This timing logic is consistent with prior premium-cabin capacity-sharing research, which finds that aggressive sharing too early can displace premium demand (Escovar-Álvarez & Belobaba, 2022).

A production implementation would update D_P whenever meaningful bookings, cancellations, schedule changes, fare changes, or availability changes occur. The result is a dynamic protection level rather than a fixed number of floated seats for the entire booking horizon.

PRACTICAL IMPLEMENTATION FOR REVENUE-MANAGEMENT TEAMS

Why the model is operationally useful

A route-level Poisson maximum-likelihood estimate collapses to zero after an all-zero training history. In a cabin-floating context that can release nearly all premium capacity even though the absence of historical requests may simply reflect a short observation window. HEPG instead interprets a zero history as evidence for a low rate while retaining a positive posterior rate learned from related markets. As exposure accumulates, the shrinkage automatically weakens.

This behavior is particularly relevant for new routes, low-frequency flights, premium fare products, and late booking-horizon buckets. It also gives analysts a transparent “how much did we borrow?” diagnostic through the shrinkage weight.

$$E(\lambda_i \mid Y_i = 0) = \frac{a}{b + S_i},$$

which decreases with zero-exposure history but remains positive. This behavior is natural for newly launched routes, infrequent services, thin fare products, or narrowly defined booking buckets.

Recommended production outputs

A production service should expose the full decision inputs, not just a point forecast. For each cell, useful outputs include the posterior mean; $P(D=0)$ and $P(D \geq 1)$; decision-relevant survival probabilities $P(D \geq y)$;

predictive quantiles; effective exposure; the shrinkage weight; an overdispersion diagnostic; and flags indicating closed inventory, no service, or censoring. For cabin floating, the system can additionally return the recommended protected and floated seats for the current fare ratio.

Operational guardrails and model segmentation

One model should not be forced across the network. HEPG is appropriate for genuinely sparse count cells with no strong residual overdispersion; A-HEPG or a fitted negative-binomial hierarchy is preferable when extra-Poisson variance persists. Hurdle or zero-inflated models should be considered only after known zero-exposure and closed-availability observations are handled. High-volume cells should remain benchmarked against mature airline forecasting methods. Cabin sharing should also be bounded by operational guardrails such as upgrade eligibility, blocked seats, accessibility requirements, through-customer constraints, and any cabin-service or product commitments.

Deployment workflow

1. Define the exact decision cell and horizon: flight/market, premium product, days-to-departure interval, and exposure.
2. Reconstruct availability and censoring before treating observed bookings as demand; no-service periods are zero exposure, not zero demand.
3. Fit pooled calendar effects and the empirical-Bayes route prior on a stable peer group; monitor shrinkage and posterior calibration.
4. Produce a predictive distribution for remaining paid premium demand and, where needed, excess economy demand.
5. Convert the distribution to a protection level using the current net economy contribution, premium fare, and operational upgrade cost.
6. Recompute late in the booking horizon as new bookings, cancellations, schedule changes, and availability changes arrive; Escovar-Álvarez and Belobaba (2022) provide evidence that sharing too early can be harmful.
7. Monitor both statistical scores and business outcomes: premium spill, unused premium seats, upgrade counts, relevant cabin revenue, and customer-experience indicators.

EMPIRICAL VALIDATION AND LIMITATIONS

The present study is deliberately simulation-based. The numerical results demonstrate internal behavior and decision mechanisms; they do not establish performance on a production airline revenue-management system. A journal extension using airline booking-request data should use rolling-origin evaluation and retain availability/censoring information. The short-data setting documented by Lee and Abdelghany (2026) provides a natural empirical benchmark, but transported-passenger totals are not a substitute for unconstrained micro-level booking demand when the intended decision is premium-seat protection.

Public traffic data can illustrate route heterogeneity and short histories, but realized passengers are not the same quantity as latent booking requests. For the cabin-floating application, the strongest validation would use anonymized or synthetic snapshots containing request counts, availability status, days to departure, premium and economy fares, cabin capacity, cancellations, and upgrade outcomes. A chronological holdout should compare both probabilistic forecast scores and realized decision regret.

Limitations

Several limitations remain. First, the simulations use stylized market heterogeneity and do not model network displacement, customer choice, price endogeneity, cancellations, no-shows, or loyalty effects. Second, A-HEPG is a moment-matched variance correction rather than a fully Bayesian negative-binomial hierarchy. Third, the global Gamma prior may be too coarse for a heterogeneous international network; hierarchical peer groups by haul length, hub status, region, or premium-demand profile may be preferable. Fourth, cabin floating can affect customer experience and operational processes in ways that a one-period revenue expression does not capture.

Extensions

Natural extensions include time-varying route intensities, peer-group or multilevel priors, explicit booking-curve effects, price and competitor covariates, a joint forecast for economy overflow, and a multi-period dynamic capacity-sharing policy. A particularly useful next step is to estimate the value of distributional forecasting under realistic network displacement and customer-choice behavior rather than a single-flight abundant-economy assumption.

CONCLUSION

Very low airline demand is a sparse count problem in which route-level estimation uncertainty can be as important as the expected demand itself. HEPG combines a Poisson arrival model with empirical-Bayes partial pooling and produces a negative-binomial posterior predictive distribution. In synthetic sparse-demand experiments, it improves probabilistic forecast scores relative to complete pooling and route-by-route Poisson estimation.

The revenue-management value emerges when that distribution is connected to a decision. A cabin-floating rule protects the y th premium seat when its expected marginal premium value exceeds the contribution from releasing that seat to economy. Because the rule depends on $P(D_P \geq y)$, parameter uncertainty and the premium/economy fare ratio directly affect the release decision. In the synthetic decision experiment, HEPG produces lower revenue regret than pooled, route-MLE, or mean-only rules across every tested fare ratio.

The managerial implication is not “always use Poisson.” It is to match the distribution to the demand regime, borrow information when individual histories are weak, preserve uncertainty through the decision layer, and release premium capacity only when the expected economy contribution justifies the risk of premium displacement. That logic is simple enough to audit and re-optimize as departure approaches, while remaining extensible to richer airline revenue-management systems.

DATA AND CODE AVAILABILITY

All numerical results are generated from synthetic data with fixed random seeds. No proprietary airline data or confidential business rules are used. Reproducibility code and simulation summaries accompany the submission package and can be deposited in a public repository with the preprint or accepted manuscript.

Clinical trial number: not applicable.

SUPPLEMENTARY TECHNICAL NOTES

Given the posterior

$$\lambda_i \mid \mathcal{D} \sim \text{Gamma}(A_i, B_i)$$

and a future count with exposure multiplier $c_{i,f}$,

$$Y_{i,f} \mid \lambda_i \sim \text{Poisson}(c_{i,f}\lambda_i),$$

integrating over the posterior uncertainty in λ_i gives

$$P(Y_{i,f} = y \mid \mathcal{D}) = \frac{\Gamma(A_i + y)}{\Gamma(A_i) y!} \left(\frac{B_i}{B_i + c_{i,f}} \right)^{A_i} \left(\frac{c_{i,f}}{B_i + c_{i,f}} \right)^y.$$

This is the negative-binomial predictive distribution used by HEPG.

Moment match for A-HEPG

Let posterior route intensity be

$$\lambda \sim \text{Gamma}(A, B)$$

and let independent multiplicative shock

$$U \sim \text{Gamma}(k, k).$$

For

$$Y \mid \lambda, U \sim \text{Poisson}(c\lambda U),$$

the mean is

$$m = \frac{cA}{B}.$$

Using independence,

$$E(\lambda^2) = \frac{A(A+1)}{B^2}, \quad E(U^2) = \frac{k+1}{k}.$$

Therefore

$$\text{Var}(Y) = m + \frac{c^2 A(k + A + 1)}{B^2 k}.$$

A negative-binomial distribution with mean m and size r has variance $m + m^2/r$. Equating the extra-variance terms gives

$$r = \frac{Ak}{k + A + 1}.$$

Reproducibility notes

The supplied Python scripts use fixed random seeds. The forecasting study uses 400 synthetic markets and 30 replications per scenario. The cabin-floating decision study uses 600 synthetic markets and 100 replications per fare-ratio scenario. CSV summaries included with the submission package correspond to the results reported in Tables 1–4 and Figures 2–4.

REFERENCES

- Aktekin, Tevfik, Refik Soyer, and Di Zhang. 2026. Bayesian forecasting of zero-inflated time-series of counts. *International Journal of Forecasting* 42 (3): 853–871. <https://doi.org/10.1016/j.ijforecast.2025.12.001>.
- Cameron, A. Colin, and Pravin K. Trivedi. 2013. *Regression Analysis of Count Data*, 2nd ed. Cambridge: Cambridge University Press.
- Croston, J. D. 1972. Forecasting and stock control for intermittent demands. *Operational Research Quarterly* 23 (3): 289–303. <https://doi.org/10.1057/jors.1972.50>.
- Escovar-Álvarez, Germán, and Peter P. Belobaba. 2022. Premium cabin capacity sharing strategies: airline RM insights. *Journal of Revenue and Pricing Management* 21 (1): 3–16. <https://doi.org/10.1057/s41272-021-00326-x>.
- Feng, Zhichao, Ang Li, Darius Walczak, and Metin Çakanyıldırım. 2024. Dynamic and flexible capacity management for flight compartments and beyond. *Production and Operations Management* 33 (12): 2401–2420. <https://doi.org/10.1177/10591478241270123>.
- Jayaram, Aparna, R. K. Amit, Amit Agarwal, and Xiaodong Luo. 2024. Elasticity-integrated pricing and allocation heuristic for airline revenue management. *Journal of Revenue and Pricing Management* 23: 305–317. <https://doi.org/10.1057/s41272-023-00454-6>.
- Jiang, Peng. 2020. A new approach to forecasting intermittent demand based on the mixed zero-truncated Poisson model. *Journal of Forecasting* 39: 1043–1060. <https://doi.org/10.1002/for.2614>.
- Korkmaz, Selçuk, O. Erhun Kundakcioglu, and Orhan Sivrikaya. 2022. A fluid approximation for the single-leg fare allocation problem with nonhomogeneous Poisson demand. *Journal of Revenue and Pricing Management* 21 (1): 81–96. <https://doi.org/10.1057/s41272-021-00284-4>.
- Lee, Kiljae K., and Ahmed F. Abdelghany. 2026. Beyond route-specific forecasting: An empirical test of two cross-series transfer learning strategies for airline demand with short-data constraints. *Journal of Air Transport Management* 133: 102981. <https://doi.org/10.1016/j.jairtraman.2026.102981>.
- Littlewood, Ken. 2005. Forecasting and control of passenger bookings. *Journal of Revenue and Pricing Management* 4 (2): 111–123. <https://doi.org/10.1057/palgrave.rpm.5170134>.
- Long, Yanbin, and Peter Belobaba. 2024. Airline revenue management with segmented continuous pricing: methods and competitive effects. *Journal of Revenue and Pricing Management* 23: 14–27. <https://doi.org/10.1057/s41272-023-00462-6>.
- Maiti, Ranjan, Atanu Biswas, and S. Das. 2015. Time series of zero-inflated counts and their coherent forecasting. *Journal of Forecasting* 34: 694–707. <https://doi.org/10.1002/for.2368>.
- Meyer, Felix, Göran Kauermann, Christopher Alder, and Catherine Cleophas. 2025. Modeling price-sensitive demand in turbulent times: an application to continuous pricing. *Journal of Revenue and Pricing Management* 24 (2): 153–177. <https://doi.org/10.1057/s41272-024-00478-6>.
- Queenan, Carrie C., Mark Ferguson, Jon Higbie, and Rohit Kapoor. 2007. A comparison of unconstraining methods to improve revenue management systems. *Production and Operations Management* 16 (6): 729–746. <https://doi.org/10.1111/j.1937-5956.2007.tb00292.x>.
- Snyder, Ralph D., J. Keith Ord, and Adrian Beaumont. 2012. Forecasting the intermittent demand for slow-moving inventories: A modelling approach. *International Journal of Forecasting* 28 (2): 485–496. <https://doi.org/10.1016/j.ijforecast.2011.03.009>.
- Swan, William M. 2002. Airline demand distributions: passenger revenue management and spill. *Transportation Research Part E: Logistics and Transportation Review* 38 (3–4): 253–263. [https://doi.org/10.1016/S1366-5545\(02\)00009-1](https://doi.org/10.1016/S1366-5545(02)00009-1).
- Talluri, Kalyan T., and Garrett J. Van Ryzin. 2004. *The Theory and Practice of Revenue Management*. New York: Springer.
- Tu, Wanzhu, and Walter W. Piegorsch. 2003. Empirical Bayes analysis for a hierarchical Poisson generalized linear model. *Journal of Statistical Planning and Inference* 111 (1–2): 235–248. [https://doi.org/10.1016/S0378-3758\(02\)00305-1](https://doi.org/10.1016/S0378-3758(02)00305-1).
- Weatherford, Lawrence R., and Peter P. Belobaba. 2002. Revenue impacts of fare input and demand forecast accuracy in airline yield management. *Journal of the Operational Research Society* 53 (8): 811–821. <https://doi.org/10.1057/palgrave.jors.2601357>.
- Weatherford, Lawrence R., and Stefan Pölt. 2002. Better unconstraining of airline demand data in revenue management systems for improved forecast accuracy and greater revenues. *Journal of Revenue and Pricing Management* 1 (3): 234–254. <https://doi.org/10.1057/palgrave.rpm.5170027>.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [illustrativepricecases.csv](#)
- [forecastoverdispersionresults.csv](#)
- [forecastpoissonresults.csv](#)
- [decisionsummary.csv](#)
- [jrpmsimulation.py](#)
- [decisionreplications.csv](#)