Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

When weekly airfare summaries are enough: Sharp risk bounds and decision certificates from aggregated quotes

Sajad Saraygord Afshari

safshari@mtroyal.ca

Mount Royal University

Case Report

Keywords: Airfare intelligence, Partial identification, Conditional value-at-risk, Interval regret, Data aggregation, Travel procurement

Posted Date: November 11th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-11529453/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: No competing interests reported.

When weekly airfare summaries are enough: Sharp risk bounds and decision certificates from aggregated quotes

Abstract

Weekly airfare feeds preserve price extrema, means and observation counts, but discard the underlying quote histories. This paper asks which historical risk comparisons remain defensible after that compression. We construct sharp bounds on empirical conditional value-at-risk by retaining both the finite quote count and the requirement that the reported extrema were observed. A majorization argument yields explicit extremal quote vectors; an integer formulation accommodates cent-denominated fares and rounded means. The bounds support itinerary-ranking certificates and exact minimax regret for selecting one alternative. Exhaustive enumeration and 1,400 independent linear-programming solves verify the implementation, followed by 6,000 synthetic aggregation experiments. A case study uses 2,000 Altus records collected over two weeks for Alaska Airlines itineraries departing San Francisco. At the upper-half tail criterion, average identification width falls from USD 14.88 under a support-and-mean relaxation to USD 4.17. Among 182 matched economy return-trip cells, 180 have a certified minimum, versus 174 under the relaxation; three selections differ from the mean-fare rule. The findings support selective investigation of ambiguous comparisons, while establishing no forecast accuracy, realised procurement saving or industry-wide effect.

Keywords: Airfare intelligence, Partial identification, Conditional value-at-risk, Interval regret, Data aggregation, Travel procurement

1. Introduction

An airfare database can contain enough information to describe a market while still containing too little information to justify a particular decision. Weekly price intelligence illustrates this distinction. An analyst may receive

the lowest, highest and average quoted fare for an itinerary, together with the number of observations, without receiving the quotes themselves. Comparing the means is straightforward. Assessing exposure to relatively expensive quotes, deciding whether an apparent ranking could reverse, or establishing whether more detailed data could change a benchmark requires a different calculation. The managerial question is therefore one of information sufficiency: what can a compressed fare record establish, and what must remain unresolved?

This question matters because the economic interpretation of price dispersion is already demanding with detailed data. Borenstein and Rose (1994) documented substantial fare inequality and related it to competition in the United States airline industry. Gerardi and Shapiro (2009) subsequently found a negative relationship between competition and dispersion using a different empirical design. More recent work distinguishes product definitions and competitive conditions: Kim et al. (2023) show that one-way and return products can exhibit different relationships between competition and dispersion, while Zhang et al. (2025) identify a reduction in intertemporal dispersion associated with competition in European airline markets. These studies explain why a pooled price range should not be read as a self-explanatory measure of market performance. They also motivate preserving product definitions when constructing comparisons.

The distinction between dispersion across products and changes over time is equally important. Alderighi et al. (2015) examine the joint effects of remaining capacity and booking time on fares. Williams (2022) model dynamic pricing in response to demand shocks and changing willingness to pay, and evaluate its welfare consequences. Aryal et al. (2024) investigate intertemporal and intratemporal discrimination in international airline markets. Gaggero and Luttmann (2025) document changes in pricing and dispersion during the COVID-19 shock. These contributions rely on temporal, transactional or structural information that cannot be recreated by assigning an arbitrary sequence to weekly minima, means and maxima. A weekly summary cannot identify when a price changed, how long a price was available, or which passengers paid it.

Airfare prediction addresses another, related decision problem. The early consumer-side work of Etzioni et al. (2003) used repeated price observations to support purchase timing. More recently, Bülbül (2025) investigates machine-learning prediction of market-level airfares. Such work concerns generalisation from observed histories or covariates to unobserved prices. That objective is

distinct from recovering a functional of an already observed but compressed quote sample. In particular, splitting a small set of weekly aggregates into random training and test subsets does not validate a within-itinerary purchase-timing forecast. The target, information available at the decision time and independent evaluation unit would all need to be specified first.

Price intelligence also enters organisations through practical decision rules. Hortaçsu et al. (2024) show that organisational structure and internal pricing procedures affect airline pricing behaviour. Their findings reinforce the value of a method that an analyst can inspect and explain, although they do not establish the performance of the procedure developed here. A transparent benchmarking rule should distinguish a preference for less expensive historical quotes from confidence that its recommendation is unaffected by information discarded upstream. The latter property can be tested even when passenger preferences and the demand system cannot be estimated.

The methodological starting point is partial identification. Rather than select one unobserved distribution, this approach reports the values consistent with explicit information and assumptions (Manski, 2003). Classical bounds on convex expectations exploit supports and moments (A. and Hochman, 1972; Bertsimas and Popescu, 2005). Tight risk bounds are also well established (Chen et al., 2011), and distributionally robust optimisation provides general frameworks for decisions under distributional ambiguity (Delage and Ye, 2010; Wiesemann et al., 2014; Rahimian and Mehrotra, 2022; Kuhn et al., 2025). Work using mean and dispersion information demonstrates the importance of specifying exactly which distributional features are known (Postek et al., 2018). Probability-bound analysis similarly avoids committing to a fully specified distribution when only limited summaries are available (Iskandar, 2021).

These foundations make clear both the opportunity and the limits of a contribution based on aggregated fares. The present paper does not propose a new theory of majorization, conditional value-at-risk, moment problems or minimax regret. Majorization supplies the relevant extremal ordering (Marshall et al., 2011); conditional value-at-risk supplies a coherent upper-tail criterion (Rockafellar and Uryasev, 2000, 2002); and regret under partial information has a substantial literature (Aissi et al., 2009; Perakis and Roels, 2008). The application-specific omission addressed here is that a weekly fare record contains more than a support and a mean. It describes a finite empirical sample, gives its size, and states that its extrema actually occurred. A generic support-and-mean ambiguity set permits arbitrary probability masses and can

ignore this additional information. Conversely, mean or maximum imputation conceals which tail-risk values the record identifies.

The resulting research question is narrow: *how much can the count and realised extrema tighten historical fare-risk bounds, and when does that tightening change a defensible itinerary benchmark?* The literature reviewed above motivates this question but does not establish that it is the first possible application of these mathematical ideas to prices. The claimed gap is an explicit connection between this commercial aggregation schema, finite-sample extremal constructions, rounding-aware implementation and decision-level certificates. Table 1 summarises that positioning.

The paper makes three bounded contributions. First, it derives constructive, sharp bounds on the historical empirical conditional value-at-risk implied by a quote count, observed minimum, observed maximum and mean, including an integer formulation for cent-valued quotes and rounded averages. Second, it connects those bounds to certified rankings, exact regret for choosing one itinerary and a decision-specific upper bound on the benefit of recovering the missing quotes. Third, it evaluates information loss using exhaustive verification, controlled synthetic experiments and a commercial-data case study, reporting both interval tightening and the limited number of decisions that change. Section 2 develops the method; Section 3 presents verification and numerical experiments; Section 4 reports the case study; Section 5 discusses managerial implications and limitations; and Section 6 concludes.

Table 1: Position of the study relative to closely related research.

Research strand	Main information and objective	Relation to the present study
Airline dispersion: Borenstein and Rose (1994); Gerardi and Shapiro (2009); Kim et al. (2023); Zhang et al. (2025)	Product or transaction data; explain dispersion and competition.	We neither estimate competition effects nor infer prices paid.
Dynamic pricing: Alderighi et al. (2015); Williams (2022); Aryal et al. (2024); Gaggero and Luttmann (2025)	Time, capacity or structural information; explain pricing dynamics.	Our target is invariant to the order of recorded quotes.
Airfare prediction: Etzioni et al. (2003); Bülbül (2025)	Histories or market features; predict or time purchases.	We bound a historical empirical functional, without a forecasting claim.
Moment and risk bounds: Chen et al. (2011); Postek et al. (2018); Kuhn et al. (2025)	Distributional restrictions; bound risk or optimise under ambiguity.	We retain finite equal quote weights, observed extrema and mean rounding.
Majorization and interval regret: Marshall et al. (2011); Aissi et al. (2009)	Extremal ordering and uncertain costs.	We apply these established tools to auditable fare comparisons.

2. Methodology

The method separates the historical quote sample from future fares. All bounds below concern the equally weighted empirical distribution of the quotes represented by one record. No sampling model, stationarity assumption or future-price support is imposed. Future use of the resulting benchmark would require separate evidence that the recorded quotes are relevant to the intended decision period.

2.1. Information preserved by one aggregate

Let a record represent $n \geq 1$ unknown quotes $x_1, \ldots, x_n$, with observed minimum a, observed maximum b and exact mean μ. Initially, prices are real valued. The compatible set is

$$\mathcal{X}(n, a, b, \mu) = \left\{ x \in [a, b]^n : \min_j x_j = a,\ \max_j x_j = b,\ \sum_{j=1}^{n} x_j = n\mu \right\}. \quad (1)$$

For $n \geq 2$, feasibility requires $(n-1)a + b \leq n\mu \leq a + (n-1)b$. A singleton requires $a = b = \mu$. If $a = b$, all quotes are known. These are ingestion checks, not assumptions to repair silently when a record is inconsistent.

The finite count has two roles. Every quote carries probability $1/n$, and each reported extremum must receive at least that mass. A support-only model merely says that quotes lie between a and b; it need not place any mass at either endpoint. These are different information sets even when the numerical endpoints are identical. Throughout, an observation count means the number of recorded quotes specified by the feed, rather than the number of seats sold, website searches by travellers or days on which a fare could be purchased.

2.2. Historical upper-tail fare criterion

For $0 \leq \alpha < 1$, define the empirical conditional value-at-risk of the recorded fares by

$$T_\alpha(x) = \min_{\eta \in \mathbb{R}} \left\{ \eta + \frac{1}{n(1-\alpha)} \sum_{j=1}^{n} (x_j - \eta)_+ \right\}. \quad (2)$$

This is the usual conditional-value-at-risk representation (Rockafellar and Uryasev, 2000, 2002), applied to fare levels as costs. With ties or discrete

data it is an average of the upper $1-\alpha$ probability mass, including a fractional contribution from the boundary observation where necessary. It is not generally the average of the observations strictly above a sample quantile.

Write the quotes in descending order as $z_1 \geq \cdots \geq z_n$, and let $h = n(1-\alpha)$, $k = \lfloor h \rfloor$ and $\theta = h - k$. Then

$$T_\alpha(x) = \frac{\sum_{j=1}^{k} z_j + \theta z_{k+1}}{h}, \tag{3}$$

with the final term omitted if $k = n$. At $\alpha = 0$, the criterion is the sample mean. At $\alpha = 0.5$, it is the average of the more expensive half of recorded quote mass, which is the primary criterion in the analysis. At sufficiently high α, it equals the maximum. The word "risk" here refers to the chosen historical tail-cost criterion. Neither its numerical level nor α is a probability that a future buyer will pay a particular fare.

2.3. Extremal quote vectors and sharp bounds

Reserve one occurrence of a and one of b. When $n > 2$, the remaining $r = n - 2$ quotes have sum $s = n\mu - a - b$ and lie in $[a, b]$. The least dispersed residual vector places each interior quote at $c = s/r$. For the most dispersed residual vector, define $d = b - a$, excess $e = s - ra$, $q = \lfloor e/d \rfloor$ and $\delta = e - qd$. Allocate q residual quotes to b, one to $a + \delta$ if $q < r$, and the remaining residual quotes to a. If $q = r$, all residual quotes equal b. Sorting is immaterial to the risk value.

Proposition 1 (Sharp finite-summary bounds)**.** *Let x^- consist of the two reserved endpoints and the balanced residual vector, and let x^+ consist of the endpoints and the endpoint-packed residual vector. For every feasible $x \in \mathcal{X}(n, a, b, \mu)$ and every $0 \leq \alpha < 1$,*

$$L_\alpha := T_\alpha(x^-) \leq T_\alpha(x) \leq T_\alpha(x^+) =: U_\alpha. \tag{4}$$

Both bounds are attained. The same extremal vectors bound any nonnegative weighted average of these empirical conditional-value-at-risk criteria.

Proof. For fixed sum and bounds, averaging a pair of unequal residual coordinates reduces dispersion in the majorization order. Repeating such transfers yields the balanced residual vector. Conversely, moving mass from a smaller interior coordinate to a larger interior coordinate until one reaches an endpoint increases dispersion; repetition leaves at most one interior residual coordinate

and gives the packed vector. Appending the same two endpoints preserves the ordering. Empirical upper-tail averages are symmetric, convex functions of the quote vector and therefore preserve this majorization order (Marshall et al., 2011). The two constructed vectors are feasible, so their risk values are attained minima and maxima. Nonnegative weighted sums preserve the same inequalities. Appendix Appendix A gives an alternative partial-sum argument. □

The result is a specialisation of classical majorization, rather than a new general extremal principle. Its practical advantage is that the commercial summary maps directly to two explicit vectors. No fitted distribution or generic nonlinear optimiser is required. A straightforward implementation constructs and sorts at most n values; in the present feed, $n \leq 7$.

Corollary 1 (Two sources of exact identification)**.** *With an exact mean, $n \leq 3$ determines the entire unordered quote sample, and hence every T_α. Independently, for any n, $T_\alpha = b$ whenever $n(1-\alpha) \leq 1$.*

For $n = 2$ the quotes are a, b; for $n = 3$ the remaining quote is $3\mu - a - b$. The second statement follows directly from Eq. (3). Neither result recovers chronology. A low count can produce a very narrow identification interval while providing weak evidence about a population or a future fare distribution.

Figure 1 illustrates the information distinction using a synthetic record with $n = 5$, $a = 100$, $b = 500$ and $\mu = 300$ USD. The balanced vector is $(100, 300, 300, 300, 500)$ and the packed vector is $(100, 100, 300, 500, 500)$. At $\alpha = 0.5$, the sharp interval is $[380, 460]$ USD. A support-and-mean relaxation gives $[300, 500]$ USD. The interval width measures how much this particular summary fails to identify; it does not measure a standard error.

2.4. Cent-valued quotes and rounded means

Treating a displayed two-decimal mean as exact can exclude feasible original quotes. The implementation therefore uses integer cents. Let $A, B, M \in \mathbb{Z}$ denote the recorded minimum, maximum and reported mean in cents. We use the conservative closed rounding band $[M - 1/2, M + 1/2]$, allowing either midpoint convention. The feasible integer quote total satisfies

$$\begin{aligned} S^- &= \max\{\lceil nM - n/2 \rceil, (n-1)A + B\}, \\ S^+ &= \min\{\lfloor nM + n/2 \rfloor, A + (n-1)B\}. \end{aligned} \tag{5}$$

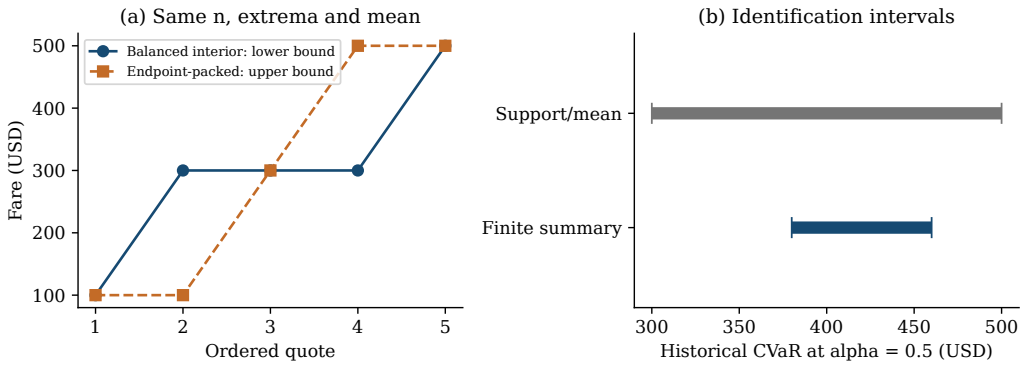


Figure 1: A synthetic example of information retained by a finite summary. Panel (a) shows the extremal unordered quote vectors; panel (b) compares their empirical conditional-value-at-risk interval with a support-and-mean relaxation. The exact mean is held fixed for this illustration.

A record is feasible if $S^- \leq S^+$, subject to the singleton condition. If the provider supplies its exact midpoint convention, an excluded boundary can be removed. The closed-band bounds are sharp for the stated admissible set, which is a conservative representation when the actual convention is unknown.

For a fixed feasible integer total S, the residual sum is $s = S - A - B$. Write $s = rq + t$, with integer q and $0 \leq t < r$. The balanced residual vector contains $r - t$ occurrences of q and t occurrences of $q + 1$. The packed residual vector uses the integer analogue of the construction above. An elementary one-cent transfer establishes the same majorization ordering on this grid. The extremal vectors are coordinatewise nondecreasing as their total increases. Hence the rounding-aware lower bound uses the balanced construction at S^-, and the upper bound uses the packed construction at S^+. All published dollar values are converted from these cent calculations.

This treatment assumes that the underlying USD quotes are cent denominated and that the mean is the unweighted arithmetic average of the n recorded quotes. The supplied dictionary describes an average and an observation count, and the records pass the resulting feasibility checks. However, this agreement does not independently establish the upstream aggregation algorithm. If the provider instead uses weighted observations, conversion before rounding, or hidden exclusions, the admissible set must be changed before the certificates are used.

2.5. Comparison with a support-and-mean relaxation

A deliberately weaker benchmark retains only the interval $[a, b]$ and exact mean μ, permitting arbitrary probabilities. Its sharp bounds are

$$L_\alpha^{\mathrm{SM}} = \mu, \qquad U_\alpha^{\mathrm{SM}} = \min\left\{b,\ a + \frac{\mu - a}{1 - \alpha}\right\}. \tag{6}$$

The lower value is attained by a point mass at μ. The upper value follows from the two-endpoint distribution and is attained by allocating the required mass to the upper tail. These extremal arguments are standard consequences of convex-order bounds; the benchmark is not presented as a novel result. With a rounded mean, the lower mean-band endpoint is used for the lower bound and the upper endpoint for the upper bound, clipped to $[a, b]$.

The finite-summary set is a subset of this relaxation. Thus its interval cannot be wider. The empirical comparison asks how much narrower it is in this feed and whether that difference resolves decisions. It is not a contest against a forecasting model, and narrowing is not evidence that the narrower interval covers future prices. An observation-weighted average across records would answer a different question and would overrepresent frequently collected products; all record-level width summaries here weight records equally.

2.6. Decision certificates, regret and information acquisition

Consider a set $\mathcal{I}$ of comparison-compatible itineraries. The historical generalised-cost criterion for itinerary i is $c_i = T_\alpha(x_i) + vt_i$, where t_i is scheduled outbound plus return duration in hours and $v \geq 0$ is an assumed value of that time. Because translation commutes with Eq. (2), its bounds are $\ell_i = L_{\alpha,i} + vt_i$ and $u_i = U_{\alpha,i} + vt_i$. The primary analysis sets $v = 0$ to isolate fare-risk information; positive values are sensitivity scenarios, not estimates of passenger preferences.

If $u_i \leq \min_{j \neq i} \ell_j$, itinerary i is a certified minimum for every compatible set of historical quote distributions. If no such option exists, use worst-case regret under the rectangular marginal ambiguity set $\prod_i [\ell_i, u_i]$. Rectangularity permits all combinations of row-level summaries; it is not an assertion of statistical independence between airlines or fares.

Proposition 2 (Exact regret for choosing one alternative)**.** *For $|\mathcal{I}| \geq 2$, the maximum regret of selecting i is*

$$R_i = \sup_{c \in \prod_j [\ell_j, u_j]} \left(c_i - \min_j c_j\right) = \max\left\{0,\ u_i - \min_{j \neq i} \ell_j\right\}. \tag{7}$$

The minimax-regret choice minimises R_i. For a singleton comparison set, regret is zero.

Proof. The expression inside the supremum equals $\max\{0, \max_{j \neq i}(c_i - c_j)\}$. Each difference is maximised by setting $c_i = u_i$ and the relevant competitor to its lower bound. Rectangularity makes this combination feasible; the attained risk bounds in Proposition 1 provide compatible quote samples for these choices. Excluding i from the competitor minimum is essential: an option cannot incur positive regret relative to itself. □

Equation (7) is the single-choice special case of interval regret, whose broader optimisation literature is reviewed by Aissi et al. (2009). Unlike joint-day realised-cost comparisons, it does not require synchronising the missing quote timestamps. It compares marginal historical tail criteria. If the decision instead concerned the probability of one itinerary being cheaper on a particular date, these summaries would not identify the required joint distribution.

Finally, let $R^* = \min_i R_i$. Recovering every missing quote could reveal an oracle minimum and eliminate regret for this retrospective criterion, but the gross improvement over the robust choice cannot exceed R^* in any compatible state. An information charge greater than R^* therefore cannot be recovered solely through improving this one historical ranking. This is a decision-specific upper bound, not an expected value of information or a valuation of forecasting data. It remains useful when no probability distribution over the compatible states is available.

3. Verification and numerical aggregation experiments

The mathematical guarantees concern information loss, so evaluation begins with problems whose underlying quote vectors are known. Verification checks exactness, while simulation illustrates the size of the intervals under transparent, assumed quote mechanisms. Neither component supplies external validation of future airfare prediction.

3.1. Independent checks of sharpness and regret

The first verification enumerates every nondecreasing vector of lengths one through seven on the six-value grid $\{0, 1, 2, 3, 4, 5\}$. These arbitrary nonnegative test units are chosen for exhaustive coverage rather than fare realism. The 1,715 vectors produce 657 distinct (n, a, b, S) summaries. For each summary

and each $\alpha \in \{0, 0.25, 0.5, 0.75, 0.9\}$, the minimum and maximum risk over all compatible enumerated vectors are compared with the constructive bounds. All 3,285 risk comparisons agree exactly to the reported numerical precision.

A second, computationally different check formulates the problem as a linear programme over sorted continuous quotes. The objective weights are obtained from Eq. (3); the constraints impose ordering, endpoints and the exact total. For 140 generated instances with counts from two to fourteen and the same five tail levels, lower and upper problems require 1,400 solves. The largest absolute difference from the analytical constructions is 7.96×10^{-13} in the test price units. A further exhaustive check of 258 rounded-summary groups confirms the integer, closed-band implementation. Finally, enumerating all interval corners in 100 single-choice problems verifies Eq. (7) with zero numerical discrepancy. These checks complement, rather than replace, the proof.

3.2. Controlled information-loss experiment

The experiment generates 250 quote vectors for each count $n \in \{2, \ldots, 7\}$ and each of four mechanisms, for 6,000 cases. Starting from $x_j = 100(200 + 300z_j)$ cents, values are rounded to integer cents and clipped below at one cent. The mechanisms are uniform $z_j \sim U(0, 1)$; right-skewed $z_j \sim \text{Beta}(1, 5)$; a two-regime mechanism $z_j = 0.8B_j + \epsilon_j$, with $B_j \sim \text{Bernoulli}(0.3)$ and $\epsilon_j \sim N(0.1, 0.04^2)$, clipped to $[0, 1]$; and an ordered mechanism with z_j equal to equally spaced points from zero to one plus $N(0, 0.04^2)$ noise. The final mechanism illustrates that chronological structure can be present even though the target statistic ignores order. The fixed random seed is 20261025.

Each original vector is compressed to n, a, b, M, with the mean rounded half-up to cents. The method then receives only that summary. The known original empirical $T_{0.5}$ is checked against the reconstructed interval. This is a deterministic information-consistency experiment: containing the original statistic is required by construction, and a 100% pass rate is not a statistical coverage estimate. We also calculate the discrepancy incurred by substituting the sample mean or maximum for the true upper-half criterion. Their respective directions are mathematical consequences of the criterion, so no algorithmic superiority is inferred from those signs.

Table 2 reports the resulting widths and the amount by which the mean falls below the historical upper-half criterion. The amount of information lost depends on the realised quote configuration, not solely on n. Figure 2 shows the count profile and the complementary errors of the two point proxies.

Table 2: Synthetic aggregation experiment at $\alpha = 0.5$. Widths and shortfalls are means in USD; "mean shortfall" is true empirical tail cost minus the original sample mean.

Quote mechanism	Cases	Finite width	Support width	Mean shortfall
monotone	1500	51.65	145.52	102.65
right skewed	1500	12.27	33.40	26.17
two regime	1500	21.33	67.28	62.09
uniform	1500	30.02	75.94	57.92

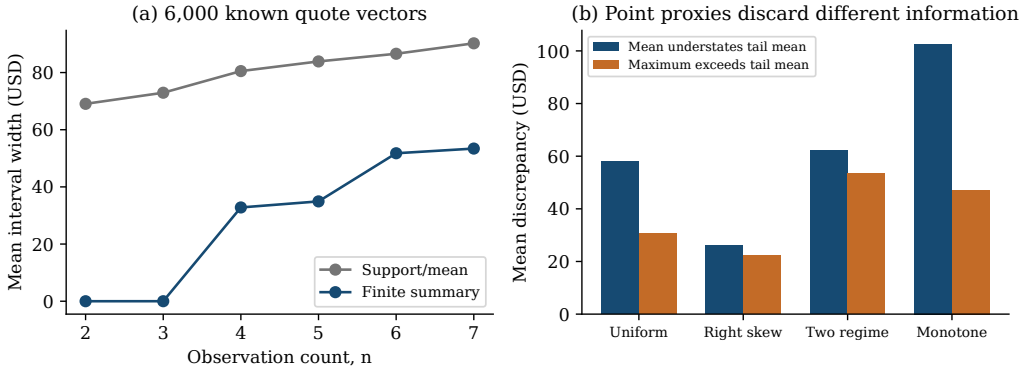


Figure 2: Information loss in 6,000 synthetic cases. Panel (a) compares identification widths; panel (b) shows mean and maximum proxy discrepancies. The experiment assesses reconstruction from summaries, not forecast accuracy or traveller savings.

Because the mechanisms are assumed and not fitted to the commercial sample, their dollar magnitudes should not be extrapolated to a market. Their role is to make the informational distinction observable under several known quote shapes.

4. Case study: weekly Altus airfare records

The case study applies the method to the supplied Altus sample without adding simulated observations to the empirical dataset. The analysis uses the supplied customer dictionary to interpret its fields and the public product description to establish the nature of the service (OAG, n.d.-a,n). This is a bounded demonstration using a customer sample, rather than a probability sample of the United States airline industry.

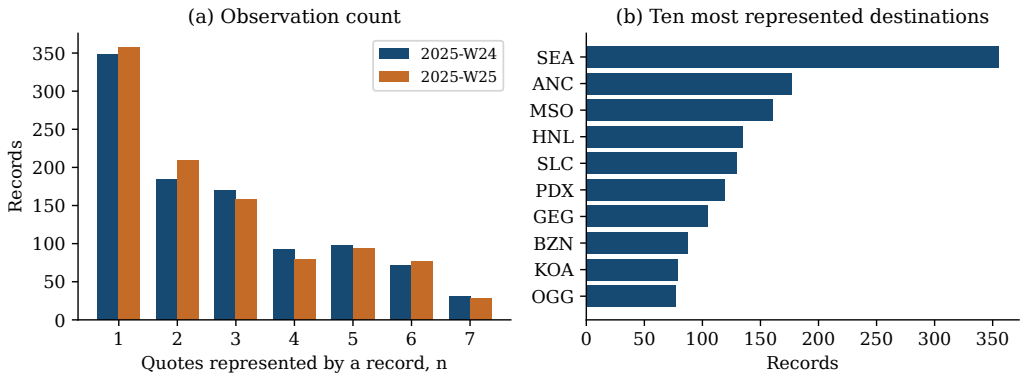


Figure 3: Composition of the supplied sample. Counts refer to aggregate records, not passengers or sales. The destination plot displays only the ten most represented destinations; all 35 are retained in analysis.

4.1. Scope, grain and data checks

The CSV contains 2,000 rows and 45 columns. Each row is an itinerary-level aggregate within its delivered observation week and provider lead-week bucket, rather than an individual transaction. Observations run from 9 to 22 June 2025, covering ISO weeks 24 and 25, with 996 and 1,004 records, respectively. All rows have origin SFO, carrier code AS and fare currency USD; 35 destinations are represented. The carrier field identifies the searched carrier, not necessarily the operating carrier of every leg. The departures span 9 June to 31 December 2025. There are 1,958 economy records and 42 first-class records, and 1,913 return trips and 87 one-way trips. Figure 3 shows the concentration by count and destination.

All record identifiers are unique. Gross minimum, average and maximum fares are positive and correctly ordered, and every record admits a cent-valued quote vector under Eq. (5). Observation counts lie between one and seven and do not exceed the inclusive first-to-last observation-date span. First and last observations remain within the same ISO week and precede or coincide with outbound departure. The one-way records have structurally missing inbound fields; return-trip fields are populated, and reported stay lengths match the difference between outbound and inbound departure dates. Local-currency and USD price fields agree throughout. Average gross fares reconcile with average net fares plus average extras to within one cent, consistent with separately rounded components. These checks establish internal consistency, not upstream collection completeness.

The dictionary includes distance and partition fields that are absent from this CSV, and uses some field names that differ from the export. The analysis uses the actual exported fields and records the aliases in the audit file; it does not invent missing distances. Minima and maxima of net fares and extras are not assumed to be synchronised components of the gross extrema. Similarly, "extras" are the taxes, surcharges and fees described in the dictionary, not a demonstrated complete basket of optional baggage and seat-selection purchases. Gross quoted price is therefore the observed fare measure, not a verified door-to-door trip cost.

The provider describes both weekly delivery and aggregation by weeks before departure. The lead-week variable is retained in comparison keys. One itinerary appears twice in the same delivered observation week under different lead-week buckets; treating those records as competing products would create a spurious alternative. Within a key that includes the lead-week bucket, the itinerary identity is unique. There are 1,999 records from the main booking-site label and one from a second-daily label. No weighting by collection frequency or interpretation of absent quotes as unavailable inventory is imposed.

4.2. What the aggregates identify

Of the 2,000 records, 706 contain one quote and 1,060 have identical minimum and maximum fares. The latter includes every singleton and some genuinely constant multi-quote samples. Consequently, only 940 records show any observed fare variation. There are 1,427 records with at most three quotes; their unordered samples are identified for a fixed mean, with at most cent-level residual uncertainty under the rounding model. Overall, the upper-half risk interval is no wider than USD 0.01 in 1,596 records, or 79.8% of the sample.

Table 3 reports widths by observation count. The mean finite-summary width is USD 4.17 across all records, compared with USD 14.88 for the support-and-mean relaxation. The reduction is 71.98%, calculated as one minus the ratio of the sums of widths over the same 2,000 records. This is not the mean of recordwise percentage reductions, which would be undefined for constant records. Restricting to the 940 varying-fare records gives mean widths of USD 8.87 and USD 31.67, respectively. The median finite-summary width is zero, but the maximum is USD 341.84. Thus the average conceals a small subset with materially incomplete tail information.

Figure 4 displays the reduction by count and the distribution of widths among varying-fare records. The zero width at $n = 2$ follows because both

Table 3: Count-specific information in the 2,000 supplied records at $\alpha = 0.5$. Range and interval-width columns are record-weighted mean USD amounts. "Varying fares" counts records with maximum above minimum.

n	Records	Varying fares	Mean range	Finite width	Support width
1	706	0	0.00	0.00	0.00
2	393	207	25.21	0.00	12.61
3	328	240	46.23	0.00	17.70
4	173	141	96.11	10.50	34.53
5	192	167	67.72	10.36	25.21
6	149	132	110.57	22.12	40.55
7	59	53	101.17	20.96	36.41

observed quotes are endpoints; it is not evidence that two quotes provide a better estimate of future fare risk than seven quotes. The larger widths for some higher-count records arise because more interior values have been discarded and because their observed ranges differ. Cross-sectional count comparisons cannot identify a causal effect of collecting additional quotes.

4.3. Matched itinerary comparisons

The decision analysis first restricts the dataset to economy return trips, leaving 1,907 records. A primary comparison cell shares observation ISO week, provider weeks-before-departure bucket, destination, outbound departure date, return departure date and stay length. Origin, searched carrier and cabin are then constant by construction. Cells with fewer than two alternatives are excluded from decision comparisons, but not from the record-level audit and width analysis. This yields 182 cells containing 388 records. These are deliberately flexible-schedule comparisons: departure times and total durations may differ, as may unobserved fare conditions. They are not asserted to be interchangeable products for every traveller.

At $\alpha = 0.5$ and $v = 0$, a finite-summary certificate identifies a minimum in 180 of the 182 cells, compared with 174 under the support-and-mean relaxation. The finite-summary minimax choice differs from the reported-mean rule in three cells, and from the support-and-mean minimax choice in two. Table 4 evaluates all selection rules using the same finite-summary worst-case regret; the numbers are retrospective criterion gaps, not realised expenditures. Mean worst-case regret is USD 0.033 for the finite-summary rule and USD 0.257 for the mean rule. The largest difference between those

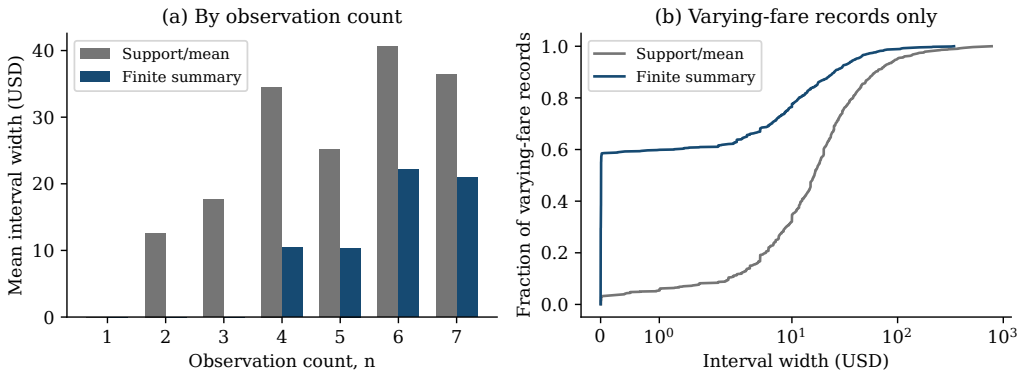


Figure 4: Historical upper-half risk identification widths. Panel (a) groups all records by count; panel (b) restricts the cumulative distribution to 940 varying-fare records. The horizontal scale in panel (b) is symmetric logarithmic to display exact zeros and a long positive tail.

Table 4: Worst-case regret of five benchmark rules in the same 182 primary cells, at $\alpha = 0.5$ and $v = 0$. Every rule is evaluated under the finite-summary ambiguity set. All values are USD; they are not observed savings.

Benchmark rule	Mean	Median	95th pct.	Maximum
Minimum fare	9.02	0.00	7.89	1224.08
Mean fare	0.26	0.00	0.00	30.38
Maximum fare	0.08	0.00	0.00	8.07
Support/mean minimax	0.09	0.00	0.00	8.07
Finite-summary minimax	0.03	0.00	0.00	4.00

two rules in one cell is USD 30.385. The finite-summary rule still has positive worst-case regret in two cells, of USD 4.00 and USD 1.937.

The minimum-fare rule has a much larger maximum regret, USD 1,224.08, driven by one cell with a large upper-tail contrast. Its mean regret is USD 9.02, whereas its median and that of every other rule are zero. Reporting only a percentage improvement over this deliberately optimistic proxy would overstate the typical managerial gain. The main result is that most rankings are already separable, while a small number of comparisons benefit from retaining the summary constraints. Figure 5(a) shows the certification comparison.

Table 5 and Fig. 5(b) show the cell with the largest excess regret under the mean rule, selected explicitly for illustration. It contains two SFO–EWR economy return itineraries departing 19 June and returning 3 July,

Table 5: Illustrative EWR cell, chosen as the largest mean-rule excess-regret example. Values are USD except the quote count. Displayed rounding can conceal sub-cent interval widths; decisions use full cent-based calculations.

Option	n	Min	Mean	Max	$L_{.5}$	$U_{.5}$	Regret
R1248	3	681.30	894.63	1056.29	1019.62	1019.63	30.38
R1536	4	814.41	943.35	998.00	989.25	998.00	0.00

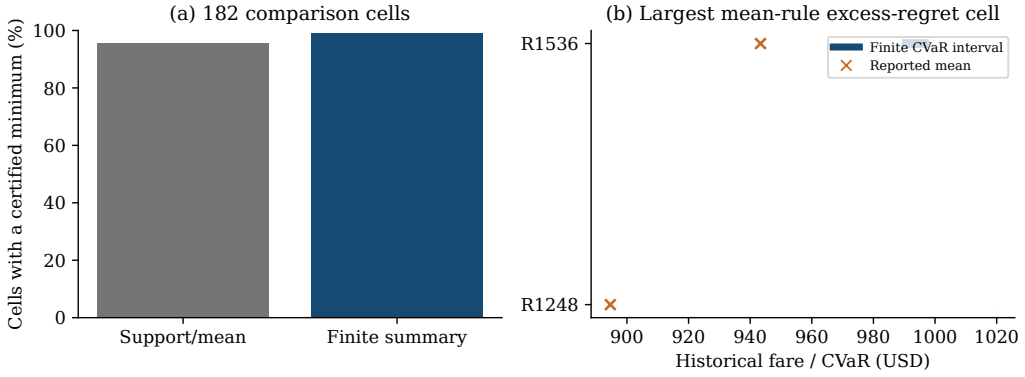


Figure 5: Decision-level consequences at $\alpha = 0.5$ and $v = 0$. Panel (a) reports the share of primary cells with a certified minimum. Panel (b) shows the explicitly selected EWR example; crosses are reported means, while horizontal marks are historical tail-risk identification intervals.

in observation week 25 and lead-week bucket zero. Option R1248 has a lower mean, USD 894.63, but its historical upper-half criterion is about USD 1,019.63. Option R1536 has a higher mean, USD 943.35, but its entire compatible tail interval lies below that of R1248. Hence R1536 is certified under the stated fare-only criterion. This is an example of a change in the chosen risk criterion, not proof that an aggregate mean was computed incorrectly.

Duration changes the interpretation of this example. R1248 has total scheduled travel time of 1,295 minutes and R1536 has 1,686 minutes. The latter therefore requires an additional 391 minutes. A value of that time above roughly USD 4.7 per hour can reverse the fare-only advantage in this example. The comparison should consequently be understood as a fare-risk benchmark, with time-sensitive procurement assessed using the generalised-cost scenarios rather than the fare-only recommendation.

Table 6: Sensitivity to comparison-cell definitions at $\alpha = 0.5$, $v = 0$. The primary key includes shared lead-week bucket and travel dates. "Same departure time bins" requires matching six-hour bins for both outbound and return departures; "same first and last observation" adds exact endpoint-date matching.

Comparison definition	Cells	Records	Certified	Mean regret
Primary: shared lead week	182	388	180	0.03
Same departure time bins	31	62	31	0.00
Same first and last observation	152	319	151	0.03
Observation week 24	89	194	88	0.04
Observation week 25	93	194	92	0.02

4.4. Sensitivity to comparison definitions and preferences

Table 6 narrows the comparison sets. Requiring both outbound and inbound departure times to fall in the same respective six-hour bins leaves 31 cells and 62 records; every cell is certified and none changes relative to its mean rule. Requiring identical first and last observation dates leaves 152 cells and 319 records; 151 are certified, and one selection changes. The date requirement aligns observed endpoints, but not necessarily the interior observation dates. The two observation weeks are also reported separately. All three mean-rule changes in the primary analysis occur in the second week. These results limit the strength of any general decision-improvement claim and demonstrate why matching choices must be disclosed.

Table 7 and Fig. 6 vary the tail level over $\{0.25, 0.5, 0.75, 0.9\}$ and the assumed value of scheduled time over $\{0, 25, 50, 100\}$ USD per hour, holding the 182 primary comparison sets fixed. Across these scenarios, between zero and six selections differ from the corresponding mean-plus-time rule. At $\alpha = 0.9$, every row's empirical tail criterion equals its recorded maximum because all counts are at most seven. The resulting zero interval widths are a consequence of Corollary 1, not evidence of highly precise estimation of a future 90th-percentile tail. More extreme risk aversion can therefore make the historical functional easier to identify while changing the economic objective substantially.

Table 7: Preference scenarios over the same 182 primary cells. Regret values are USD. [a]"Changed" counts selections differing from the reported-mean-plus-time rule at the same assumed v; it does not count changes from the $v = 0$ recommendation.

α	v (USD/h)	Cells	Certified	Changed[a]	Mean regret
0.25	0	182	180	1	0.02
0.25	25	182	182	0	0.00
0.25	50	182	182	3	0.00
0.25	100	182	182	0	0.00
0.50	0	182	180	3	0.03
0.50	25	182	182	1	0.00
0.50	50	182	182	4	0.00
0.50	100	182	182	1	0.00
0.75	0	182	182	2	0.00
0.75	25	182	181	2	0.00
0.75	50	182	181	5	0.20
0.75	100	182	181	3	0.07
0.90	0	182	182	2	0.00
0.90	25	182	182	2	0.00
0.90	50	182	182	6	0.00
0.90	100	182	182	4	0.00

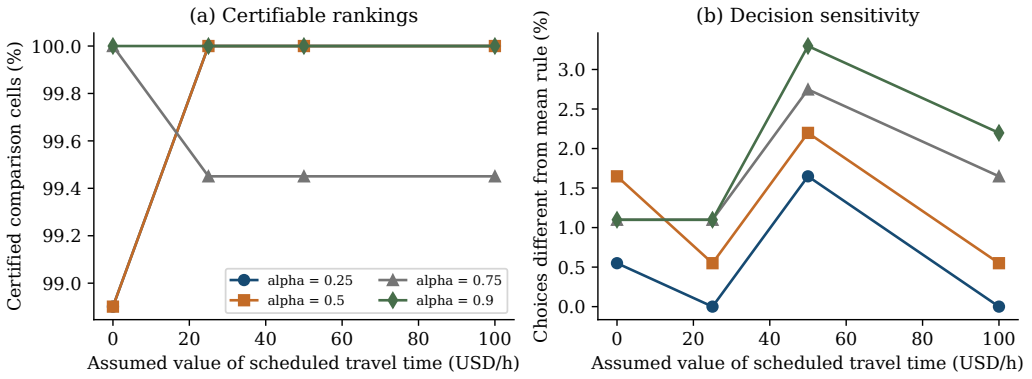


Figure 6: Sensitivity of historical ranking certificates and departures from the mean-plus-time rule. Values of scheduled time are assumed scenarios; no passenger preference parameter is estimated. Tail levels change the criterion rather than a confidence level.

5. Managerial implications and limitations

The case study supports a restrained interpretation: a commercial summary can preserve enough information for many historical comparisons, and its adequacy can be tested rather than assumed. This section translates that result into possible uses while identifying the evidence needed before deployment.

5.1. Use certificates to allocate analytical effort

A practical workflow begins by defining comparable products and the intended criterion, then calculating the two risk bounds. A certified comparison can proceed without choosing an invented within-week fare distribution. An unresolved comparison can be inspected for fare conditions, schedule restrictions and collection differences before more complex modelling is attempted. The positive-regret cells in this sample are few. Recovering quote detail solely to resolve the primary historical ranking could improve the criterion by no more than USD 4.00 in one unresolved cell and USD 1.94 in the other, given the stated ambiguity model. Those are caps on this narrowly defined benefit; detailed data may have additional monitoring, compliance, service-design or forecasting value that the calculation does not address.

The procedure also explains when added analytical complexity is unlikely to help. A zero-regret certificate establishes that no compatible reconstruction changes the preferred historical criterion. In this sample, adopting the robust calculation throughout the primary cells changes only 1.65% of selections relative to the mean rule. Even that percentage falls to zero under the tighter time-bin definition. A manager should therefore view the result as an auditable exception screen, rather than as evidence that a new optimisation system will generate a large procurement return. The method's strongest output can be a defensible decision to retain a simple existing benchmark.

5.2. Preserve count, extrema and product definitions in data products

For data producers, the results identify information that can be retained at low schema complexity. Counts make equal empirical masses explicit, and observed extrema distinguish a realised range from a merely assumed support. Together with the mean they permit stronger conclusions than a mean-only export. The feed should also specify the arithmetic used to produce the mean, its rounding convention, the treatment of repeated observations and the relationship between delivery week and lead-week bucket. These definitions

determine the admissible set and therefore belong to the data contract, rather than to undocumented analyst judgement.

Low identification width should not become a data-quality score. Singletons trivially determine their empirical tail criterion, while providing no evidence about unobserved intraweek variation. A responsible reporting surface should display the count, observation dates, product key, empirical range and certificate together. It should distinguish "the historical ranking is fixed by this summary" from "the fare risk is well estimated." The former is a mathematical result conditional on the record; the latter requires a collection design and an inferential model.

For aviation management research and teaching, this distinction connects quantitative methods to the evaluation of purchased intelligence. Analysts can examine whether a decision needs a full quote history, what a particular aggregate can justify and which additional fields would change that judgement. The contribution is relevant to the governance of analytical evidence in transport organisations, without claiming that the present data establish organisational behaviour or human decision-making effects.

5.3. Limits of the evidence and a stronger validation design

Several limits constrain publication claims. The sample contains one origin, one searched carrier and two observation weeks, with no documented probability sampling design. Record-level averages consequently describe this file alone. The dataset does not support causal statements about competition, consumer discrimination or pricing strategy, nor does it establish effects for Alberta or Canadian aviation. No passenger outcome, willingness-to-pay measure, transaction, seat inventory or realised procurement decision is observed. Cross-sectional relationships with booking horizon would also mix travel dates, product attributes and selection mechanisms; they should not be described as causal booking curves.

The mathematical guarantees are conditional on the summary semantics. Cent-valued unweighted quotes, observed extrema and rounded arithmetic means are explicit assumptions, supported by internal consistency but not verified against raw provider logs. The marginal Cartesian set is conservative when shared pricing mechanisms constrain alternatives jointly. Conversely, an invalid matching key could make apparently strong certificates irrelevant to the actual traveller. Missing fare-family restrictions, baggage allowances, change conditions and preferred departure windows must be considered before turning a benchmark into procurement advice.

The close observation windows are another material limit. The largest illustrated mean-rule reversal compares a three-quote record ending on 18 June with a four-quote record ending on 19 June. Its certificate is valid for their two recorded marginal samples; it is not a guarantee about a common daily choice opportunity. The endpoint-matched sensitivity reduces this problem, but equal first and last dates do not prove that all missing interior quotes were synchronised. A study of realised same-day choice, relative price probabilities or buy-versus-wait policies would need the underlying timestamps and availability.

The controlled experiments establish computational correctness and illustrate information loss under stipulated mechanisms. They do not address representativeness or future calibration. In particular, reducing uncertainty relative to a relaxation is guaranteed by set inclusion, and minimising the chosen worst-case regret is optimal by definition. These properties are reasons to use the method for its stated target, not independent evidence of business effectiveness. The economically relevant empirical questions are how often rankings become certifiable, whether recommendations change, and whether the comparison sets correspond to an actual decision. The modest changes and sensitivity to matching are therefore central results.

A stronger next-stage validation would obtain dated raw quotes paired with the provider's weekly aggregates across several origins, carriers and seasons. The aggregation audit would first verify count, weighting, price units and rounding. Historical reconstruction would then compare each proposed interval with the exact empirical criterion computed from the hidden raw sample, using product-level or market-level separation where estimation is introduced. A separate prospective procurement study would log feasible alternatives, decision-time information, traveller constraints, actual purchase prices and subsequent outcomes. Only that latter design could test realised savings or future-price calibration. It should evaluate simple rules alongside the proposed screen, and keep preference elicitation separate from forecast assessment.

6. Concluding remarks

Weekly airfare summaries need not be treated either as complete quote histories or as information too coarse to support rigorous comparisons. Their count, realised extrema and mean imply an explicit set of compatible empirical samples. Classical majorization reduces upper-tail risk identification to two

constructive vectors; an integer formulation accounts for cent-valued fares and mean rounding. The resulting bounds support ranking certificates and exact single-choice regret without assuming a parametric price distribution.

In the supplied two-week Altus sample, this treatment reduces mean upper-half risk interval width by approximately 72% relative to a support-and-mean relaxation. It certifies a minimum in 180 of 182 primary itinerary cells, but changes only three selections relative to a mean-fare rule. Tighter schedule matching removes those changes. The contribution is therefore a verifiable test of when a compressed fare feed is sufficient for a historical criterion, together with an explicit account of when it is not. Broader economic and operational benefits remain questions for richer, prospective data.

Data and code availability

The analysis uses the supplied commercial Altus customer sample and its accompanying customer dictionary. These inputs should not be described as an open dataset: public redistribution permission has not been established. The accompanying private reproduction package contains the supplied CSV, source checksum, analysis code, derived results and figure-generation scripts. Public release of the raw sample or derived record-level extracts is subject to the data owner's terms. The synthetic experiments and exhaustive verification can be reproduced independently of commercial-data access.

Declaration of generative AI use

OpenAI ChatGPT was used for enhancing the literature discovery, and manuscript proofreading. Responsibility for checking and approving the final scientific content remains with the submitting author.

Appendix A. Partial-sum proof and integer implementation

This appendix records additional details needed to audit the finite-summary construction. Let $y \in [a, b]^r$ be a sorted residual vector with sum s, and write y^{bal} and y^{pack} for its balanced and packed extremes. For any k residual coordinates, the sum of the largest k coordinates cannot exceed either kb or $s - (r - k)a$. The packed vector attains the smaller of these two quantities simultaneously for every integer k. Conversely, among residual vectors with fixed sum, the balanced vector minimises every largest-k sum.

On the real grid that minimum is ks/r; on the integer-cent grid it is attained by distributing the remainder across adjacent integers.

Appending the same endpoints to two vectors preserves their majorization order: a largest-k sum of the appended vector is the maximum of the corresponding sums formed by allocating the k positions between the common appended coordinates and the original vector. Therefore the residual partial-sum inequalities extend to full quote vectors. The empirical upper-tail value in Eq. (3) is a nonnegative interpolation of adjacent largest-k sums, so the same two vectors attain its extrema for every α.

With integer residual sum $s = rq+t$, $0 \leq t < r$, the least dispersed residual vector is q repeated $r - t$ times and $q + 1$ repeated t times. Increasing s by one raises exactly one coordinate of its sorted vector. For the packed residual vector, increasing the excess above ra fills the current remainder to b before starting the next coordinate; its sorted coordinates are also nondecreasing. Monotonicity of empirical conditional value-at-risk then implies that the smallest feasible total in Eq. (5) gives the lower bound and the largest gives the upper bound. This avoids interchanging a minimisation over thresholds with a supremum over distributions.

Appendix B. Reproduction protocol and numerical conventions

The private package contains a single source CSV with SHA-256 checksum

`346d3a078bbeede023f31f680d79c548fc8038ccbd0c53ca7271b68f24e1105b`.

The analysis reads record identifiers as strings and converts monetary summaries to integer cents using decimal arithmetic. The input has 45 columns; additional time-bin and risk columns in derived tables are calculated fields. Gross USD price is the sole fare outcome used by the method. No missing fare is imputed, and structural return-field missingness is retained for one-way records. Derived row labels such as R1248 identify the original CSV row position, starting at one.

Running `python code/verify.py` performs the independent mathematical checks. Running `python code/reproduce.py` recreates the case-study and synthetic result CSVs, generated table fragments and all six figures. The documented reference environment uses Python 3.12.14, NumPy 2.3.5, pandas 2.2.3 and Matplotlib 3.10.8; SciPy provides the independent linear-programming verification. The synthetic generator is NumPy's default generator with seed 20261025. Floating-point tolerances are used only in numerical

comparisons and display; the admissible quote totals themselves use exact integer-cent constraints.

The minimax calculation evaluates every primary cell, using stable original-row ordering for exact ties. The same comparison cells and duration valuation are used for each benchmark within a scenario. The support-and-mean policy is selected using its relaxed interval, then evaluated using the tighter finite-summary regret, so the comparison does not mix evaluation criteria. All decision-level averages weight cells equally. Counts, widths and regrets are descriptive quantities for the supplied sample; no independence assumption, population confidence interval or inferential p-value is attached to them.

References

A., B. T., & Hochman, E. (1972). More bounds on the expectation of a convex function of a random variable. *Journal of Applied Probability, 9*(4), 803–812. https://doi.org/10.2307/3212616

Aissi, H., Bazgan, C., & Vanderpooten, D. (2009). Min-max and min-max regret versions of combinatorial optimization problems: A survey. *European Journal of Operational Research, 197*(2), 427–438. https://doi.org/10.1016/j.ejor.2008.09.012

Alderighi, M., Nicolini, M., & Piga, C. A. (2015). Combined effects of capacity and time on fares: Insights from the yield management of a low-cost airline. *Review of Economics and Statistics, 97*(4), 900–915. https://doi.org/10.1162/REST_a_00451

Aryal, G., Murry, C., & Williams, J. W. (2024). Price discrimination in international airline markets. *Review of Economic Studies, 91*(2), 641–689. https://doi.org/10.1093/restud/rdad037

Bertsimas, D., & Popescu, I. (2005). Optimal inequalities in probability theory: A convex optimization approach. *SIAM Journal on Optimization, 15*(3), 780–804. https://doi.org/10.1137/S1052623401399903

Borenstein, S., & Rose, N. L. (1994). Competition and price dispersion in the U.S. airline industry. *Journal of Political Economy, 102*(4), 653–683. https://doi.org/10.1086/261950

Bülbül, K. G. (2025). Airfare prediction: Leveraging market data for better decision-making. *Research in Transportation Business & Management, 61*, 101408. https://doi.org/10.1016/j.rtbm.2025.101408

Chen, L., He, S., & Zhang, S. (2011). Tight bounds for some risk measures, with applications to robust portfolio selection. *Operations Research, 59*(4), 847–865. https://doi.org/10.1287/opre.1110.0950

Delage, E., & Ye, Y. (2010). Distributionally robust optimization under moment uncertainty with application to data-driven problems. *Operations Research, 58*(3), 595–612. https://doi.org/10.1287/opre.1090.0741

Etzioni, O., Tuchinda, R., Knoblock, C. A., & Yates, A. (2003). To buy or not to buy: Mining airfare data to minimize ticket purchase price. In *Proceedings of the Ninth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* (pp. 119–128). ACM. https://doi.org/10.1145/956750.956767

Gaggero, A. A., & Luttmann, A. (2025). How does COVID-19 affect intertemporal price discrimination and price dispersion? Evidence from the airline industry. *International Journal of Industrial Organization, 101*, 103171. https://doi.org/10.1016/j.ijindorg.2025.103171

Gerardi, K. S., & Shapiro, A. H. (2009). Does competition reduce price dispersion? New evidence from the airline industry. *Journal of Political Economy, 117*(1), 1–37. https://doi.org/10.1086/597328

Hortaçsu, A., Natan, O. R., Parsley, H., Schwieg, T., & Williams, K. R. (2024). Organizational structure and pricing: Evidence from a large U.S. airline. *The Quarterly Journal of Economics, 139*(2), 1149–1199. https://doi.org/10.1093/qje/qjad051

Iskandar, R. (2021). Probability bound analysis: A novel approach for quantifying parameter uncertainty in decision-analytic modeling and cost-effectiveness analysis. *Statistics in Medicine, 40*(29), 6501–6522. https://doi.org/10.1002/sim.9195

Kim, M., Shen, L., & Ge, Q. (2023). Does competition increase or decrease price dispersion? Insights from one-way vs. round-trip airfares. *Oxford Bulletin of Economics and Statistics, 85*(2), 435–455. https://doi.org/10.1111/obes.12488

Kuhn, D., Shafiee, S., & Wiesemann, W. (2025). Distributionally robust optimization. *Acta Numerica, 34*, 579–804. https://doi.org/10.1017/S0962492924000084

Manski, C. F. (2003). *Partial Identification of Probability Distributions*. Springer. https://doi.org/10.1007/b97478

Marshall, A. W., Olkin, I., & Arnold, B. C. (2011). *Inequalities: Theory of majorization and its applications* (2nd ed.). Springer. https://doi.org/10.1007/978-0-387-68276-1

OAG (n.d.-a). *Altus Data: Historical flight price database*. Retrieved October 25, 2026, from https://www.oag.com/altus-data.

OAG (n.d.-b). *Altus Standard V2: Weekly feed (customer edition)*. Unpublished technical specification supplied with the dataset.

Perakis, G., & Roels, G. (2008). Regret in the newsvendor model with partial information. *Operations Research, 56*(1), 188–203. https://doi.org/10.1287/opre.1070.0486

Postek, K., Ben-Tal, A., den Hertog, D., & Melenberg, B. (2018). Robust optimization with ambiguous stochastic constraints under mean and dispersion information. *Operations Research, 66*(3), 814–833. https://doi.org/10.1287/opre.2017.1688

Rahimian, H., & Mehrotra, S. (2022). Frameworks and results in distributionally robust optimization. *Open Journal of Mathematical Optimization, 3*, 1–85. https://doi.org/10.5802/ojmo.15

Rockafellar, R. T., & Uryasev, S. (2000). Optimization of conditional value-at-risk. *The Journal of Risk, 2*(3), 21–41. https://doi.org/10.21314/JOR.2000.038

Rockafellar, R. T., & Uryasev, S. (2002). Conditional value-at-risk for general loss distributions. *Journal of Banking & Finance, 26*(7), 1443–1471. https://doi.org/10.1016/S0378-4266(02)00271-6

Wiesemann, W., Kuhn, D., & Sim, M. (2014). Distributionally robust convex optimization. *Operations Research, 62*(6), 1358–1376. https://doi.org/10.1287/opre.2014.1314

Williams, K. R. (2022). The welfare effects of dynamic pricing: Evidence from airline markets. *Econometrica, 90*(2), 831–858. https://doi.org/10.3982/ECTA16180

Zhang, H., Czerny, A. I., van Ommeren, J., & Niemeier, H. M. (2025). Intertemporal price dispersion: The role of competition. *Economics of Transportation, 43*, 100424. https://doi.org/10.1016/j.ecotra.2025.100424

Supplementary Files

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

- illustration.tex
- sensitivity.tex
- bycount.tex
- policies.tex
- matching.tex
- synthetic.tex
- CodesandData.zip