

Impact by design: translating *Lead times in flux* into an R handbook with code

Harrison Katz
Forecasting, Data Science, Airbnb
Harrison.Katz@airbnb.com

November 20, 2025

Abstract

This commentary translates the central ideas in *Lead times in flux* into a practice ready handbook in R. The original article measures change in the full distribution of booking lead times with a normalized L1 distance and tracks that divergence across months relative to year over year and to a fixed 2018 reference. It also provides a bound that links divergence and remaining horizon to the relative error of pickup forecasts. We implement these ideas end to end in R, using a minimal data schema and providing runnable scripts, simulated examples, and a prespecified evaluation plan. All results use synthetic data so the exposition is fully reproducible without reference to proprietary sources.

1 What the original article established

Let $L_t(k)$ denote the discrete distribution of lead times for arrival month t on integer days $k = 0, 1, \dots, \Delta_{\max}$. The normalized L1 divergence between two monthly distributions L_a and L_b is

$$D(L_a, L_b) = \frac{1}{2} \sum_{k=0}^{\Delta_{\max}} |L_a(k) - L_b(k)|. \quad (1)$$

If $D = 0.25$ then one quarter of the probability mass has effectively shifted on the support. The article constructs D_t year over year by comparing month t with month $t - 12$, and also relative to the same calendar month in 2018, and decomposes D_t with STL to separate trend and seasonality (Cleveland et al., 1990; Katz et al., 2025). This distributional view detects changes that means or medians may miss.

The article also provides a bound on the relative error of a pickup forecast at horizon Δ :

$$|\epsilon| \leq \frac{2D(1 - \Delta/\Delta_{\max})}{C_{\text{hist}}(\Delta)}, \quad (2)$$

where $C_{\text{hist}}(\Delta)$ is the historical pickup fraction at horizon Δ and D is the divergence between the historical and actual lead time distributions. The bound falls as the service date approaches and grows when divergence is high or when the historical pickup fraction is small at the chosen horizon. This yields a transparent risk index that can sit next to standard pickup reporting (Katz et al., 2025). For probability metrics background, see Gibbs and Su (2002).

Basic properties. The normalized L_1 distance equals the total variation distance on a finite support, so $D \in [0, 1]$ and $D = 0$ iff $L_a \equiv L_b$. For a fixed (D, Δ_{\max}) , the right side of (2) is strictly decreasing in Δ and equals zero at $\Delta = \Delta_{\max}$. **Proof sketch.** The factor $(1 - \Delta/\Delta_{\max})$ decreases linearly in Δ while $C_{\text{hist}}(\Delta)$ is nondecreasing in Δ , so the quotient decreases.

Related modeling work for lead times When analysts need a parametric forecast of the entire lead time vector as a composition, the Bayesian Dirichlet auto regressive moving average model (B-DARMA) models compositional time series with a Dirichlet likelihood and VARMA dynamics in additive log ratio space (Katz et al., 2024). The IJF article develops B-DARMA for forecasting recognition shares across future intervals and is applicable when a distributional forecast of $L_t(k)$ is desired in addition to monitoring.

2 Why a handbook with code

Lead time behavior is nonstationary. Most organizations already track pickup curves and a few summary metrics. Extending that reporting with two elements from the article gives leverage with minimal lift. The first element is a month by month divergence series that flags when the full lead time distribution has moved. The second element is the pickup risk bound in Equation (2), which converts divergence and remaining horizon into a horizon specific risk number. The handbook implements both, along with simple decision templates and an evaluation plan for use in operational settings.

3 The R handbook

The package `leadtimefluxR`¹ implements signals that matter for practice and runs on local CSV files. It uses base R and common tidyverse packages and does not rely on external services.

3.1 Minimal dataset and cohort construction

Use one row per booking with the fields:

Listing 1: Minimal fields

```
arrival_date, booking_ts, stay_nights, channel, segment, origin,
price_at_booking, cancelled, property_id
```

Arrival dates are normalized to dates. Lead time is computed in whole days and negative values are dropped. A safe default Δ_{\max} is chosen from the support and may be capped by the user. Discrete histograms $L_t(k)$ are constructed by month of arrival and by any grouping variable such as property or market. Cumulative pickup $C_{\text{hist}}(\Delta)$ is the cumulative sum of $L_t(k)$ up to Δ .

Support and censoring. Pick Δ_{\max} so that at least 95% of historical mass lies on $[0, \Delta_{\max}]$ for the group, and include a right-censored bin $\Delta_{\max}+$. When a nontrivial right tail exists, compute D on the augmented support with a final bin $\Delta_{\max}+$ to avoid artificial spikes at the cap (cf. Figure 3). For efficiency, daily bins can be aggregated beyond 28 days into weekly bins without affecting the interpretation of D at horizons under 21 days.

¹<https://github.com/harrisonekatz/leadtimefluxR>

3.2 Divergence series and STL decomposition

For each group and month the package computes D_t either year over year or relative to 2018 by calendar month. The divergence series can be decomposed with `stats::stl` to separate trend and seasonality in a robust manner (Cleveland et al., 1990). This supports monitoring and alerting that is aligned with calendar effects.

Baseline choice. Adjacent-month $D(L_t, L_{t-1})$ is sensitive to local movements; year-over-year $D(L_t, L_{t-12})$ controls for seasonality; fixed-year baselines (e.g., 2018) stabilize long-run comparisons but can drift if the baseline regime differs structurally. We recommend monitoring both adjacent and YoY series, decomposed by STL, and reporting the 90th percentile of YoY D as the reference level for risk mapping when at least 13 months are available.

3.3 Pickup bound and decision templates

The right hand side of Equation (2) defines a horizon specific risk index. The package includes templates that map this risk to actions such as pricing cadence, advance purchase buffers, and staffing buffers (Talluri and van Ryzin, 2004). Thresholds are explicit and can be tuned to local constraints. The approach complements model based forecasting, including B-DARMA when a parametric compositional forecast is desired (Katz et al., 2024) and aligns with established revenue-management forecasting practice (Weatherford and Kimes, 2003).

3.4 Uncertainty for D_t and risk indices

Because D_t is computed from monthly histograms, finite counts can introduce variability. We recommend a nonparametric bootstrap at the booking level within each (group, t) to obtain confidence intervals for D_t and for the bound in (2). For each bootstrap replicate $b = 1, \dots, B$, resample bookings with replacement within the month, rebuild $L_t^{(b)}$, recompute $D_t^{(b)}$ and the corresponding bound. Percentile or basic bootstrap intervals then summarize uncertainty. This also supports alerting rules that require D_t to exceed a threshold and its lower confidence limit to exceed a smaller guardrail.

3.5 Evaluation utilities

The package includes MASE and sMAPE for forecast evaluation and pinball loss for quantile forecasts (Hyndman and Koehler, 2006). Metrics are reported by horizon so that monitoring and decision rules can focus on the last two or three weeks where changes are most material.

4 Quick start in R

Listing 2: End to end example on synthetic data

```
library(leadtimefluxR)

df <- generate_synthetic_bookings(start_date = "2021-01-01",
                                  end_date   = "2022-12-31",
                                  avg_bookings_per_day = 20,
                                  properties = 3,
                                  max_lead_days = 60,
                                  compression_level = 0.4,
```

```

seed = 123)

Lk <- leadtime_histograms(df, group_cols = c("property_id"), max_lead_days
  = 60)
Lk_pickup <- pickup_curve(Lk, group_cols = c("property_id"))

# Divergence: adjacent always available; YoY when sample >= 13 months
D_adj <- adjacent_divergence_series(Lk, group_cols = c("property_id"))
D_yoy <- try(yoy_divergence_series(Lk, group_cols = c("property_id")),
  silent = TRUE)

D_est <- if (inherits(D_yoy, "try-error") || nrow(D_yoy) == 0) {
  safe_divergence_quantile(D_adj, probs = 0.90, default = 0.20)
} else {
  as.numeric(stats::quantile(D_yoy$D, probs = 0.90, na.rm = TRUE))
}

# Risk at 14 days for the latest cohort of the first property
pid <- dplyr::first(Lk_pickup$property_id)
latest <- max(Lk_pickup$month)
cohort <- subset(Lk_pickup, property_id == pid & month == latest)
Chist_14 <- cohort$Chist[cohort$k == 14][1]

bound <- relative_error_bound(D = D_est, delta = 14, delta_max = 60, Chist
  _delta = Chist_14)
actions <- recommend_actions(bound)
print(list(latest_month = latest, D_est = round(D_est, 3),
  Chist_14 = round(Chist_14, 3), bound = round(bound, 3), actions
    = actions))

```

5 Results on synthetic data

All artifacts are produced by the script `scripts/make_paper_artifacts.R` and saved under `paper_artifacts/`.

5.1 Adjacent month divergence

Figure 1 shows the adjacent month normalized L1 divergence $D(L_t, L_{t-1})$ by property over 2021 to 2022 with quarterly x axis ticks for readability.

The series stay in a narrow band between roughly 0.12 and 0.20. Property P001 exhibits a late year rise, property P002 shows a dip followed by recovery, and property P003 trends downward into late 2022. These patterns are the distributional changes that a mean or a single percentile can fail to register. The 90th percentile and related summary statistics are reported in Table 1, pulled directly from the generated CSV.

5.2 Pickup and the risk bound

Figure 2 shows $C_{\text{hist}}(\Delta)$ for the latest cohort of property P001. About half of total pickup occurs by two weeks before arrival, which is consistent with the heavy short horizon mass in the corresponding histogram in Figure 3. Combining a typical divergence in the upper teens with a 14 day horizon and $C_{\text{hist}}(14)$ near one half yields a risk bound near one half by Equation (2). The exact values used in the manuscript are read from the CSV in Table 2.

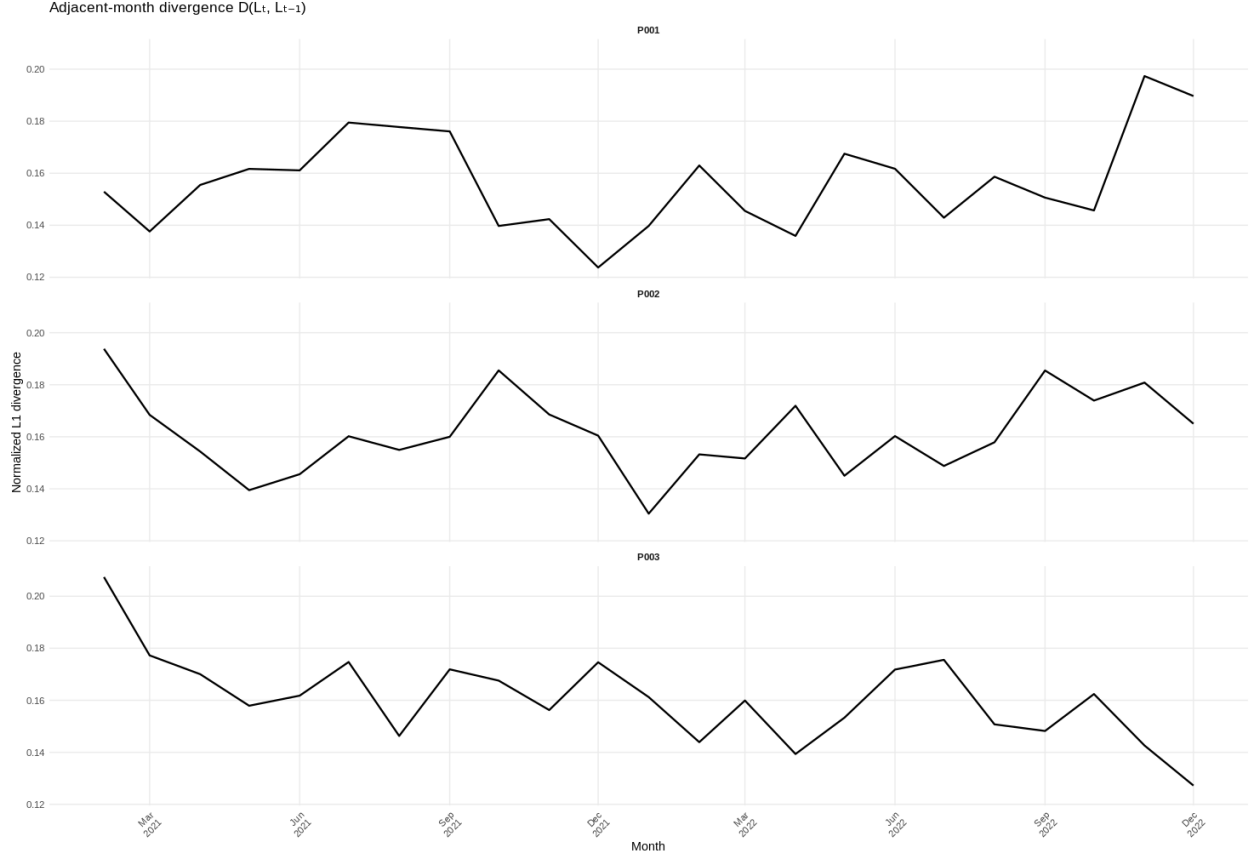


Figure 1: Adjacent month divergence $D(L_t, L_{t-1})$ by property. A value of 0.20 indicates that 20% of the mass in the lead time distribution moved relative to the prior month. Properties differ in both amplitude and pattern, which is expected under heterogeneous demand and policy environments.

The table reports the bound at 7, 14, and 21 days together with a simple, auditable mapping to operations. Months and properties with larger divergence and smaller $C_{\text{hist}}(\Delta)$ at the chosen horizon produce higher risk values and stricter actions. As the service date approaches the $(1 - \Delta/\Delta_{\text{max}})$ term contracts the bound monotonically.

6 Synthetic case studies

To keep everything reproducible, we include calibrated simulations that mimic the empirical patterns reported in the article. A compression parameter $c \in [0, 1]$ moves probability mass toward short horizons, which raises the divergence series and the pickup risk index in predictable ways. The R scripts generate arrivals across multiple properties and segments, compute D_t and $C_{\text{hist}}(\Delta)$, and apply the decision templates.

6.1 Data generating process

Lead times are drawn from a two component lognormal mixture. The compression parameter c shifts weight toward the short horizon component and therefore increases mass inside the last two weeks. Seasonality and event weeks can be introduced as multiplicative factors so that the

Table 1: Divergence summary by property (adjacent month). Values read from `paper_artifacts/tables/tbl3.divergence.summary.csv`.

Property	Months	Mean D	Median D	P90 D
P001	23	0.157	0.155	0.179
P002	23	0.162	0.160	0.185
P003	23	0.161	0.161	0.175

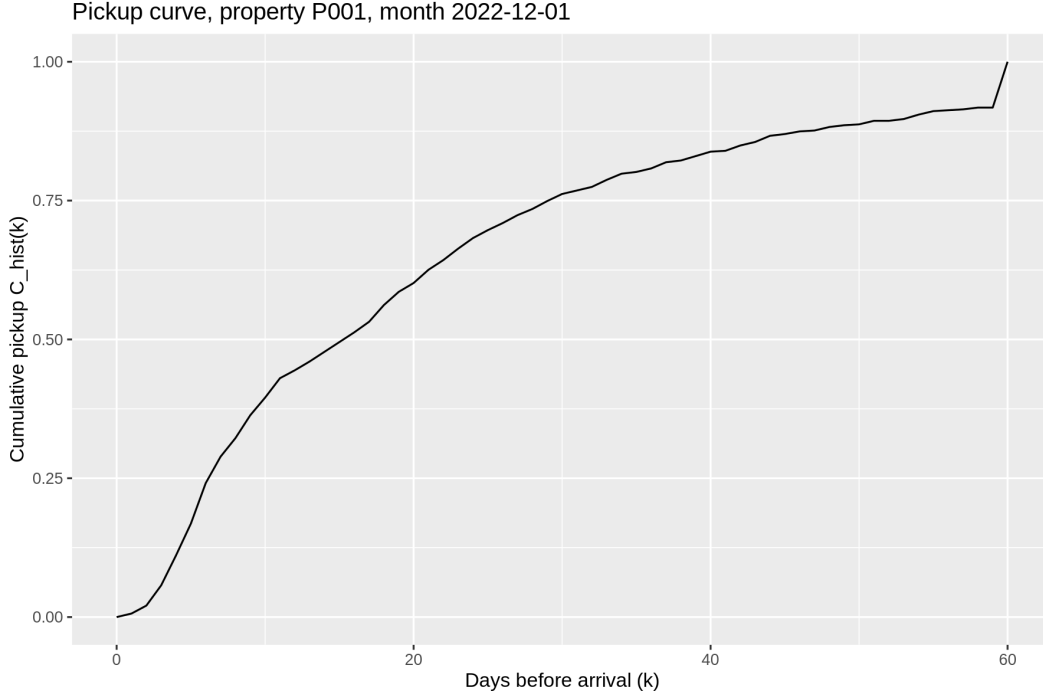


Figure 2: Cumulative pickup $C_{\text{hist}}(\Delta)$ for property P001, latest month.

divergence series reflects both steady changes and occasional spikes. Segment heterogeneity is introduced as random effects. All generators accept seeds so that figures are exactly reproducible.

6.2 Illustrative results

As compression rises, the divergence series increases and change points appear earlier in the period. The pickup risk index is largest when divergence is high and the historical pickup fraction is low at the chosen horizon. The templates therefore suggest a higher pricing cadence and looser advance purchase buffers in those months. As the service date approaches the bound contracts by construction and the recommended actions revert to a less intensive cadence.

7 Evaluation plan for operational settings

Effects can be estimated with standard designs that are robust to nonstationarity.

- **Difference in differences.** Compare units that adopt the monitoring and decision templates against matched units on a rolling basis, using horizon specific forecast metrics as outcomes

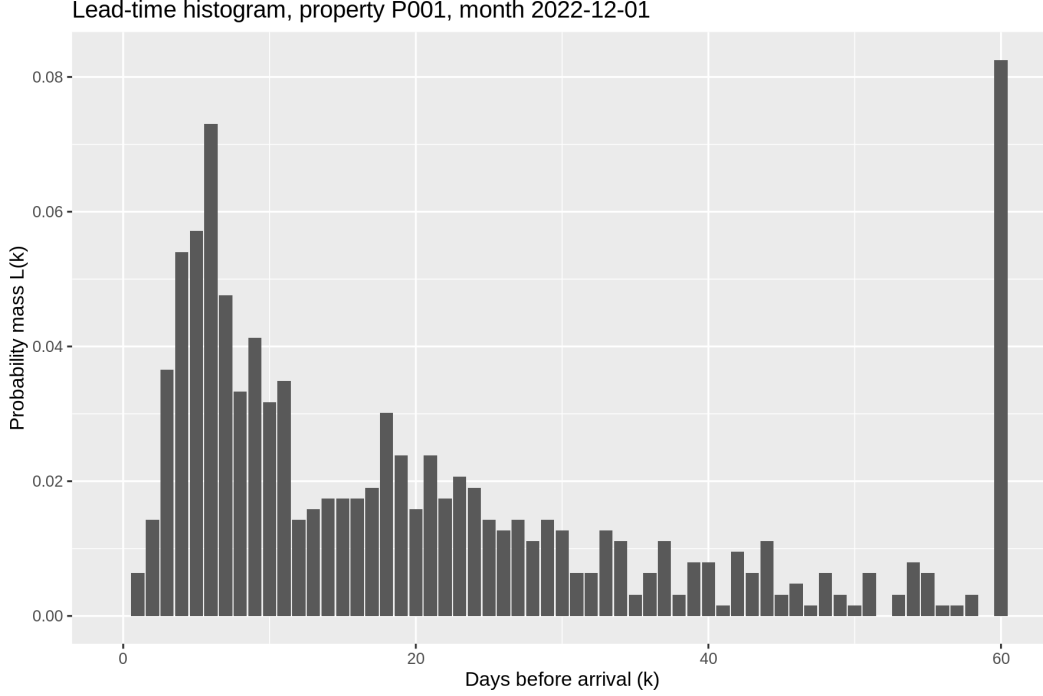


Figure 3: Lead time histogram $L_t(k)$ for property P001, latest month. The tall bar at $k = 60$ is a truncation artifact from capping Δ_{\max} at 60 in the simulation. In practice either increase Δ_{\max} or treat a right censored bin 60+ separately.

(Callaway and Sant’Anna, 2021).

- **Interrupted time series with synthetic control.** When the number of units is small, construct a counterfactual series for a treated unit and evaluate level and variance changes after the intervention (Abadie et al., 2010).
- **Experimentation when feasible.** For pricing cadence or advance purchase rules, randomize policies at the unit level and analyze horizon specific metrics.

Primary forecast outcomes are MASE and sMAPE at 0 to 7, 8 to 14, and 15 to 21 days. Operational outcomes include spoilage, denied service, and the variance of ADR or RevPAR.

8 Governance and reproducibility

All analysis runs locally on booking event data. No external connections are required. The package ships with a synthetic dataset and scripts that regenerate all figures. A brief executive summary reports the divergence series, the pickup risk index at standard horizons, and any actions taken. Versioned code and seeds allow third parties to reproduce results and extend simulations.

9 Limits

The divergence metric captures global shifts in the distribution shape. It does not attribute causes and it does not replace demand models. The pickup bound is conservative, especially when divergence is large and historical pickup is small at the chosen horizon. Users who require parametric

Table 2: Risk index and mapped actions at standard horizons for the latest month of each property. Values read from `paper_artifacts/tables/tbl2_risk_latest_month.csv`.

Property	Month	Delta (days)	C_hist	Bound	Price cadence	AP buffer (days)	Staffing buffer
P001	2022-12	7	0.289	1.086	intraday	0	15%
P001	2022-12	14	0.478	0.570	intraday	0	15%
P001	2022-12	21	0.625	0.369	intraday	0	15%
P002	2022-12	7	0.276	1.138	intraday	0	15%
P002	2022-12	14	0.521	0.522	intraday	0	15%
P002	2022-12	21	0.634	0.364	intraday	0	15%
P003	2022-12	7	0.260	1.208	intraday	0	15%
P003	2022-12	14	0.479	0.568	intraday	0	15%
P003	2022-12	21	0.611	0.377	intraday	0	15%

forecasts of the full lead time vector can adopt B-DARMA or related compositional time series models while still benefiting from divergence monitoring and the risk index (Katz et al., 2024). The decision templates are intentionally simple and are intended as safe defaults that can be tuned over time.

10 Conclusion

By turning two ideas from the article into a minimal R artifact that analysts can run on their own files, it becomes easier to detect distributional change, to translate that change into a clear risk index, and to connect the index to concrete actions. The same artifact defines how to measure effects in operational settings. This closes the loop between research and decision and sets the stage for comparative evaluation across markets and segments.

Declarations

Funding None declared.

Competing interests The authors declare no conflicts of interest and that all work and opinions are their own and that the work is not sponsored or endorsed by Airbnb.

Data availability All results are generated from simulations. A synthetic dataset and an R package archive are provided as supplementary material.

Code availability The R package `leadtimefluxR` and example scripts are provided as supplementary files and are available at <https://github.com/harrisonekatz/leadtimefluxR>.

Ethics approval, consent, and permissions Not applicable.

References

- Katz, H., Savage, E., Coles, P. (2025). Lead times in flux: Analyzing Airbnb booking dynamics during global upheavals (2018–2022). *Annals of Tourism Research Empirical Insights*, 6(2), 100185.
- Katz, H., Brusch, K. T., Weiss, R. E. (2024). A Bayesian Dirichlet auto regressive moving average

- model for forecasting lead times. *International Journal of Forecasting*, 40(4), 1556–1567. <https://doi.org/10.1016/j.ijforecast.2024.01.004>
- Cleveland, R. B., Cleveland, W. S., McRae, J. E., Terpenning, I. (1990). STL: A seasonal trend decomposition procedure based on loess. *Journal of Official Statistics*, 6(1), 3–73.
- Hyndman, R. J., Koehler, A. B. (2006). Another look at measures of forecast accuracy. *International Journal of Forecasting*, 22(4), 679–688.
- Gibbs, A. L., Su, F. E. (2002). On choosing and bounding probability metrics. *International Statistical Review*, 70(3), 419–435.
- Weatherford, L. R., Kimes, S. E. (2003). A comparison of forecasting methods for hotel revenue management. *International Journal of Forecasting*, 19(3), 401–415.
- Talluri, K. T., van Ryzin, G. J. (2004). *The Theory and Practice of Revenue Management*. Springer.
- Callaway, B., Sant’Anna, P. H. C. (2021). Difference in differences with multiple time periods. *Journal of Econometrics*, 225(2), 200–230.
- Abadie, A., Diamond, A., Hainmueller, J. (2010). Synthetic control methods for comparative case studies. *Journal of the American Statistical Association*, 105(490), 493–505.