Uncertainty and Assessment
Volume 1 · Number 1 · 2026

Shared Stations, Shared Predictability? A Route-Edge Count Study of Jersey City Bike Trips

Robert Ward[1] · Austin Reed[2] · Paul Lewis[3]

[1] Romania · [2] Bulgaria · [3] Serbia

Original publication: October 9, 2026 | Article 572

Abstract

A shared station can connect two routes without making their recent flows useful for forecasting. We converted January 2020 Jersey City trip records into hourly counts for thirty routes selected from the training period. Adding a lagged shared-endpoint neighbor summary to a Poisson regression produced a positive training association, but its seven-day holdout MAE was 0.37015 versus 0.36993 for the own-route baseline. The paired hourly difference interval crossed zero. Mean Poisson deviance moved slightly in the opposite direction. This case separates evidence of network association from evidence of an improvement on a declared forecasting loss.

Public-data analysis draft. Scientific interpretation, authorship and submission decisions require human review.

Shared Stations, Shared Predictability? A Route-Edge Count Study of Jersey City Bike Trips

Public-data analysis draft. Scientific interpretation, authorship and submission decisions require human review.

Abstract

A shared station can connect two routes without making their recent flows useful for forecasting. We converted January 2020 Jersey City trip records into hourly counts for thirty routes selected from the training period. Adding a lagged shared-endpoint neighbor summary to a Poisson regression produced a positive training association, but its seven-day holdout MAE was 0.37015 versus 0.36993 for the own-route baseline. The paired hourly difference interval crossed zero. Mean Poisson deviance moved slightly in the opposite direction. This case separates evidence of network association from evidence of an improvement on a declared forecasting loss.

1. Why a line graph needs an incremental-value test

Many transport observations naturally belong to connections. A trip from one docking station to another contributes to a directed route count, while a station-level total combines several different connections. Two routes sharing a station can face common demand, calendar rhythms, or availability constraints. Yet a shared endpoint does not establish that one route's recent count adds predictive information once the focal route's own memory and schedule are already represented.

We examine that incremental question using the operator's released Citi Bike trip histories (Citi Bike, 2026). The particular file contains Jersey City starts during January 2020. The broader bike-sharing synthesis in Fishman et al., 2013 explains why usage patterns have several possible behavioral and operational sources. Our analysis concerns recorded route counts and a specified short-horizon prediction task. It does not infer the number of people who wished to travel or the number prevented from doing so.

The research question is whether a lagged shared-station route-neighbor statistic improves prediction beyond own-route lags, route intercepts, and calendar structure. We distinguish two answers: whether its conditional training coefficient differs from zero, and whether the resulting frozen model improves held-out absolute error. A forecasting claim should rest on the second comparison rather than being inferred from the first.

The target paper, Yan and Kuruoglu, 2023, presents line-graph transformation as a way to represent time-varying edge signals with node-based estimation tools. Its public version discusses transport and meteorological applications. We apply the representation idea to route counts, using a conventional Poisson conditional-mean model instead of the target's adaptive least-mean-square procedure. The distinction keeps our empirical contribution specific: an incremental-information audit of a transparent edge-neighborhood feature.

2. Translating route connections into a predictor

A directed start-end pair is an edge in the station network. In our line graph, each selected directed route becomes a node. Two such nodes are adjacent when the routes share any endpoint; the relation is undirected even though each underlying route remains directed. The construction includes opposite-direction routes when they meet this rule. It ignores travel distance and does not treat a station map as a road-routing graph.

The line-graph discussion in Evans and Lambiotte, 2009 shows why representing links as nodes changes the object on which a network procedure operates. The broader graph-signal treatments in Shuman et al., 2012 and Ortega et al., 2017 likewise make the chosen graph part of the measurement definition. Here a shared station is an operationally intelligible relation, but it does not identify the direction of passenger transfer, bicycle redistribution, or a causal flow between routes.

Existing traffic forecasting work offers more elaborate spatial-temporal representations. Li et al., 2017 uses diffusion convolution with recurrent prediction, and Yu et al., 2017 develops a spatiotemporal graph-convolution framework. Those examples motivate asking whether a graph carries useful information; they do not establish that our simpler line-neighbor feature must help. No learned graph network or published traffic-model implementation is evaluated in this article.

We define P1 as the conditional association of the lagged line-neighbor summary with the next observed route-hour count in the training sample. P2 is the held-out augmented-minus-baseline MAE difference. These propositions use different units of evidence: fitted route-hour counts for P1, and paired hourly averages of prediction losses across the same thirty routes for P2. Figure 1 makes that distinction explicit.

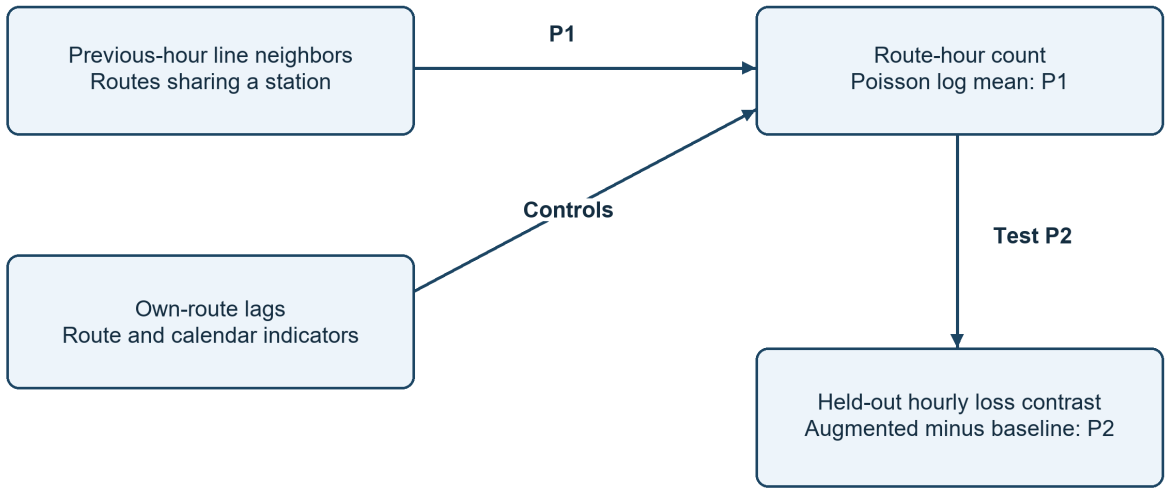

Figure 1. P1 adds the previous-hour shared-endpoint route mean to an own-route baseline. P2 evaluates the two frozen predictors on identical later route-hours. Calendar and route indicators are included in both training models.

3. From a monthly release to a defined count panel

The downloaded archive contains 26,020 trip rows. We read the start timestamp and start/end station identifiers for this analysis. We verify that every retained start belongs to January 2020 and assign it to its clock hour. An edge key preserves direction, so a route from station A to B and its reverse have distinct counts. We count starts rather than trip completions or minutes in transit.

Route selection uses only starts before 25 January. There are 20,093 rows in this selection period. We choose the thirty most frequent directed pairs, resolving frequency ties by a stable station-key ordering. The selected routes account for 6,938 starts across the full month. This focus supplies more repeated observations per route than an unrestricted panel, while conditioning the findings on relatively active connections in this particular month.

For those thirty routes, we form all 744 hours of January. An absent route-hour count becomes zero because the observation is the number of recorded starts in that cell. It does not become evidence that the station was open and able to serve every potential rider. Missing-data considerations in Little and Rubin, 2002 remain relevant to the interpretation of unrecorded operational conditions, even though the constructed count panel has no missing outcome cells.

The selected line graph has 146 undirected adjacencies, and each route has between two and fifteen neighbors. These are relationships among the selected thirty routes, not among every route appearing anywhere in the monthly file. We save the selected edge list and its construction rule so that a later expansion of the network cannot be mistaken for a harmless relabeling of the same predictor.

Taking a neighbor mean rather than a neighbor sum limits a mechanical degree effect: routes with more connections do not automatically receive larger feature values simply because more counts are added. The normalization does not make different neighborhoods equally informative. A mean based on two other routes has a different composition and potential sensitivity from a mean based on fifteen. Nor does a shared endpoint ensure that the neighboring starts precede the focal rider's decision. The previous-hour lag establishes data availability for prediction, while the adjacency relation establishes a descriptive network connection.

The first twenty-four days serve two roles: defining the active-route subset and estimating its count model. Consequently, the coefficient intervals are conditional on this selected route universe. They do not include uncertainty about which routes would be selected in another month. The held-out comparison still preserves its temporal boundary because no start from the final week chooses a route or a graph edge. A future test could add a separate earlier selection period, but that would answer a revised protocol with a different amount of training information. The present result should retain its actual two-stage use of the first part of January rather than being presented as a fully prespecified network sampled independently of the recorded flows.

4. Two count models on the same training rows

The baseline uses the previous-hour own-route count and the count at the same hour on the preceding day. Both enter as log one plus count. Route, hour-of-day, and day-of-week indicators provide stable route differences and calendar adjustment. We then add the log one plus previous-hour arithmetic mean count over the focal route's line neighbors. Every component uses observations available before the predicted hour.

Count regression provides a natural way to preserve a nonnegative conditional mean (Cameron and Trivedi, 1998). We fit a Poisson model with a log link. The response itself is not log-transformed, so observed zeros remain in the analysis. The pseudo-likelihood perspective discussed by Silva and Tenreyro, 2006 also cautions against treating a convenient logarithmic transformation of a nonnegative response as automatically equivalent to modeling its conditional mean.

Dropping the first twenty-four hours needed for the daily lag leaves 16,560 training route-hours, from 2 through 24 January. Both models use precisely those rows. Coefficients describe a conditional log mean. Because the predictors are log one plus counts, a coefficient is not a constant elasticity with respect to the untransformed count at every value. It should especially not be described as the effect of adding one bicycle or one trip to a neighboring route.

The models converge under the specified iteration limit. Their training Pearson dispersion ratios are approximately 1.30 and 1.28, so a literal equal-mean-and-variance count distribution is not an adequate universal description of these observations. We retain the conditional-mean comparison and calculate

covariance clustered by training day, rather than treating the model-based Poisson variance as a complete uncertainty estimate. Cluster guidance in Cameron and Miller, 2015 and heteroskedasticity analysis in White, 1980 support making that choice explicit.

5. Holding out a week rather than recycling a coefficient

The test period is 25 through 31 January, comprising 5,040 route-hours and 168 clock hours. Route selection, graph structure, and regression coefficients are frozen. At each test hour, actual observations from earlier test hours are available for lag construction. This is sequential one-hour-ahead prediction, not a seven-day recursive forecast in which earlier predicted counts are fed back as observations.

Temporal ordering follows the dependence-sensitive validation principles in Roberts et al., 2017. The model-selection concern in Varma and Simon, 2006 matters if the same week is repeatedly used to choose graphs or calendar specifications. We do not search such alternatives in this study. The short holdout remains a single realized week, however, and cannot establish stable performance across seasons or future operational changes.

We report MAE and mean Poisson deviance separately. Accuracy summaries require a stated loss and comparison unit (Hyndman and Koehler, 2006), while the scoring perspective in Gneiting and Raftery, 2007 explains why different losses can rank predictors differently. Neither criterion measures station availability, passenger welfare, or the quality of a generated transport video. A count-forecast metric should retain its own operational meaning.

For P2, we subtract baseline absolute error from augmented absolute error at every route-hour, then average the difference across routes within each hour. An intercept-only regression of these 168 hourly contrasts uses HAC covariance with twenty-four lags, a finite-sample adjustment, and t reference intervals (Newey and West, 1987). This treatment acknowledges hourly dependence without turning the 5,040 route-hour cells into independent replicates. Seven days still provide limited information about longer dependence patterns.

Table 1. Training coefficients are Poisson log-mean coefficients, with route, hour and day-of-week indicators; covariance clusters by 23 training days and uses t(22) intervals. Fit for count models is deviance explained. The final row model is an OLS mean of 168 held-out hourly MAE differences with HAC(24); its R-squared is zero by construction.

Model	Term	Estimate (SE)	95% CI	N
Own-route lags: fit	Deviance explained = 0.4073	-	-	16560
Own and line-neighbor lags	Log(1 + line-neighbor lag): P1	0.2781 (0.1244)	[0.0202, 0.5360]	16560
Own and line-neighbor lags: fit	Deviance explained = 0.4081	-	-	16560
Held-out mean augmented minus baseline MAE	Augmented minus baseline test MAE: P2	0.000224 (0.000456)	[-0.000675, 0.001124]	168
Held-out mean augmented minus baseline MAE: fit	R-squared = 0.0000	-	-	168

6. A training association with no clear MAE improvement

The augmented line-neighbor coefficient is 0.2781, with standard error 0.1244 and 95% interval from 0.0202 to 0.5360. Own-hour and own-day lag coefficients also remain positive in the augmented fit. Training deviance explained increases from 0.4073 to 0.4081. These quantities describe conditional association and fit on the stated training panel. They do not themselves answer whether the additional feature is useful in the held-out week.

Baseline holdout MAE is 0.36993 starts per route-hour, compared with 0.37015 for the augmented predictor. The augmented-minus-baseline difference is 0.000224, with HAC standard error 0.000456 and interval from -0.000675 to 0.001124. This uncertainty does not support a clear improvement on MAE. Failure to distinguish the contrast from zero also does not prove the predictors equivalent under a prespecified operational tolerance.

Mean Poisson deviance decreases slightly, from 0.71463 to 0.71375. We retain this different direction rather than choosing the favorable metric after seeing the result. Agreement analysis in Bland and Altman, 1986 helps explain why correlated predictions can still differ in their errors. Reliability concepts in Koo and Li, 2016 additionally require actual replication: the same monthly panel scored twice does not supply repeated independent deployment tests.

7. Explanations compatible with the observed result

Shared endpoints may capture a common commuting rhythm that the baseline already represents through route memory and calendar indicators. Other mechanisms remain plausible. The land-use analysis in Faghih-Imani et al., 2014 treats bicycle flows as related to their urban setting, and Gebhart and Noland, 2014 examines weather-dependent bikeshare use. We join neither land-use nor weather measurements here. Their possible roles therefore explain omitted information, not estimated mechanisms in our fitted coefficients.

Measured counts also combine demand with the ability to start a trip. Measurement-error principles in Carroll et al., 2006 discourage treating an operational proxy as a perfect latent demand measure. Causal-graph reasoning (Greenland et al., 1999) makes the shared causes of neighboring routes relevant to interpretation. Transport considerations in Stuart, 2010 further limit moving a January Jersey City comparison to another city merely because both systems use docking stations.

A future influence audit could use the diagnostic logic in Cook, 1977 to examine exceptional training days without suppressing inconvenient observations. If many graph definitions and lag combinations are subsequently tried, the inferential family changes. The different dependence conditions in Benjamini and Hochberg, 1995 and Benjamini and Yekutieli, 2001 would matter for a large collection of significance claims. Our present report supplies effect estimates and intervals for the declared comparison rather than selecting a collection of positive route-specific findings.

8. Reusable evidence and conclusion

The retained archive, hourly aggregation, graph definition, exact route selection, model specification and test predictions permit a direct recalculation. The documentation approach in Gebru et al., 2021 encourages preserving the difference between recorded starts and latent demand. The Wilkinson et al., 2016 principles support keeping source identifiers and reusable analysis metadata alongside those calculations. Neither documentation framework can extend the empirical scope beyond the release that was actually analyzed.

A lagged line-neighbor statistic has a positive conditional training association in this selected route network. Its addition does not clearly improve held-out MAE, and a small deviance change points in a different direction. The contribution is therefore an explicit separation of network association from demonstrated predictive value. A richer line-graph estimator remains worth evaluating under its own protocol, with longer holdouts, declared operational losses, and comparable information access.

References

1. Cook, R. Dennis. (1977). Detection of Influential Observation in Linear Regression. Technometrics. https://doi.org/10.1080/00401706.1977.10489493.
2. Greenland, Sander, Judea Pearl, and James M. Robins. “Causal Diagrams for Epidemiologic Research.” *Epidemiology*, vol. 10, no. 1, 1999, pp. 37-48. https://doi.org/10.1097/00001648-199901000-00008.
3. Li, Yaguang; Yu, Rose; Shahabi, Cyrus; Liu, Yan. (2017). Diffusion Convolutional Recurrent Neural Network: Data-Driven Traffic Forecasting. arXiv:1707.01926. https://arxiv.org/abs/1707.01926.
4. Gneiting, Tilmann; Raftery, Adrian E. (2007). Strictly Proper Scoring Rules, Prediction, and Estimation. Journal of the American Statistical Association. https://doi.org/10.1198/016214506000001437.
5. Little, Roderick J. A., and Donald B. Rubin. *Statistical Analysis with Missing Data*. Wiley, 2002. https://doi.org/10.1002/9781119013563.
6. Evans, T. S.; Lambiotte, R.. (2009). Line graphs, link partitions, and overlapping communities. Physical Review E. https://doi.org/10.1103/physreve.80.016105.
7. White, Halbert. “A Heteroskedasticity-Consistent Covariance Matrix Estimator and a Direct Test for Heteroskedasticity.” *Econometrica*, vol. 48, no. 4, 1980, pp. 817-838. https://doi.org/10.2307/1912934.
8. Silva, J. M. C. Santos; Tenreyro, Silvana. (2006). The Log of Gravity. The Review of Economics and Statistics. https://doi.org/10.1162/rest.88.4.641.
9. Benjamini, Yoav and Daniel Yekutieli. “The control of the false discovery rate in multiple testing under dependency.” *The Annals of Statistics*, vol. 29, no. 4, 2001, pp. 1165-1188. https://doi.org/10.1214/aos/1013699998.
10. Carroll, Raymond J., David Ruppert, Leonard A. Stefanski, and Ciprian M. Crainiceanu. “Measurement Error in Nonlinear Models.” 2006. https://doi.org/10.1201/9781420010138.
11. Gebru, Timnit, et al.. "Datasheets for datasets" *Communications of the ACM*, 2021. https://doi.org/10.1145/3458723.
12. Hyndman, Rob J. and Anne B. Koehler. “Another look at measures of forecast accuracy.” *International Journal of Forecasting*, vol. 22, no. 4, 2006, pp. 679-688. https://doi.org/10.1016/j.ijforecast.2006.03.001.
13. Faghih-Imani, Ahmadreza; Eluru, Naveen; El-Geneidy, Ahmed M.; Rabbat, Michael; Haq, Usama. (2014). How land-use and urban form impact bicycle flows: evidence from the bicycle-sharing system (BIXI) in Montreal. Journal of Transport Geography. https://doi.org/10.1016/j.jtrangeo.2014.01.013.
14. Newey, Whitney K., and Kenneth D. West. “A Simple, Positive Semi-Definite, Heteroskedasticity and Autocorrelation Consistent Covariance Matrix.” *Econometrica*, vol. 55, no. 3, 1987, pp. 703-708. https://doi.org/10.2307/1913610.
15. Stuart, Elizabeth A. "Matching Methods for Causal Inference: A Review and a Look Forward." *Statistical Science*, vol. 25, no. 1, 2010, pp. 1-21. https://doi.org/10.1214/09-sts313.
16. Wilkinson, Mark D.; Dumontier, Michel; Aalbersberg, IJsbrand Jan; Appleton, Gabrielle; Axton, Myles; Baak, Arie; Blomberg, Niklas; Boiten, Jan-Willem; da Silva Santos, Luiz Bonino; Bourne, Philip E.; Bouwman, Jildau; Brookes, Anthony J.; Clark, Tim; Crosas, Mercè; Dillo, Ingrid; Dumon, Olivier; Edmunds, Scott; Evelo, Chris T.; Finkers, Richard; Gonzalez-Beltran, Alejandra; Gray, Alasdair J.G.; Groth, Paul; Goble, Carole; Grethe, Jeffrey S.; Heringa, Jaap; ’t Hoen, Peter A.C; Hooft, Rob; Kuhn, Tobias; Kok, Ruben; Kok, Joost; Lusher, Scott J.; Martone, Maryann E.; Mons, Albert; Packer, Abel L.; Persson, Bengt; Rocca-Serra, Philippe; Roos, Marco; van Schaik, Rene; Sansone, Susanna-Assunta; Schultes, Erik; Sengstag, Thierry; Slater, Ted; Strawn, George; Swertz, Morris A.; Thompson, Mark; van der Lei, Johan; van Mulligen, Erik; Velterop, Jan; Waagmeester, Andra; Wittenburg, Peter; Wolstencroft, Katherine; Zhao, Jun; Mons, Barend. (2016). The FAIR Guiding Principles for scientific data management and stewardship. Scientific Data. https://doi.org/10.1038/sdata.2016.18.
17. Bland, J. Martin, and Douglas G. Altman. "Statistical Methods for Assessing Agreement between Two Methods of Clinical Measurement." *The Lancet*, vol. 327, no. 8476, 1986, pp. 307-310. https://doi.org/10.1016/S0140-6736(86)90837-8.

18. Fishman, Elliot; Washington, Simon; Haworth, Narelle. (2013). Bike Share: A Synthesis of the Literature. Transport Reviews. https://doi.org/10.1080/01441647.2013.775612.

19. Koo, Terry K., and Mae Y. Li. "A Guideline of Selecting and Reporting Intraclass Correlation Coefficients for Reliability Research." *Journal of Chiropractic Medicine*, vol. 15, no. 2, 2016, pp. 155-163. https://doi.org/10.1016/j.jcm.2016.02.012.

20. Benjamini, Yoav and Yosef Hochberg. "Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing." *Journal of the Royal Statistical Society Series B: Statistical Methodology*, vol. 57, no. 1, 1995, pp. 289-300. https://doi.org/10.1111/j.2517-6161.1995.tb02031.x.

21. Ortega, Antonio; Frossard, Pascal; Kovačević, Jelena; Moura, José M. F.; Vandergheynst, Pierre. (2017). Graph Signal Processing: Overview, Challenges and Applications. arXiv:1712.00468. https://arxiv.org/abs/1712.00468.

22. Varma, Sudhir and Richard Simon. "Bias in error estimation when using cross-validation for model selection." *BMC Bioinformatics*, vol. 7, no. 1, 2006. https://doi.org/10.1186/1471-2105-7-91.

23. Roberts, David R., et al. "Cross-validation strategies for data with temporal, spatial, hierarchical, or phylogenetic structure." *Ecography*, vol. 40, no. 8, 2017, pp. 913-929. https://doi.org/10.1111/ecog.02881.

24. Cameron, A. Colin, and Douglas L. Miller. "A Practitioner's Guide to Cluster-Robust Inference." *Journal of Human Resources*, vol. 50, no. 2, 2015, pp. 317-372. https://doi.org/10.3368/jhr.50.2.317.

25. Cameron, A. Colin; Trivedi, Pravin K.. (1998). Regression Analysis of Count Data. Cambridge University Press. ISBN 9780521632010. https://www.cambridge.org/9780521632010.

26. Shuman, David I; Narang, Sunil K.; Frossard, Pascal; Ortega, Antonio; Vandergheynst, Pierre. (2012). The Emerging Field of Signal Processing on Graphs: Extending High-Dimensional Data Analysis to Networks and Other Irregular Domains. arXiv:1211.0053. https://arxiv.org/abs/1211.0053.

27. Citi Bike, . (2026). System Data: Citi Bike Trip Histories. January 2020 Jersey City archive; accessed 6 November 2026. https://citibikenyc.com/system-data.

28. Yu, Bing; Yin, Haoteng; Zhu, Zhanxing. (2017). Spatio-Temporal Graph Convolutional Networks: A Deep Learning Framework for Traffic Forecasting. arXiv:1709.04875. https://arxiv.org/abs/1709.04875.

29. Gebhart, Kyle; Noland, Robert B.. (2014). The impact of weather conditions on bikeshare trips in Washington, DC. Transportation. https://doi.org/10.1007/s11116-014-9540-7.

30. Yan, Yi; Kuruoglu, Ercan Engin. (2023). Adaptive Spatio-temporal Estimation on the Graph Edges via Line Graph Transformation. arXiv:2311.00656. https://arxiv.org/abs/2311.00656.