# Distance Decay in Online Deliveries and the Sustainability Impact of Pricing Incentives

Javad Eshtiyagh[1,3], Anupriya Anupriya[1], Daniel Hörcher[1,2], Paolo Santi[3,4], and Daniel J. Graham*[1]

[1]Transport Strategy Centre, Imperial College London, London, UK
[2]Institute for Advanced Studies, Corvinus University of Budapest, Budapest, Hungary
[3]Senseable City Lab, Massachusetts Institute of Technology, Cambridge MA, USA
[4]Istituto di Informatica e Telematica del CNR, Pisa, Italy

## Short Summary

The rapid expansion of last-mile delivery has reshaped urban logistics, yet how delivery flows vary with distance remains poorly understood. Using approximately 48 million delivery records from multiple platforms across 12 cities on four continents, we provide the first large-scale, multi-city empirical estimation of flow-distance relationships in last-mile delivery. Semi-parametric spline and Gaussian process regression models reveal a universal exponential decay of express delivery flows with distance, while scheduled parcel deliveries exhibit near-uniform distance distributions. We further show that delivery fee discounts increase average ordering distances by 14%, with lower-income customers responding most strongly. These findings demonstrate that the spatial and environmental externalities of urban delivery are behaviorally responsive to pricing strategies, with significant implications for sustainable urban logistics policy.
**Keywords**: distance decay; gravity model; last-mile delivery; pricing incentives; sustainability; urban logistics

## 1 Introduction

The sustained growth of e-commerce has made express last-mile delivery a defining feature of contemporary cities. Instant and same-day delivery services have expanded rapidly, driven by intensified platform competition and rising consumer expectations. While these services provide substantial convenience, they generate mounting pressure on urban road networks, with consequences for spatial planning, emission control, and congestion management (Ranieri et al., 2018; Bosona, 2020). The World Economic Forum projects that carbon emissions from urban deliveries will increase by approximately 60% globally by 2030, potentially accounting for up to 54% of transport-sector emissions and around 13% of total city emissions (Gibbs et al., 2024). Understanding and governing these impacts requires reliable, empirically grounded models of delivery flows.

Despite the growing importance of express delivery, our knowledge of how delivery flows are distributed across distances, and how they scale with distance and travel time in urban environments, remains limited. The theoretical literature has modeled interactions among customers, couriers, platforms, and vendors, emphasizing pricing, matching, and waiting times as key decision variables (Bai et al., 2019; Ke et al., 2024). A key limitation of this work is its predominantly theoretical nature and its treatment of order travel distance as an exogenous factor determined solely by the geographic distribution of customers and vendors.

The empirical literature on online delivery demand has improved our understanding of who orders online and where demand concentrates (Su et al., 2025; Barua et al., 2023). However, it does not provide generalizable models of delivery flows or quantify how far goods travel within cities. This stands in contrast to passenger mobility research, where large-scale data have revealed stable exponential or power-law decay patterns across cities and trip purposes (González et al., 2008; Schläpfer et al., 2021). Such findings suggest that aggregate urban flows may be governed by stable spatial interaction mechanisms regardless of mode.

This paper addresses two interrelated objectives. First, we estimate the functional form governing the relationship between delivery flows and distance across different delivery types and urban

contexts. Second, we quantify the extent to which delivery fee pricing incentives alter these spatial patterns, and assess the implied environmental consequences. Together, these contributions provide a new empirical foundation for modeling and governing the growth of express delivery in cities.

## 2 Methodology

### *Data*

This study draws on four primary datasets totalling approximately 48 million delivery records spanning 12 major cities across West Asia, East Asia, Africa, and North America.
*Platform #1* consists of approximately 39 million food, grocery, and parcel delivery orders from a major on-demand delivery platform in Dubai (2023). The dataset includes precise vendor and customer coordinates, timestamps, fees, vendor- and platform-offered discounts, and user identifiers, covering approximately 1.5 million unique customers and 12,000 unique vendors.
*Platform #2* provides approximately 5 million online grocery orders from one of the world's largest supermarket chains, across six cities: Dubai, Abu Dhabi, Doha, Riyadh, Cairo, and Nairobi (2022). Orders are matched to the nearest store branch prior to spatial aggregation.
*Platform #3* draws on the LaDe dataset (Wu et al., 2025), comprising approximately 10.7 million express and scheduled package deliveries from Shanghai, Hangzhou, and Chongqing over a six-month period.
*Platform #4* uses a publicly available Amazon dataset (Merchán et al., 2024) containing 9,100 scheduled parcel deliveries in Boston, Chicago, and Los Angeles (2018), used primarily for comparison with express delivery patterns.
Several auxiliary datasets supply socioeconomic and urban form controls, including population, median rent, vendor density, and neighborhood income distributions (Bazant, 2022; OpenStreetMap, n.d.). A key derived variable is vendor entropy, computed as the Shannon entropy of the vendor distribution within each spatial unit: $-\sum_v p_v \log p_v$, where $p_v$ denotes the share of orders associated with vendor $v$. All variables are harmonized to a consistent spatial resolution using Uber's H3 hexagonal grid (Uber, 2018).

### *Gravity Model and Empirical Specification*

The analysis is grounded in the gravity model of spatial interaction (Haynes & Fotheringham, 1984), which posits that the flow between two locations decreases with distance and increases with the attractiveness of origins and destinations. We adopt a semi-parametric direct estimation approach that allows the shape of the distance decay to be estimated from data without imposing a prior functional form. The log-linearized specification is:

$$\log(Q_{ijt}) = f(d_{ijt}) + \sum_m \beta_m^{\text{origin}} \phi_{im} + \sum_n \beta_n^{\text{dest}} \psi_{jn} + u_i + v_j + \epsilon_{ijt} \tag{1}$$

where $Q_{ijt}$ is the observed flow from origin cell $i$ to destination cell $j$ at time $t$; $d_{ijt}$ is the Euclidean distance between vendor and customer; $f(\cdot)$ is a smooth, non-parametric distance decay function estimated via penalized spline or Gaussian process regression; $\phi_{im}$ and $\psi_{jn}$ are origin and destination covariates; and $u_i$, $v_j$ are spatial random effects capturing unobserved heterogeneity. Origin covariates include vendor count, vendor diversity, mean item prices, and average delivery fee; destination covariates include local population, median rent, and nearby vendor diversity. Estimation focuses on the 1–20 km range, which captures the vast majority of express orders while avoiding sparse tails. Results are normalized as shares of total flow to allow comparison across cities with differing order volumes.
Robustness is assessed through alternative travel cost metrics (Euclidean, network distance, travel time), varying spline knot counts, and alternative smoothing approaches including B-splines, natural cubic splines, and Gaussian process regression with a Matérn kernel. Across all specifications, findings remain qualitatively and quantitatively consistent.

### *Impact of Pricing Incentives*

To quantify how delivery fee discounts affect ordering distance, we conduct a follow-up analysis on a subset of 361,938 frequent Platform #1 users who placed both discounted and non-discounted orders during the study period. Analysis is restricted to vendor-offered (non-targeted) discounts

to avoid endogeneity from platform-level algorithmic targeting. A customer fixed-effects regression is estimated:

$$\log(d_{zk}) = \alpha_0 + \alpha_1 \kappa_z + \sum_l \alpha_l^{\text{order}} \eta_{zl} + \omega_k + \epsilon_{zk} \tag{2}$$

where $d_{zk}$ is the distance traveled by order $z$ from customer $k$; $\kappa_z$ is a binary indicator for a vendor-offered delivery fee discount; $\eta_{zl}$ captures order-level controls including transaction value, average delivery fee for comparable orders, temporal dummies, and product category; and $\omega_k$ denotes customer fixed effects. The coefficient of interest, $\alpha_1$, measures the average within-customer percentage change in delivery distance attributable to fee discounts. Heterogeneity is examined by repeating the analysis across income quartiles and baseline ordering distance quartiles.

## 3 Results and Discussion

### *Spatial Structure of Express Delivery Flows*

Across all nine express delivery datasets (three platforms, up to 12 cities), distance is a robust and highly significant predictor of flow intensity, jointly significant at the 1% level across all model specifications. After controlling for origin and destination characteristics and spatial random effects, the adjusted $R^2$ reaches 0.87 in the Dubai dataset, the largest in the sample at over 39 million orders.

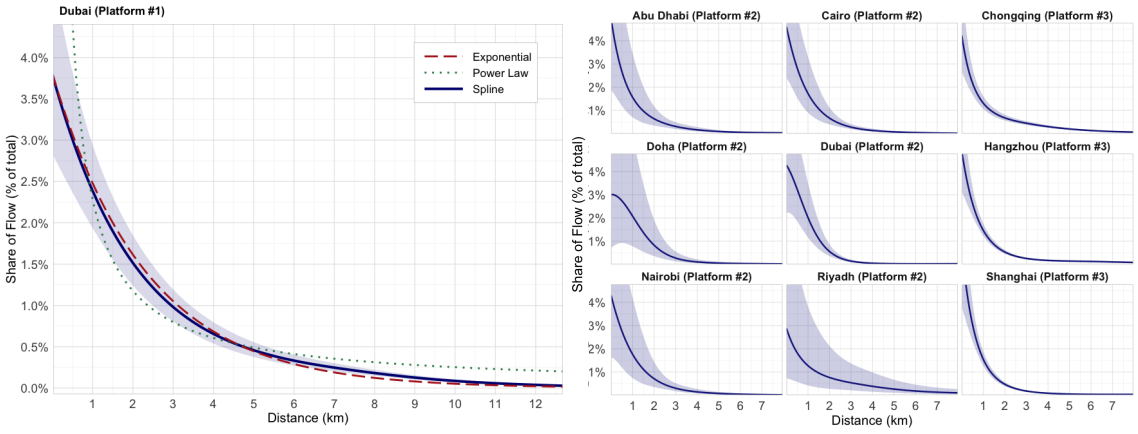



Figure 1: Estimated flow distribution of express orders with 95% confidence intervals, derived from a semi-parametric regression of the distance function without prior assumptions about its functional form. The share of flow represents the percentage of orders traveling at each distance. The left panel shows results for the full model specification estimated on the comprehensive dataset from Platform #1 in Dubai, along with fitted exponential and power-law curves. The smaller panels show results for Platforms #2 and #3 in other cities.

Without imposing any prior functional form, the estimated distance decay function $f(d)$ consistently takes the shape of an exponential decay across all cities and platforms. Compared to power-law alternatives, the exponential form provides a markedly better fit across the 1–20 km range. This pattern implies a roughly constant incremental cost associated with additional travel distance, and is consistent with exponential decay patterns documented for passenger mobility across diverse global contexts (Schläpfer et al., 2021). While the exponential form is universal, decay rates vary across cities, reflecting differences in urban form, market penetration, platform type, and competitive environment.

Within Dubai, lower-income neighborhoods exhibit steeper decay parameters after controlling for other covariates, suggesting higher price sensitivity among lower-income residents. Areas with greater nearby vendor variety, measured by cuisine-type entropy, also exhibit faster decay, indicating that customers are more likely to order from nearby when a wide range of options is locally

available. These patterns hold consistently across different seasons, days of the week, and times of day, confirming the robustness of the spatial interaction structure.
Scheduled parcel deliveries (Platform #4, Amazon US) display a fundamentally different pattern. Applying the same semi-parametric gravity framework, the flow-distance relationship for scheduled deliveries is approximately uniform with respect to distance after controlling for spatial covariates. This reflects the centralized fulfillment model, in which two to five peripheral warehouses each serve entire metropolitan areas regardless of customer proximity. Unlike express orders, which originate from thousands of vendors distributed throughout the city, scheduled deliveries structurally decouple consumer demand from travel distance, systematically extending average trip lengths and diffusing spatial externalities across the urban area.

***Impact of Pricing Incentives***

The presence of vendor-offered delivery fee discounts significantly alters the spatial structure of express orders. Applying the flow-distance estimation separately to discounted and non-discounted orders shows that discounted orders display a rightward shift in the distance distribution, with the share of orders traveling more than 2.2 km increasing and the effect becoming statistically significant beyond 5 km.
The fixed-effects regression confirms that delivery fee discounts increase average ordering distance by 14% ($\alpha_1 > 0$, statistically significant at the 1% level, robust across all specifications). This finding directly contradicts the common assumption in the delivery literature that order travel distance is exogenously determined by geography alone.
Heterogeneity analysis reveals important differences across income groups. Customers in medium- and high-income neighborhoods increase their ordering distance by 10.2–14.7% in response to discounts, while customers in the bottom income quintile extend their distances by an average of 19.3%. This indicates that lower-income customers are considerably more price-sensitive and place higher relative value on delivery fee savings. Additionally, customers who habitually order from far locations respond less to discounts (approximately 5.9% increase in the top distance quartile), while the effect for shorter-range habitual buyers ranges from 12.8% to 18.8%.

***Environmental Implications***

The behavioral response to fee incentives carries measurable environmental consequences. Scaling the 14% distance increase from Platform #1 to Dubai's broader delivery market by platform market share, and applying it to approximately 24% of orders (those receiving discounts), implies 8.2–10.3 million additional vehicle-kilometers per year citywide. Combining direct fuel consumption, road congestion externalities, and lifecycle fleet emissions, and applying conservative emission factors from the transportation literature, this translates to a minimum increase of 1,376 tonnes of $CO_2$ per year in Dubai alone, equivalent to the annual carbon absorption of approximately 62,000 mature trees.
These estimates highlight a critical and underappreciated pathway through which pricing strategies interact with urban sustainability. When platforms or vendors obscure delivery costs through fee subsidies or "free delivery" embedded in product prices, they weaken consumers' natural distance sensitivity, encouraging more spatially dispersed delivery patterns that amplify vehicle-kilometers traveled and associated emissions.

## 4 Conclusions

This study provides the first large-scale, multi-city empirical analysis of flow-distance relationships in last-mile delivery, drawing on approximately 48 million records across 12 cities and four continents. The key findings are threefold.
First, express delivery flows exhibit a universal exponential decay with distance across all studied cities and platforms. This pattern is robust to differences in urban form, socioeconomic composition, and platform type, and mirrors spatial interaction regularities documented in passenger mobility. Decay rates vary across cities, but the functional form is universal.
Second, scheduled parcel deliveries from centralized facilities exhibit a fundamentally different, near-uniform distance distribution. This structural difference reflects the decoupling of consumer demand from travel distance in centralized logistics systems, generating unpriced and often overlooked environmental and infrastructural burdens.

Third, and most critically for policy, delivery fee incentives significantly reshape spatial ordering patterns. Fee discounts increase average ordering distances by 14%, with the largest responses observed in lower-income neighborhoods (19.3%). This demonstrates that the spatial and environmental externalities of last-mile delivery are, at least in part, endogenous to pricing strategies rather than fixed by geography.

Together, these findings point to a central policy challenge: market competition drives platforms to subsidize and obscure delivery costs, while the resulting spatial patterns impose unpriced externalities on cities. A regulatory approach that acknowledges this behavioral elasticity is essential. Fees calibrated to reflect the true social cost of distance, incorporating emissions and congestion externalities, could harness consumers' price sensitivity to reduce vehicle-kilometers and associated emissions. Future work will extend this analysis to additional cities and explore optimal pricing mechanisms for internalizing spatial externalities in urban delivery networks.

## References


Bai, J., So, K. C., Tang, C. S., Chen, X., & Wang, H. (2019, July). Coordinating Supply and Demand on an On-Demand Service Platform with Impatient Customers. *Manufacturing & Service Operations Management*, *21*(3), 556–570. doi: 10.1287/msom.2018.0707

Barua, L., Zou, B., Zhou, Y., & Liu, Y. (2023, April). Modeling household online shopping demand in the U.S.: a machine learning approach and comparative investigation between 2009 and 2017. *Transportation*, *50*(2), 437–476. Retrieved 2025-10-11, from `https://doi.org/10.1007/s11116-021-10250-z` doi: 10.1007/s11116-021-10250-z

Bazant, S. (2022, October). A Geospatial Study of High Net Worth Individuals in Dubai. *Webster Pacific*. Retrieved 2025-12-08, from `https://websterpacific.com/october-18,-2022:-a-geospatial-study-of-high-net-worth-individuals-in-dubai/`

Bosona, T. (2020, January). Urban Freight Last Mile Logistics—Challenges and Opportunities to Improve Sustainability: A Literature Review. *Sustainability*, *12*(21), 8769. Retrieved 2025-11-11, from `https://www.mdpi.com/2071-1050/12/21/8769` doi: 10.3390/su12218769

Gibbs, A., Wylie, J., Arora, P., Baweja, V., Fröbel, M., & Risgallah, V. (2024). Transforming Urban Logistics in Cities. *World Economic Forum*. (https://reports.weforum.org/docs/WEF_Transforming_Urban_Logistics_2024.pdf)

González, M. C., Hidalgo, C. A., & Barabási, A.-L. (2008, June). Understanding individual human mobility patterns. *Nature*, *453*(7196), 779–782. Retrieved 2025-10-15, from `https://www.nature.com/articles/nature06958` doi: 10.1038/nature06958

Haynes, K. E., & Fotheringham, A. S. (1984). *Gravity and Spatial Interaction Models*. Sage. (Chapter 6)

Ke, J., Wang, C., Li, X., Tian, Q., & Huang, H.-J. (2024, April). Equilibrium analysis for on-demand food delivery markets. *Transportation Research Part E: Logistics and Transportation Review*, *184*, 103467. Retrieved 2025-02-28, from `https://www.sciencedirect.com/science/article/pii/S1366554524000577` doi: 10.1016/j.tre.2024.103467

Merchán, D., Arora, J., Pachon, J., Konduri, K., Winkenbach, M., Parks, S., & Noszek, J. (2024, January). 2021 Amazon Last Mile Routing Research Challenge: Data Set. *Transportation Science*, *58*(1), 8–11. Retrieved 2025-11-14, from `https://pubsonline.informs.org/doi/10.1287/trsc.2022.1173` doi: 10.1287/trsc.2022.1173

OpenStreetMap. (n.d.). *OpenStreetMap*. Retrieved 2025-12-08, from `https://www.openstreetmap.org/` (https://www.openstreetmap.org/)

Ranieri, L., Digiesi, S., Silvestri, B., & Roccotelli, M. (2018, March). A Review of Last Mile Logistics Innovations in an Externalities Cost Reduction Vision. *Sustainability*, *10*(3), 782. Retrieved 2025-11-11, from `https://www.mdpi.com/2071-1050/10/3/782` doi: 10.3390/su10030782

Schläpfer, M., Dong, L., O'Keeffe, K., Santi, P., Szell, M., Salat, H., ... West, G. B. (2021, May). The universal visitation law of human mobility. *Nature*, *593*(7860), 522–527. Retrieved 2025-10-16, from `https://www.nature.com/articles/s41586-021-03480-9` doi: 10.1038/s41586-021-03480-9

Su, R., Eshtiyagh, J., Santi, P., Hu, S., Duarte, F., Mazzarello, M., & Ratti, C. (2025, September). Determinants of the localized behavior of individual food delivery choices. *Transportation Research Part C: Emerging Technologies*, *178*, 105185. Retrieved 2025-10-11, from `https://www.sciencedirect.com/science/article/pii/S0968090X25001895` doi: 10.1016/j.trc.2025.105185

Uber. (2018, June). *H3: Uber's Hexagonal Hierarchical Spatial Index.* (https://www.uber.com/en-GB/blog/h3/)

Wu, L., Wen, H., Hu, H., Mao, X., Xia, Y., Shan, E., ... Wan, H. (2025, January). *LaDe: The First Comprehensive Last-mile Delivery Dataset from Industry.* arXiv. Retrieved 2025-11-14, from `http://arxiv.org/abs/2306.10675` doi: 10.48550/arXiv.2306.10675