# The Consumer City Revisited: Consumption Responses to Real-Time Population Presence*

David Bounie[1], Chloé Breton[1,2], John Galbraith[3], and Gabrielle Gambuli[1,2]

[1]Télécom Paris, CREST
[2]Insee, SSPLab
[3]McGill University and CIRANO

March 20, 2026

**Abstract**

This paper combines high-frequency mobile phone location data with card transaction records, to examine the relationship between the number of individuals present in a zone ('real-time population') and economic transactions. Using data from the metropolitan area of Lyon, France's second-largest urban area, and a Poisson Pseudo-Maximum Likelihood estimator, we estimate the elasticity of transactions with respect to a real-time count of individuals present in each of over one thousand zones. We find (1) substantial temporal and geographical heterogeneity in consumption responsiveness, with elasticities ranging from 1.04 in the urban core to 0.84 in peripheral areas and peaking at 1.08 during Saturday midday, with elasticities declining to 0.98 on Sunday; (2) persistent spatial frictions, with transaction flows declining by about 2% for a 1% increase in distance from home; (3) a goods-services dichotomy, where goods, especially essentials such as retail food purchases, exhibits near-unit elasticity while experiential services show significantly lower responsiveness.

*Keywords*: Transaction data, mobile phone data, mobility, real-time population

*JEL code:* R12

*This research is supported by the ANR MobiTIC project (grant number ANR-19-CE22-0010), a collaborative initiative involving Orange, Université Gustave Eiffel, and Insee. This research has also been conducted within the Research Chair Digital Finance under the aegis of the Risk Foundation, a joint initiative by Groupement des Cartes Bancaires CB, Insee, Caisse des Dépôts, Télécom Paris and University Paris-Panthéon-Assas. We are grateful to Romain Lesur, Mélina Hillion, Loÿs Moulin, Katérina Levallois, and Samuel Willy for their insightful comments on earlier drafts. We also thank participants at the Barcelona School of Economics Summer Forum (2025), the International Workshop in Spatial Econometrics and Statistics (SEW, 2025), the Association Française de Sciences Économiques (AFSE, 2025), Eurostat's MNO-MINDS Final Conference (OECD, 2025) and the Association Française d'Économie Numérique (AFREN, 2025) for their valuable feedback.

# 1 Introduction

The economic advantages of urban density have long been studied through two complementary lenses. The first, rooted in the seminal work of Duranton and Puga (2004), focuses on supply-side productivity gains—infrastructure sharing, worker-firm matching, and knowledge spillovers—as the primary drivers of urban concentration. Yet this production-centered perspective overlooks a fundamental aspect of urban life: households choose locations not only for employment opportunities but also for access to goods, services, and amenities that enhance well-being. This second perspective, formalized by Glaeser et al. (2001) through the "consumer city" concept, positions consumption externalities as a core mechanism of urban agglomeration. While the productivity benefits of density are now well-documented (Ahlfeldt and Pietrostefani, 2019), the consumption-side channels through which density generates economic activity remain poorly understood (Henderson and Thisse, 2024), a critical gap recently emphasized by Donovan et al. (2024), who argue that "to develop a fuller understanding of urban economic advantages, we advocate for more research into agglomeration economies in consumption."

This paper provides direct empirical measurement of the degree to which real-time population presence (observed presence of individuals) translates into economic transactions via on-site commercial consumption,[1] and the way in which spatial frictions shape these interactions. We combine two novel datasets at an unprecedented level of granularity: (1) mobile phone records capturing real-time population across 1,151 Iris zones in the Lyon metropolitan area, with 30-minute resolution over a 28-day period in September 2022;[2] and (2) comprehensive card transaction records documenting the exact timing, value, and location of economic activity.[3] By merging these datasets at the 30-minute interval and with fine geographic resolution, we construct a unique panel that links consumer presence to economic outcomes with both spatial and temporal precision.

To estimate the elasticity of transactions with respect to population presence, we employ a Poisson Pseudo-Maximum Likelihood (PPML) estimator (Gourieroux et al., 1984; Silva and Tenreyro, 2006). This approach is particularly well suited to our data for three reasons. First, it accommodates the zeroes in our high-frequency data, where many zone-time combinations exhibit

[1] We study on-site commercial consumption of goods and services: e.g. shopping in stores, dining in restaurants and bars, and attending theaters and other cultural venues.

[2] Lyon is France's second-largest functional urban area, characterized by a dense central city, intermediate suburbs, and rural outer suburbs. Iris zones, defined below, are similar to census block groups.

[3] These data were made available through a partnership with Cartes Bancaires CB (hereafter 'CB'), with all analyses conducted in compliance with EU General Data Protection Regulation (Article 89).

no transactions. Second, it remains consistent even when the error distribution deviates from the Poisson assumption, addressing potential heteroskedasticity in our panel data. Third, the PPML framework allows us to incorporate comprehensive fixed effects, including zone and date-hour fixed effects, to control for unobserved heterogeneity across spatial units and temporal periods.

We further extend this framework to a micro-gravity model, adapting the traditional gravity equation (Anderson and van Wincoop, 2003; Head and Mayer, 2014) to quantify the degree to which distance and travel time deter economic interactions at the intra-urban scale. This dual approach, estimating both presence-transaction elasticities and spatial decay patterns, enables us to address two critical questions: (1) How does the presence of individuals generate economic activity in real time? and (2) How sensitive are economic transactions to intra-urban travel costs?

Our analysis yields five key findings that reshape our understanding of urban economic dynamics. First, both transaction count and total value exhibit near-unity elasticities with respect to presence (estimated as 0.99 and 0.93, respectively), with overlapping confidence intervals. This suggests that agglomeration influences the extensive margin of economic activity similarly, whether measured as number or value of transactions; we do not find a statistically significant difference in response to presence counts.

Second, our analysis uncovers substantial temporal heterogeneities in consumption responsiveness to urban density. The elasticity of transactions to population presence reaches 1.08 during Saturday midday peaks, double the 0.54 observed during morning commuting hours. Conversely, we estimate that Sundays exhibit 13% lower responsiveness than weekdays, indicating a shift toward non-commercial activities. By documenting substantial temporal heterogeneities in consumption responsiveness to population density, we quantify for the first time the fluctuations in the relation between urban density and commercial transactions with the rhythm of daily and weekly cycles, a dimension entirely missing from previous aggregate analyses. In particular, the large intra-day differences introduce an important qualification to previous results, computed on daily or weekly data, that find mobility data to be a good proxy for economic activity.

Third, we provide direct empirical evidence of diminishing response of transactions to counts of individual presence across spatial dimensions. Consumption elasticities decline systematically from 1.04 in the urban core to 0.84 in rural suburbs. This spatial gradient is compatible with the idea that central locations offer greater variety and more efficient consumer-provider matching, while peripheral areas suffer from thinner markets and higher search costs, and is consistent with spatial heterogeneity in consumption responsiveness. This spatial gradient corrobo-

rates the monocentric city model of Alonso-Muth-Mills (Henderson and Thisse, 2024), where agglomeration benefits are theoretically predicted to concentrate in central areas. By extending this framework to consumption behavior, we provide a missing link between spatial economics and consumer-level outcomes.

Fourth, our estimates point to a dichotomy between goods and service sectors. We find that goods-oriented sectors (e.g., food retailing, with an elasticity near unity) tend to convert population presence into economic transactions to a higher degree than do service sectors such as entertainment and cultural activities. This unexpected result suggests that goods sectors reduce consumers' search costs more effectively through greater variety and proximity, whereas service sectors may suffer from capacity constraints (e.g., limited seating) and congestion effects (e.g., waiting times), which diminish their ability to capitalize on density advantages. This finding re-conceptualizes the mechanisms of agglomeration economies, indicating that access to goods, rather than services, emerges as the primary channel through which urban density stimulates economic activity. It contradicts the prevailing assumption in the literature, notably by Glaeser et al. (2001), which posits that service sectors benefit disproportionately from urban density due to their reliance on social interactions.

Finally, we quantify persistent spatial frictions at the intra-urban scale, showing that a 1% increase in distance from home (or travel time) reduces transaction flows by 2.1% (2.4%). This result emerges from a Poisson Pseudo-Maximum Likelihood (PPML) estimator (with origin-day and destination-day fixed effects), a methodological approach uniquely suited to our zero-inflated, high-frequency mobility data. By leveraging this framework, we find evidence that intra-urban economic interactions remain highly sensitive to physical proximity, even in the digital age, a finding that complete the broader mobility trends observed at larger geographic scales (Bounie et al., 2026). Our micro-gravity analysis uncovers a critical but overlooked dimension of urban economic activity: localized travel frictions, nonlinear costs, and the availability of close substitutes collectively dampen transaction flows within cities, even over short distances. This persistence of spatial constraints underscores that physical proximity continues to dominate economic interactions at the micro-geographic level. These findings highlight the need for policies targeting localized barriers to economic interaction, an aspect often neglected in macro-level analyses but crucial for understanding the uneven economic performance of urban neighborhoods.

The remainder of the paper is organized as follows. Section 2 reviews the relevant literature on consumer mobility measurement and urban economic analysis and the links of the present paper

with this literature. Section 3 describes the unique dataset and study area characteristics. Section 4 presents the econometric framework for estimating transaction elasticities, and the resulting empirical analysis of presence-transaction relationships and their systematic variations. Section 5 examines the dynamic patterns of mobility through the gravity model framework and compares mobility patterns from both data sources. Section 6 concludes.

## 2 Related and Contributions to the Literature

This section reviews the literature in three complementary domains that inform our analysis: consumer mobility, agglomeration economies, and micro-gravity models of economic interaction. Each of these bodies of literature has developed separately, with distinct methodological approaches and empirical challenges. The present study bridges these domains by integrating high-frequency, high-granularity mobility and transaction data, to provide a framework for analysis of the interaction between real-time population presence and economic activity at micro-geographic scales.

**Consumer Mobility: From Patterns to Economic Impacts.** The empirical analysis of consumer mobility has undergone a paradigm shift with the advent of high-frequency geolocated data, moving beyond the static snapshots provided by traditional travel surveys. Early studies relied on costly, infrequent surveys that captured only aggregate mobility patterns, but recent advances in mobile phone and transaction data have enabled more dynamic analyses (Bonnetain et al., 2021). However, with only a few exceptions, prior work has typically examined mobility patterns and consumption behavior in isolation, leaving unanswered the critical question of how these dimensions jointly shape economic activity.

A growing body of literature leverages mobile phone data to analyze urban mobility and to predict economic activity. Kreindler and Miyauchi (2023) exploit cell phone records of daily commuting flows to estimate economic outcomes, showing that destination fixed effects predict local income with high accuracy. While powerful, this approach is limited to commuting patterns and cannot capture intra-day and non-commuting movements. In contrast, Miyauchi et al. (2025) uses smartphone GPS data to examine trip chains in Japan and study how mobility may generate consumption externalities across locations. However, while their study identifies mobility patterns, it does not measure the economic transactions resulting from these movements. Thus, while these studies provide important insights into mobility, they typically cannot capture the ways in which

overall daily movements (including short, intra-day, and non-commuting trips) translate into economic activity.

In parallel, a distinct literature employs transaction data to study consumption patterns. Agarwal et al. (2017) pioneered the use of credit card transactions to analyze local geographical mobility, while Carpio-Pinedo et al. (2022) and Dolfen et al. (2023) map urban expenditure patterns and quantify e-commerce's consumer surplus. More recently, Bounie et al. (2023, 2026) use billions of French card transactions to study inter-regional trade, revealing how COVID-19 altered mobility and spending. These transaction-based studies focus primarily on the geographical distribution of purchases rather than the real-time relationship between population presence and economic activity; the present study addresses the latter objective.

Two notable exceptions, Campos-Vazquez and Esquivel (2021) and Cepparulo (2025), attempt to bridge this gap by combining transaction data with mobility metrics. Campos-Vázquez and Esquivel (2021) use data from Mexico to estimate the elasticity of point-of-sale (POS) expenditures with respect to mobility, finding a near-unity relationship. However, their analysis is constrained by three critical limitations: (1) their POS data cannot distinguish between online and in-store transactions, conflating substitution effects between physical and digital consumption; (2) their mobility data (Google/Apple) measures visits to broad categories (e.g., "Retail and Recreation") rather than sector-specific presence, obscuring the alignment between mobility and transaction contexts; and (3) their state-week level aggregation masks sub-regional heterogeneities in consumption responsiveness. Cepparulo (2025) advances the literature by separating online and in-store spending in the UK, revealing that mobility is a strong predictor of in-store—but not online—consumption. Yet Cepparulo's district-level analysis lacks sectoral decomposition and relies on mobility categories that do not align precisely with transaction-specific contexts.

Our analysis builds on and extends prior work in four ways. First, we disaggregate consumption by sector, estimating elasticities for 12 economic sectors. This allows us to examine differences in responsiveness between goods and services, a distinction that earlier, more aggregated studies could not explore in detail. Second, we ensure consistency between mobility and transaction contexts by using the same aggregate presence measures (by zone and time interval) for all transaction sectors. While mobility data are not sector-specific, this approach allows us to compare the ways in which different economic sectors respond to overall population presence within each zone and time interval. Third, we estimate the model using Poisson Pseudo–Maximum Likelihood (PPML), which is well suited to our data given the prevalence of zeros and potential heteroskedasticity. As

in previous specifications, we include two-way fixed effects, though at a much finer level (zone × date-hour) to absorb location- and time-specific unobserved heterogeneity. Finally, we assess how consumption responsiveness varies across time of day, day of week, and urban subgroups (core vs. periphery). This decomposition highlights systematic patterns tied to urban density, sectoral composition, and daily economic rhythms, dimensions that coarser, lower-frequency data may obscure.

**Agglomeration Economies: Empirical Gaps.** The literature on agglomeration economies provides the theoretical framework for understanding the ways in which spatial concentration contributes to economic activity. Our work builds directly on the "consumer city" hypothesis of Glaeser et al. (2001), which argued that urban success increasingly depends on cities' ability to provide attractive consumption amenities. However, Glaeser et al. (2001) relied on indirect proxies (e.g., restaurant density, housing price growth) to infer the demand for urban amenities. This paper provides the most fine-scaled measurement of the way real-time population presence generates economic transactions in the literature, addressing three critical gaps in these analyses: (1) the lack of micro-geographic data to test consumption-based agglomeration mechanisms; (2) the inability to distinguish between goods and service sectors; and (3) the absence of dynamic, high-frequency evidence on the ways that urban density translates into economic activity.

Our analysis reveals three key findings that challenge conventional assumptions about urban consumption. First, we uncover a systematic goods-services dichotomy, where goods sectors (e.g., retail food) exhibit near-unity elasticities with respect to population presence, while service sectors (e.g., entertainment) show significantly lower responsiveness, a pattern that prior aggregate studies could not detect. Second, we quantify persistent spatial frictions at micro-geographic scales, estimating that transaction flows significantly decrease with distance from home, even within a single metropolitan area. Third, we document pronounced temporal heterogeneities, with consumption responsiveness varying systematically by time of day and day of week. Elasticities peak during Saturday midday and decline in the afternoon and on Sundays, reflecting shifts toward non-commercial activities and reduced retail activity. These results extend the "consumer city" framework by providing evidence that urban density stimulates economic activity not only through reduced search costs (as hypothesized by Glaeser et al. (2001) but also through sector-specific and temporally dynamic mechanisms. By aligning mobility with transaction contexts and employing robust econometric methods (PPML with comprehensive fixed effects), we provide the first micro-geographic evidence on how spatial and temporal patterns of population presence

shape urban economic outcomes.

**Micro-Gravity Models: Bridging Macro and Micro Perspectives.** The literature on gravity models has evolved from its origins in international trade analysis, pioneered by Tinbergen (1962) and later formalized by Anderson and van Wincoop (2003), to more recent applications in inter-regional trade, where studies such as Combes et al. (2005) and Head and Mayer (2014) have shown that spatial frictions can shape economic linkages within countries. This work has consistently shown that distance remains a key determinant of trade patterns, even as digital technologies reduce some traditional barriers. However, while inter-regional studies have advanced our understanding of sub-national trade flows, they typically rely on coarse spatial units (regions) or annual data, limiting their ability to capture the micro-geographic dynamics of economic interactions between or within cities. For instance, while Bounie et al. (2026) examine inter-regional trade using a rich data set consisting of card transactions made on approximately 85 million bank cards in France, they focus on annual aggregates, rather than the intra-daily and intra-weekly variations that we are able to measure using high-frequency counts of individuals.

In this paper, we construct a micro-gravity framework using mobile phone data with geo-located card transactions at very fine scale, and we quantify the spatial decay of economic interactions with unprecedented precision. Our results show that sectoral composition, temporal alignment with consumption cycles, and localized spatial frictions all affect the relationship between density and economic activity.

## 3 Data Sources and Study Area Characteristics

This section describes the data used to analyze real-time presence and consumption patterns in the Lyon functional urban area in September 2022. We combine two main sources: mobile phone data that measure population presence and mobility and card transaction data that capture retail consumption, both at high spatial and temporal resolution.

**Mobile Phone Data.** Real-time population data –that is, counts of individuals present in a zone during a time interval– are provided by the Orange mobile network operator. The dataset records counts of individuals detected within each Iris zone[4] at 30-minute intervals over a 28-day period,

[4]The Iris zones (Ilots Regroupés pour l'Information Statistique) are sub-municipal geographic units defined by geographic and demographic criteria. These zones are primarily used in municipalities with at least 10,000 inhabitants, with a significant proportion in those having between 5,000 and 10,000 inhabitants. Their population typically ranges from 1,800 to 5,000 inhabitants.

from September 5th to October 2nd, 2022. Presence counts are derived from SIM card detections and have been adjusted by Orange to better reflect actual population presence.[5]

The geographical coverage corresponds to the Lyon *Aire d'Attraction de la Ville*, consistent with the EU–OECD definition of a functional urban area (FUA) (Djikstra et al., 2019). To avoid truncating mobility patterns near the metropolitan fringe, the perimeter is extended by incorporating adjacent *bassins de vie*[6] (urban services zones). The resulting study area comprises 1,151 Iris zones.[7]

The first dataset used in Section 4 captures these presence counts and is organized as an $N \times T$ matrix, where $N = 1{,}151$ Iris zones and $T = 896$ time intervals (28 days at 30-minute intervals between 6 a.m. and 10 p.m.). Each row $i$ corresponds to an Iris zone and each column $t$ records the number of individuals present in zone $i$ during interval $t$.

For the mobility analysis in Section 5, we use a second dataset capturing real-time presence in destination zones $d \in N$, segmented by individuals' home locations $o \in M$. The data are organized as $M \times N$ matrices ($M = 492$, $N = 1{,}151$),[8] with one matrix per 30-minute interval. For each individual, the residence zone is inferred as the most frequent overnight location on the previous day based on mobile phone activity. These matrices therefore describe origin–destination mobility flows linking individuals' home zones to the zones they visit throughout the day.

To analyze mobility distances, we compute road-network distances and free-flow travel times between zones using the `metric.osrm` R package, which queries an OSRM (Open Source Routing Machine) routing server. Distances and travel times are first calculated between the centroids of all 1,151 Iris zones, producing a complete 1,151 × 1,151 origin–destination matrix. Because residence zones in the mobile phone data are larger than Iris zones, we aggregate Iris-level distances using population weights derived from the French population census. For the remainder of this paper, we will use the terms 'presence' and 'presence count' to indicate these counts of the number of individuals observed to be present in a particular zone and 30-minute time interval. These observations are of French residents; we will sometimes use the term 'domestic' presence to emphasize this feature.

[5] Appendix A.1 provides additional details on the mobile phone data and the real-time presence indicator.

[6] A bassin de vie is another form of zone defined by Insee, which "constitutes the smallest territory where residents have access to the most common services and facilities."

[7] Insee is the French national statistical agency, Institut national de la statistique et des études économiques.

[8] Night-time residence zones are larger than Iris zones and typically aggregate around three Iris zones each.

**Card Transaction Data.** To measure consumption, we use transaction-level data provided by the Groupement des Cartes Bancaires (CB).[9] France has a mature bank card market, and the CB scheme—created by French banks in 1984—is the largest payment networks in the country. As of 2023, the system included around 72 million cards and 1.8 million affiliated merchants.

The dataset records each transaction with detailed information including its value, the date and time of the transaction (to the second), and the geographical location of the merchant. For the purposes of this study, we restrict attention to point-of-sale transactions (those made in person), excluding online purchases so that consumption activity can be directly compared with mobility patterns captured in the mobile phone data.

A key advantage of the dataset is its broad coverage: it represents approximately 48% of household consumption expenditure excluding fixed charges, making it highly representative of consumption behavior in France (see Bounie et al. (2023) for a detailed description of the data).

To facilitate comparison with the mobile phone data, we aggregate transactions to the same spatial and temporal resolution as the mobile phone dataset. The resulting data are organized at the Iris $\times$ 30-minute interval level and cover the same observation period and study area. The construction of the analytical dataset is described in detail in Appendix A, while Appendix A.3.2 provides further details on the sectoral segmentation.

For the estimation of gravity models in Section 5, we also construct an origin–destination matrix linking the postcode of cardholders' residences to the Iris zones where expenditures occur. Residence locations are identified when cardholders provide billing or shipping addresses in online purchases. These postcode-based residence zones are larger than those defined in the mobile phone data, encompassing around eight Iris zones on average, resulting in 150 residence zones within the Lyon FUA. The resulting matrix therefore has dimensions $L \times N \times T = 150 \times 1{,}151 \times 896$. Further details are provided in Appendix A.3.3.

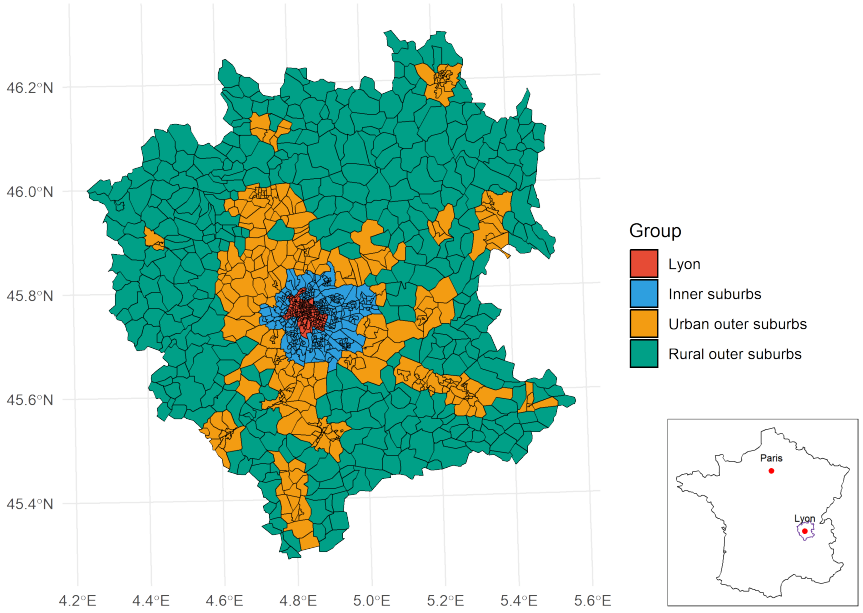
**Spatial Classification of the Study Area.** To characterize the spatial structure of the Lyon metropolitan area and to detect heterogeneity in the relationship between presence and economic transactions, we use the municipal density grid produced by Insee (Beck et al., 2022). This classification distinguishes zones according to their level of urbanization and allows us to group Iris zones into

[9] We use the abbreviation 'CB' to denote transactions processed through the Cartes Bancaires network. CB cards include immediate-debit cards, deferred-debit cards, and credit cards, although merchants accepting CB typically do not distinguish between these types.

four categories: the city centre of Lyon, the inner suburbs, urban outer suburbs, and rural outer suburbs.

Figure 1 presents the study area and the classification of Iris zones according to this density grid. According to the 2020 population census, the Lyon functional urban area contains approximately 2.7 million inhabitants and covers 6,953 km$^2$.

Figure 1: Classification of Iris zones within the functional urban area of Lyon




Notes: This figure displays the map of the study region, Lyon's functional urban area, and the Iris zones classified into four different groups. These Iris groups are defined according to Insee's municipal density grid: the city centre of Lyon, the rest of the urban core (the inner crown), and the hinterland (the outer crown). The hinterland is further divided into two categories, urban zones and rural zones, again based on the municipal density grid established by Insee. The axes of the figure indicate latitude and longitude.

## 4 Transaction Responses to Population Presence

This section examines the relationship between counts of individual presence and economic transactions. We begin by presenting comprehensive descriptive statistics that reveal key stylized facts about urban activity patterns in the Lyon metropolitan area. Following this analysis, we outline our econometric strategy, detailing the Poisson Pseudo-Maximum Likelihood (PPML) estimation approach. The section then presents our core empirical results.

### 4.1 Descriptive Statistics: Key Patterns in Urban Activity

Descriptive statistics for card transaction data and mobile phone presence measurements provide essential context for our empirical analysis. The combination of these datasets enables us to estimate the elasticity of economic transactions with respect to consumer presence in specific zones and time intervals.

The mobile phone data cover a 28-day period in September-October 2022 (from September 5 to October 2, 2022), with observations collected every 30 minutes between 06:00 and 22:00 across 1,151 Iris geographic zones. This results in a total of 1,031,296 ($= 896 \times 1{,}151$) zone-period observations, providing exceptionally rich temporal and spatial detail to capture urban dynamics. At any given timeslot, up to 3.18 million unique individuals are detected across the Lyon functional urban area (7,000 $km^2$). Mobility within the area is substantial and varies over the course of the day and the week. On average, within 30-minute timeslots, individuals are observed in between 1.2 Iris zones (at 6 a.m. on weekends, when mobility is lowest) and 2.4 zones (around 7:30 a.m. and 5:30 p.m. on weekdays, when mobility peaks).

The card transaction dataset complements this picture by capturing the consumption behavior of present individuals. It records approximately 20.5 million transactions totaling €708 million, carried out by 2.38 million unique cards across 18,220 distinct commercial establishments (SIRETs).[10] This represents an average of 8.6 transactions per active card over the month. However, consumer behavior exhibits substantial heterogeneity in card usage patterns, with a mix of occasional and intensive card users. While more than 45% of cards are used on only a single day, a large share of transactions is generated by cards observed repeatedly throughout the period (around 2.74 million cards).

**Spatial distribution.** Table 1 reveals notable spatial disparities in average presence, transaction counts, and transaction values per 30-minute interval across the Lyon functional urban area and its Iris zone groups.

In Panel A, the city of Lyon stands out with the highest real-time presence, with median values around 4,579 people per interval and peaks exceeding 26,000 in the busiest zones. Inner and urban outer suburbs show slightly lower medians but display a similar range of variability, while rural outer suburbs register much smaller medians and maximum values, reflecting their lower activ-

[10] A SIRET (Système d'Identification du Répertoire des Établissements) is a unique 14-digit identifier assigned to each physical business establishment.

ity. Across all 1,151 zones, the mean presence reaches 4,549 people per interval, with a standard deviation of 3,871, highlighting considerable heterogeneity between zones.

The average and median presence counts do not by themselves reveal large differences between Lyon, the inner suburbs, and the outer suburbs. This is partly because Iris zones vary in size: more urbanized areas tend to have smaller zones, each designed to encompass roughly the same number of residents. However, when results are aggregated by zone group, the differences in overall presence and population density become clearer.

Across the entire Lyon functional urban area (7,000 $km^2$), up to 3.18 million unique individuals are detected on average.[11] Within Lyon city, the number of detected individuals reaches 644,000 across 48 $km^2$, corresponding to a density of 13,400 individuals per $km^2$–meaning that roughly 20% of total presence is concentrated in just 0.7% of the Lyon FUA. The inner suburbs host up to 853,000 individuals over 300 $km^2$ (around 2,900 per $km^2$), so 27% of presence is concentrated in only 4.3% of the FUA. The outer urban and rural suburbs account for 1,006,000 and 678,000 individuals, respectively, across 1,500 $km^2$ and 5,000 $km^2$, with densities of 620 and 140 per $km^2$. In other words, 32% of presence is located in the urban zones of the outer suburbs covering about 20% of the FUA, while 21% is in rural outer suburbs covering about 75% of the FUA.

Panel B, which reports transaction counts, highlights stark contrasts in consumer behavior across zones and zone groups. Many zones, particularly in rural areas, record zero transactions per interval, whereas central urban zones reach medians of 17–25 transactions and occasional extremes up to 580 in Lyon. Inner and urban outer suburbs display moderate activity, with median transactions around 3–10 per interval and occasional high-volume spikes.

The spatial distribution of transactions mirrors the one of population presence across the FUA, but transactions are even more strongly concentrated in urban areas. About 30% of transactions occur in Lyon city on average, 26% in the inner suburbs, 36% in the urban areas of the outer suburbs, and only 8% in the rural zones.

Panel C, covering transaction values, follows a similar pattern. Central Lyon zones display the

[11] This estimate is computed by summing, within each timeslot, the difference between presence and entries across zones. Presence refers to the total number of individuals detected in a zone during a given timeslot, while entries correspond to individuals detected in that zone during the timeslot who were not present there in the previous one. For each timeslot, we then aggregate these counts to estimate the number of unique individuals across zones, implicitly assigning mobile individuals to the first zone they occupy at the start of the timeslot. The figure of 3.18 million corresponds to the average daily peak, typically observed around 6 p.m., over the period from September 5 to October 2, 2022. In the measure of real-time presence used in the tables and estimations, mobile individuals are counted in each Iris zone they occupy within a given timeslot. As a result, the average presence per zone is artificially inflated due to the mobility of individuals.

Table 1: Spatial disparities in presence, transaction count, and value

| Group | # Zones | Min | Q1 | Median | Mean | Q3 | Max | SD |
|---|---|---|---|---|---|---|---|---|
| ***Panel A: Presence*** | | | | | | | | |
| Lyon | 185 | 475.50 | 2,835.42 | 4,578.63 | 5,590.86 | 6,766.22 | 26,321.61 | 4,065.89 |
| Inner suburbs | 281 | 403.36 | 2,208.72 | 4,033.89 | 5,115.84 | 6,663.04 | 29,038.02 | 4,403.43 |
| Urban outer suburbs | 298 | 328.57 | 2,661.80 | 4,345.57 | 5,421.67 | 7,193.13 | 22,912.53 | 3,904.02 |
| Rural outer suburbs | 387 | 84.75 | 1,086.50 | 2,190.51 | 2,966.45 | 3,856.21 | 20,504.81 | 2,666.63 |
| All | 1,151 | 84.75 | 1,849.74 | 3,531.61 | 4,548.68 | 5,966.36 | 29,038.02 | 3,870.71 |
| ***Panel B: Transaction count*** | | | | | | | | |
| Lyon | 185 | 0 | 6.40 | 16.77 | 34.08 | 42.43 | 580.55 | 57.98 |
| Inner suburbs | 281 | 0 | 3.04 | 8.04 | 19.58 | 19.95 | 266.34 | 33.82 |
| Urban outer suburbs | 298 | 0 | 3.00 | 9.87 | 25.28 | 31.57 | 276.54 | 39.31 |
| Rural outer suburbs | 387 | 0 | 0.15 | 0.75 | 4.51 | 3.75 | 97.38 | 10.85 |
| All | 1,151 | 0 | 0.74 | 5.56 | 18.32 | 20.09 | 580.55 | 37.04 |
| ***Panel C: Transaction value*** | | | | | | | | |
| Lyon | 185 | 0 | 140.92 | 420.50 | 978.07 | 978.74 | 21,699.30 | 2,124.36 |
| Inner suburbs | 281 | 0 | 91.29 | 216.07 | 736.75 | 569.18 | 12,746.71 | 1,653.20 |
| Urban outer suburbs | 298 | 0 | 94.46 | 323.24 | 1,061.51 | 1,218.67 | 14,101.46 | 1,904.35 |
| Rural outer suburbs | 387 | 0 | 8.08 | 27.92 | 180.59 | 140.60 | 4,058.55 | 462.30 |
| All | 1,151 | 0 | 28.07 | 167.56 | 672.62 | 588.95 | 21,699.30 | 1,591.39 |

Notes: This table shows the distribution of Iris zones' average presence, transaction count, and transaction value across 30-minute intervals in our 28-day sample (6a.m. to 10p.m.), by Iris zone groups and for the entire functional urban area of Lyon. For example, in Lyon, the zone with the lowest average presence recorded approximately 475 people per timeslot, whereas the zone with the highest average presence recorded around 23,321 people per timeslot.

highest median values (€421 per interval) and extreme peaks exceeding €21,000, while inner and urban outer suburbs show moderate medians with substantial variation. Rural zones, in contrast, experience much lower transaction values. Across the entire urban area, economic activity varies widely, with a mean transaction value of €673 per interval and a standard deviation of €1,591.

Results show interesting differences in average basket size. While Lyon city sees the highest number of transactions, the average spending per transaction is slightly lower (€25) compared with outer urban (€33) and rural zones (€37). This suggests that in rural areas, even though transaction frequency is low, individual transactions tend to be larger, likely reflecting bulk purchases or fewer but more substantial shopping trips. Conversely, in central urban areas, the high volume of transactions corresponds to smaller, more frequent purchases.

In terms of spatial distribution across the Lyon FUA, spending broadly mirrors the pattern observed for transaction counts but underscores the relatively larger contribution of urban zones in the outer suburbs. Approximately 23% of total spending occurs in Lyon city, 27% in the inner suburbs, 41% in the urban areas of the outer suburbs, and only 9% in rural zones.

Figure 2 presents the same statistics (average presence, transaction counts, and transaction values per 30-minute interval) as maps, with one map each for presence, transaction count, and trans-

action value. Visually, presence, transaction count, and transaction value appear to be strongly correlated. Again, presence, transaction count and transaction value per zone and timeslot do not visually reveal substantial differences between the urban core and the outer suburbs due to Iris zones being smaller in more urbanized areas (466 zones in the urban core, covering only 5% of the total FUA area, compared with 685 zones in the outer suburbs, covering 95% of the total FUA area).

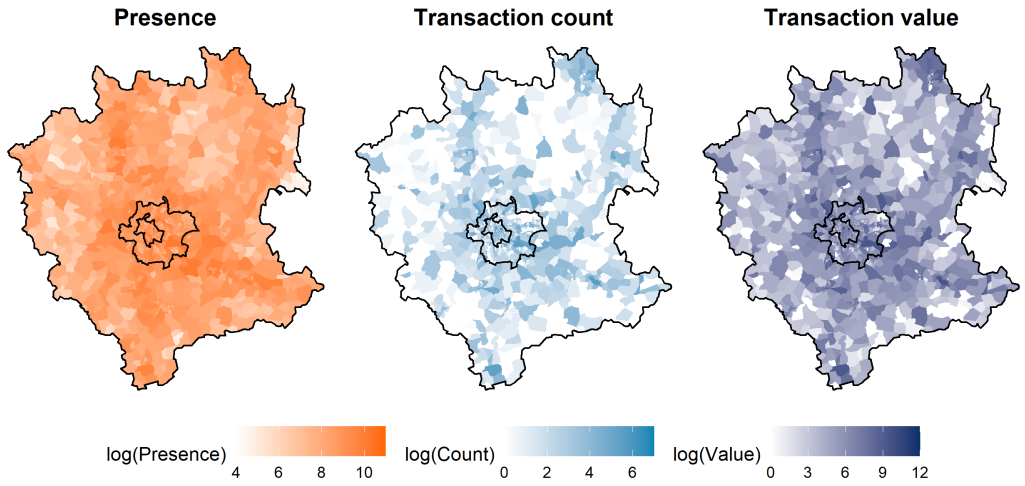



Figure 2: Spatial distribution of average presence, transaction count and transaction value

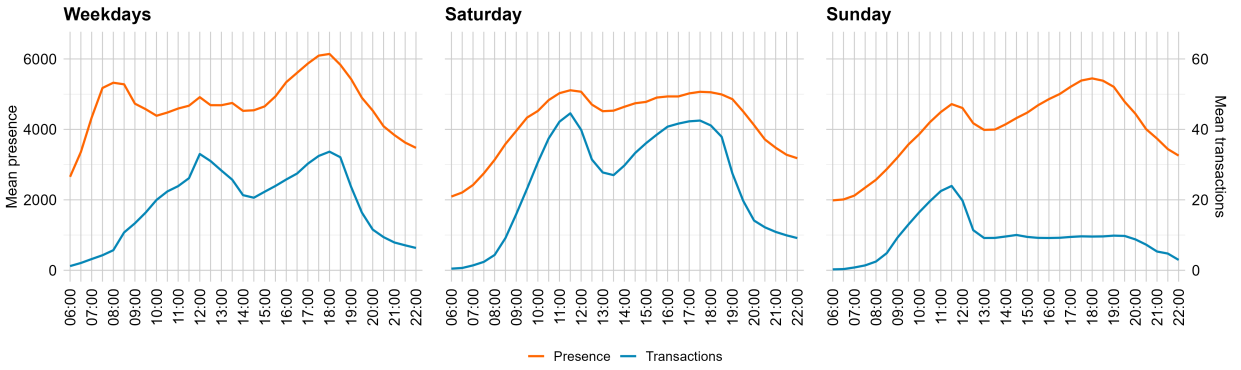
Notes: This figure presents spatial distributions of average population presence, transaction count, and transaction value across Iris zones in the Lyon metropolitan area. The maps visualize 30-minute interval averages over our 28-day study period (6:00 to 22:00). To enhance visual comparability, all values are displayed on a logarithmic scale, with zero values assigned to zones where transaction counts or values fell below 1.

**Temporal distribution.** Figures 3 and 4 compare the hourly dynamics of real-time population presence with card transaction activity. Figure 3 highlights temporal patterns across the day for weekdays, Saturdays, and Sundays, while Figure 4 adds a spatial dimension by distinguishing between zone groups, as defined in Figure 1. Note that, in the mobile phone data, average presence counts are inflated by individuals moving across multiple Iris zones within the same timeslot, meaning that peaks in presence primarily reflect mobility rather than an increase in the total number of unique individuals present in the FUA. By contrast, peaks in transactions correspond to periods when more transactions are carried out.

Figure 3 highlights systematic differences between presence and consumption over the daily cycle. Peaks in presence and transactions do not occur at exactly the same times. On weekdays, mobile phone data capture pronounced morning peaks in average presence per zone that are

not mirrored by transaction activity, reflecting commuting flows that precede most consumption events. Around lunchtime and in the early evening, presence and transactions evolve more synchronously, indicating a closer alignment between population mobility and consumption opportunities. During weekends, Saturday patterns are broadly similar, whereas on Sundays, transaction activity declines sharply in the afternoon despite sustained presence, reflecting retail opening restrictions in France.

Figure 3: Comparison of real-time population presence and card transactions across hours and days



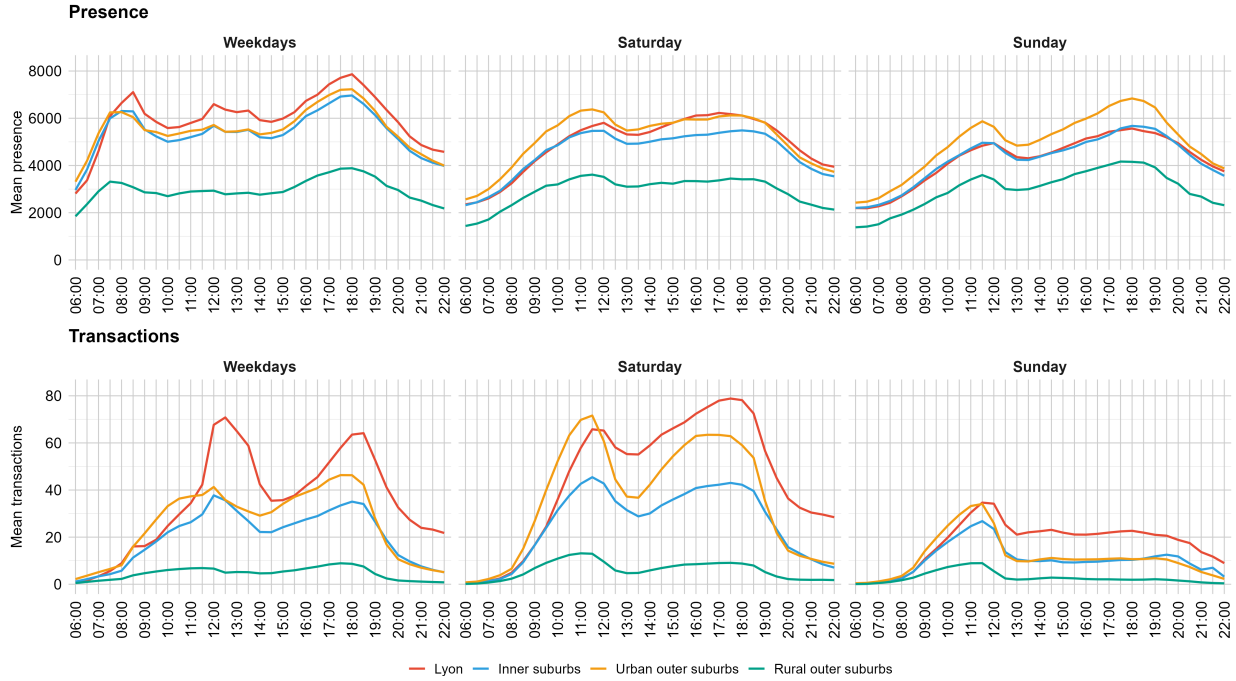

Notes: Mean presence counts (left axis) and mean card transactions (right axis) by 30-minute intervals between 06:00 and 22:00. The transaction series is scaled to ensure comparability across panels. All panels share the same y-axis range, calibrated on weekday peak presence.

Introducing the geographical dimension in Figure 4 shows that these temporal patterns vary across space. Certain zones consistently stand out across days, with Lyon city and the inner suburbs exhibiting the highest levels of both presence and transactions. Spatial concentration and mobility are especially pronounced on Saturday mornings, when presence and transaction activity peak simultaneously in these areas—likely reflecting grocery shopping trips to large retail outlets, such as hypermarchés, the largest class of supermarkets.

**Within-zone variations.** The main source of identification in our main regression relies on within-zone variation in real-time population presence. To facilitate the interpretation of our results, we compute two measures in Table 2: within-zone variation relative to the local mean (Panel A) and the gap between low- and high-activity periods within zones (Panel B).

Within-zones variation relative to the local mean are computed using the zones' average presence over the period and express each 30-minute observation as a percentage deviation from this mean. This enables to assess how much presence fluctuates across timeslots within a given

Figure 4: Comparison of real-time population presence and card transactions across hours, days, and zone groups




Notes: The figure plots mean presence counts (top row) and mean card transactions (bottom row) by 30-minute intervals between 06:00 and 22:00. Columns correspond to weekdays, Saturdays, and Sundays. Lines are colored by zone (Lyon, Inner suburbs, Urban outer suburbs, Rural outer suburbs). Within each row, panels share a common y-axis scale to facilitate comparisons across day types.

zone. Results show that dispersion is important. Across all zones and timeslots, the interquartile range lies roughly between -15% and +15%, indicating that presence typically varies by about 15% around its local mean. Extreme values are much larger, with maximum increases exceeding +700% in some rural zones and drops to -99%. Overall, standard deviations around 25% point to significant variability. Differences across zone types remain relatively limited, although rural areas exhibit slightly more negative medians, as well as larger extremes due to lower average presence counts.

On weekdays, variations are generally moderate, with mean deviations close to zero from Monday to Thursday, suggesting that most zones operate near their average activity levels. Friday stands out with the highest weekday mean deviation, reflecting an overall increase in activity. On weekends, the pattern shifts: both Saturday and Sunday show negative mean deviations, indicating that many zones fall below their typical weekday activity levels. Within the day, zones tend to be near their mean during midday and evening, while the morning shows substantial negative deviations, reflecting lower-than-average activity in many areas. Conversely, the afternoon

exhibits strong positive deviations, suggesting that some zones experience above-average activity.

Within-zones gap between low- and high-activity periods are computed using the difference between the 75th and 25th percentiles of presence for each zone and day, expressed as a percentage of the lower quartile. This measure reflects how much activity increases from relatively quiet to busy periods within a day, for each Iris zone. The results indicate that, on average, presence in a given zone is about 30% higher during high-activity periods than during low-activity periods. While most zones display moderate variation, the upper tails reveal substantial heterogeneity, with some zones experiencing increases of over 300% in the inner suburbs.

Daily and hourly patterns show that both the day of the week and the time of day strongly shape fluctuations in population presence, with weekends and mornings standing out as the most variable. On weekdays, the median difference between high- and low-activity periods ranges from 23% to 27%, with Wednesday slightly more variable and Monday the calmest. Weekends are noticeably more uneven: Saturday reaches about 32%, while Sunday climbs to 38%. Across the day, mornings (06:00–22:00) exhibit the largest variations, 40–45% between high- and low-activity periods, likely reflecting commuting flows, whereas midday, afternoon, and evening are steadier, with variations around 18–22%.

Table 2: Within-zone variation in presence (%)

| Group | Min | Q1 | Median | Mean | Q3 | Max | SD |
|---|---|---|---|---|---|---|---|
| ***Panel A: Relative to the mean*** | | | | | | | |
| All | -99.43 | -14.91 | 0.28 | 0 | 15.36 | 756.59 | 25.15 |
| Lyon | -97.01 | -16.59 | 1.26 | 0 | 17.93 | 199.21 | 27.26 |
| Inner suburbs | -99.43 | -13.77 | 0.77 | 0 | 15.22 | 402.20 | 24.36 |
| Urban outer suburbs | -98.80 | -14.00 | 0.75 | 0 | 14.70 | 197.82 | 23.55 |
| Rural outer suburbs | -98.45 | -15.70 | -0.86 | 0 | 14.76 | 756.59 | 25.85 |
| Monday | -98.98 | -13.82 | 0.30 | -0.21 | 14.07 | 206.93 | 21.75 |
| Tuesday | -99.43 | -13.20 | 1.26 | 1.72 | 16.29 | 239.21 | 23.17 |
| Wednesday | -99.24 | -13.50 | 0.23 | 0.63 | 14.92 | 756.59 | 25.39 |
| Thursday | -98.24 | -11.35 | 1.53 | 2.46 | 16.97 | 172.43 | 21.74 |
| Friday | -97.34 | -6.24 | 6.54 | 7.42 | 21.25 | 244.15 | 23.14 |
| Saturday | -99.43 | -20.77 | -2.47 | -4.54 | 12.05 | 183.39 | 25.89 |
| Sunday | -98.64 | -27.58 | -8.42 | -7.43 | 10.38 | 402.20 | 30.84 |
| Morning | -99.43 | -32.44 | -10.03 | -12.14 | 7.33 | 172.43 | 27.01 |
| Midday | -97.73 | -11.01 | -0.54 | 0.13 | 10.28 | 253.01 | 18.16 |
| Afternoon | -98.80 | -3.56 | 9.88 | 10.68 | 24.33 | 756.59 | 23.04 |
| Evening | -96.32 | -16.85 | -0.60 | 1.14 | 17.49 | 402.20 | 25.86 |
| ***Panel B: Between daily low- and high-activity periods*** | | | | | | | |
| All | 11.51 | 23.29 | 28.08 | 30.32 | 34.41 | 311.01 | 13.16 |
| Lyon | 14.44 | 21.74 | 27.63 | 29.34 | 34.31 | 67.58 | 9.84 |
| Inner suburbs | 11.51 | 21.98 | 25.50 | 28.63 | 32.28 | 311.01 | 19.06 |
| Urban outer suburbs | 13.08 | 23.16 | 27.73 | 28.92 | 33.62 | 63.81 | 7.97 |
| Rural outer suburbs | 11.80 | 24.94 | 30.67 | 33.11 | 38.16 | 94.47 | 12.05 |
| Monday | 9.04 | 18.43 | 22.62 | 25.89 | 28.70 | 763.30 | 25.09 |
| Tuesday | 8.56 | 18.97 | 23.28 | 26.59 | 29.50 | 584.63 | 20.62 |
| Wednesday | 11.10 | 22.00 | 26.65 | 29.40 | 33.16 | 710.73 | 22.54 |
| Thursday | 9.91 | 20.07 | 24.94 | 26.90 | 31.53 | 201.49 | 10.75 |
| Friday | 9.52 | 19.93 | 24.41 | 26.44 | 30.18 | 90.33 | 9.92 |
| Saturday | 8.14 | 22.10 | 31.89 | 35.63 | 43.99 | 142.47 | 18.89 |
| Sunday | 9.17 | 26.17 | 37.97 | 41.43 | 52.99 | 135.77 | 19.39 |
| Morning | 10.35 | 28.18 | 39.54 | 45.11 | 55.64 | 669.10 | 30.48 |
| Midday | 7.98 | 14.21 | 17.94 | 21.52 | 23.19 | 332.14 | 15.75 |
| Afternoon | 8.52 | 14.84 | 18.53 | 21.79 | 23.54 | 325.57 | 15.07 |
| Evening | 7.33 | 14.66 | 18.14 | 20.88 | 23.74 | 324.76 | 13.31 |

Notes: This table reports within-zone variation in presence. Panel A measures variation relative to the zone-specific mean: for each Iris zone, presence at the 30-minute level is expressed as the percentage deviation from its average daily level, capturing how much presence fluctuates across timeslots within a day. These variations are summarized across all Iris zones in the Lyon functional urban area, by zone group, by day of the week, and by time of day (morning: 06:00–10:00; midday: 10:00–14:00; afternoon: 14:00–18:00; evening: 18:00–22:00). By construction, the mean variation is zero when computed across all zones or within zone groups, because deviations are measured relative to each zone's average. When aggregating by day of the week or time of day, however, the mean can be positive or negative, reflecting that some periods consistently exhibit higher or lower activity than the zone's overall average. Panel B captures the gap between low- and high-activity periods within zones. For each zone and day, we compute the percentage difference between the 75th and 25th percentiles of presence, and then average this measure across days. Results are reported for the full sample, by zone group, and by day of the week. In addition, we compute analogous differences within zones and timeslots, and average them by time of day. All statistics are expressed in percentage terms.

## 4.2 Empirical Strategy

The estimation of transaction elasticities with respect to counts of individuals present, in a zone and time period, presents several distinctive econometric challenges that shape our analytical approach. We discuss first the estimation challenges and our methodological approach, and second, we present the specifications.

### 4.2.1 Estimation Challenges and Methodological Approach

Our high-frequency, spatially disaggregated data exhibits three key characteristics that require specialized estimation methods: the prevalence of zero transactions in many zone-time combinations, the structural absence of certain commercial amenities in specific zones, and the likely presence of heteroskedastic error terms across different spatial units and temporal periods.

The frequent occurrence of zero transactions constitutes our most immediate challenge. In many zone-time windows, particularly during off-peak hours or in less commercial zones, we observe no transactions at all. This zero inflation makes conventional logarithmic transformations inappropriate for our analysis. While traditional solutions like adding constants to zero values or dropping zero observations might seem appealing, both introduce significant biases: the former arbitrarily alters the data distribution while the latter excludes potentially informative observations about periods or locations with no economic activity.

A second challenge arises from the structural absence of certain commercial amenities in particular zones. Some Iris zones in our study area may completely lack specific types of establishments. For example, rural zones may have no museums, concert venues, or high-end restaurants. In such cases, zero transactions do not reflect a lack of demand but rather the absence of supply opportunities. This distinction between ”structural zeroes” (where the amenity does not exist) and ”sampling zeroes” (where no transactions occur despite the amenity’s presence) requires careful handling to avoid conflating these fundamentally different phenomena in our elasticity estimates.

The potential for heteroskedasticity in our error terms presents a third significant challenge. Given the substantial variation in commercial activity across different zones and time periods, we reasonably expect non-constant error variance. This heteroskedasticity can lead to biased and inefficient estimates when using conventional methods that assume homoskedastic errors. The combination of zero-inflated data, structural zeroes, and potential heteroskedasticity thus necessitates an estimation approach that can simultaneously handle all three challenges while providing consistent parameter estimates.

To address these interconnected challenges, we employ a Poisson Pseudo-Maximum Likelihood (PPML) estimator. This approach offers several decisive advantages for our context: it naturally accommodates zero observations without ad-hoc adjustments, remains consistent even when errors are not Poisson-distributed (Gourieroux et al., 1984), and allows for fixed effects to control for unobserved heterogeneity. While we use PPML as our primary method, we also present OLS estimates for robustness checks (in Table 15 in Appendix), interpreting these cautiously due to their known limitations with zero-inflated, heteroskedastic data (Silva and Tenreyro, 2006).

#### 4.2.2 Estimating Equations

Our primary estimating equation for the overall elasticity of card transactions with respect to population presence takes the following form:

$$T_{z,t} = \exp\big(\alpha_z + \gamma_t + \beta \log(P_{z,t})\big) + \varepsilon_{z,t}, \tag{1}$$

where $z$ indexes spatial units (Iris zones), $t$ indexes half-hour time slots, $\alpha_z$ are zone fixed effects, $\gamma_t$ are date-hour fixed effects, $P_{z,t}$ is the real-time population presence counts derived from mobile phone data in zone $z$ at time $t$, and $T_{z,t}$ is either the number or total value of card transactions. Zone fixed effects control for time-invariant unobserved heterogeneity across zones (e.g., infrastructure, demographics, commerce density), while date-hour fixed effects capture temporal patterns such as daily and hourly cycles in consumer behavior. Identification of $\beta$ therefore relies on deviations in zone-level presence from both each zone's average and the contemporaneous cross-zone mean at each date–hour. Intuitively, it captures whether zones experiencing higher-than-expected presence at a given time also exhibit proportionally higher transactions, holding constant persistent spatial differences and common temporal shocks. The model is estimated by PPML with the two-way fixed effects given above, and the coefficient $\beta$ represents the elasticity of transactions with respect to the count of individuals present.

To examine heterogeneous effects across different dimensions, we estimate several extensions of this baseline specification. First, we decompose the elasticity by day of the week using:

$$T_{z,t} = \exp\left(\alpha_z + \gamma_t + \sum_{d \in \{\text{Mon},\ldots,\text{Sun}\}} \beta_d \, \mathbf{1}\{\text{day}(t) = d\} \cdot \log(P_{z,t})\right) + \epsilon_{z,t}. \tag{2}$$

This specification allows the elasticity to vary by day, capturing weekly consumption cycles and distinguishing weekday patterns from weekend behavior. In this specification, the identifica-

tion of $\beta_d$ relies on deviations of presence counts from the contemporaneous cross-zone mean, for each date–hour that falls on day $d$.

To analyze spatial heterogeneity, we modify our baseline to include urban area group effects:

$$T_{z,t} = \exp\left(\alpha_z + \gamma_t + \sum_{g \in \mathcal{G}} \beta_g \, \mathbf{1}\{\text{group}(z) = g\} \cdot \log(P_{z,t})\right) + \epsilon_{z,t}, \tag{3}$$

with the set of four urban area groups defined as $\mathcal{G} = \{\text{Lyon city, inner suburbs, outer suburbs, rural outer suburbs}\}$ and where $\mathbf{1}\{.\}$ is the indicator function.

Other variants of this model are used to investigate differences among economic sectors such as food stores, restaurants, hotels, entertainment, and so on. As noted above, establishments corresponding with one or another sector may not be present in a particular zone, so that assigning zero as the transaction value for the unrepresented sector would tend to bias downward the effect of presence of individuals on the number of transactions to be expected in that sector. In these cases we eliminate zones in which there is no establishment corresponding with a sector, from the data set used to examine that sector's transaction elasticity. For each sector, we restrict the sample to zones where at least one establishment of that sector is present, ensuring that zero transactions reflect true demand rather than structural absence. The estimates of each of these models are presented in the next section.

All models are estimated using the PPML estimator with comprehensive fixed effects and standard errors clustered at the Iris zone level to account for correlation of residuals within zones over time.[12] This approach provides robust elasticity estimates that account for the complex realities of our zero-inflated, heteroskedastic, spatially disaggregated data.

### 4.3 Results

This section presents our comprehensive empirical analysis of how population presence translates into transaction activity, examining this fundamental relationship across three critical dimensions. We begin by establishing baseline aggregate elasticities that characterize the overall proportionality between presence and transactions. We then explore temporal decompositions that reveal how this relationship varies across days of the week and times of day. Finally, we examine spatial and sectoral heterogeneities that uncover systematic variations in consumption responsiveness.

[12] For computational implementation, we use R's fixest package with the *fepois* function, clustering standard errors by both zone and date-hour to account for potential spatial and temporal dependence in the data.

**Result 1: Near-unity baseline elasticities with balanced extensive and intensive margins.** Table 3 establishes the foundational relationship between population presence and transaction activity in Lyon's metropolitan area. Using PPML estimation with zone and date-hour fixed effects to control for unobserved spatial and temporal heterogeneity, we find that both transaction count and value elasticities center remarkably close to unity. Specifically, a 1% increase in presence is associated with a 0.992% increase in transaction count and a 0.934% increase in transaction value, both statistically significant at the 1% level.[13]

Table 3: Elasticity of transactions with respect to presence count

| | Transaction count (1) | Transaction value (2) |
|---|---|---|
| log(Presence) | 0.992*** | 0.934*** |
| | (0.082) | (0.085) |
| | [0.831; 1.154] | [0.768; 1.101] |
| Fixed Effects | | |
| Zone | Yes | Yes |
| Date-Hour | Yes | Yes |
| Model Fit | | |
| Observations | 963,187 | 963,187 |
| Squared Correlation | 0.900 | 0.901 |
| Pseudo-$R^2$ | 0.860 | 0.897 |
| BIC | 7,528,193.020 | 229,748,404.598 |

Notes: This table presents baseline estimates of transaction elasticities with respect to real-time population presence in the Lyon metropolitan area. The dependent variables are transaction count and total transaction value. The key independent variable is the logarithm of population presence, measured as individuals detected in a zone during 30-minute intervals. Estimates use a PPML estimator with standard errors clustered by zone and date–hour (shown in parenthesis). 95% confidence intervals are reported in square brackets. Zone and date-hour fixed effects control for unobserved heterogeneity. Significance levels: *** 1%, ** 5%, * 10%. Additional robustness results, to clustering of the standard error can be seen in Appendix table 14

This near-unity proportionality indicates that population presence scales similarly with both the extensive margin (number of transactions) and the intensive margin (value of transactions). While the elasticity for transaction count is marginally higher, the 0.058 difference is not statistically significant, although a similar margin in the estimates occurs across a number of differing samples. The high explanatory power of our models (pseudo-R² of 0.86–0.90) and the comprehensive fixed effects structure lend robustness to these estimates.

**Result 2: Pronounced temporal heterogeneity with Saturday super-proportionality.** Our temporal decomposition reveals pronounced variations in consumption responsiveness across the week and within daily cycles. Table 4 decomposes the baseline elasticities by day of the week,

[13] A joint likelihood ratio test ($\chi^2(1) = 0.59, p = 0.44$) does not reject equality of these coefficients.

revealing significant temporal heterogeneity in consumption responses. Weekday elasticities remain close to unity (ranging from 0.95 to 1.00 for transaction counts), indicating a stable relationship between presence and consumption during standard business hours. Saturday emerges as a clear outlier with elasticities exceeding unity for both transaction count (1.08) and value (1.04), suggesting that additional presence on Saturdays generates more-than-proportional increases in economic activity. This Saturday effect likely reflects the concentration of planned shopping trips, leisure activities, and social outings that characterize weekend consumption patterns.

In contrast, Sunday exhibits the lowest elasticities in our sample (0.980 for count, 0.877 for value), suggesting that Sunday outings are less commercially oriented. This pattern likely reflects the composition of Sunday activities, which often include non-market engagements (such as family visits or cultural events) or lower-value transactions (such as café purchases or market shopping). The particularly low elasticity value on Sundays indicates that while presence may generate additional transactions, these tend to be of modest economic significance.

Table 4: Elasticity of transactions with respect to presence count, by day of week

| | Transaction count (1) | Transaction value (2) |
|---|---|---|
| Monday × log(Presence) | 1.000*** | 0.959*** |
| | (0.081) | (0.080) |
| | [0.842; 1.159] | [0.802; 1.116] |
| Tuesday × log(Presence) | 0.950*** | 0.889*** |
| | (0.080) | (0.079) |
| | [0.793; 1.108] | [0.734; 1.044] |
| Wednesday × log(Presence) | 0.966*** | 0.895*** |
| | (0.081) | (0.079) |
| | [0.807; 1.124] | [0.739; 1.051] |
| Thursday × log(Presence) | 0.971*** | 0.906*** |
| | (0.080) | (0.079) |
| | [0.814; 1.127] | [0.751; 1.061] |
| Friday × log(Presence) | 0.958*** | 0.892*** |
| | (0.080) | (0.078) |
| | [0.801; 1.115] | [0.738; 1.046] |
| Saturday × log(Presence) | 1.076*** | 1.044*** |
| | (0.085) | (0.087) |
| | [0.909; 1.244] | [0.873; 1.215] |
| Sunday × log(Presence) | 0.980*** | 0.877*** |
| | (0.105) | (0.096) |
| | [0.774; 1.186] | [0.688; 1.066] |
| Fixed Effects | | |
| Zone | Yes | Yes |
| Date-Hour | Yes | Yes |
| Model Fit | | |
| Observations | 963,187 | 963,187 |
| Squared Correlation | 0.905 | 0.909 |
| Pseudo-$R^2$ | 0.860 | 0.898 |
| BIC | 7,504,851.694 | 228,098,746.532 |

Notes: This table reports day-of-week decompositions of transaction elasticities with respect to population presence. The dependent variables are transaction count and total transaction value. The key independent variables are day-of-week dummies interacted with the logarithm of population presence. All specifications include zone and date-hour fixed effects. Standard errors are clustered by zone and date–hour and shown in parentheses. 95% confidence intervals are reported in square brackets. Significance levels: *** 1%, ** 5%, * 10%.

**Result 3: Clear intra-daily consumption cycles with morning commuting disconnect.** Table 5 provides further granularity by examining intra-daily variations in consumption responses. The morning period (6:00-10:00) exhibits the lowest elasticities (0.536 for count, 0.458 for value), consistent with commuter-dominated hours where presence primarily reflects transit rather than commercial activity. The midday period (10:00-14:00) shows a reversal, with elasticities exceeding unity for transaction count (1.08), aligning with lunch-hour dining and commercial activity. Afternoon hours (14:00-18:00) return to low elasticities (0.523-0.620), potentially reflecting, again, a substantial number of counts of individuals attributable to passage through a zone while commuting. The evening period (18:00-22:00) then demonstrates a resurgence with elasticities approach-

ing unity (0.984 for count, 1.01 for value), consistent with leisure-oriented consumption such as dining and entertainment.

Table 5: Elasticity of transactions with respect to presence count, by time of day

Panel A: Morning (06:00–10:00)

| | Transaction Count (1) | Transaction Value (2) |
|---|---|---|
| log(Presence) | 0.536*** | 0.458*** |
| | (0.071) | (0.075) |
| | [0.396; 0.676] | [0.311; 0.606] |
| Fixed Effects | | |
| Zone | Yes | Yes |
| Date-Hour | Yes | Yes |
| | | Model Fit |
| Observations | 226,591 | 226,591 |
| BIC | 1,052,743.922 | 29,206,073.829 |

Panel B: Midday (10:00–14:00)

| | Transaction Count (1) | Transaction Value (2) |
|---|---|---|
| log(Presence) | 1.080*** | 0.924*** |
| | (0.082) | (0.089) |
| | [0.920; 1.241] | [0.749; 1.098] |
| Fixed Effects | | |
| Zone | Yes | Yes |
| Date-Hour | Yes | Yes |
| Model Fit | | |
| Observations | 235,615 | 235,615 |
| BIC | 1,991,381.891 | 55,855,129.684 |

Panel C: Afternoon (14:00–18:00)

| | Transaction Count (1) | Transaction Value (2) |
|---|---|---|
| log(Presence) | 0.523*** | 0.620*** |
| | (0.075) | (0.091) |
| | [0.376; 0.670] | [0.441; 0.799] |
| Fixed Effects | | |
| Zone | Yes | Yes |
| Date-Hour | Yes | Yes |
| Model Fit | | |
| Observations | 240,127 | 240,127 |
| BIC | 1,679,779.646 | 62,919,048.535 |

Panel D: Evening (18:00–22:00)

| | Transaction Count (1) | Transaction Value (2) |
|---|---|---|
| log(Presence) | 0.984*** | 1.005*** |
| | (0.238) | (0.168) |
| | [0.519; 1.450] | [0.677; 1.334] |
| Fixed Effects | | |
| Zone | Yes | Yes |
| Date-Hour | Yes | Yes |
| Model Fit | | |
| Observations | 236,803 | 236,803 |
| BIC | 1,806,753.382 | 49,711,150.144 |

Notes: This table presents time-of-day decompositions of transaction elasticities with respect to population presence. Each panel represents a four-hour time block. The dependent variables are transaction count and total transaction value. The key independent variable is the logarithm of population presence. All specifications include zone and date-hour fixed effects. Standard errors are clustered by zone and date–hour and are shown in parenthesis. 95% confidence intervals are reported in square brackets. Significance levels: *** 1%, ** 5%, * 10%.

These intra-daily patterns underscore the dynamic nature of urban economic activity, where the predictive power of population presence varies dramatically across the diurnal cycle. The temporal patterns also provide empirical support for Eaton and Lipsey (1982)'s specialization theory extended to temporal dimensions, where different periods serve distinct economic functions within the urban system. We note however that findings in the literature to the effect that mobility predicts transactions (again Campos-Vazquez and Esquivel (2021), Cepparulo (2025)) do not appear to extend to higher-frequency measures, where we observe considerable variation in elasticity during the daily cycle.

**Result 4: Spatial gradient in consumption responsiveness across urban subgroups.** Turning from the temporal to the geographical dimension, Table 6 examines spatial heterogeneity across urban subgroups within the Lyon metropolitan area. The analysis reveals a clear monocentric pattern: the core city of Lyon exhibits the highest elasticities (1.04 for count, 1.01 for value), reflecting its dense commercial infrastructure and concentration of economic activity. Rural outer suburbs show substantially lower elasticities (0.842 for count, 0.732 for value), suggesting that additional presence in these areas generates proportionally fewer transactions of lower average value. This 0.20 point difference (1.04-0.84) represents a 19.2% reduction in consumption responsiveness per unit of population presence as one moves from core to periphery. This spatial gradient likely reflects decreasing commercial density and increasing transportation costs as one moves from the urban core to the periphery, and provides a first quantification of the economic significance of urban agglomeration effects.

Urban outer suburbs present an intermediate case with elasticities (1.00 for count, 0.953 for value) approaching those of the city center, as these areas often contain shopping malls and retail hubs that serve as secondary commercial centers. Inner suburbs display elasticities between these extremes (0.942 for count, 0.864 for value), consistent with their role as transitional zones between the commercial core and residential periphery.

Table 6: Elasticity of transactions with respect to presence count, by urban sub-grouping

| Urban Sub-Groups | Transaction Count (1) | Transaction Value (2) |
|---|---|---|
| Lyon | 1.039*** | 1.005*** |
| | (0.086) | (0.107) |
| | [0.869; 1.208] | [0.795; 1.215] |
| Inner suburbs | 0.942*** | 0.864*** |
| | (0.124) | (0.107) |
| | [0.699; 1.185] | [0.654; 1.073] |
| Urban outer suburbs | 1.001*** | 0.953*** |
| | (0.090) | (0.089) |
| | [0.824; 1.178] | [0.779; 1.127] |
| Rural outer suburbs | 0.842*** | 0.732*** |
| | (0.114) | (0.109) |
| | [0.619; 1.064] | [0.519; 0.946] |
| Fixed Effects | | |
| Zone | Yes | Yes |
| Date-Hour | Yes | Yes |
| Model Fit | | |
| Observations | 963,187 | 963,187 |
| Squared Correlation | 0.900 | 0.902 |
| Pseudo-$R^2$ | 0.860 | 0.898 |
| BIC | 7,525,627.873 | 229,561,602.681 |

Notes: This table presents urban sub-group decompositions of transaction elasticities with respect to population presence. The dependent variables are transaction count and total transaction value. The key independent variables are urban sub-group dummies interacted with the logarithm of population presence. All specifications include zone and date-hour fixed effects. Standard errors are clustered by zone and date–hour and are shown in parenthesis. Significance levels: *** 1%, ** 5%, * 10%. 95% confidence intervals are reported in square brackets.

**Result 5: Goods-services dichotomy with essential sectors showing highest responsiveness.** We might expect that transaction elasticities would also differ across retail sectors: that some sectors might be more sensitive to individuals' presence than others. In this section we investigate this through an aggregate division of establishments into goods or services, and also by a more granular analysis into twelve retail sectors, described in Appendix B.3. Table 7 gives results at the level of the goods/services distinction, suggesting higher elasticities for goods purchases; however, the estimated coefficients for the two cases are approximately one standard error apart.

Table 7: Elasticity of transactions with respect to presence count, by two-sector classification

| | Transaction count | Transaction value |
|---|---|---|
| Goods × log(Presence) | 1.00*** | 0.954*** |
| | (0.074) | (0.076) |
| Services × log(Presence) | 0.959*** | 0.870*** |
| | (0.077) | (0.079) |
| Fixed-effects | | |
| Zone | Yes | Yes |
| DateHour | Yes | Yes |
| Fit statistics | | |
| Observations | 8,601,619 | 8,601,619 |
| Log-Likelihood | -31,209,848.5 | -1,100,327,153.9 |
| AIC | 62,423,643 | 2,200,658,254 |
| BIC | 62,451,201 | 2,200,685,812 |

Notes: This table reports estimates of transaction elasticities with respect to population presence, distinguishing between goods and services sectors. Standard errors (in parentheses) are two-way clustered by zone and date-hour (30-minute intervals). The regression specification includes zone and date-hour fixed effects to control for unobserved spatial and temporal heterogeneity. "Goods" includes all retail trade categories (food, general merchandise, automotive, etc.), while "Services" comprises accommodation, food services, arts, entertainment, and other service-oriented activities. Significance levels: *** 1%, ** 5%, * 10%.

Table 8 provides a more granular analysis, decomposing elasticities by retail sector and suggesting substantial sectoral distinctions. Goods sectors (food retail, clothing and beauty, general retail) tend to exhibit higher elasticities near unity, while service sectors (particularly sports and recreation, arts and entertainment) tend to show lower responsiveness. Apart from the anomalous Gambling sector, Food retail tends to show the highest responsiveness to presence, reflecting the essential nature of food purchases; additional presence reliably generates proportional increases in transactions. In contrast, arts and entertainment displays the lowest elasticities, indicating a relatively low rate of translation of incidental presence into purchases in this sector; this result may be unsurprising given that some such events are ticketed and therefore may require advance planning, as opposed to being available to incidental visitors to a zone.

The sectors also display somewhat different patterns across the extensive and intensive margins of consumption. In relatively discretionary service sectors such as restaurants and bars, entertainment, cultural sites and sports, the extensive margin (transaction count elasticity of 0.976) tends to exceed the intensive margin (transaction value elasticity of 0.869); that is, additional presence of consumers is associated with more frequent transactions but with lower average spending per transaction. (The gambling sector is again an exception, where the intensive margin exceeds unity and the extensive margin elasticity. Note that the spatial coverage of establishments offering gambling is very low across the Lyon metropolitan area, which may partly explain the unexpected

high elasticity of gambling transactions with respect to population presence.) For retail purchases of goods, the extensive and intensive margin elasticities tend to be closer to each other. We note however that the standard errors of estimation generally imply sufficiently wide confidence intervals that such conclusions must be tentative.

Table 8: Elasticity of transactions with respect to presence count, by retail sector (twelve-sector disaggregated NAF base)

| | Transac count (1) | Transac value (2) |
|---|---|---|
| log(Presence) × Accommodation | 0.777*** | 0.828*** |
| | (0.076) | (0.077) |
| log(Presence) × Arts and Entertainment | 0.763*** | 0.683*** |
| | (0.079) | (0.078) |
| log(Presence) × Automotive | 0.899*** | 0.911*** |
| | (0.075) | (0.077) |
| log(Presence) × Bars and Drinks | 0.944*** | 0.838*** |
| | (0.076) | (0.077) |
| log(Presence) × Clothing and Beauty Retail | 0.939*** | 0.948*** |
| | (0.082) | (0.085) |
| log(Presence) × Food Retail | 1.05*** | 0.993*** |
| | (0.073) | (0.074) |
| log(Presence) × Gambling | 0.967*** | 1.10*** |
| | (0.109) | (0.115) |
| log(Presence) × General Retail | 0.921*** | 0.909*** |
| | (0.077) | (0.078) |
| log(Presence) × Health and Wellness Retail | 0.908*** | 0.819*** |
| | (0.077) | (0.080) |
| log(Presence) × Museums and Cultural Sites | 0.956*** | 0.859*** |
| | (0.086) | (0.089) |
| log(Presence) × Restaurants | 1.01*** | 0.905*** |
| | (0.078) | (0.080) |
| log(Presence) × Sports and Recreation | 0.776*** | 0.705*** |
| | (0.084) | (0.082) |
| Fixed-effects | | |
| Zone | Yes | Yes |
| DateHour | Yes | Yes |
| Fit statistics | | |
| Observations | 8,601,619 | 8,601,619 |
| Log-Likelihood | -27,782,863.4 | -1,033,257,331.6 |
| AIC | 55,569,692.9 | 2,066,518,629.2 |
| BIC | 55,597,390.4 | 2,066,546,326.7 |

Notes: This table presents estimates of transaction elasticities with respect to population presence across 12 distinct retail sectors. Standard errors (in parentheses) are two-way clustered by zone and date-hour (30-minute intervals). All specifications include zone and date-hour fixed effects. Statistical significance is indicated by: *** 1%, ** 5%, * 10%.

These results, taken collectively, suggest that presence counts do not predict economic activity uniformly. Instead, the impact of changes in the number of individuals in an area varies systematically with the temporal context (peak vs. off-peak periods, day of the week), spatial embeddedness (central vs. peripheral locations), and sectoral characteristics (particular goods and services).

# 5 Mobility and Urban-Area Gravity

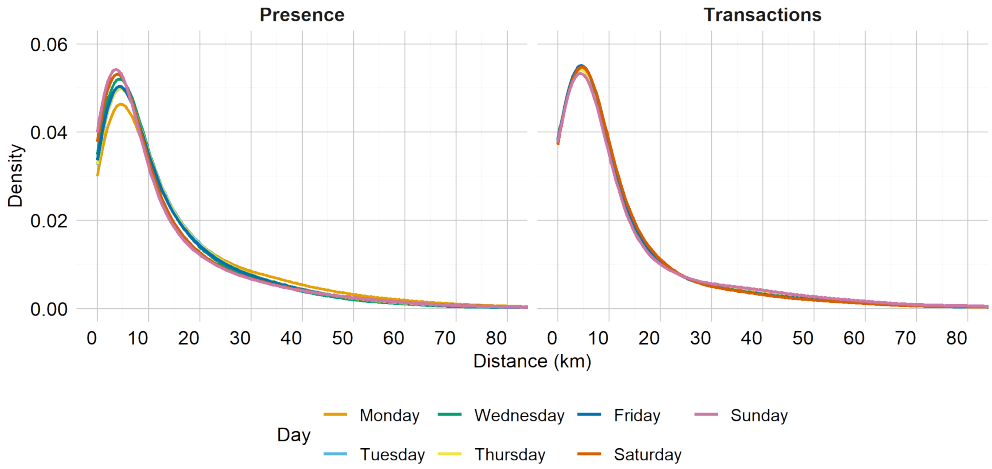
The preceding sections highlighted how real-time population presence is associated with economic transactions, revealing significant temporal and sectoral variations. This section extends the analysis by examining how spatial frictions shape these interactions through a micro-gravity framework. We begin with a detailed examination of mobility patterns, followed by our econometric strategy, and conclude with the estimation results structured around five key findings that quantify the persistent role of physical proximity in urban economic activity, even at micro-geographic scales.

## 5.1 Descriptive Analysis of Mobility Patterns and Measurement Considerations

This section analyzes raw distributions of traveled distances, comparing patterns derived from mobile phone presence data and card transaction records. Figure 5 presents kernel density estimates of distances from residential locations for both datasets, disaggregated by day of the week. These distributions reveal several key characteristics of urban mobility in the study area

The most prominent feature is the pronounced right-skewness in distance distributions, indicating that most trips occur within relatively short radii of residential locations. Mobile phone data capturing all detected movements show a median trip distance of approximately 10 kilometers, while consumption-related trips identified through transaction data exhibit even greater spatial concentration with a median distance of only 5 kilometers (see also Figure 14 in the Appendix D). This 50% difference in median distances suggests that while individuals may travel farther for non-commercial activities (e.g., commuting, social visits), consumption behavior remains significantly more localized, aligning with the monocentric model's prediction that commercial activity concentrates near residential areas to minimize travel costs.

Figure 5: Kernel density functions of distances from home, by day of the week




Notes: This figure presents kernel density estimates of population-weighted road-network distances from residence zones to destination Iris zones, stratified by day of the week. Distances are computed using the Metric-OSRM routing engine, which calculates shortest-path routes between zone centroids. Each curve represents a distinct day of the week, with distances aggregated across all 30-minute intervals from 6:00 to 22:00. The population-weighting procedure adjusts for differences in zone sizes and resident distributions, as detailed in Appendix A.3.3 (Equation 5). The analysis covers the full range of daily mobility patterns in the Lyon metropolitan area.

A further decomposition by day of the week reveals additional heterogeneity in mobility patterns. While card transaction distances remain relatively stable across weekdays, mobile phone data show greater variation, with Wednesdays exhibiting the highest concentration of short-distance movements and Mondays showing comparatively longer average trip distances. This variation likely reflects differences in daily routines, with midweek errands concentrated near residential areas and Monday commutes extending to more distant employment centers. Both datasets indicate that traveled distances are more concentrated at short ranges on weekends, particularly Sundays, although this weekend compression is more pronounced in the mobile phone data.

Before presenting our estimation results, it is crucial to address two measurement issues that may affect our analysis. First, mobile phone data may record individuals in multiple zones as they move within a 30-minute interval, potentially compressing measured distances through multiple counting of the same individual across different zones. To mitigate this bias, we focus our analysis on zone entries and exits during relatively short intervals, reducing the impact of transient movements. Second, spatial units vary in size across our study area, with Iris zones being smaller in the urban core and larger in peripheral areas, and residence zones in the transaction data being coarser than those in the mobile phone data.

Our methodological approach addresses these challenges through three key strategies: population-

weighted distance calculations that account for varying zone sizes and resident distributions, comprehensive fixed effects that control for unobserved heterogeneity, and sensitivity analyses that verify consistency across alternative spatial aggregations. Despite these limitations, the persistent difference between the distance distributions from both data sources underscores the more spatially concentrated nature of consumption activities compared to general mobility, directly motivating the formal gravity analysis in the following part.

### 5.2 Empirical Strategy: the Urban Gravity Relation

To quantify the relationship between spatial separation and economic activity, we specify a micro-gravity model that extends the PPML framework introduced in Section 4.2 to explicitly incorporate origin-destination distances:

$$Y_{odt} = \exp[\beta \log(\text{distance}_{od}) + \alpha_{ot} + \gamma_{dt} + \epsilon_{odt}], \tag{4}$$

where $Y_{odt}$ denotes either population flows or transaction flows from residence zone $o$ to destination Iris $d$ on weekday $t \in \{\text{Mon}, \ldots, \text{Sun}\}$. The parameters $\beta$, captures the elasticity of flows with respect to distance. The model incorporates two sets of fixed effects to control for unobserved heterogeneity: origin-day-specific effects ($\alpha_{ot}$) and destination-day-specific effects ($\gamma_{dt}$).

The Poisson Pseudo-Maximum Likelihood estimator is particularly well-suited for this analysis due to its ability to handle the zero-inflated and heteroskedastic nature of high-frequency origin-destination matrices. Standard errors are clustered at the origin-time level to account for potential spatial and temporal correlation in the error structure. This specification represents a micro-geographic adaptation of the gravity model tradition with three key innovations: the use of 30-minute intervals to capture fine temporal variations, estimation at the Iris-zone level for precise spatial resolution, and the simultaneous inclusion of both distance and travel time to distinguish between physical and time-based barriers to mobility.

Spatial separation between zones is measured using two complementary metrics. Road-network distance is calculated as the shortest path between zone centroids using the Open Source Routing Machine, providing an accurate representation of actual travel routes. Travel time represents the estimated free-flow travel duration between the same points, capturing the temporal cost of spatial separation. To address the varying sizes of urban zones and ensure that distance measurements reflect the actual distribution of residents, we implement population-weighted distance calculations:

$$\text{distance}_{od} = \frac{\sum_{k \in \mathcal{O}(o)} N_k \, dist_{kd}}{\sum_{k \in \mathcal{O}(o)} N_k}, \tag{5}$$

where $N_k$ represents the population of sub-zone $k$ within origin zone $o$, $N_o$ is the total population of zone $o$, and $dist_{kd}$ is the road distance between sub-zone $k$'s centroid and destination zone $d$'s centroid. This weighting scheme is particularly important for larger peripheral zones that may contain multiple population centers, ensuring that distance measurements accurately represent the spatial distribution of residents within each origin zone.

To ensure the robustness of the results, the estimation procedure includes several verification steps. The model is estimated using both distance and travel time as alternative explanatory variables to test the sensitivity of results to different measures of spatial separation. Separate specifications are estimated with and without intra-zonal flows to isolate the effects of distance on trips that involve actual displacement between distinct geographic areas. The consistency of estimates is verified across alternative time windows, with particular attention to periods of high economic activity identified in previous sections. Potential measurement biases arising from the aggregation of origin and destination zones are addressed through sensitivity analyses that vary the level of spatial aggregation.

### 5.3 Estimation Results: Key Findings on Spatial Decay Patterns

Our micro-gravity analysis yields five key findings that fundamentally reshape our understanding of spatial frictions in urban economic activity.

**Result 6: Distance and travel time strongly shape population mobility.** Table 9 presents the baseline estimation results for the gravity model specified in Equation 4, including all zero flows. For population presence flows, a 10% increase in distance is associated with an 8.6% decrease in inter-zonal mobility (column 1, Panel A). Introducing an indicator for remaining within the residence zone (column (2)) reduces the elasticity in absolute value to $-0.74$, reflecting the large share of within-zone presence. Excluding home-to-home flows altogether (column (3)) yields a much larger elasticity of $-1.77$, indicating that distance plays a substantially stronger role for trips that actually leave the residential zone.

**Result 7: Consumption is more distance-sensitive than population presence.** Transaction flows exhibit even greater sensitivity to distance. The elasticity of transaction counts with respect to distance is -2.09 in the baseline specification (column 1, Panel B), indicating that a 10% increase in distance reduces transaction flows by 20.9%. This effect is nearly 2.5 times larger than the corresponding elasticity for population presence, suggesting that consumption activity is significantly more localized than general mobility. When considering transaction values rather than counts, the elasticity remains substantial at -2.03, confirming that both the frequency and monetary value of transactions decline sharply with distance.

Table 9: Gravity model estimates of population and transaction flows: distance vs. travel time effects

| Sample : | All | All | $d \notin o$ | All | All | $d \notin o$ |
|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) |
| ***Panel A: Population flows*** | | | | | | |
| log(Distance) | -0.863*** | -0.738*** | -1.770*** | | | |
| | (0.008) | (0.008) | (0.007) | | | |
| log(Travel time) | | | | -0.930*** | -0.718*** | -2.240*** |
| | | | | (0.009) | (0.010) | (0.009) |
| Home | | 0.984*** | | | 1.48*** | |
| | | (0.047) | | | (0.043) | |
| Fit statistics | | | | | | |
| Observations | 3,911,151 | 3,911,151 | 3,902,180 | 3,911,151 | 3,911,151 | 3,902,180 |
| Squared Correlation | 0.22951 | 0.27644 | 0.64003 | 0.20781 | 0.28396 | 0.61074 |
| Pseudo $R^2$ | 0.59202 | 0.59659 | 0.71232 | 0.55858 | 0.56848 | 0.70980 |
| BIC | 76,508,766.0 | 75,653,520.9 | 39,870,332.4 | 82,764,683.1 | 80,912,362.0 | 40,218,140.9 |
| ***Panel B: Transaction flows in count*** | | | | | | |
| log(Distance) | -2.09*** | -1.89*** | -2.06*** | | | |
| | (0.012) | (0.016) | (0.014) | | | |
| log(Travel time) | | | | -2.61*** | -2.41*** | -2.64*** |
| | | | | (0.015) | (0.020) | (0.019) |
| Home | | 0.518*** | | | 0.410*** | |
| | | (0.026) | | | (0.024) | |
| Fit statistics | | | | | | |
| Observations | 1,086,061 | 1,077,270 | 1,062,344 | 1,086,061 | 1,077,270 | 1,062,344 |
| Squared Correlation | 0.87500 | 0.88038 | 0.86471 | 0.87718 | 0.88202 | 0.85469 |
| Pseudo $R^2$ | 0.87468 | 0.87822 | 0.81517 | 0.87660 | 0.87895 | 0.81720 |
| BIC | 3,128,955.4 | 3,025,958.7 | 2,417,829.8 | 3,082,727.9 | 3,008,533.5 | 2,392,436.7 |
| ***Panel C: Transaction flows in value*** | | | | | | |
| log(Distance) | -2.03*** | -1.95*** | -2.12*** | | | |
| | (0.013) | (0.017) | (0.015) | | | |
| log(Travel time) | | | | -2.54*** | -2.48*** | -2.72*** |
| | | | | (0.016) | (0.021) | (0.020) |
| Home | | 0.230*** | | | 0.132*** | |
| | | (0.030) | | | (0.028) | |
| Fit statistics | | | | | | |
| Observations | 1,086,061 | 1,077,270 | 1,062,344 | 1,086,061 | 1,077,270 | 1,062,344 |
| Squared Correlation | 0.85798 | 0.85776 | 0.83469 | 0.85677 | 0.85756 | 0.82523 |
| Pseudo $R^2$ | 0.89252 | 0.89331 | 0.86127 | 0.89365 | 0.89403 | 0.86381 |
| BIC | 94,722,212.6 | 93,521,615.9 | 70,461,333.5 | 93,733,730.3 | 92,893,235.2 | 69,171,833.1 |

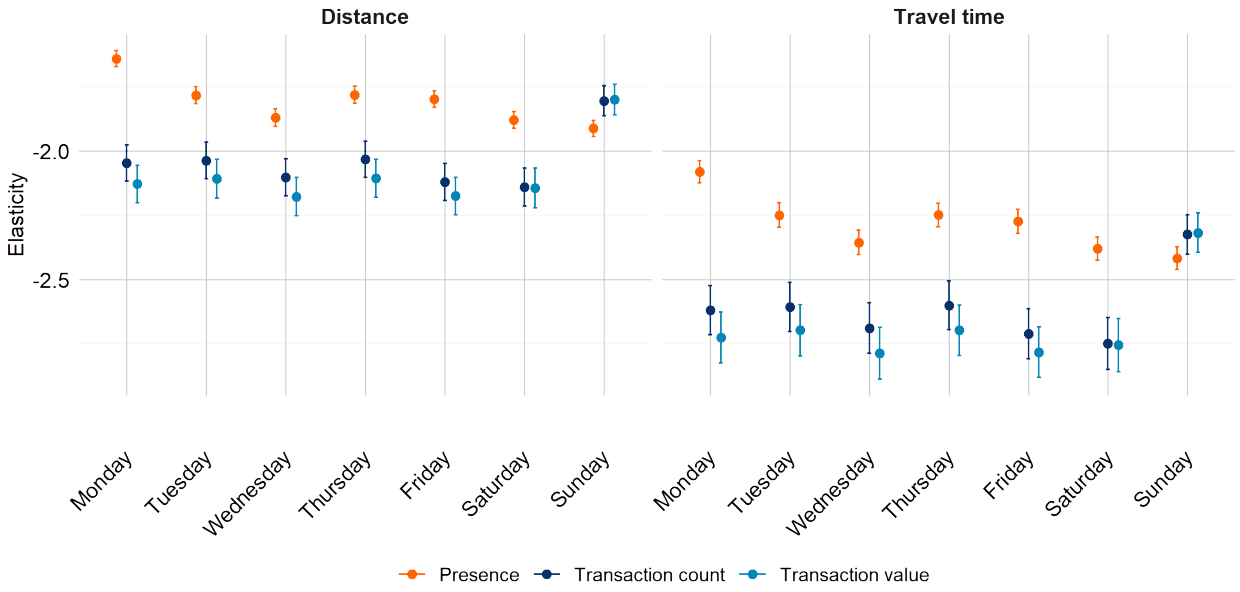
Notes: This table reports Poisson Pseudo-Maximum Likelihood (PPML) estimates of gravity models explaining population and transaction flows (counts and values) as functions of log(Distance) and log(Travel time). Columns (1)-(3) present models using log(Distance), while columns (4)-(6) use log(Travel time). Columns 3 and 6 exclude within-zone trips ($d \notin o$), while other columns include all trips. The "home" dummy variable indicates whether the destination is included in the residence zone. Significance levels are denoted by ***: 0.01, **: 0.05, : 0.1. Standard errors, clustered at the residence-day level, are shown in parentheses. All specifications include residence-day and destination-day fixed effects. Indices $o$ and $d$ denote origin and destination zones, respectively. Further details on the sample construction, the mobility-based dataset, and the estimation methodology are provided in Appendix A.3.3.

The comparison between distance and travel time specifications (columns 4-6) reveals that travel time exerts a slightly stronger deterrent effect than physical distance. For population flows, the travel time elasticity of -2.24 (column 6, Panel A) exceeds the distance elasticity of -1.77, sug-

gesting that the temporal cost of travel may be a more significant barrier than pure physical separation. This pattern holds for transaction flows as well, with travel time elasticities of -2.64 for counts and -2.72 for values.

Figure 6 decomposes the baseline elasticities by day of the week, revealing systematic temporal variations in distance and travel time sensitivity.

Figure 6: Day-of-week variations in distance and travel time elasticities for population and transaction flows




Note: This figure presents Poisson Pseudo-Maximum Likelihood (PPML) estimates of the elasticity of population and transaction flows (counts and values) with respect to route distance and travel time, stratified by day of the week. The 95% confidence intervals are shown. All models include residence-day and destination-day fixed effects to control for unobserved heterogeneity.

**Result 8: Mobility frictions vary by day of the week.** The day-specific elasticities reveal strong evidence of temporal heterogeneity in spatial consumption behavior. Most notably, Sunday exhibits a counterintuitive divergence between activity types: while population flows show significantly increased distance sensitivity (elasticity of -1.9 vs. the weekday average of -1.8, a 10% increase), transaction flows demonstrate markedly reduced sensitivity (-1.8 vs. the weekday average of -2.2, a 16% decrease). This inverse relationship suggests that Sunday consumption trips become more destination-focused despite reduced transport options, while general mobility localizes. Moreover, the remarkable stability of weekday elasticities (±3% variation around the mean) indicates that routine consumption patterns are less affected by temporal factors, likely reflecting consistent work-related mobility and commercial activity during standard business hours. Re-

garding travel time effects, the elasticities systematically exceed those for physical distance across all days, with the gap averaging 25% (e.g., -2.2 vs. -1.8 for population flows). This consistent pattern underscores that temporal costs act as a stronger deterrent than physical distance.

**Result 9: Clear spatial gradient in consumption responsiveness across urban groups.** We allow now the distance elasticity to vary by the type of zone involved in the trip. Zones are classified into four mutually exclusive categories-Lyon city, inner suburbs, urban outer suburbs, and rural outer suburbs-based on urban structure and connectivity. Table 10 reports estimates from specifications in which $\log(\text{distance}_{ij})$ is interacted either with the type of the residence zone (column (1)) or with the type of the destination zone (column (2)), revealing a clear spatial gradient. For population flows, the distance elasticity is -1.59 for trips originating in the urban core but increases to -2.29 for trips from rural outer suburbs, reflecting greater travel costs and fewer local alternatives in less dense areas. Transaction flows show an even steeper spatial gradient, with elasticities ranging from -1.81 in the core to -2.52 in rural areas, suggesting that consumption activity is particularly concentrated in central locations where commercial density and variety reduce search costs.

While Table 10 documents systematic differences by origin and destination types separately, mobility frictions may also depend on their combination. For instance, trips from rural areas toward urban centers may respond differently to distance than trips within the dense urban core or between suburban areas. We therefore estimate specifications in which the distance elasticity is allowed to vary by origin–destination pair type, controlling for residence–day, destination–day, and pair type–day fixed effects.

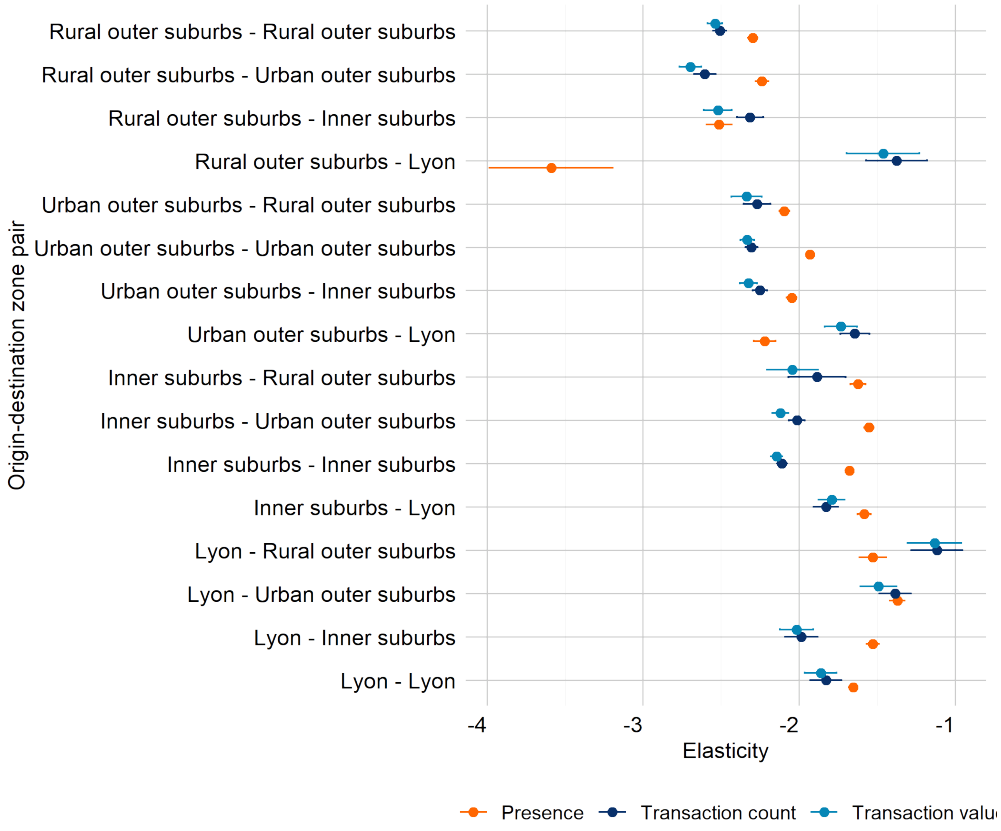
Table 10: Elasticity of population and transaction flows to distance by origin and destination zone types

| Elasticity to distance | from residence zone type (1) | to destination zone type (2) |
|---|---|---|
| ***Panel A: Population flows*** | | |
| log(Distance) × Lyon | -1.59*** | -1.69*** |
| | (0.010) | (0.012) |
| log(Distance) × Inner suburbs | -1.65*** | -1.73*** |
| | (0.010) | (0.008) |
| log(Distance) × Urban outer suburbs | -1.99*** | -1.92*** |
| | (0.011) | (0.009) |
| log(Distance) × Rural outer suburbs | -2.29*** | -2.17*** |
| | (0.016) | (0.011) |
| Fit statistics | | |
| Observations | 3,901,032 | 3,901,032 |
| Squared Correlation | 0.65656 | 0.65676 |
| Pseudo $R^2$ | 0.74050 | 0.73777 |
| BIC | 35,146,428.3 | 35,514,970.9 |
| ***Panel B: Transaction flows in count*** | | |
| log(Distance) × Lyon | -1.81*** | -1.74*** |
| | (0.042) | (0.026) |
| log(Distance) × Inner suburbs | -2.01*** | -2.14*** |
| | (0.017) | (0.014) |
| log(Distance) × Urban outer suburbs | -2.27*** | -2.30*** |
| | (0.020) | (0.015) |
| log(Distance) × Rural outer suburbs | -2.52*** | -2.32*** |
| | (0.028) | (0.027) |
| Fit statistics | | |
| Observations | 1,062,344 | 1,062,344 |
| Squared Correlation | 0.86640 | 0.86742 |
| Pseudo $R^2$ | 0.82091 | 0.82086 |
| BIC | 2,347,828.9 | 2,348,474.0 |
| ***Panel C: Transaction flows in value*** | | |
| log(Distance) × Lyon | -1.85*** | -1.75*** |
| | (0.045) | (0.027) |
| log(Distance) × Inner suburbs | -2.06*** | -2.20*** |
| | (0.019) | (0.015) |
| log(Distance) × Urban outer suburbs | -2.32*** | -2.35*** |
| | (0.022) | (0.017) |
| log(Distance) × Rural outer suburbs | -2.61*** | -2.38*** |
| | (0.028) | (0.031) |
| Fit statistics | | |
| Observations | 1,062,344 | 1,062,344 |
| Squared Correlation | 0.83813 | 0.83701 |
| Pseudo $R^2$ | 0.86793 | 0.86774 |
| BIC | 67,083,420.1 | 67,177,922.3 |

Notes: This table reports Poisson Pseudo-Maximum Likelihood (PPML) estimates of the elasticity of population and transaction flows with respect to distance, stratified by origin and destination zone types. Columns (1) and (2) present elasticities for models where the log of distance is interacted with the type of the residence zone and the type of the destination zone, respectively. Standard errors, clustered at the residence-day level, are shown in parentheses. Significance levels are denoted by ***: 0.01, **: 0.05, : 0.1. All specifications include fixed effects for residence-day, destination-day and pair type-day.

**Result 10: Mobility frictions depend on origin–destination pair types.** Figure 7 reports the resulting elasticities for trips outside the residence zone ($j \notin i$), revealing systematic spatial heterogeneity across origin-destination pair types. The results show that economic interactions are most distance-sensitive in peripheral-peripheral pairs (e.g., Rural-Rural with elasticities of -2.5), where residents face both higher travel costs and limited local alternatives, while core-core interactions (Lyon-Lyon, -1.8) benefit from agglomeration economies and superior connectivity. Notably, core-periphery flows exhibit asymmetric elasticities, with periphery-core trips showing 10-15% higher distance sensitivity than core-periphery trips, reflecting the urban core's role as a central hub for essential services and employment. Table 16 in Appendix D provides the underlying descriptive statistics of population and transaction flows that support these elasticity patterns.

Figure 7: Spatial heterogeneity in distance elasticities by origin-destination pair types




Note: This figure presents Poisson Pseudo-Maximum Likelihood (PPML) estimates of distance elasticities, stratified by origin–destination zone pair types. The econometric specification controls for time-varying unobserved heterogeneity through residence–day, destination–day, and pair type–day fixed effects, ensuring identification relies on within-pair, within-day variations in distance–flow relationships. The analysis is restricted to inter-zonal trips ($j \notin i$), excluding within-residence-zone flows to focus on actual spatial displacement. Error bars represent 95% confidence intervals clustered at the residence–day level to account for spatial and temporal dependence. Distances are calculated using population-weighted road-network metrics (see Appendix D). Descriptive statistics of underlying population and transaction flows, including pair-type-specific means, standard deviations, and zero-flow proportions, are reported in Table 16 in Appendix D.

# 6 Conclusion

This study provides a systematic quantification of the connection between real-time population presence and economic transactions through both extensive and intensive margins, exploiting exceptionally high granularity to provide fundamental new insights into urban economic dynamics. By integrating high-frequency mobile phone location data with comprehensive card transaction records, we have established several key findings that challenge conventional assumptions about urban economic activity and offer important implications for academic research.

At larger scales, our analysis reveals that the elasticities of transaction count and value are both close to unity. This suggests that urban density affects the extensive margin (number of transactions) and the intensive margin (value of transactions) in a balanced manner, without a pronounced divergence between the two. The temporal decomposition further uncovers large variations in consumption responsiveness, with Saturday midday periods exhibiting super-proportional elasticities that exceed morning periods by more than double, while Sundays show significantly lower responsiveness. These temporal patterns suggest that the economic returns to population presence vary substantially throughout the day and week, with important implications for the usefulness of presence counts as an indicator of economic activity (as in, for example, Campos-Vazquez and Esquivel (2021)). The spatial and sectoral decompositions provide equally significant insights. We estimate 19% higher consumption responsiveness in urban cores compared with peripheral areas, quantifying for the first time the economic significance of agglomeration effects at micro-geographic scales. The persistent spatial concentration patterns, with transaction flows declining by approximately 2% for each 1% increase in distance, provide further evidence that economic interactions remain spatially constrained even at fine geographic scales, challenging optimistic assessments about digital technologies' ability to overcome geographic barriers. As well, we identify a goods-services dichotomy whereby goods sectors tend to show a higher elasticity of transactions with respect to measured presence of individuals in a zone, implying that different economic activities derive uneven benefits from urban density.

While our findings offer substantial advances in understanding urban economic activity, several limitations should be acknowledged. First, our mobile phone data, while comprehensive, may miss short-duration trips, potentially affecting our presence measurements. Second, our transaction data only captures card payments, excluding cash transactions that may be more prevalent in certain sectors or among specific population segments. Finally, the 28-day observation period, while providing rich temporal variation, may not capture seasonal effects or longer-term trends

in consumption patterns.

For future research, several promising directions emerge from this study. The development of more sophisticated models that incorporate these heterogeneities could improve our understanding of urban economic dynamics. Comparative analyses across different cities could reveal how these patterns vary with urban structure and size. Longer-term studies could examine how these relationships evolve over time, particularly with ongoing digital transformation. And the integration of additional data sources, such as public transport usage or online transaction records, could provide a more comprehensive picture of urban economic activity. By establishing a framework for analyzing the real-time relationship between population presence and economic transactions, this study lays the groundwork for more nuanced and effective approaches to urban economic analysis and policy design.

## Acknowledgements

This research is funded by the ANR MobiTIC project (grant number ANR-19-CE22-0010), a collaborative initiative involving Orange, Université Gustave Eiffel, and Insee. This research has also been conducted within the Research Chair Digital Finance under the aegis of the Risk Foundation, a joint initiative by Groupement des Cartes Bancaires CB, Insee, Caisse des Dépôts, Télécom Paris and University Paris-Panthéon-Assas.

## References

Agarwal, S., J. B. Jensen, and F. Monte (2017, July). Consumer Mobility and the Local Structure of Consumption Industries. Technical Report w23616, National Bureau of Economic Research, Cambridge, MA.

Ahlfeldt, G. M. and E. Pietrostefani (2019). The economic effects of density: A synthesis. Journal of Urban Economics 111, 93–107.

Anderson, J. E. and E. van Wincoop (2003). Gravity with gravitas: A solution to the border puzzle. American Economic Review 93(1), 170–192.

Beck, S., M.-P. De Bellefon, J. Forest, M. Gerardin, and D. Levy (2022). La grille communale de densité à 7 niveaux. Documents de travail Insee (18).

Bonnetain, L., A. Furno, N.-E. El Faouzi, M. Fiore, R. Stanica, Z. Smoreda, and C. Ziemlicki (2021, September). TRANSIT: Fine-grained human mobility trajectory inference at scale with mobile network signaling data. Transportation Research Part C: Emerging Technologies 130, 103257.

Bounie, D., Y. Camara, and J. W. Galbraith (2023). Consumer mobility and expenditure during the COVID-19 containments: Evidence from French transaction data. European Economic Review 151, 104326.

Bounie, D., Y. Camara, and J. W. Galbraith (2026). Consumer mobility, offline and online regional trade: Evidence from high-frequency transaction data. International Economic Review 67(2).

Campos-Vazquez, R. M. and G. Esquivel (2021, June). Consumption and geographic mobility in pandemic times. Evidence from Mexico. Review of Economics of the Household 19(2), 353–371.

Carpio-Pinedo, J., G. Romanillos, D. Aparicio, M. S. H. Martín-Caro, J. C. García-Palomares, and J. Gutiérrez (2022, December). Towards a new urban geography of expenditure: Using bank card transactions data to analyze multi-sector spatiotemporal distributions. Cities 131, 103894.

Cepparulo, B. (2025, September). Is mobility a good proxy for consumption? Economics Letters 255, 112454.

Combes, P.-P., M. Lafourcade, and T. Mayer (2005, May). The trade-creating effects of business and social networks: evidence from France. Journal of International Economics 66(1), 1–29.

Djikstra, L., H. Poelman, and P. Veneri (2019, December). The EU-OECD definition of a functional urban area. OECD Regional Development Working Papers 2019/11. Series: OECD Regional Development Working Papers Volume: 2019/11.

Dolfen, P., L. Einav, P. J. Klenow, B. Klopack, J. D. Levin, L. Levin, and W. Best (2023, January). Assessing the Gains from E-Commerce. American Economic Journal: Macroeconomics 15(1), 342–370.

Donovan, S., T. de Graaff, H. L. F. de Groot, and C. C. Koopmans (2024). Unraveling urban advantages—a meta-analysis of agglomeration economies. Journal of Economic Surveys 38(1), 168–200.

Duranton, G. and D. Puga (2004). Micro-foundations of urban agglomeration economies. Handbook of regional and urban economics 4, 2063–2117.

Eaton, B. C. and R. G. Lipsey (1982). An economic theory of central places. The Economic Journal 92(365), 56–72.

Glaeser, E. L., J. Kolko, and A. Saiz (2001). Consumer city. Journal of Economic Geography 1(1), 27–50.

Gourieroux, C., A. Monfort, and A. Trognon (1984). Pseudo Maximum Likelihood Methods: Applications to Poisson Models. Econometrica 52(3), 701–720. Publisher: [Wiley, Econometric Society].

Head, K. and T. Mayer (2014). Chapter 3 - gravity equations: Workhorse,toolkit, and cookbook. In G. Gopinath, E. Helpman, and K. Rogoff (Eds.), Handbook of International Economics, Volume 4 of Handbook of International Economics, pp. 131 – 195. Elsevier.

Henderson, J. V. and J.-F. Thisse (2024). Urban and spatial economics after 50 years. Journal of Urban Economics 144, 103711.

Kreindler, G. E. and Y. Miyauchi (2023, 07). Measuring commuting and economic activity inside cities with cell phone records. The Review of Economics and Statistics 105(4), 899–909.

Miyauchi, Y., K. Nakajima, and S. J. Redding (2025, 08). The economics of spatial mobility: Theory and evidence using smartphone data*. The Quarterly Journal of Economics 140(4), 2507–2570.

Silva, J. S. and S. Tenreyro (2006). Log-transforms and the ppml estimator for gravity equations. The Review of Economics and Statistics 88(4), 642–652.

Tinbergen, J. (1962). Shaping the World Economy: Suggestions for an International Economic Policy. New York: Twentieth Century Fund.

## Supplementary Materials (intended for online posting)

This set of appendices contains four comprehensive sections that provide detailed supplementary material supporting our empirical analysis of the relationship between population presence and economic transactions. **Data Appendix A** describes our two primary datasets: mobile phone location data and card transaction records. It covers data generation, processing, validation against official benchmarks, and the construction of integrated datasets. **Sectoral Heterogeneity Appendix B** examines variations in consumption responsiveness across sectors, days of the week, and urban zones. It presents temporal patterns (weekdays, Saturday, Sunday) and spatial patterns (Lyon center, inner suburbs, urban outer suburbs, rural outer suburbs), along with our 12-sector classification system. **Robustness Tests Appendix C** validates our findings through sensitivity analyses. It shows stability across alternative clustering schemes and econometric specifications, confirming the reliability of our PPML estimation approach. Finally, **Gravity Relation Appendix D** provides additional visual and quantitative evidence supporting our gravity model. It includes scatter plots, cumulative distribution functions, and flow statistics that illustrate spatial decay patterns in economic interactions.

# A Data Appendix

This appendix provides a detailed account of the data sources, processing methodologies, and validation procedures that underpin the empirical analysis presented in this paper. The analysis relies on two high-frequency datasets: mobile phone location data capturing real-time population presence and card transaction records documenting economic activity at fine spatial and temporal scales. The appendix is structured as follows: Section A.1 describes the mobile phone data, including its generation, processing, and validation against official census benchmarks, while Section A.2 outlines the card transaction data, emphasizing sectoral selection, spatial-temporal alignment, and representativeness assessments. Finally, Section A.3 describes the construction of the integrated datasets used in the empirical analysis, which merge the two data sources.

## A.1 Mobile phone data

Mobile phone network data offer high-frequency, large-scale information on population presence and mobility, enabled by nationwide antenna coverage and the pervasive use of mobile devices in daily life. For this study, we use 30-minute presence counts for each Iris zone, France's finest census block group, in the Lyon Functional Urban Area. The dataset spans 28 consecutive days, from Monday, September 5th to Sunday, October 2nd, 2022, and is provided by Orange, one of France's primary mobile network operators. The following subsection details the data generation and processing pipeline.

### A.1.1 Data generation and processing

Orange generates real-time presence indicators from signaling data collected through its network infrastructure. Each network event, whether user-initiated such as calls, SMS, or data usage, or automatic such as location-area updates or periodic pings occurring every three hours, produces a location record linked to the cell tower handling the communication. On average, one event occurs every seven minutes per device, amounting to roughly 205 events per device per day. All

records associated with the same SIM card are linked through a unique anonymized identifier. Devices are then probabilistically assigned to Iris zones using a signal-propagation model that incorporates antenna layout, built density, and local topography.

The operator aggregates these assignments to geolocate active devices in real time at the Iris–30-minute level. A reweighting procedure is subsequently applied to address three key aspects: Orange's market share to ensure representativeness across operators, differences in penetration rates across demographic groups such as age or socio-economic status, and alignment with Insee census population totals. The resulting presence indicators provide an approximation of the real-time population in each zone. The analysis focuses on presence measures constructed from national SIM cards only, excluding foreign SIM cards to maintain consistency with the card transaction data, which primarily reflect domestic consumption patterns and avoid additional heterogeneity related to short-term international roaming behavior.

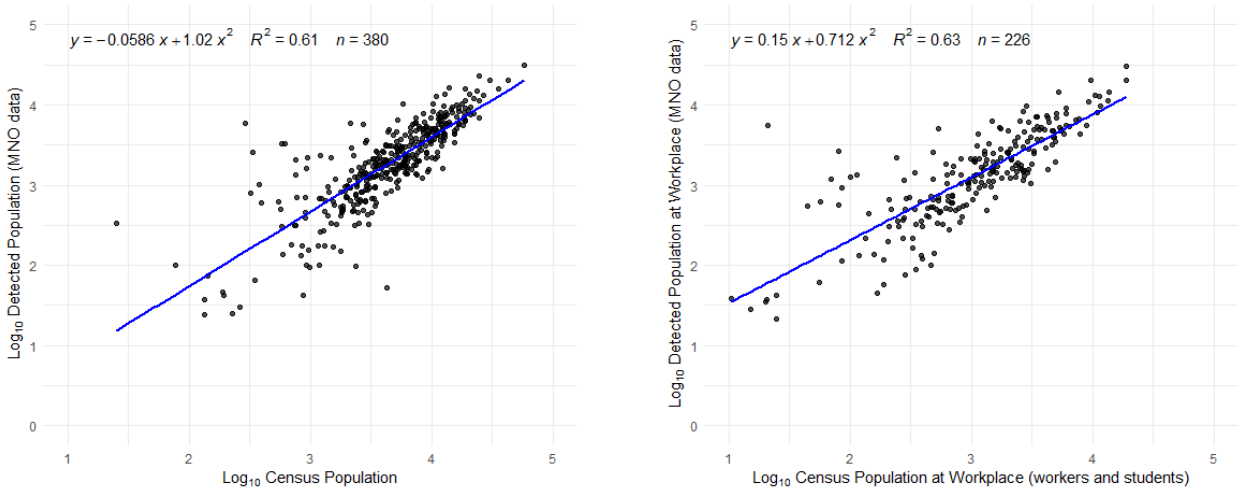
Orange applies strict anonymization protocols. Raw individual-level location data are inaccessible, and only aggregated presence counts are delivered. Masking rules suppress observations with fewer than 20 individuals to preserve privacy.

### A.1.2 Data representativeness

To assess the representativeness of the mobile phone data, we compare origin–destination flows from residential zones to zones of primary activity, such as work or study, with Insee's residence–workplace census matrix. Primary activity zones are inferred from the locations where devices spend the most time during daytime hours over the 15 days preceding the study period. The origin–destination zones aggregate multiple Iris units, with an average of three Iris per zone.

Figure 8 illustrates the strong correlation between detected trips and census benchmarks. Panel A plots the logarithm of the average daily number of individuals detected in the origin–destination data, conditional on cross-zone trips, against the logarithm of the resident population from the official census. Panel B compares the logarithm of the average daily number of individuals with respect to their primary activity zone against the logarithm of workers and students according to their place of study or work in the census. In both panels, the determination coefficient exceeds 60 percent, with an almost one-to-one correspondence between the average daily departures and arrivals and the census data, confirming the alignment of the mobile phone data with official benchmarks.

Figure 8: Mobile phone data representativeness and comparison with the population census




(a) residence (b) workplace

Notes: Panel A plots the logarithm of the average daily number of individuals detected in the origin–destination data, conditional on cross-zone trips, against the logarithm of the resident population from Insee's census. Panel B compares the logarithm of the average daily number of individuals with respect to their primary activity zone against the logarithm of workers and students according to their place of study or work in the census. Lines show log-OLS regression fits.

## A.2 CB card transaction data

### A.2.1 Data generation and processing

The card transaction data are sourced from the Groupement des Cartes Bancaires (CB), France's leading domestic card payment scheme. The analysis focuses on proximity payments, specifically in-store transactions, while excluding online payments to ensure comparability with the mobile phone presence data. Each transaction record includes the monetary value of the payment, a timestamp precise to the second, the geographical location of the merchant, and an anonymized card identifier that is persistent over time but unlinked to external information.

To ensure that card transactions correspond to situations where consumers are physically present at the point of sale, we restrict the sample to merchant activity sectors where in-person payments are standard, such as retail and personal services. We exclude sectors where card data may not reliably reflect local activity, such as vending machines or centralized payment systems, where a single merchant identifier may aggregate transactions across many physical locations. This sectoral restriction improves the consistency between card transaction data and mobile phone presence measures. The list of merchant activity sectors and their corresponding NAF classification codes is provided in Table 13.

For spatial harmonization, transactions are assigned to Iris zones using a spatial join based on merchant coordinates. Transactions with missing or invalid coordinates are excluded from the aggregation. Temporally, the data are aggregated into 30-minute intervals to match the resolution of the mobile phone data, yielding a panel indexed by Iris zone and time interval that is directly comparable to the mobile phone presence indicators. More precisely, for each Iris zone $i$ and time interval $t$, we compute: (i) the number of transactions, (ii) the total transaction value (in euros),

and (iii) the number of active cards, defined as the number of distinct anonymized card identifiers observed in that zone–time interval. The resulting dataset is a panel indexed by $(i, t)$, directly comparable to the mobile phone presence indicators.

### A.2.2 Data representativeness

We assess the representativeness of the CB card transaction data through three complementary validation exercises. First, we infer residential locations for cards based on their proximity payment behavior, identifying the Iris zone where a card is most frequently used over a rolling one-year window centered on the study period. Second, we exploit information from online card transactions, using the most frequently observed billing address over the same period to cross-validate residence assignments. Third, we define residential locations based on the most frequent delivery address from online transactions.

**Residential location inferred from proximity payments.**

In a first approach, we infer a residential location for cards based on their proximity payment behavior. For each anonymized card identifier, we identify the Iris zone in which the card is most frequently used for proximity payments over a rolling one-year window centered on the study period (six months before and six months after). This Iris is interpreted as the card's likely place of residence.

We restrict attention to cards that are observed at least once within the Lyon functional urban area during the study period. Aggregating inferred residential cards by Iris, we compare the logarithm of the number of resident cards with the logarithm of the resident population from the Insee population census (RP). Figure 9 presents the corresponding log–log relationship together with the fitted regression line.

**Residential location inferred from online billing addresses.**

In a second validation exercise, we exploit information from online card transactions. The anonymized card identifier is common across proximity and online payments, allowing us to link the two transaction types. For cards that made at least one proximity payment in the study area and period, and for which online transaction information is available, we infer a residential location based on the billing address most frequently used over the same rolling one-year window.

We then aggregate cards by inferred billing-address Iris and compare the logarithm of the number of cards to the logarithm of the census population. This sample is smaller than in the first approach, as not all cards are used for online payments, but provides a more direct measure of residential location. The corresponding relationship is shown in Figure 9.

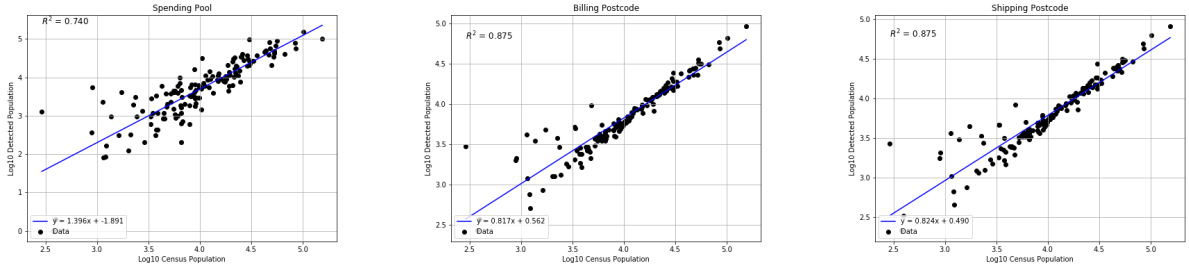
**Delivery location from online transactions.**

Finally, we perform a similar exercise using delivery addresses associated with online card transactions. For each eligible card, we define the delivery Iris as the address most frequently used over the rolling one-year window. Aggregating cards by delivery Iris, we again compare the logarithm of the number of cards with the logarithm of census population. This approach captures a different dimension of spatial attachment and provides an additional robustness check. Results are reported in Figure 9.

Figure 9 compares the logarithm of card-based residence assignments to Insee census population logs for each method. All three approaches show strong positive correlations, providing

reassurance that CB card activity is spatially well-aligned with underlying population distributions. However, the estimated elasticities differ meaningfully across methods. The proximity-based assignment yields a slope of approximately 1.40, above unity, suggesting that larger Iris zones attract disproportionately more card activity than their residential population alone would predict—likely reflecting the fact that high-density zones concentrate both residents and non-resident consumers, so that proximity payment frequency overestimates residential attachment in dense areas. The billing and delivery address approaches yield slopes closer to 0.82, somewhat below unity, consistent with a mild attenuation bias arising from the smaller and potentially selected sample of cards with observable online transaction records. Taken together, the three methods bracket the true population–card correspondence and collectively support the spatial representativeness of the data, while highlighting that proximity-based inference may amplify the weight of dense urban zones.

Figure 9: Representativeness of CB card data




(a) Proximity payments (b) Online billing address (c) Online delivery address

Notes: Each panel plots the logarithm of the number of cards assigned to an Iris zone against the logarithm of the resident population from the Insee census, using alternative definitions of card location.

## A.3 Joint Data Base

This section describes the construction of the integrated datasets used in the empirical analysis, which merge mobile phone presence data with card transaction records at high spatial and temporal resolution. The joint datasets enable direct measurement of the relationship between real-time population presence and economic activity, supporting both aggregate and sector-specific analyses.

### A.3.1 Baseline joint CB $\times$ presence dataset

The baseline dataset used in the empirical analysis combines card transaction data and mobile phone presence data at a common spatial and temporal resolution. The unit of observation is an Iris zone observed in 30-minute intervals over the study period, yielding a balanced panel at the Iris $\times$ date $\times$ time level.

For each Iris zone $i$ and 30-minute interval $t$, the dataset includes aggregate measures of card transaction activity—namely the number of transactions, total transaction value, and number of active cards—as well as measures of real-time population presence derived from mobile phone data. The two data sources are merged using common spatial and temporal identifiers, ensuring full alignment between consumption and presence measures.

The resulting dataset contains observations for all Iris zones and all 30-minute intervals between 06:00 and 22:00 over the 28-day study period, for a total of 963,187 Iris–time observations. Because the data are observed at high temporal frequency, many Iris–time cells record zero card transactions. These zero observations correspond to genuine periods without observed consumption activity and are an inherent feature of high-frequency transaction data. They are retained in the analysis, as they provide meaningful information about short-run variation in local consumption intensity.

This baseline joint dataset constitutes the main input for the empirical analysis of the relationship between real-time population presence and consumption activity. Sector-level analyses and mobility-based extensions rely on separate datasets constructed specifically for those purposes, as described in subsequent sections.

Table 11: Structure of the baseline joint dataset (illustrative)

| Iris zone | Date | Time | Transaction count | Transaction value | Presence |
|---|---|---|---|---|---|
| Iris_001 | 2022-09-05 | 08:00–08:30 | 3 | 124.50 | 31 |
| Iris_001 | 2022-09-05 | 08:30–09:00 | 0 | 0.00 | 20 |
| Iris_001 | 2022-09-05 | 09:00–09:30 | 7 | 312.80 | 50 |
| Iris_245 | 2022-09-05 | 08:00–08:30 | 1 | 18.90 | 38 |
| Iris_245 | 2022-09-05 | 08:30–09:00 | 0 | 0.00 | 21 |
| ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | |

*Notes:* This table illustrates the structure of the baseline joint dataset. Each row corresponds to an Iris zone observed in a 30-minute time interval. Reported values are illustrative and do not correspond to actual observations.

### A.3.2 Sector-level CB × presence dataset

To study heterogeneity in the relationship between presence and consumption across economic activities, we construct a sector-level version of the joint dataset. Card transactions are classified into 12 broad activity sectors (`Sector_12`) based on merchants' NAF codes. The mapping from NAF codes to sector categories is reported in Table 13.

The sector-level dataset is constructed at the Iris × time × sector level. For each Iris zone $i$, 30-minute interval $t$, and sector $s$, we compute sector-specific measures of consumption activity, including the number of transactions and the total transaction value. These sectoral transaction aggregates are then merged with total population presence measured at the Iris–time level.

A key feature of the construction concerns the treatment of zero observations. Sector–zone–time observations are included in the dataset only for Iris–sector pairs in which the sector is observed at least once over the sample period. For such pairs, all time intervals are retained, and transaction counts are allowed to be zero when no transactions occur in a given interval. By contrast, we do not generate observations for sector–zone combinations in which no merchant activity from that sector is ever observed.

This construction ensures that zero transaction observations correspond to economically meaningful situations—namely, periods with no observed demand despite the presence of relevant supply—while avoiding spurious zeroes driven by the absence of sectoral activity in a given location.

The resulting sector-level joint dataset contains 67,576,332 Iris–time–sector observations, reflecting the combination of high-frequency temporal resolution, fine spatial granularity, and sectoral disaggregation.

Table 12: Structure of the sector-level joint dataset (illustrative)

| Iris zone | Sector | Date | Time | Transaction count | Transaction value | Presence |
|---|---|---|---|---|---|---|
| Iris_001 | Food Retail | 2022-09-05 | 08:00–08:30 | 2 | 45.60 | 312 |
| Iris_001 | Food Retail | 2022-09-05 | 08:30–09:00 | 0 | 0.00 | 298 |
| Iris_001 | Restaurants | 2022-09-05 | 12:00–12:30 | 5 | 124.30 | 421 |
| Iris_245 | General Retail | 2022-09-05 | 17:30–18:00 | 1 | 32.10 | 187 |
| Iris_245 | General Retail | 2022-09-05 | 18:00–18:30 | 0 | 0.00 | 179 |
| ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ |

Notes: This table illustrates the structure of the sector-level joint dataset. Each row corresponds to an Iris zone observed in a 30-minute time interval for a given sector. Transaction values and presence figures are illustrative and do not correspond to actual observations.

### A.3.3 Mobility-based CB dataset

This appendix describes the construction of the mobility-based payment card dataset used in the gravity analysis, which links inferred residential locations of cardholders to the locations where their consumption takes place. The objective of this dataset is to identify origin–destination consumption flows at a fine spatial scale, enabling the estimation of distance-based frictions in intra-urban economic activity.

**Sample selection and identification of residence.** The analysis relies on a restricted but well-identified sample of anonymized cardholders. We retain cards that satisfy two conditions: (i) they are observed making at least one proximity (in-store) payment within the Lyon Functional Urban Area (FUA) during the study period, and (ii) they are associated with at least one online transaction for which a billing or delivery address is available.

For each retained card, a residential location is inferred from the addresses reported in online transactions. Specifically, we define the cardholder's residence as the Iris zone corresponding to the most frequently observed billing or delivery address over a rolling one-year window centered on the study period. This procedure yields a stable and time-invariant residential location for each cardholder.

**Construction of residence–destination flows.** Using the inferred residential Iris, we track where cardholders conduct proximity payments during the study period. Transactions are assigned to destination Iris zones based on the geolocation of the merchant. Aggregating across cards, we construct residence–destination flows that measure consumption activity from residential zone $i$ to destination zone $j$.

These flows can be aggregated at different temporal resolutions depending on the analysis. In the gravity estimations, transaction flows are aggregated to the residence–destination–weekday level by averaging total daily flows across the four occurrences of each weekday in the observation window. This aggregation smooths high-frequency volatility while preserving systematic spatial patterns.

**Relation to mobile phone residence–destination matrices.** The mobility-based CB dataset is conceptually analogous to the residence–destination matrices constructed from mobile phone data, which link nighttime residence zones to zones of observed presence. While the two data sources differ in coverage and zoning definitions, both generate origin–destination representations of individual mobility within the Lyon FUA.

In the payment card data, residence zones are defined at the postcode level and are therefore larger than the nighttime residence zones used in the mobile phone data. On average, a postcode-based residence zone encompasses approximately eight Iris units, compared with about three Iris units for nighttime zones in the mobile phone data. These differences are taken into account when computing distances and interpreting spatial frictions.

**Distance and travel time computation.** Distances and travel times between residence and destination zones are computed using the routing engine Metric–OSRM, which provides road-network distances (in kilometers) and free-flow travel times (in minutes) between geographic coordinates. These measures are first computed at the Iris-to-Iris level, using the centroids of all Iris zones within the Lyon FUA, yielding a complete origin–destination matrix.

Because residence zones in both the mobile phone and payment card data are defined over spatial units that aggregate multiple Iris zones, distances from an origin zone $o$ to a destination zone $d$ are not observed at a unique point. To account for this spatial imprecision, we compute origin–destination distances as weighted averages of Iris-level distances. Specifically, for each origin zone $o$ and destination zone $d$, the distance is defined as the population-weighted average of the distances between all Iris zones composing origin $o$ and destination zone $d$, where weights correspond to census population counts in the residence Iris zones.

This aggregation procedure ensures that distance and travel time measures reflect the spatial distribution of origins within each origin zone, rather than relying on a single centroid that may poorly represent large or heterogeneous origin areas. The same procedure is applied consistently to both distance and travel time measures and across data sources.

Formally, let $\mathcal{O}(o)$ denote the set of Iris zones composing origin zone $o$, and let $d$ denote a destination Iris zone. Let $dist_{kd}$ denote the Metric–OSRM distance between the centroids of Iris zone $k \in \mathcal{O}(o)$ and destination Iris $d$, and let $N_k$ denote the census population of Iris zone $k$.

The residence–destination distance between zones $o$ and $d$ is defined as the population-weighted average:

$$\text{distance}_{od} = \frac{\sum_{k \in \mathcal{O}(o)} N_k \, dist_{kd}}{\sum_{k \in \mathcal{O}(o)} N_k}. \tag{6}$$

**Scope and interpretation.** The resulting dataset captures consumption-related mobility patterns for a selected subset of cardholders for whom residential locations can be reliably identified. Although this sample is smaller than the universe of payment cards, it enables a direct and explicit linkage between place of residence and place of consumption. As such, it is particularly well suited for analyzing spatial spillovers, distance decay, and gravity-type relationships in intra-urban consumption behavior.

Throughout the paper, indices $i$ and $j$ denote residence and destination zones, respectively. Additional details on distance computation, zoning harmonization, and comparability with mobile phone data are provided in the surrounding appendix sections.

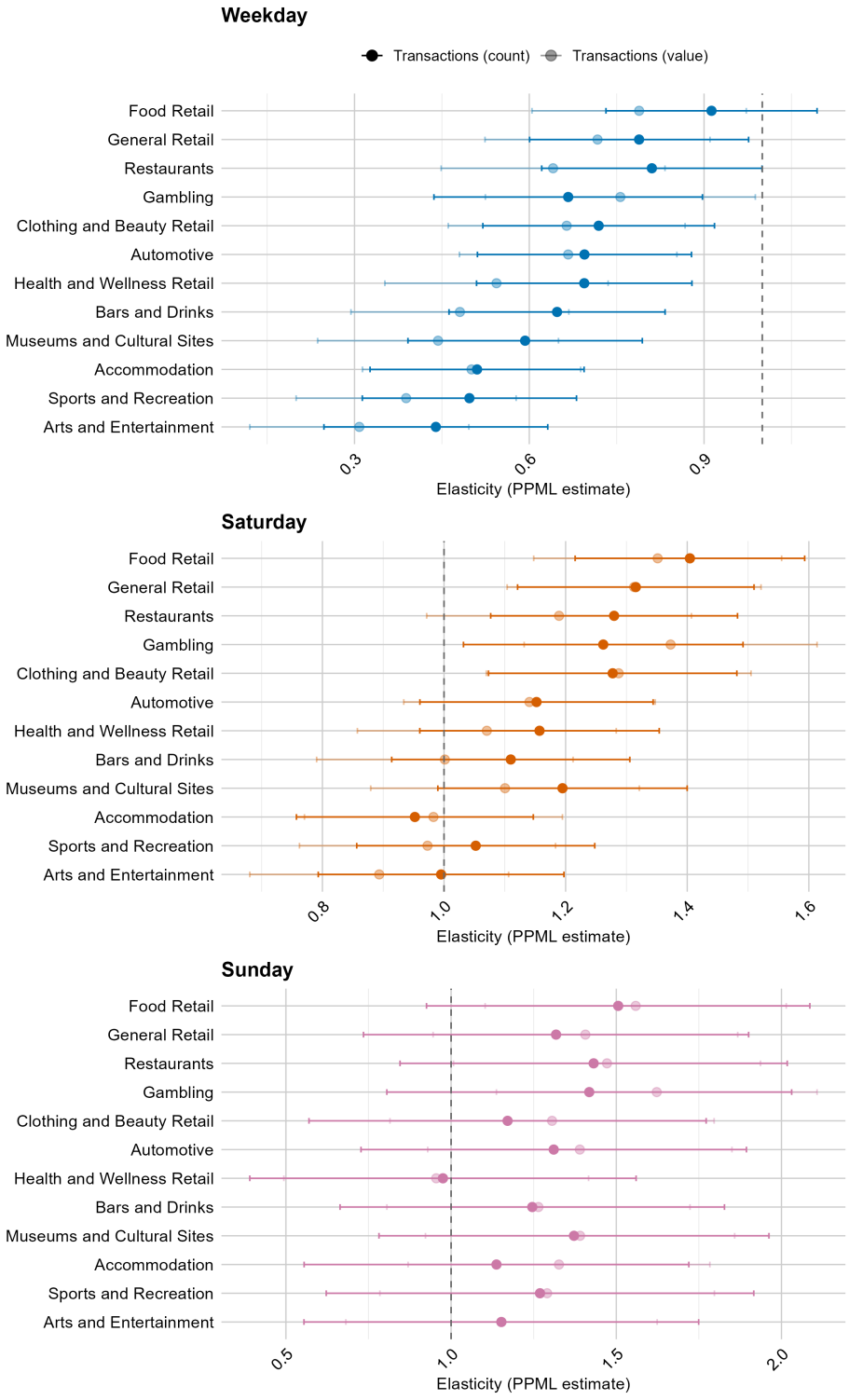
# B Appendix: Sectoral Heterogeneity in Consumption Responsiveness

This appendix presents supplementary analyses that extend our core findings on the relationship between population presence and economic transactions, focusing on how these relationships vary across economic sectors, days of the week, and urban zone types.

## B.1 Temporal Variations in Sectoral Elasticities

The relationship between population presence and transaction activity exhibits significant temporal heterogeneity, as demonstrated in Figure 10. Panel (a) establishes the baseline consumption responsiveness across all sectors during weekdays, revealing a clear goods-services dichotomy where essential retail sectors demonstrate near-unity elasticities while discretionary service sectors exhibit lower responsiveness. Panels (b) and (c) illustrate how this relationship transforms during weekend days. Saturday patterns (Panel b) show increased responsiveness across most sectors, with a narrowing of the goods-services gap as both categories benefit from planned shopping and leisure activities. Sunday patterns (Panel c) differ markedly from other days, demonstrating a distinct elasticity profile characterized by reduced overall responsiveness. While essential sectors like food retail maintain relatively high elasticities, many service sectors show significantly lower responsiveness on Sundays, suggesting a shift toward more destination-specific consumption patterns and reduced commercial activity on this day. This arrangement of weekday, Saturday, and Sunday patterns provides a comprehensive view of how consumption responsiveness varies systematically across the week, reflecting the changing rhythms of urban economic activity.

Figure 10: Elasticities of transactions with respect to presence by retail sector and day type.

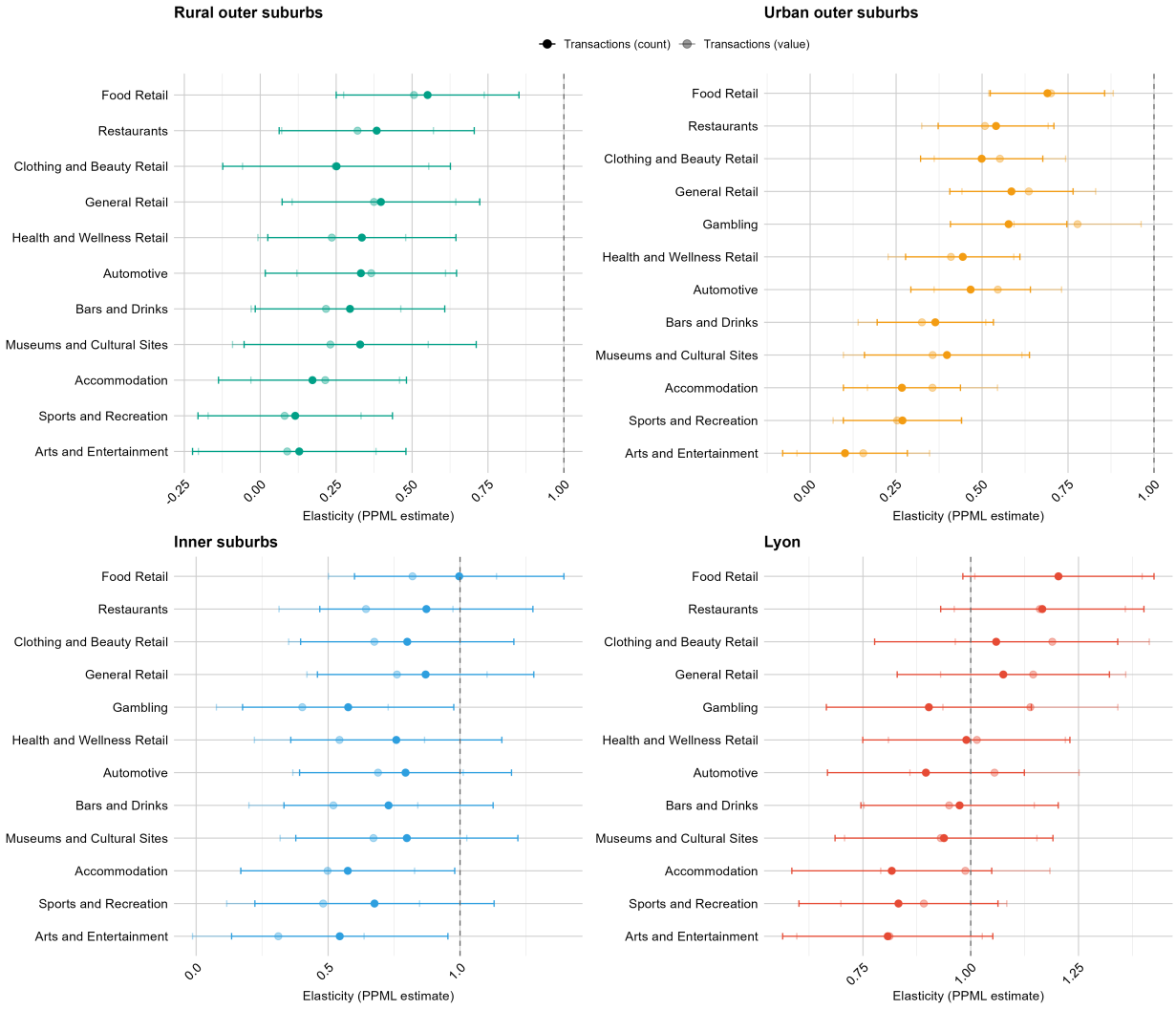



Notes: Each panel shows sector-specific elasticities differentiated by day type. From top to bottom: weekdays (Monday–Friday), Saturday, and Sunday.

## B.2 Spatial Patterns in Sectoral Elasticities

The spatial dimension of consumption responsiveness is examined in Figure 11, which decomposes sector-specific elasticities across four distinct urban zone types. This analysis reveals systematic spatial heterogeneities that reflect the varying economic structures and consumption opportunities across the metropolitan area. The urban core (Lyon city center) exhibits the highest elasticities across most sectors, reflecting its dense commercial infrastructure and efficient consumer-provider matching. Inner suburbs show moderate elasticities with some sectoral variation, particularly in service sectors where capacity constraints may limit responsiveness. Urban outer suburbs demonstrate lower overall elasticities but maintain concentration in retail hubs, while rural outer suburbs show the lowest responsiveness across most sectors, indicating thinner markets and higher search costs in peripheral areas.

Figure 11: Elasticities of transactions with respect to presence by retail sector, across urban groups




Notes: Each panel shows sector-specific elasticities estimated for a distinct spatial group: Lyon, inner suburbs, urban outer suburbs, and rural outer suburbs.

### B.3 Sector Classification System

The sector classification organizes 73 NAF codes into 12 broad economic sectors, grouped into two categories. The Goods category (46 codes) covers automotive sales, clothing and beauty retail, food retail, general retail, and health and wellness retail. The Services category (27 codes) covers accommodation, arts and entertainment, automotive repair, bars and drinks, gambling, museums and cultural sites, restaurants, and sports and recreation. The full mapping is presented in Table 13.

Table 13: NAF codes by sector and Goods/Services classification

| Sector | NAF Codes |
|---|---|
| ***Panel A: Goods (46 codes)*** | |
| Automotive | 4511Z, 4519Z, 4532Z, 4730Z |
| Clothing and Beauty Retail | 4771Z, 4772A, 4772B, 4775Z, 4777Z |
| Food Retail | 4711A, 4711B, 4711C, 4711D, 4711E, 4711F, 4719A, 4719B, 4721Z, 4722Z, 4723Z, 4724Z, 4725Z, 4726Z, 4729Z |
| General Retail | 4741Z, 4742Z, 4743Z, 4751Z, 4752A, 4752B, 4753Z, 4754Z, 4759A, 4759B, 4761Z, 4762Z, 4763Z, 4764Z, 4765Z, 4776Z, 4778B, 4778C, 4779Z |
| Health and Wellness Retail | 4773Z, 4774Z, 4778A |
| ***Panel B: Services (27 codes)*** | |
| Accommodation | 5510Z, 5520Z, 5530Z, 5590Z |
| Arts and Entertainment | 9001Z, 9002Z, 9003A, 9003B, 9004Z |
| Automotive Repair | 4520A |
| Bars and Drinks | 5630Z |
| Gambling | 9200Z |
| Museums and Cultural Sites | 9102Z, 9103Z, 9104Z |
| Restaurants | 5610A, 5610B, 5610C, 5621Z, 5629A, 5629B |
| Sports and Recreation | 9311Z, 9312Z, 9313Z, 9319Z, 9321Z, 9329Z |

Notes: NAF codes follow the French nomenclature of activities (Nomenclature d'Activités Françaises), consistent with NACE Rev. 2 at the 5-digit level. Sector 4540Z and the "Other/Unknown" category are excluded from the analysis. The Automotive sector appears in both panels: vehicle sales (Panel A) and repair services (Panel B).

# C Appendix : Additional Empirical and Robustness Tests

This appendix presents comprehensive robustness checks that validate the reliability of our baseline empirical findings regarding the relationship between real-time population presence and economic transactions. We examine two critical dimensions of robustness: 1. the sensitivity of our results to alternative clustering schemes for standard errors, and 2. the stability of our estimates across different econometric specifications and functional forms. These tests address potential concerns about spatial-temporal dependencies in our high-frequency panel data and the appropriateness of our Poisson Pseudo-Maximum Likelihood (PPML) estimation approach.

**Clustering of the standard error.** The robustness of our baseline elasticity estimates to alternative clustering schemes for standard errors is demonstrated in Table 14. When comparing our primary specification—where standard errors are clustered by both zone and date-hour—to alternative specifications clustering by zone only or date-hour only, we observe complete stability in the point estimates for both transaction count (0.992) and transaction value (0.934) elasticities. This invariance confirms that our core findings regarding the relationship between population presence and economic transactions are not sensitive to the method of accounting for within-group correlation in the errors. However, the standard errors vary meaningfully across specifications, with the two-way clustering yielding the most conservative estimates (SE = 0.082 for counts, 0.085 for values), while date-hour only clustering produces substantially smaller standard errors (SE = 0.040). This pattern underscores the importance of our baseline approach, which accounts for both spatial and temporal dependencies in the high-frequency panel data, thereby providing the most reliable inference. The stability of these results across clustering schemes strengthens our confidence in the estimated elasticities and supports our conclusion that real-time population presence has a robust and economically meaningful impact on urban economic activity.

Table 14: Robustness to Alternative Clustering of Standard Errors

| | Transaction count | | Transaction value | |
|---|---|---|---|---|
| Clustering | Elasticity | SE | Elasticity | SE |
| Zone + DateHour (baseline) | 0.992 | (0.082) | 0.934 | (0.085) |
| Zone only | 0.992 | (0.075) | 0.934 | (0.076) |
| DateHour only | 0.992 | (0.040) | 0.934 | (0.040) |
| Fixed Effects: Zone and Date–Hour | | | | |

Notes: This table reports PPML estimates of the elasticity of card transactions with respect to real-time population presence under alternative clustering schemes for standard errors. The baseline specification clusters standard errors two ways, by zone and by date–hour (30-minute interval). Alternative specifications cluster by zone only or by date–hour only. All models include zone and date–hour fixed effects

**Alternative Estimation Methods** To assess the robustness of our baseline results to alternative econometric specifications, Table 15 presents estimates using different estimators and functional forms for both transaction counts (Panel A) and transaction values (Panel B). All models maintain the same fixed effects structure (zone and date-hour) and standard error clustering, ensuring comparability with our primary PPML specification.

Table 15: Robustness: alternative estimators and transformations

| | $\hat{\beta}$ | SE | 95% CI lo | 95% CI hi | $N$ | sq.cor | pr2 | BIC |
|---|---|---|---|---|---|---|---|---|
| ***Panel A: Transaction count*** | | | | | | | | |
| PPML | 0.992 | 0.082 | 0.831 | 1.154 | 963187 | 0.900 | 0.860 | 7528193.0 |
| OLS: log(1+y) | 0.615 | 0.055 | 0.507 | 0.723 | 1024558 | 0.799 | 0.427 | 2231336.3 |
| OLS: asinh(y) | 0.676 | 0.061 | 0.557 | 0.796 | 1024558 | 0.794 | 0.389 | 2575804.9 |
| OLS: asinh(y), asinh(pres) | 0.676 | 0.061 | 0.557 | 0.796 | 1024558 | 0.794 | 0.389 | 2575804.6 |
| NegBin (counts) | 0.765 | 0.067 | 0.633 | 0.897 | 963187 | 0.845 | 0.235 | 5098086.1 |
| ***Panel B: Transaction value*** | | | | | | | | |
| PPML | 0.934 | 0.085 | 0.768 | 1.101 | 963187 | 0.901 | 0.897 | 229748404.6 |
| OLS: log(1+y) | 0.855 | 0.081 | 0.696 | 1.014 | 1024558 | 0.742 | 0.268 | 3815773.0 |
| OLS: asinh(y) | 0.895 | 0.087 | 0.725 | 1.065 | 1024558 | 0.735 | 0.253 | 4039301.9 |
| OLS: asinh(y), asinh(pres) | 0.895 | 0.087 | 0.725 | 1.065 | 1024558 | 0.735 | 0.253 | 4039301.7 |
| Gamma (value) | 0.616 | 0.066 | 0.488 | 0.745 | 656986 | 0.821 | 0.062 | 24682447.2 |

Notes: All models include fixed effects for Zone and Date $\times$ Hour; standard errors clustered by Zone and Date $\times$ Hour. PPML is Poisson pseudo-maximum likelihood. OLS models use $\log(1 + y)$ or $\text{asinh}(y)$ as the dependent variable transformation. The specification $\text{asinh}(y) \sim \text{asinh}(\text{pres})$ is not reported separately as it yields numerically identical estimates to $\text{asinh}(y) \sim \log(\text{pres})$ once fixed effects are included, since $\text{asinh}(\text{pres}) \approx \log(2\,\text{pres})$ over the observed support, with the additive constant absorbed by the fixed effects.

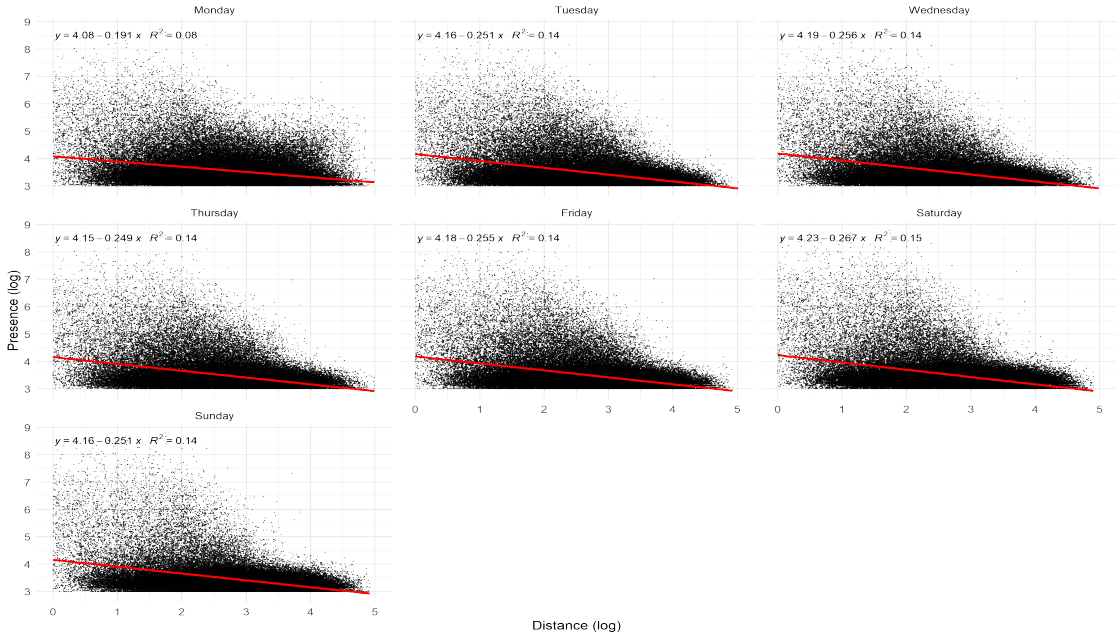
The robustness analysis confirms the reliability and economic significance of our baseline findings regarding the relationship between population presence and economic transactions. For transaction counts (Panel A), our preferred PPML specification yields an elasticity estimate of 0.992 with tight confidence intervals (0.831-1.154), indicating a near-proportional relationship between population presence and transaction volume. While alternative estimators produce systematically lower elasticities—ranging from 0.615 for log-transformed OLS to 0.765 for negative binomial models—all specifications maintain positive and statistically significant coefficients, reinforcing the fundamental economic relationship. The transaction value models (Panel B) exhibit a similar pattern, with PPML estimating an elasticity of 0.934 compared to 0.616-0.895 for alternative specifications. This consistency across diverse econometric approaches—including transformations to handle zeroes (log(1+y), asinh(y)) and distributional alternatives (negative binomial, Gamma)—demonstrates that our core finding of substantial consumption responsiveness to population presence is not sensitive to functional form or estimator choice. The superior model fit of PPML (pseudo-$R^2$ of 0.860-0.897 versus 0.427 for alternatives) further validates our baseline approach, while the conservative estimates from alternative specifications provide reassuring bounds on the true effect size. These results collectively confirm that real-time population presence generates economically meaningful increases in both the frequency and value of urban economic transactions, with near-unity elasticities suggesting that agglomeration effects operate primarily through the extensive margin of more frequent interactions rather than substantially higher spending per transaction.

# D Appendix: Gravity Relation: Visual

This appendix provides complementary visual and quantitative analyses that extend our gravity model findings from Section 5, offering additional empirical support for the spatial patterns of population mobility and transaction flows in the Lyon metropolitan area. The figures and table presented here work synergistically to illustrate how distance from residential locations systematically shapes economic interactions at micro-geographic scales.

**Visual Evidence of Spatial Decay Patterns.** Figure 12 illustrates the fundamental relationship between population presence counts and distance from residential locations. Each data point represents a residence-destination pair observed during peak activity periods (10:00-11:30 and 14:00-15:30), averaged by day of the week. The clear negative relationship visible in this scatter plot provides visual confirmation of our gravity model estimates, showing how population presence systematically declines with increasing distance from home locations. The pattern reveals both the overall distance decay effect and potential day-of-week variations in mobility patterns.

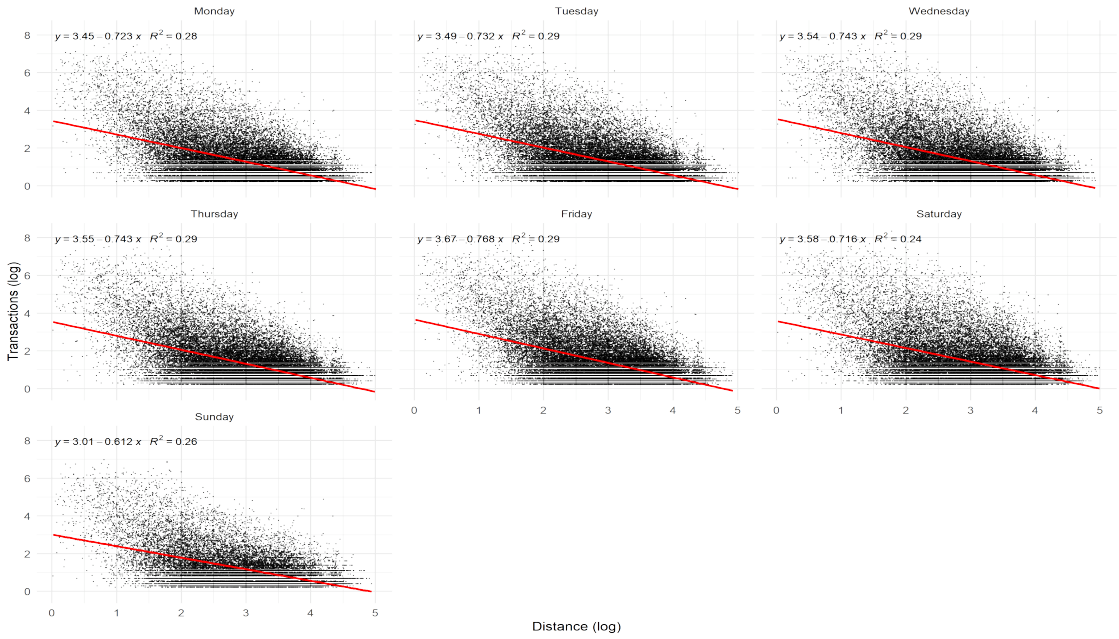
Figure 12: Relationship between presence counts and distance from home.




Notes: Each point represents a residence–destination pair on a given day of the week averaged over 10:00-11:30 and 14:00-15:30 time intervals.

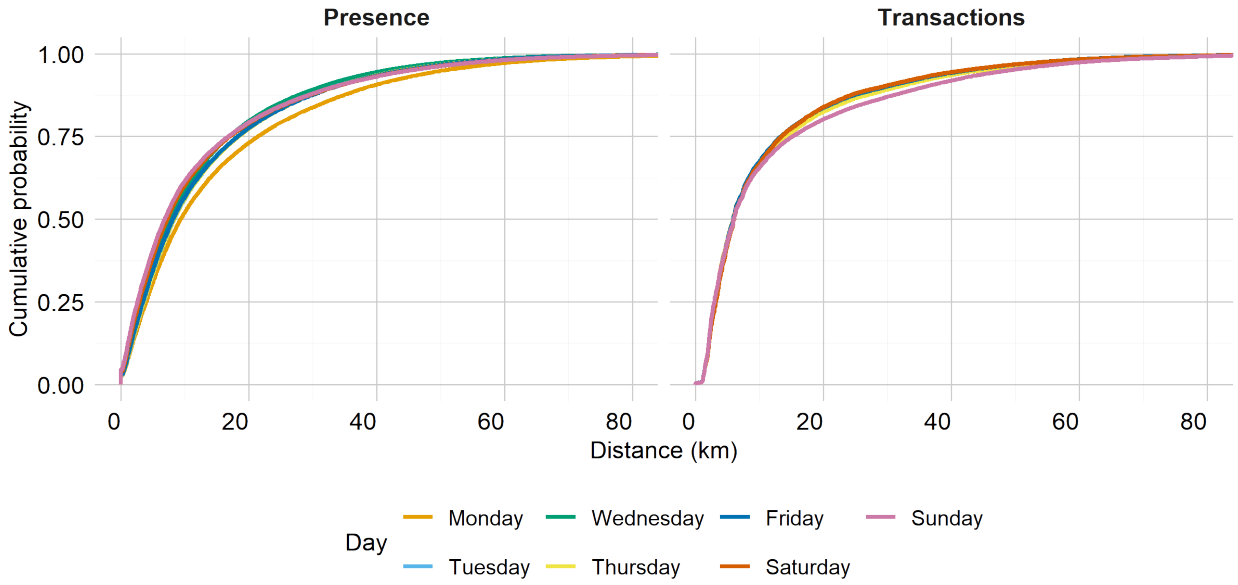
Figure 13 extends this analysis to economic transactions, plotting total daily transaction counts against distance from residential locations. This visualization demonstrates that transaction activity exhibits an even stronger distance decay pattern than general population presence. The steeper negative slope suggests that consumption activities are more spatially concentrated than overall mobility, with most transactions occurring within short distances from residential zones. This pattern aligns with our gravity model results showing higher distance elasticities for transactions compared to population flows.

Figure 13: Relationship between total daily transaction counts and distance from home. Each point represents a residence–destination pair on a given day of the week.




**Cumulative Distance Distributions.** Figure 14 presents cumulative distribution functions of traveled distances from residential locations, offering a comprehensive view of the spatial distribution of mobility patterns. The CDF curves allow direct comparison of how different types of activities (general presence vs. transactions) and different days of the week exhibit distinct distance distributions. The steeper curves for transaction-related distances compared to general presence confirm our quantitative finding that economic activities are more spatially concentrated. Day-of-week variations in the CDFs further illustrate temporal patterns in mobility, with weekend distributions typically showing different characteristics than weekdays.

Figure 14: Cumulative distribution functions of distances from home




**Quantitative Decomposition of Flow Patterns.** Table 16 provides a detailed quantitative decomposition of population mobility and transaction flows across different origin-destination zone pairs in the Lyon metropolitan area, offering empirical validation for our gravity model's spatial decay patterns. The statistics reveal striking differences between general population movements (Panel A) and economic transactions (Panel B) across four urban zone types. For population flows, we observe a clear monocentric pattern where Lyon-Lyon pairs dominate with mean flows of 36.04, while peripheral interactions show substantially lower volumes (rural-rural mean = 8.29) and higher variability, confirming our gravity model's distance decay estimates. Transaction flows exhibit even stronger spatial concentration, with Lyon-Lyon pairs averaging 55.92 transactions - nearly five times higher than the next most active pair - and 75% of peripheral zone pairs showing zero transactions. The data demonstrates that while population mobility follows a gradual distance decay pattern, economic transactions are highly concentrated in central zones, with sharp drop-offs in peripheral areas. This spatial disparity between general mobility and transaction activity underscores the persistent importance of physical proximity for economic interactions, even at micro-geographic scales. The statistics particularly reinforce our finding that transaction flows exhibit stronger distance sensitivity than population flows, with the urban core serving as the primary economic hub while peripheral zones show signs of economic isolation.

Table 16: Summary statistics of daily population and transaction flows by pair group

| Pair group | Min | Q1 | Median | Mean | Q3 | Max | N |
|---|---|---|---|---|---|---|---|
| ***Panel A: Population flows (MNO)*** | | | | | | | |
| Lyon - Lyon | 0 | 0 | 20 | 36.04 | 31.00 | 7,178.46 | 108,780 |
| Lyon - Inner suburbs | 0 | 0 | 0 | 10.77 | 22.00 | 2,084.79 | 165,228 |
| Lyon - Urban outer suburbs | 0 | 0 | 0 | 4.08 | 0.00 | 442.96 | 175,224 |
| Lyon - Rural outer suburbs | 0 | 0 | 0 | 1.81 | 0.00 | 538.12 | 227,556 |
| Inner suburbs - Lyon | 0 | 0 | 0 | 14.55 | 25.00 | 1,153.17 | 157,990 |
| Inner suburbs - Inner suburbs | 0 | 0 | 0 | 25.96 | 26.09 | 6,033.50 | 239,974 |
| Inner suburbs - Urban outer suburbs | 0 | 0 | 0 | 5.77 | 0.00 | 1,310.50 | 254,492 |
| Inner suburbs - Rural outer suburbs | 0 | 0 | 0 | 2.05 | 0.00 | 883.00 | 330,498 |
| Urban outer suburbs - Lyon | 0 | 0 | 0 | 5.42 | 0.00 | 399.96 | 189,070 |
| Urban outer suburbs - Inner suburbs | 0 | 0 | 0 | 7.08 | 0.00 | 1,292.50 | 287,182 |
| Urban outer suburbs - Urban outer suburbs | 0 | 0 | 0 | 21.59 | 0.00 | 6,617.75 | 304,556 |
| Urban outer suburbs - Rural outer suburbs | 0 | 0 | 0 | 6.57 | 0.00 | 4,287.54 | 395,514 |
| Rural outer suburbs - Lyon | 0 | 0 | 0 | 1.32 | 0.00 | 275.12 | 178,710 |
| Rural outer suburbs - Inner suburbs | 0 | 0 | 0 | 1.98 | 0.00 | 298.62 | 271,446 |
| Rural outer suburbs - Urban outer suburbs | 0 | 0 | 0 | 6.05 | 0.00 | 1,902.62 | 287,868 |
| Rural outer suburbs - Rural outer suburbs | 0 | 0 | 0 | 8.29 | 0.00 | 4,689.79 | 373,842 |
| ***Panel B: Transaction flows (CB)*** | | | | | | | |
| Lyon - Lyon | 0 | 2.5 | 9.75 | 55.92 | 35.25 | 3,147.00 | 11,655 |
| Lyon - Inner suburbs | 0 | 1.0 | 2.00 | 7.96 | 5.50 | 778.25 | 17,703 |
| Lyon - Urban outer suburbs | 0 | 0.0 | 1.00 | 2.27 | 2.25 | 275.75 | 18,774 |
| Lyon - Rural outer suburbs | 0 | 0.0 | 0.00 | 0.48 | 1.00 | 66.75 | 24,381 |
| Inner suburbs - Lyon | 0 | 1.0 | 2.00 | 7.00 | 4.75 | 3,002.75 | 36,260 |
| Inner suburbs - Inner suburbs | 0 | 0.0 | 1.00 | 11.46 | 2.75 | 3,272.75 | 55,076 |
| Inner suburbs - Urban outer suburbs | 0 | 0.0 | 0.00 | 1.96 | 1.50 | 1,045.75 | 58,408 |
| Inner suburbs - Rural outer suburbs | 0 | 0.0 | 0.00 | 0.32 | 0.00 | 404.00 | 75,852 |
| Urban outer suburbs - Lyon | 0 | 0.0 | 1.00 | 1.72 | 2.00 | 158.75 | 80,290 |
| Urban outer suburbs - Inner suburbs | 0 | 0.0 | 0.00 | 1.57 | 1.00 | 1,021.75 | 121,954 |
| Urban outer suburbs - Urban outer suburbs | 0 | 0.0 | 0.00 | 7.15 | 1.00 | 4,237.75 | 129,332 |
| Urban outer suburbs - Rural outer suburbs | 0 | 0.0 | 0.00 | 0.78 | 0.00 | 1,000.75 | 167,958 |
| Rural outer suburbs - Lyon | 0 | 0.0 | 0.00 | 0.59 | 1.00 | 68.00 | 64,750 |
| Rural outer suburbs - Inner suburbs | 0 | 0.0 | 0.00 | 0.39 | 0.00 | 138.25 | 98,350 |
| Rural outer suburbs - Urban outer suburbs | 0 | 0.0 | 0.00 | 1.53 | 1.00 | 710.25 | 104,300 |
| Rural outer suburbs - Rural outer suburbs | 0 | 0.0 | 0.00 | 1.12 | 0.00 | 1,432.25 | 135,450 |