



**HAL**  
open science

## On the Spatial Scope of Warehouse Activity: An Exploratory Study in France

David Guerrero, Jean Paul Hubert, Martin Koning, Nicolas Roelandt

► **To cite this version:**

David Guerrero, Jean Paul Hubert, Martin Koning, Nicolas Roelandt. On the Spatial Scope of Warehouse Activity: An Exploratory Study in France. *Journal of Transport Geography*, 2022, 99, 25p. 10.1016/j.jtrangeo.2022.103300 . hal-03551270

**HAL Id: hal-03551270**

**<https://hal.science/hal-03551270v1>**

Submitted on 1 Feb 2022

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# On the Spatial Scope of Warehouse Activity: An Exploratory Study in France

Guerrero D.<sup>1</sup>, Hubert J.-P.<sup>2</sup>, Koning M.<sup>1</sup>, Roelandt N.<sup>†</sup>

<sup>1</sup>Université Gustave Eiffel, IFSTTAR, AME-SPLOTT, 14-20 Bd Newton, 77420 Champs-sur-Marne

<sup>2</sup>Université Gustave Eiffel, IFSTTAR, AME-DEST, 14-20 Bd Newton, 77420 Champs-sur-Marne

<sup>†</sup>Université Gustave Eiffel, IFSTTAR, AME, 14-20 Bd Newton, 77420 Champs-sur-Marne

## Highlights

- Four warehouse location categories are defined based on a population potential
- Survey data describing the spatial scope and operational characteristics of warehouses is exploited
- Location categories and spatial scope are related using a logistic regression
- The results reveal a significant relationship between our location categories, spatial scope and some operational characteristics
- Differences were identified between the outbound scope of metropolitan and core categories

## Keywords

Warehouse location, spatial scope, freight transportation, population potential.

## Abstract

This article analyses the relationship between the location of warehouses, the spatial scope of their activities and some of their operational characteristics. It uses the results of a French national survey providing detailed characteristics for 1,974 warehouses of more than 5,000 square metres. A typology of four spatial categories is built based on a population potential with different values of friction ( $\alpha$ ). Applying logistic regression, the four categories are related to survey data describing the spatial scope of the activities of warehouses, controlling for variations in other characteristics such as seasonality and vehicle movements. The results show a significant relationship between the location of warehouses and their spatial scope. As expected, warehouses with local origins are more frequent outside the core (inner Paris) and metropolitan location categories (outer Paris and other large urban areas). The findings suggest there may be significant differences in the spatial scope of warehouse activity even within the same urban area.

## 1. Introduction

The purpose of this research is to analyse the relationship between the location of warehouses and the spatial scope of freight movements. It aims to contribute to a fuller understanding of the transport chains to which the warehouses belong, allowing an informed analysis of their location.

By the term "spatial scope" we mean a specialization of a warehouse on certain truck movements in space, e.g. deliveries or pick-ups in dense urban areas, between peripheral areas within a region or long journeys within the highway network. The spatial scope can be correlated to a range of vehicle

---

<sup>1</sup> Corresponding author : david.guerrero@ifsttar.fr

sizes and particular positions in the supply chain. Concerning the warehouse location, we consider different geographical spaces to which it gives access: close space, more or less dense, and far space, more or less connected by the highway network.

In recent decades, there have been significant shifts in the geography of warehouses (Cidell, 2010; Dablanc et al., 2014). As in other sectors, one of the main drivers of change in the logistics industry has been the search for scale economies and decreased transport costs (McKinnon, 2009; Hesse and Rodrigue, 2004). With the use of more extensive facilities, firms have been seeking to obtain higher efficiency levels, both in warehousing and truck operations (Kohn and Brodin, 2008). Logistics activities are being moved further and further away from the urban centres in search of large areas of affordable land (see, for example, Hesse 2004; Bowen, 2008).

Centrography studies, comparing the average distance of warehouses to a barycentre, have confirmed this decentralization trend in several large urban areas in many parts of the world (Dablanc and Rakotonarivo, 2010; Allen et al., 2012; Dablanc et al., 2014; Sakai et al., 2015; Giuliano and Kang, 2018). It is sometimes suggested that the increase in distances may result in an increase truck travel (Dablanc et al., 2014; Heitz and Dablanc, 2015; Aljohani & Thompson, 2016) which could limit this centrifugal movement. But this trend may also be the sign of an evolution of the logistic chains leading the warehouses to operate differently according to a different spatial scope, making logistics decentralization a potential area for policy measures. Recent studies, made possible by comprehensive and periodic surveys in the Tokyo Metropolitan Area, have shown that the type of truck traffic generated by warehouses may vary considerably depending on the geographical scope, suggesting that the balance between local and non-local shipments is an important factor (Sakai et al 2015, 2020a). Data on truck movements is needed to establish causal links (Giuliano and Kang, 2018). Unfortunately, the high cost of such types of survey is a barrier to their generalization to other countries and urban areas.

The aim of this paper is to help to fill this gap in the literature, by (a) building a typology of warehouse locations based on population potential with different values of friction ( $\alpha$ ); (b) by analysing the spatial orientation of freight flows; (c) by analysing the influence of the operational characteristics of warehouses on their locations. The results of a warehouse survey are analyzed, focusing on a selection of qualitative indicators related to the spatial scope of the warehouse's activity. Since the survey results are not significant at the level of urban areas, they were aggregated to form four location categories. The main contribution of this paper to the existing literature is to provide an informed analysis of the link between warehouse operational characteristics, including their spatial scope, and location patterns. More precisely, this research argues that the location of warehouses is tightly related to the spatial scope of freight movements.

As a preview of our results, we find that the spatial scope of warehouses varies depending on their location. This is particularly clear within the Paris region, where the warehouses located at the core are more focused on local deliveries than those located at the fringes. Conversely, the warehouses located at the fringes (and in four other large urban areas) are more oriented to non-local destinations. This result is non-trivial and calls into question former interpretations suggesting that warehouse decentralization leads to more truck travel.

The rest of the paper is organized as follows. The next section presents a short literature review on warehouse location, focusing on research that takes account of freight flows. The third section

presents the data and method and introduces the typology of location categories based on proximity to the population. The fourth section analyses the relationships between location categories and the spatial scope through logistic regression. The fifth section discusses the results, and the last section concludes the paper.

## 2. Literature review

### 2.1 Firms' location behaviour and accessibility

The subject of this paper, namely the influence of transport flows on facility location, has its roots in the pioneering work of Weber (1909). Weber's main hypothesis was that for manufacturing activities requiring inputs from more than one supplier, the location of a plant depends on the relative transportation costs arising from the locations of the various suppliers and the market. Solving this problem involves finding the point in a plane where the sum of the transportation costs of raw materials and the final product is a minimum (Smith et al, 2009). Weber's contribution set the ground for modern location theory and the p-median<sup>2</sup> model (Hakimi, 1965, ReVelle and Swain, 1970). Despite its simplicity<sup>3</sup>, the framework remains relevant for industries where minimizing the total transportation costs is essential, such as distribution and logistics (Ghosh and Craig, 1984, Farahani et al., 2010).

However, it is difficult to find precise information on the customers and suppliers, especially when there are many, and change over time. In such cases, a probabilistic approach is more appropriate. Location decisions can be analysed with discrete choice models relating the firm's characteristics and the attributes of each alternative location in terms of production factor costs and accessibility (McFadden, 1974, Arauzo-Carod et al. 2010).

The impact of accessibility on firms' location decisions remains a subject of discussion, because accessibility can have different meanings. The classical view, based on Weber's work, considers that industrial location decisions are driven by the minimization of transport cost. An alternative perspective suggests that business location decisions are more complex and decision-makers take account of other factors in addition to transportation accessibility. More precisely, it can be measured as (a) the distance to the most efficient transport infrastructure such a highway entrance or a train station (accessibility to network) or, as (b) a gravity type measure in which the opportunities at each possible origin/destination within a certain time or distance range which can be estimated by a distance decay (or friction) function.

Holguin-Veras et al. (2005) analysed the impact of accessibility to transport networks on the location decisions of firms moving to New Jersey. Using a discrete choice model, the authors found that firms place a higher value on areas with good highway access. The results also suggest that firms place a different value on accessibility at different ranges depending on the type of activity. Those selling goods and services to consumers place a premium on non-local (regional scale) accessibility (New York

---

<sup>2</sup> The p-median problem, which is similar to Weber's problem but in a network instead of a plane, consists in finding the locations of p points in a network to minimize the sum of the demand-weighted distances to the nearest facility (Daskin, 2008).

<sup>3</sup> The idea that industry locates to minimize its costs has been criticized for its lack of realism. "Industries try to make money in some combined short and long-run perspective, and if it takes money to make money, they spend it" (Alonso, 1990, p.27).

and Philadelphia metropolitan areas), while firms performing manufacturing or other services do not. Local, or short range, accessibility (within New Jersey) is also an important factor of choice.

De Bok and Sanders (2005) analysed the impact of accessibility on the location of firms in the Netherlands. They found that accessibility to population appears only to be of minor importance in the locational preferences of firms, compared to wage levels or the availability of space. However, locations near a highway or with a train station are very attractive for specific industries.

Most of these works conclude that location patterns of warehouses are mostly related to the characteristics of the areas in which they locate. The works on the operational characteristics of warehouses are rare, in particular due to data limitation.

## 2.2 Changes in the location of logistics activities

The location determinants of logistics activities have been less extensively studied than manufacturing activities. In recent decades, logistics activities have decentralized from their historic locations in the urban cores to the urban fringes. In most developed economies, this resulted not only from major changes in the global logistics industries (Hesse and Rodrigue, 2004) but also other factors common to all suburban growth (Cidell, 2010). These include the generalization of road transport and the resulting congestion in urban centres (Dablanc et al., 2014). An increasing share of new logistic facilities is, therefore, built in the periphery of urban areas, where large parcels of land close to highway interchanges are available at lower prices (Verhetsel et al., 2015; Jakubicek and Woudsma, 2019). Another reason for the observed logistics decentralization in large urban areas may be that the higher density of pick-ups and deliveries favours transport efficiency and make less necessary for warehouses to be close to the urban centre (Combes, 2019).

Within the specific context of France, several studies have examined spatial changes in logistics activities. Most of them are focused on the Paris region (*Ile-de-France*) or the extended Paris region (*Bassin Parisien*) (Dablanc and Rakotonarivo, 2010; Heitz and Dablanc, 2015). They highlight a trend towards decentralization, affecting not only large urban areas but also smaller ones. However, the movement of logistics activities to fringes is not straightforward: its intensity can vary considerably between urban areas and even within a single metropolitan area over different periods. Guerrero and Proulhac (2016) showed that the most frequent decentralization pattern in French urban areas is initial sprawl, followed by the emergence of a few suburban centres. In some cases, a period of decentralization is followed by recentralization (Guerrero and Proulhac, 2016; Heitz and Dablanc, 2015), sometimes due to the adoption of more conservative strategies on the part of warehouse investors and developers (Raimbault, 2016).

Heitz and Dablanc (2015) conducted a centrography study of the logistics firms in the extended Paris region between 2000 and 2012. They suggested that the development of logistics activities at the fringes of the Paris region might increase the distance travelled by distribution trucks (Heitz and Dablanc, 2015), which could be proved only with data on truck movements. Disaggregated information on truck movements, at the individual warehouse level, which could give a precise description of their spatial scope, is unavailable outside the Tokyo Metropolitan Area.

The geographical movements of logistics activities are explained by many factors, of which access to transport infrastructure is one (Appendix 1 provides an overview of this topic). Investigation of the influence of the individual characteristics of warehouses has so far been somewhat deficient compared

to other aspects (Masson and Petiot, 2012). The essential characteristics that could influence location decisions include the origins and destinations of freight flows. One study has recognized this factor as critical and highlighted the need for detailed analysis (Sakai et al, 2020b).

### 2.3 Linking the location of logistics facilities with freight flows

The findings suggest an association between decentralization and an increase in truck travel because of a rise in the distance to origins and destinations (Sakai et al., 2015). However, most of the increase is due to a wider distance range of origins and destinations of shipments, which means a change of the spatial scope, with only a marginal contribution from warehouse decentralization. Moreover, when the optimum<sup>4</sup> locations of warehouses for different types of commodities were compared, it was found that being close to the urban centre is not systematically optimal (Sakai et al., 2020b). The authors explored the relationship between location of warehouses and the transport distances covered by the shipments. They found that the warehouses handling local shipments were often located close to the optimal location, where the number of truck-kilometres could be reduced. In contrast, the warehouses handling a substantial proportion of non-local shipments were generally located a long way from the optimal locations, in the periphery but close to highways.

A small-scale study performed in Hamburg, Germany, suggested that a concentration of logistics close to the urban centre would help limit truck travel (Wagner, 2010). However, this analysis is limited to a specific type of facility and assumes that shipments are handled at the same warehouses regardless of their origins and destinations. An interesting contribution to this topic has been done recently within the context of the Paris Metropolitan Area. Although it is based on the case of a single company (DB Schenker), it provides an interesting reflection on the optimal location of truck terminals with regard to the spatial organization of pick-ups and deliveries (Robichet and Niérat, 2020). This study compares the current location of the truck terminals with the optimum locations obtained by using the p-median model. The results show that when one company controls several truck terminals, only some terminals serve the city centre. Therefore, being located far from the city centre does not necessarily mean more truck-kilometres. In the case of the focal company (DB Schenker) its spatial organization is close to the optimal one, and hence the truck-kilometres can hardly be reduced.

The present work aims to contribute to this discussion by analyzing the spatial scope of warehouse activities with a 2010 survey on French warehouses providing an approximate picture of the origins and destinations of freight flows for one year. With regard to the work of Sakai et al 2020 focused on Tokyo MA, our analysis develops a nation-wide study taking into account other operational characteristics of warehouses such as seasonality and vehicle movements. We will also analyse the warehouse location in terms of population potentials (gravity-type accessibility) for different distance ranges.

## 3. Data and method

### 3.1 A detailed dataset on warehouse characteristics

We used data from a survey carried out by the French Ministry of Transport in 2010: *“Enquête sur l'activité des entrepôts et des établissements logistiques”*. This database contains information on

---

<sup>4</sup> The optimum location of a warehouse is usually determined by minimizing the trucking costs/distance to the points of pick-up/delivery. The optimum location varies depending on the number of authorized warehouses (Robichet and Niérat, 2020). More complex models can take into account both the warehouse and transport costs and find a solution that minimizes both.

individual establishments. It is not publicly available and can only be accessed under a confidentiality agreement. It provides qualitative information on the spatial location of origins and destinations, which it divides into three modalities: local (within the same NUTS2 region), national, and international. Other useful information for our study is the geographical location of the warehouse (xy coordinates), its floor size, number of employees, and characteristics of the activity. The data was collected through questionnaires sent to warehouse managers by regular mail or e-mail. The resulting dataset contains 1974 records from warehouses with a floor size of at least 5,000 square metres (CGDD, 2012a). This sample is representative of a total population of 4470 warehouses<sup>5</sup>. The sample broadly reflects the structure of the national population in terms of floor size and the type of activity. This survey has only been carried out once, so we are unable to discuss the location dynamics since 2010. However, due to the 2008/2009 crisis and recent sluggish economic activity, fewer new warehouses have been built than during the previous two decades. We can assume that the location patterns of warehouses have not radically changed in the last decade. One limitation of this database is that it does not make a clear distinction of the functions of logistics facilities, particularly between storage and cross-docking activities. Even if the operator's name is available in the database, the distinction cannot be inferred from it, since it is common that a same logistics provider operates two types of warehouses: (a) cross-docking facilities which are part of its own parcel delivery network, and, (b) dedicated warehouses for shippers, the main activity of a warehouse can hardly be inferred from the warehouse's owner or user. We use the word "warehouse", as a translation for the French "entrepôt", for the ease of reading, even if the logistics facilities may perform different functions, which may affect the location determinants.

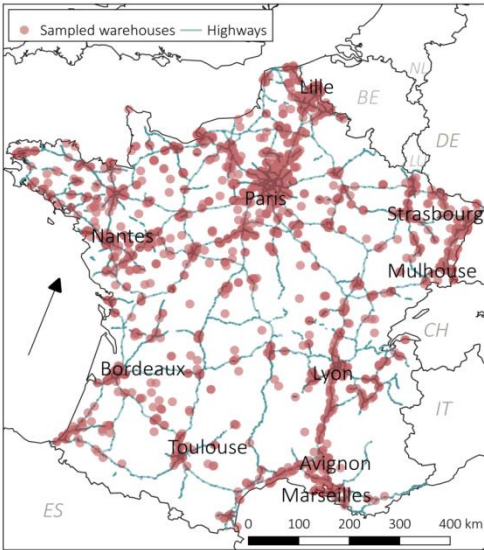


Figure 1. Sampled warehouses

Figure 1 maps the geographical distribution of the sampled warehouses. Most of them are located along highways with high concentrations around Paris and, to a lesser extent, a few other large urban areas. Given the high sampling rate of the survey, this map is expected to reflect well the real distribution of the warehouses.

---

<sup>5</sup> The survey sample is representative of the warehouses characteristics at the NUTS1 level but not at the level of urban areas. Therefore it could be misleading to conclude, for a single urban area, that there is an overrepresentation or underrepresentation of certain warehouses characteristics. This shortcoming has prompted us to conduct aggregated spatial categories.

### 3.2 Selected variables in the analysis of spatial scope and type of warehouse

We have selected fifteen variables from the survey database.

The characteristics related to the geographical scope (local, national, international) were available for origins and destinations. They were analysed as such in the descriptive analysis (Section 4.1). In the regression analysis (Section 4.2) they were grouped in two categories, simply distinguishing if the activity is local (within the same NUTS-2 region) or not. The table 1 reports the descriptive statistics for the sampled warehouses. It should be noted that data on the number of employees was only available for 74% of the observations (n=1460).

Three variables were built to characterize the magnitude of warehouse activity. Two dummy variables reflect large numbers of light vehicle (*van\_xl*) and truck movements (*heavy\_xl*), when the respective thresholds of 20 and 10 vehicles per day are met. In addition to the latter, our specification included the number of employees (*jobs*).

Several dummy variables were used to capture the site characteristics and type of activity: if the activity has a seasonal peak in summer (*temp\_sum*), if the warehouse is leased (*leased*), if the warehouse is within a logistics park (*logpark*), and if the warehouse is connected to the rail network (*railspur*).

Another possible source of variation in warehouse location is the proximity to markets and other activities. To capture the characteristics of the neighborhood of the warehouse we considered an extended version of the model by adding two variables related to the characteristics of the employment zone in which the warehouse is located: median household income (*Local\_HouseholdMedian\_inc*) and share of manufacturing in employment (*Local%EmployMNF*). The results of this model are presented in the appendix 5.

	Mean	Std Dev	Min	Max
<i>jobs</i>	97.38	162.75	1	4174
<i>van_xl</i>	0.11	0.32	0	1
<i>heavy_xl</i>	0.44	0.50	0	1
<i>temp_summer</i>	0.56	0.50	0	1
<i>loc_ori</i>	0.15	0.36	0	1
<i>loc_dest</i>	0.25	0.43	0	1
<i>leased</i>	0.48	0.50	0	1
<i>logpark</i>	0.63	0.48	0	1
<i>railspur</i>	0.16	0.37	0	1
<i>Local_Household_inc</i>	20 398	1 647	16 471	27 749
<i>Local%EmployMNF</i>	0.21	0.08	0	1

Table 1. Descriptive statistics of the variables used in the model

### 3.3 Considering location through population potentials

To study the locational characteristics of warehouses we explored classical approaches such as measuring the distance to air or rail terminals for different transport modes (e.g. Bowen, 2008,



Verhetsel et al., 2015). In the context of our study, conducted on the scale of the continental French territory as a whole, this approach has proven to be less relevant for several reasons:

In the case of air freight, 80% of the throughput of French airports (in tonnes) is processed by a single airport (Paris Charles-de-Gaulle)<sup>6</sup>. Although we are aware that the latter may attract logistics activities, it is difficult to isolate its effect from other characteristics of the capital region.

In the case of maritime transport, more than two thirds of the French foreign trade (in tonnage) are handled by three ports, of which only two are in France<sup>7</sup>. Curiously, only few of the surveyed warehouses are close to these two ports. For surface transport the roles of rail and waterway are marginal. In France the modal share of the road for goods transport is 89% (in tonnes-kilometre)<sup>8</sup>.

Given this context, and for the sake of simplicity, we decided to focus exclusively on road transport. We initially measured the time needed to reach an expressway ramp but it has not proven to be a discriminating criterion for the large warehouses investigated in the survey. Hence, we decided to use a potential, based on time-distance on truck, on the road network. This allowed taking into account the fact that some parts of the expressway network provided better access to markets than others.

We choose the population as a measure of market potential. It could be argued that large warehouses are more concerned with the location of firms or employment than with the places of residence. However, to a certain extent, firms tend to be close to highly populated areas to have access to clients, suppliers and employees. Moreover, the categories used to measure firms and employment greatly vary from one country to another, making cross-border comparisons difficult. This is particularly problematic in the context of France because many of the warehouses are concentrated in border areas such as Lille and Strasbourg. Hence, population seemed a good trade-off between the accuracy and availability of data, on one hand, and the relevance for the subject of study on the other hand. It should be noted that to compute population potentials of French warehouses we also used the populations and road networks of the neighbouring countries.

The *concept of accessibility* inspired how the population potential was considered in this work (Hansen, 1959). The population potential consists of two elements: one is related to the attractiveness (in terms of its population) of potential origins/destinations and the effort (time) required to reach them (See, for example, Rich (1978) and Handy (1992) for compelling explanations). In this paper, warehouse locations near the population are considered to be more attractive than distant ones. Following this philosophy, we used a classical function to estimate the proximity to the population at the municipal level (LAU2). The potential equation is:

$$P_i = \sum_j \frac{M_j}{d_{ij}^\alpha}$$

---

<sup>6</sup> Union des aéroports français

<sup>7</sup> Guerrero (2019)

<sup>8</sup> Ministère de la Transition Ecologique (2020)

Where  $P_i$  is the population potential at the warehouse location  $i$ ,  $M_j$  is the population (attractiveness) of municipality  $j$ , and  $d_{ij}$  is the time-distance by road (off-peak hour) between  $i$  and  $j$ . In the equation,  $\alpha$  is a friction parameter reflecting how the population is weighted by time-distance.

This definition implies substitutability between the transportation system (time-distance) and the population for determining the attractiveness of a warehouse location. It should be noted that the same score can result from very different trade-offs between population size and proximity. For example, a warehouse 5 minutes away from a small urban area may have the same score as another warehouse 2 hours away from a large urban area. But, with a higher value of friction, the time-distance of 2 hours will reduce the impact of the large city much more than for the small city à 5 minutes. Hence, two different friction parameters can provide a more discriminating description of warehouses locations.

For each warehouse, a potential is calculated by summing up the population of all the municipalities in France and of the neighboring countries, weighted by the travel time to go there. The metric used is the minimum travel times on a road network at off-peak hours. Table 2 provides an indication on how the weighting varies for different friction values ( $\alpha$ ) and travel times. It shows that when the friction ( $\alpha$ ) is -0.06, only 49% of the population of a municipality located 30 minutes away from the warehouse will be included in the population potential. For a lower value of friction (-0.016), a municipality located at 2 hours would have slightly the same contribution (50%).

Travel time	30 min	60 min	90 min	2h
Friction value ( $\alpha$ )	-0.06	-0.022	-0.016	-0.012
Weighting (share of the municipal population)	49%	42%	50%	56%

Table 2. Example on how the weighting varies depending on travel time and friction values

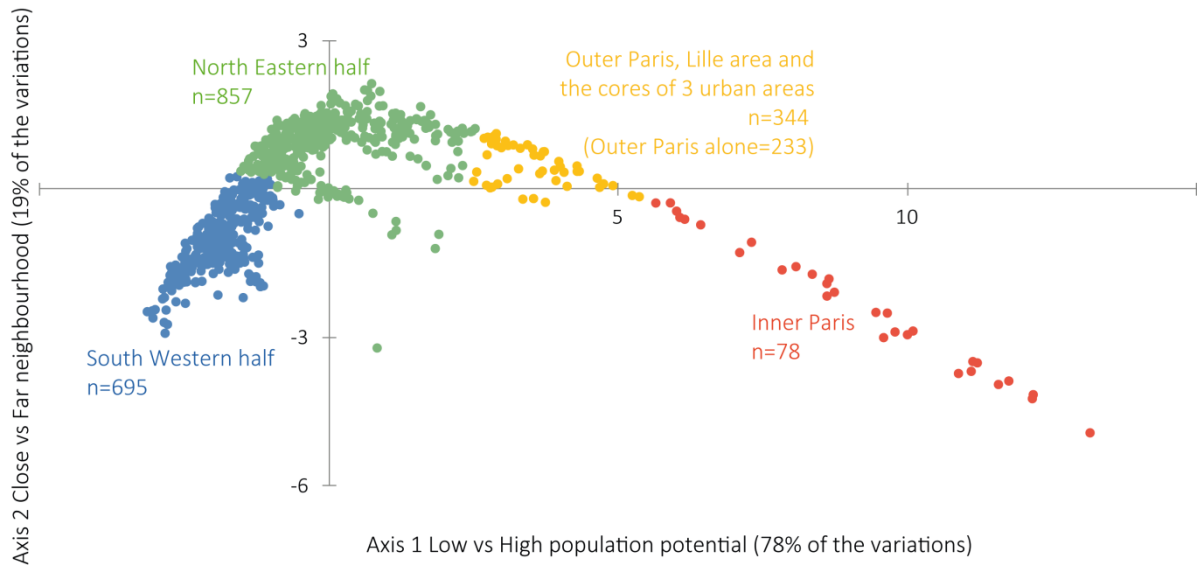
### 3.4 Location categories of warehouses and geographical interpretation

Since there is no pre-established definition of what correspond to the “close” or “far” neighbourhoods (spaces) accessible from a given warehouse, we have tested a population potentials using seven values of friction( $\alpha$ ), ranging from 0.005 to 0.15. As one would expect, these measures are strongly correlated. But what is interesting in choosing a location is to take advantage of slight discrepancies in this general correlation; e.g. lower potential in the close neighbourhood - i.e. low density - in an environment of high far population potential, etc.

On first analysis, we find that the first two axes of a Principal Component Analysis (PCA) allow this type of situation to be reflected. The first axis (78% of the variance) quantitatively translates the level of population potential, with the problem of undifferentiation mentioned above. But the second axis (19%) contrasts locations where the close neighbourhood contributes more or less strongly to the population potential.

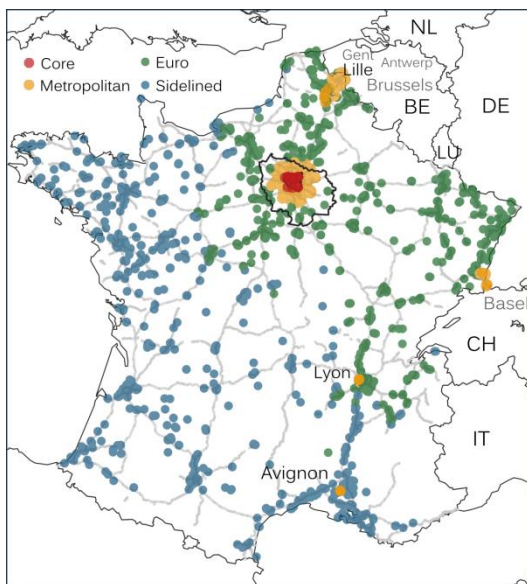
The resulting scatter plot has an inverted u-shape (Figure 2). The warehouses at both ends of the distribution owe most of their general population potential (combining all distance ranges/values of friction [ $\alpha$ ]) to their close neighbourhood, while the warehouses in the middle owe their score to their potential to remote surrounding metropolises. Four location categories were identified with respect

to both axes, using the k-means method. The red and yellow spots are characterized by high population potentials. This analysis is robust and does not depend on the clustering methods or distances used.

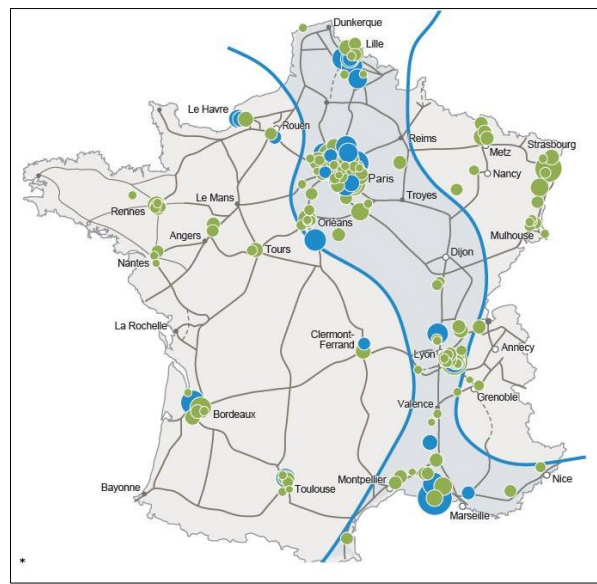


\* Categories resulting from a k-means typology and Principal Component Analysis. Raw data: Population potential with  $\alpha$ -values ranging from 0.005 to 0.15.

**Figure 2. Location categories of warehouses based on different population neighbourhoods**



**Figure 3. Map of the location categories**



**Figure 4. Map of the “logistics backbone” (circle: building permits, BNP Paribas, 2016)**

Figure 3 maps the resulting categories, which have been named as follows:

- A. **Core** (inner Paris, n=78) These warehouses are located in the central area of a highly populated region, which has a central position in France and western Europe. The contribution of the close neighbourhood (1-hour radius) to the general population potential is comparatively high. This area includes the international airports of Charles-de-Gaulle and Orly and Gennevilliers, which is the main river port in the country.
- B. **Metropolitan** [outer Paris, Lille area, and cores of 3 other urban areas, n=344 (outer Paris, n=233)]. These warehouses have a high population potential that results either from a

location in the periphery of Paris metropolitan area, or in the core of the metropolitan areas of the next size down after Paris (Lille, Lyon), or inside an area of large cities (Avignon, periphery of Basel). Belonging to this class is the result of a combination of population potentials where the far population neighbourhood can be preponderant (Avignon), as much as the close population neighbourhood.

- C. **Euro (North-East half, n=857)** A high far population neighbourhood is generally the key to membership of this class which is spread in the North East part of France.
- D. **Sidelined (South-West half, n=695)** The far population neighbourhood is relatively weaker due to an off-centre location in France and Europe. Therefore, the categories Euro and Sidelined are more separated along the vertical axis. No clear hierarchy exists between metropolitan areas where spots of categories Euro and Sidelined are found.

The map in Figure 4 shows the “logistics backbone,” a schematic representation of a large area of logistics hot spots published by a real estate investor (BNP Paribas, 2016). The comparison between Figures 3 and 4 is insightful. While the categories Metropolitan and Core are mostly located in the “logistics backbone,” this applies to only part of the Euro category. The main differences are in the East and South-East of France. The two maps share some common features, suggesting that our categories, which were defined on the basis of population potentials, are not disconnected from the view of logistics investors and developers. The last two clusters cover large logistics regions and mask a significant internal diversity, which may be related to the size of the cities and their position in the European corridors. In this first analysis, at nation-wide scale, these discrepancies in population potential are less significant than those in the region around Paris. But it would be useful to highlight them in subsequent studies, which focus more on the regions and less on the French capital.

## 4. Relationship between location and warehouse characteristics

### 4.1 Descriptive analysis of variables related to the spatial scope

Our hypothesis is that the close and far population neighbourhoods matter in the location of warehouses. One can expect that the far population neighbourhood would matter more for warehouses serving distant origins and destinations (non-local). Conversely, the close population neighbourhood is expected to matter more for warehouses oriented to near origins and destinations (local).

Two questions in the survey were associated with the spatial scope of inbound and outbound flows, distinguishing between three levels: local, national and international. Figure 5 schematically presents the resulting nine modalities. A series of tests were carried out to ascertain whether there was dependency between the rows and columns (Appendix 2). The results were all statistically significant and confirmed the relationship between the location categories and the modalities related to spatial scope.

The most frequent modality involves international origins and destinations (“international-international”), representing about 40% of the responses. Other frequent modalities are “international-national” and “national-national,” involving 17.7% and 12.4% of respondents, respectively. Local destinations were less commonly reported: “national-local” and “international-local” involved respectively 9% and 6.5% of warehouses.

To reflect the distribution of modalities between the location categories, we measured the deviations from the average profile. These were calculated as the share of a modality at a location category divided by the share of the same modality within the national profile.

The analysis of deviations in Figure 5 shows some regularity. The warehouses with local origins or local destinations are 1.4-1.7 times more frequent in the category Sidelined than in the average profile. The warehouses in the category Euro are close to the average profile, with a slight under-representation of local origins. Warehouses with national destinations are frequent in the Metropolitan profile. The last observation is that warehouses with international origins and destinations are particularly well represented in the Core category. This is consistent with previous studies, highlighting the importance of the gateway function of three logistics clusters in the inner suburbs of Paris: around the port of Gennevilliers to the north, Charles-de-Gaulle airport to the northeast, and the Orly–Rungis centre to the south (Heitz and Dablanc, 2015, Sakai et al., 2020a). Given the critically high number of observations of the modalities “international-national,” “international-international”, the relationships with the categories Core and Metropolitan are particularly significant (long vectors), as shown by the results of a Factor Analysis (Appendix 3). It clearly shows a contrast between local origins and destinations on the one hand and national and international ones on the other.

These results support the hypothesis of a link between the population neighbourhood of the warehouse and the spatial scope of freight flows. They reveal on one hand, the non-local destinations of the warehouses belonging to the Euro and Metropolitan categories, and, on the other hand, the local origins and destinations of the category Sidelined.

O \ D	Local n=363 (18.4%)	National n=692 (35.1%)	Int'l n=919 (46.6%)
Local n=221 (11.2%)	Loc_Loc Deviations to the average profile 	Loc_Nat 	Loc_Int 
Nation n=484 (24.5%)	Nat_Loc 	Nat_Nat 	Nat_Int 
Int'l n=1269 (64.3%)	Int_Loc 	Int_Nat 	Int_Int 

Note for the reader  
 - The modality international-International (Int\_Int) is 1.4 times (or 40%) more frequent in the **Category Core**  
 10% less (0.9) frequent in the **Category Sidelined**

Figure 5. Survey results. The spatial scale of OD flows (asterisks \* when the number of observations < 30)

#### 4.2 More comprehensive analysis of the relationship between location profiles and warehouse characteristics

The question is whether the preference for a type of location can be inferred from the characteristics of the spatial scope and the mode of operation of the warehouse. Some of these relationships between location categories and spatial scope may hide relationships with non-spatial factors such as size, seasonality or vehicle movements. We applied a logistic regression to controlling for these variations.

Our analysis aims to explain the location of warehouses in one of the four categories using several characteristics.

Because these characteristics are invariant across alternatives, the multinomial logit is an appropriate model. This model, which is also called the baseline-category logit model, simultaneously uses all the pairs of categories (Agresti, 2019). The category Euro was considered as the baseline to which the other categories are compared. The baseline category logits are:

$$\log\left(\frac{\pi_j}{\pi_c}\right), j = 1, \dots, c - 1.$$

Where  $\pi$  is the probability of belonging to the location category  $j$ , and  $c$  denotes the total number of categories.

For  $c=4$ , the model uses  $\log\left(\frac{\pi_1}{\pi_3}\right)$ ,  $\log\left(\frac{\pi_2}{\pi_3}\right)$  and  $\log\left(\frac{\pi_4}{\pi_3}\right)$ . Conditional on the response falling into category  $j$  or category  $c$ ,  $\log\left(\frac{\pi_j}{\pi_c}\right)$  is the log odds that the response is  $j$ .

The model has  $c-1$  equations, with separate parameters for each. The effects vary according to the category paired with the baseline. With  $p$  explanatory variables, the model is:

$$\log\left(\frac{\pi_j}{\pi_c}\right) = \alpha_j + \beta_{j1}x_1 + \beta_{j2}x_2 + \dots + \beta_{jp}x_p, j = 1, \dots, c - 1.$$

Where  $\beta_{j1}, \dots, \beta_{jp}$  are  $p$  vectors of unknown regression parameters, each of which is different.  $x_1, \dots, x_p$  are explanatory variables and are constant across alternatives. For each explanatory variable, different logits have different effects. For this model structure, the odds of being on one category over another do not depend on the other alternatives in the choice set or on their values for the explanatory variables. This property of the model is called independence from irrelevant alternatives (IIA). This hypothesis seems unrealistic in the case of warehouses for which alternatives are not always clearly distinct. However, it has been shown that the violation of the IIA assumption is not necessarily problematic when the model is not predictive (Train, 2013), which is the case in this research.

Moreover, the sense of causality between the operational characteristics of warehouses and the location may not be straightforward. Warehouses' characteristics such the number of employees or vehicle movements may be influenced by the location of the firm, and not the opposite. A possible solution to this endogeneity bias is the instrumental variables approach, which has been already been implemented in other studies on the location of logistics activities (see, for exemple, Holl and Mariotti, 2018b). Unfortunately in our case data availability was an issue, and our tests and conceptual arguments failed

to identify a useful instrument. Hence, we proceeded without.

	Core		Metropolitan		Sideline	
	Est	Std Error	Est	Std Error	Est	Std Error
Intercept	-3.5957***	0.4244	-1.7333***	0.2188	-0.3618**	0.1512
jobs	0.0069***	0.00202	0.00531***	0.00145	0.000942	0.0013
van_xl	1.4857***	0.3421	0.8629***	0.2212	-0.0394	0.2217
heavy_xl	0.6184	0.4146	0.7589***	0.2008	0.0741	0.1656
temp_summer	-0.3658	0.3071	-0.6166***	0.1606	0.0263	0.1249
loc_ori	-0.119	0.4967	-1.1408***	0.3494	0.6083***	0.1608
loc_dest	0.7655**	0.3276	-0.2394	0.2084	0.8923***	0.1405
leased	0.2349	0.3102	0.322**	0.1625	-0.1747	0.124
logpark	0.4825	0.3437	0.6341***	0.1844	-0.1728	0.1275
railspur	-1.7388**	0.7376	-0.6466***	0.2281	-0.301*	0.1678
jobs*heavy_xl	-	0.00249	-	0.0015	-0.00152	0.00142
N Obs	1460					

**Table 3. Multinomial logit results. Base model. The Euro profile was used as a reference (Significant at 10%\*/5%\*\*/1%\*\*\* level)**

Table 3 shows the results of the base model in which the dependent variable is the location category and the independent variables are dummy variables related to the warehouses' operational characteristics. Warehouses having large number of employees are more frequent in the categories Core and Metropolitan. Massive light vehicle traffic (van\_xl) is significant and positively correlated with the probability of being located within the categories Core and Metropolitan. There is also a positive association between massive truck traffic but only with the category Metropolitan. We found a significant negative interaction between jobs and heavy\_xl, which means that the warehouses generating massive truck traffic tend to have fewer employees than the rest.

Warehouses with a seasonal peak in the summer (temp\_summer) tend to be located at some distance from the core and premium categories. The latter would have lower average levels of utilization over the year. Therefore they are expected to be located in less central places than activities ensuring high levels of utilization during all the year.

The relationship between logistics parks and category Metropolitan is significant and positive. This could be related to conservative attitudes among public and private developers, who tend to give priority to locations with good access to populated areas but where land is still available. Warehouses with a direct connection to the rail network (rail spur) are underrepresented in all the categories as compared to the baseline category (Euro).

The characteristics related to the spatial scope of warehouse activity confirm some of the expectations mentioned in the descriptive analysis. The model confirms the cross table OD scopes and location categories, except for the non-local origins in the class "Core". This does not mean that gateways are poorly represented within the category Core, but rather that the gateway function of inner Paris is somewhat attenuated by the fact that the warehouses in the baseline category (Euro) are highly oriented towards international origins and destinations.

All in all, the logistic regression confirms two major results related to the scope of warehouses. The Metropolitan profile (orange) is negatively associated with local origins. The positive association between the category Core (red) and local destinations has been confirmed.

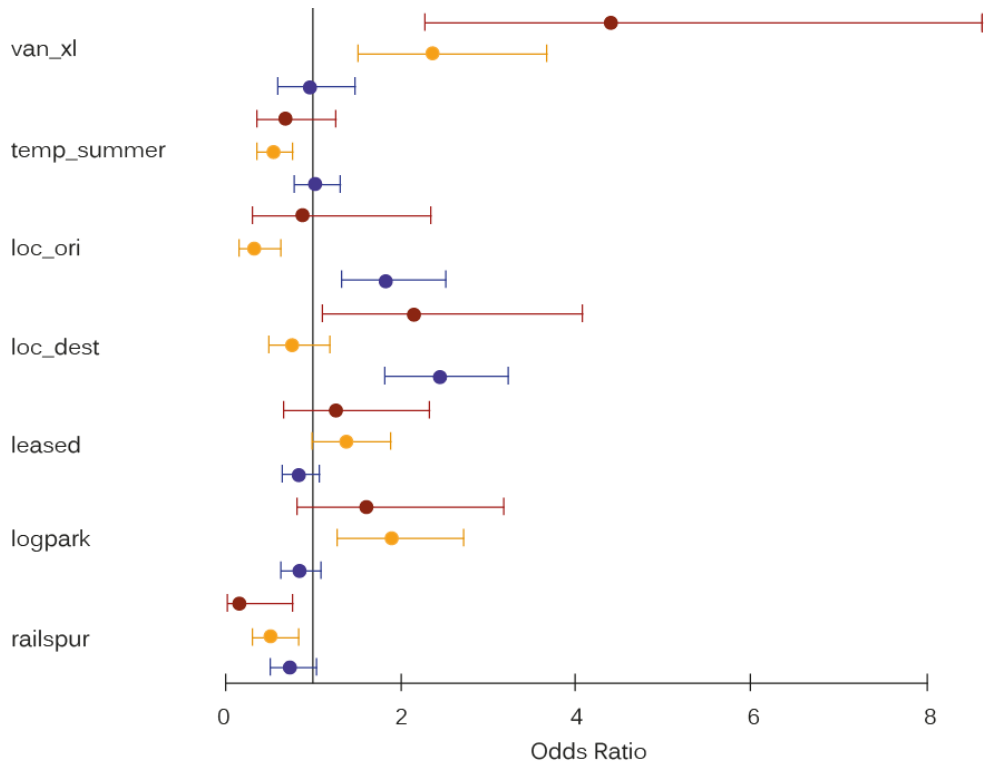


Figure 6. Odds ratio with 95% Wald Confidence limits

### 4.3 Discussion

In general, the results of the model support the hypothesis of a relationship between the location categories and the spatial scope of freight flows. Figure 6 shows the odds ratio measures for the main variables. It shows that the odds of being in the profile Sidelined are considerably higher (by a factor of 1.8-2.4) for warehouses having local origins and destinations.

The warehouses belonging to the categories Core and Metropolitan share some common characteristics. The odds of being in these two categories are higher for warehouses generating large amounts of light vehicle traffic (by a factor of between 2.4-4.4) or that are not affected by summer seasonality (by a factor of about 0.5-0.7). Conversely, for the warehouses that are connected to the rail network (railspur), the odds of being located in the two former profiles are considerably lower (by a factor of between 2 and 4). To sum up, characteristics such as seasonality, the relative absence of rail facilities, and the movement of light vehicles, point to more intensive utilization of warehouses in Metropolitan and Core locations. This is in line with previous studies on the relationship between location and intensity of warehouse use (Kang, 2018). The high cost of space in these areas may limit the use of railways.

However, there are important differences between the categories Core and Metropolitan regarding the spatial scope of outbound flows. While the former is more focused on local deliveries<sup>9</sup> (stronger

<sup>9</sup> The Figure 6 shows that, for the warehouses located in the category Core, the odds of having local destinations are higher than those in the baseline category (Euro). This does not mean that warehouses in the Core category



association with `van_xl` in Core than in Metropolitan) the latter is more oriented to non-local deliveries. The relationship between the profile Metropolitan and the large use of heavy vehicles (`heavy_xl`) supports this hypothesis. Heavy vehicles are generally used for non-local shipments. This is consistent with truck size limitations, and increasing trip frequency and costs that apply in densely populated areas (Giuliano et al., 2018).

Overall, the results suggest that the trips generated by warehouses in the category Metropolitan are larger in scope than those generated by the warehouses in the category Core, which seems to be more local. These differences may also result from differences in the types of logistics activities between inner Paris and outer Paris and other urban areas. While the first category probably tends to be more specialized in wholesale trade, partly oriented towards retailers close to the urban centre, the second may be more focused on other logistics activities (such as freight forwarding, warehousing, and freight transport), requiring better access to highways (Strale, 2019). All in all, these results support the hypothesis put forward by Sakai et al. (2020b) that the location of warehouses on the fringes of metropolitan areas are not necessarily suboptimal with regard to minimizing truck travel.

#### 4.4 Robustness analyses

We assessed the relationship between warehouses characteristics and their location. Our statistical analysis, however, may suffer from an omitted variables bias. To limit this problem we analyzed the influence of the corporate ownership of the warehouse on location choice. Another potential problem of our study is that location decisions may also be affected by the characteristics of the local areas. To check their influence we tested the influence of local variables in an extended version of the model.

Warehouses are not located individually, and large corporate entities seek to optimize their entire network (Robichet and Niérat, 2020). Therefore, the location choices may differ depending on the number of warehouses controlled by a single entity. In our dataset, some warehouses do belong to the same firm and the same manager may have decided where to locate various warehouses. To check for the whether the characteristics of the owner impact or not the results of the model, we divided our sample in two subgroups: a same corporate entity manages one (single ownership) or several warehouses (multi-ownership).

First, from a purely descriptive point of view, it should be first noticed that the distributions of the two subgroups between the location categories are similar. We also looked at the regional distribution at the NUTS1 level (8 regions) only minor deviations were found. Second, we have run the base model on these two sub-samples and we do not observe any difference in both sign, size and significance of parameters (see appendix 4). Based on these results we may conclude that localization choices follow similar patterns<sup>10</sup>.

---

are predominantly local in their destinations, but that the latter are more frequent than in the baseline category. Therefore, there is no major contradiction with the results presented in the Figure 5, in which it can be seen that the warehouses located in the category Core are predominantly international both in their origins and destinations.

<sup>10</sup> In order to question this intuition at the individual level, we have considered the sample of warehouses that belong to multi-establishments groups and we have introduced dummy variables that characterize these groups. These individual effects could control for the propensity of the group manager to locate different warehouses in similar places, because of common and unobserved factors. Unfortunately, the results of this model and its overall significance were highly questionable due to the large amount of new parameters to be estimated.

Finally, to take into account variations on the characteristics of the local area in which the warehouse is located (“zone d’emploi”), the extended model includes two proxies related to the wealth (*Local\_HouseholdMedian\_inc*) and the economic specialization in manufacturing (*Local%EmployMNF*). The appendix 5 shows the corresponding results. Unsurprisingly, both variables are significant. However the signs and statistical significance of the other variables are overall similar to those of the base model.

## 4 Conclusion

In this paper, we have tested the hypothesis that the locations of warehouses can explain some of the heterogeneity of the spatial scope of their activity, that is to say the origins and destinations of the freight handled in the warehouses. We have used the results of a nation-wide survey providing qualitative information on the spatial scope of the outbound and inbound flows of warehouses. To analyse the results of the survey from a spatial perspective, we have developed a typology of four location categories based on proximity to the population. This typology provides a stimulating schematization of the complexity of the spatial patterns of population potentials at different scales. It can be reproduced in other contexts and deepened, for example by calculating potentials for other economic entities (industries, retailers...) with which the warehouses interact instead of the simple population that serves as a proxy.

At its scale of analysis and the limitations of the data, the study found correlations between qualitative differentiations in the spatial profile and mode of operation of warehouses on the one hand, and combinations of population potentials at different distance ranges on the other. It extends for the French context concordant results obtained with more detailed data on warehouses in the Tokyo metropolitan area. Our descriptive analysis of the survey results provides empirical evidence of the varying spatial scopes of the warehouses depending on their location. It shows that local origins and destinations are more frequent in the category Sidelined. Conversely, the overrepresentation of warehouses generating large movements of heavy trucks in the category Metropolitan reveals a stronger orientation towards non-local destinations.

The main results from the descriptive analysis have been confirmed by logistic regression controlling for variations in other characteristics. The activities of the warehouses belonging to the categories Core and Metropolitan generally produce more intensive vehicle movements than the other categories. A significant finding of this work is the different spatial scope of destinations between warehouses in the categories Core (inner Paris) and Metropolitan (outer Paris and four other urban areas). While the former are more oriented towards local destinations, the latter are more focused on non-local destinations. This difference is further confirmed by the information on light vehicles and trucks. Given the contribution of outer Paris to the category Metropolitan (67% of the warehouses), our findings suggest significant variations in the spatial scopes of the activities of warehouses within the Paris region. This backs up the findings of previous studies, tending to confirm that interpretations of geographical shifts in warehouse location cannot be exclusively considered through the prism of distance from a barycentre.

This discussion on the location patterns of large warehouses can also be updated and extended by developing case studies for specific firms, such for example within the field of e-commerce. Recent case studies focused on Amazon in the US showed that the location of large facilities such fulfilment

centres is highly selective, to maximize regional market accessibility (Rodrigue, 2020, Blanquart et al, 2019). A similar trend occurs in France (Berthelot, 2021), where all the main Amazon's fulfilment centres are located in the North and the East of the country, which correspond to the Euro and Metropolitan categories proposed in this study (Appendix 6). This concentration is even higher for the volumes, with the three fulfilment centres in the Lille and Paris regions (Metropolitan category) handling two thirds of the parcels (Appendix 6). Of course, fulfilment centres are only part of a broader system of logistics facilities that could be less spatially concentrated. The inclusion of the different layers of the distribution system (including small size facilities) in future analyses could reveal more specific location factors than the sole market accessibility. The results of the present study are useful for policymakers. They suggest that the re-location of warehouses in the core of highly congested megacities may only be optimal for activities primarily oriented towards the urban core. For other logistics activities, the scope of activity should be examined very closely and thoroughly, with finer typologies notably within the "Euro" and "Sidelined" regional types.

## References

- Agresti, A. (2019). *An introduction to categorical data analysis*. John Wiley & Sons.
- Aljohani, K., Thompson, R.G. (2016). Impacts of logistics sprawl on the urban environment and logistics: taxonomy and review of literature. *Journal of Transport Geography*, 57, 255–263.
- Allen, J., Browne, M., Cherrett, T. (2012). Investigating relationships between road freight transport, facility location, logistics management, and urban form. *Journal of Transport Geography*, 24, 45–57.
- Alonso, W. (1990). From Alfred Weber to Max: the shifting style of regional policy. In *Dynamics and conflict in regional structural change* (pp. 25-41). Palgrave Macmillan, London.
- Arauzo-Carod, J. M., Liviano-Solis, D., & Manjón-Antolín, M. (2010). Empirical studies in industrial location: an assessment of their methods and results. *Journal of Regional Science*, 50(3), 685-711.
- Blanquart, C., Liu, Z., Maudhuit, N., & Zérroual, T. (2019). Les entrepôts du e-commerce: quelles spécificités? Une étude du cas Amazon. *Logistique & Management*, 27(1), 44-54.
- BNP Paribas (2016) *Immobilier Tertiaire. Marché des entrepôts en France*. <https://www.bnppre.fr/actualite/marche/20160722/immobilier-tertiaire-marche-des-entrepots-5-000-m2-en-france-au-t2-2016-102.html> (checked in December 2019)
- Bowen Jr, J. T. (2008). Moving places: the geography of warehousing in the US. *Journal of Transport Geography*, 16(6), 379-387.
- CGDD (2012a). Les entrepôts et leur activité en 2010 (Wrap-up of the results of the national warehouse survey), *SOeS Chiffres et statistiques*, July, 12 p.
- CGDD (2012b). Bilan qualité de l'enquête sur l'activité des entrepôts et des plates-formes logistiques en 2010 (Quality assessment of the national warehouse survey), *SOeS Chiffres et statistiques*, August, 13 p.

- Cidell, J. (2010). Concentration and decentralization: The new geography of freight distribution in US metropolitan areas. *Journal of Transport Geography*, 18(3), 363-371.
- Combes, F. (2019). Equilibrium and Optimal Location of Warehouses in Urban Areas: A Theoretical Analysis with Implications for Urban Logistics. *Transportation Research Record*, 2673(5), 262-271.
- Dablanc, L., Ogilvie, S., Goodchild, A. (2014). Logistics sprawl: differential warehousing development patterns in Los Angeles, California, and Seattle, Washington. *Transportation Research. Record* 2410, 105–112.
- Dablanc, L., Rakotonarivo, D. (2010). The impacts of logistics sprawl: How does the location of parcel transport terminals affect the energy efficiency of goods' movements in Paris, and what can we do about it? *Procedia Social Behavioral Sciences* 2 (3), 6087–6096.
- Daskin, M. S. (2008). What you should know about location modeling. *Naval Research Logistics (NRL)*, 55(4), 283-294.
- De Bok, M., & Sanders, F. (2005). Firm relocation and accessibility of locations: Empirical results from the Netherlands. *Transportation research record*, 1902(1), 35-43.
- Farahani, R. Z., SteadieSeifi, M., & Asgari, N. (2010). Multiple criteria facility location problems: A survey. *Applied mathematical modelling*, 34(7), 1689-1709.
- Gingerich, K., & Maoh, H. (2019). The role of airport proximity on warehouse location and associated truck trips: Evidence from Toronto, Ontario. *Journal of Transport Geography*, 74, 97-109.
- Giuliano, G., Kang, S. (2018). Spatial dynamics of the logistics industry: Evidence from California. *Journal of Transport Geography*, 66, 248-258
- Giuliano, G., Kang, S., Yuan, Q. (2018). Using proxies to describe the metropolitan freight landscape. *Urban Studies*, 55(6), 1346-1363.
- Guerrero, D., Proulhac, L. (2016). La dynamique spatiale des activités logistiques dans les aires urbaines françaises. *Cybergeo : European Journal of Geography*, 773, DOI : 10.4000/cybergeo.27517
- Guerrero, D. (2019). Impacts of transport connections on port hinterlands. *Regional Studies*, 53(4), 540-549.
- Hakimi, S. L. (1965). Optimum distribution of switching centers in a communication network and some related graph theoretic problems. *Operations research*, 13(3), 462-475.
- Handy, S. L. (1992). Regional versus local accessibility: neo-traditional development and its implications for non-work travel. *Built Environment (1978-)*, 253-267.
- Heitz, A., Dablanc, L. (2015). Logistics Spatial Patterns in Paris. Rise of Paris Basin as Logistics Megaregion, *Transportation Research Record*, 2477, p. 76-84.
- Hansen, W. G. (1959). How accessibility shapes land use. *Journal of the American Institute of planners*, 25(2), 73-76.

- Hesse, M. (2004). Land for logistics: locational dynamics, real estate markets and political regulation of regional distribution complexes. *Tijdschrift voor Economische en Sociale Geografie*, 95 (2), 162–173.
- Hesse, M. (2008). *The City as a Terminal: The Urban Context of Logistics and Freight Transport*. Ashgate Publishing Company, Hampshire, England.
- Hesse, M., Rodrigue, J.P. (2004). The transport geography of logistics and freight distribution. *Journal of Transport Geography*, 12 (3), 171–184.
- Holguin-Veras, J., Xu, N., Levinson, H. S., Paaswell, R. E., Mcknight, C. E., Weiner, R. D., ... & Ozmen-Ertekin, D. (2005). An investigation on the aggregate behavior of firm relocations to New Jersey (1990–1999) and the underlying market elasticities. *Networks and Spatial Economics*, 5(3), 293-331.
- Holl, A., Mariotti, I. (2018a). The geography of logistics firm location: the role of accessibility. *Networks and Spatial Economics*, 18(2), 337-361.
- Holl, A., & Mariotti, I. (2018b). Highways and firm performance in the logistics industry. *Journal of Transport Geography*, 72, 139-150.
- Jakubicek, P., Woudsma, C. (2011). Proximity, land, labor, and planning? Logistics industry perspectives on facility location. *Transp. Lett.* 3 (3), 161–173.
- Kang, S. (2018). Warehouse location choice: A case study in Los Angeles, CA. *Journal of Transport Geography*, <https://doi.org/10.1016/j.jtrangeo.2018.08.007>
- Kang, S. (2020). Relative logistics sprawl: Measuring changes in the relative distribution from warehouses to logistics businesses and the general population. *Journal of Transport Geography*, 83, 102636.
- Kohn, C., Brodin, M.H. (2008). Centralised distribution systems and the environment: how increased transport work can decrease the environmental impact of logistics. *Int J Log Res Appl* 11 (3), 229–245.
- Masson, S., Petiot, R. (2012). Attractivité territoriale, infrastructures logistiques et développement durable. *Les Cahiers Scientifiques du transport*, 61, 63-90.
- McFadden, D. (1974). Conditional logit analysis of qualitative choice behaviour. In: Zarembka, P. (Ed.), *Frontiers in Econometrics*. Academic Press, 105–142.
- McKinnon, A. (2009). The present and future land requirements of logistical activities. *Land Use Policy*, 26, 293-301.
- Ministère de la Transition Ecologique (2020) Chiffres clés du transport – Edition 2020, <https://www.statistiques.developpement-durable.gouv.fr/sites/default/files/2020-03/datalab-66-chiffres-cles-transport-edition-2020-mars2020.pdf>
- Raimbault, N. (2016). Ancrer le capital dans les flux logistiques: la financiarisation de l'immobilier logistique. *Revue d'Économie Régionale Urbaine* 2016 (1), 131–154.
- ReVelle, C. S., & Swain, R. W. (1970). Central facilities location. *Geographical analysis*, 2(1), 30-42.

- Rich, D. C. (1978). Population potential, potential transportation cost and industrial location. *Area*, 222-226.
- Robichet, A., & Nierat, P. Consequences of logistics sprawl: Order or chaos?-the case of a parcel service company in Paris metropolitan area. *Journal of Transport Geography*, 90, <https://doi.org/10.1016/j.jtrangeo.2020.102900>
- Rodrigue, J.-P. (2004). Freight, Gateways and Mega-Urban Regions: The Logistical Integration of the BostWash Corridor. *Tijdschrift voor Economische en Sociale Geografie*, 95(2), 147–161.
- Rodrigue, J. P. (2020). The distribution network of Amazon and the footprint of freight digitalization. *Journal of transport geography*, 88, 102825.
- Sakai, T., Kawamura, K., Hyodo, T. (2015). Locational dynamics of logistics facilities: Evidence from Tokyo. *Journal of Transport Geography*, 46, 10-19.
- Sakai, T., Kawamura, K., & Hyodo, T. (2020a). Logistics facilities for intra and inter-regional shipping: Spatial distributions, location choice factors, and externality. *Journal of Transport Geography*, 86, 102783.
- Sakai, T., Beziat, A., & Heitz, A. (2020b). Location factors for logistics facilities: Location choice modeling considering activity categories. *Journal of Transport Geography*, 85, 102710.
- Smith, H. K., Laporte, G., & Harper, P. R. (2009). Locational analysis: highlights of growth to maturity. *Journal of the Operational Research Society*, 60(sup1), S140-S148.
- Strale, M. (2019). Logistics sprawl in the Brussels metropolitan area: Toward a socio-geographic typology. *Journal of Transport Geography* (in press) <https://doi.org/10.1016/j.jtrangeo.2018.12.009>
- Train, K. E. (2009). *Discrete choice methods with simulation*. Cambridge university press.
- Union des aéroports Français (2020) Résultats d'activité des aéroports français 2019, Report , [https://www.aeroport.fr › rapport\\_2019.pdf](https://www.aeroport.fr › rapport_2019.pdf).
- Berthelot, B. (2021) Ces régions où Amazon France compte ouvrir de nouveaux entrepôts, *Capital*, 21/04/2021, <https://www.capital.fr/entreprises-marches/ces-regions-ou-amazon-france-compte-ouvrir-de-nouveaux-entrepots-1400778>
- Verhetsel, A., Kessels, R., Goos, P., Zijlstra, T., Blomme, N., Cant, J. (2015). Location of logistics companies: a stated preference study to disentangle the impact of accessibility. *Journal of Transport Geography*, 42, 110-121.
- Wagner, T. (2010). Regional traffic impacts of logistics-related land use. *Transport Policy*, 17(4), 224–229.
- Weber, A. (1909). *Theory of the location of industries*. Translation by Friedrich, C. J. (1929). University of Chicago Press.

## Appendices

### Appendix 1. Related works

Author(s)	Data, country and time period	Aim	Methodology	Main Results
Holguin-Veras et al (2005)	1017 establishments which relocated into New Jersey, US, 1990-1999	To analyze the impact of accessibility on firms' relocation decision.	Discrete choice model.	The effect of non-local accessibility (distance to NY and Philadelphia) varies depending on the industry. Local accessibility (intra-state) matters.
Bowen (2008)	Warehousing establishments (CBP, US census, 2007) , US counties ,1998,2005	To examine the changing geography of warehouses in the US between 1998 and 2005.	Correlation analysis.	The number of establishments in 2005 and the growth in were strongly correlated with distance to airports and highways and to a lesser extent railways.
Verhetsel et al. (2015)	Top 200 logistics companies, and 235 large logistics sites, Flanders, Belgium	To investigate the location factors of logistics companies.	Discrete choice model.	Cost of land was the most important location factor, followed by accessibility to a port and a motorway, location in a business park and in an inland navigation terminal.
Sakai et al. (2015)	4109 logistics establishments (Tokyo Metropolitan Freight Survey – TMFS) , 1980-2003	To analyze logistics sprawl and its negative impacts on society due to increasing truck trip distances.	Proximity measures for shipment distance.	Logistics facilities have migrated outwards, although much less than in the US and Europe.
Holl & Mariotti (2018a)	8959 logistics establishments , Spain, municipal level,2007	Impact of accessibility on the location of logistics establishments.	Poisson regression model, count data analysis.	Access to transportation infrastructure and proximity to large urban areas are important location determinants
Kang (2018)	5364 warehousing facilities, (CoStar) 1951-2016, LA, US	To analyze changes in location determinants.	Discrete choice model	Accessibility to inter-regional markets is more important for recent logistics facilities
Gingerich and Maoh (2019)	489 warehouses, Toronto MA, Canada, GPS dataset on truck movements	To analyze the relationship between warehouse location and length of truck movements.	Discrete choice model	Industrial land use, distance to infrastructure, level of urbanization and land price influence warehouse location. Warehouses locating close to the airport generate longer truck trips
Woudsma and Jakubicek, (2019)	5 Canadian cities, EPOI dataset, 2002,2012	To analyze the outward migration of logistics.	Location mapping, barycentre	The speed of outward migration varies depending on the city
Sakai et al. (2020a)	826 logistics facilities established during 2003-2013	To investigate the determinants of location choices for different types of warehouses	Discrete choice model	Zoning regulations (e.g logistics parks) and accessibility affect location, but differently depending on the type of warehouse
Sakai et al. (2020b)	1707 logistics facilities reporting inbound and outbound shipments, 2013 TMFS, Tokyo, JP	To compare the observed and optimal locations for facilities with intra- and inter-regional scope	Discrete choice model	The location of local distribution facilities almost optimal one (trip-distance-minimization). The location of facilities serving extra-regional shipments is not so easily predictable.
Robichet and Niérat (2020)	4 terminals of a parcel company (DB Schenker) within the Paris Metropolitan Area reporting 600K	To compare the observed and optimal locations of terminals	Discrete choice model (p-median)	The spatial organization of the warehouses of DB Schenker is close to the optimal one (trip distance minimization).

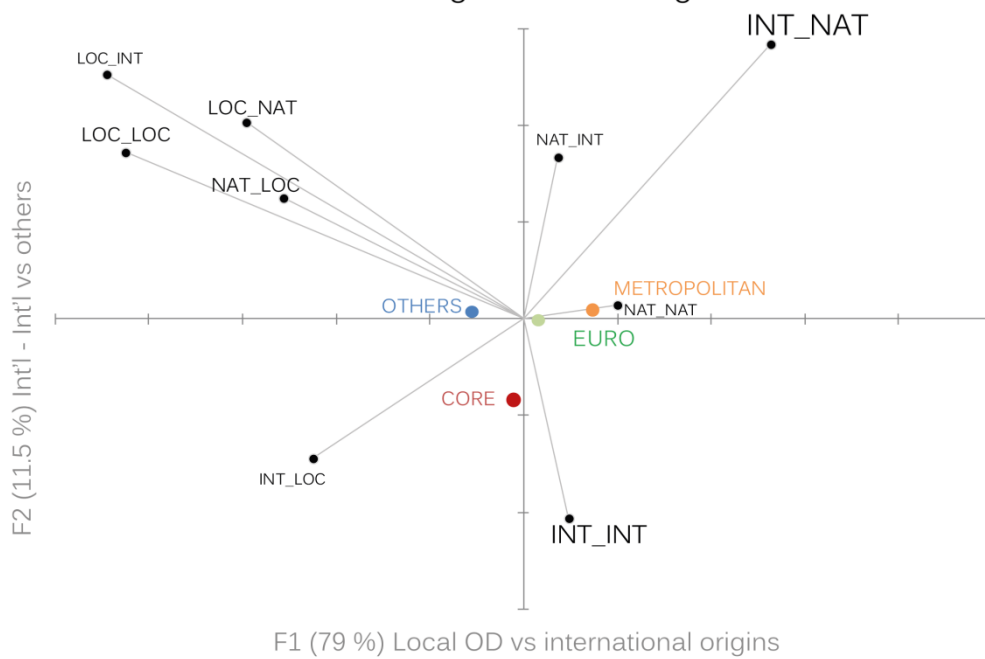
inputs including pick-ups and deliveries.

**Appendix 2. Tests comparing the observed and theoretical proportions (Table 6)**

	Chi-square	Wilks G <sup>2</sup>	Montecarlo
Observed value	132.241	143.927	132.241
Critical value	36.415	36.415	36.539
DF	24	24	3
p-value	< 0.0001	< 0.0001	< 0.0001
alpha	0.05	0.05	0.05

**Appendix 3. Factor analysis of the location categories vs. survey results on the spatial scope of OD**

Location categories vs main goods' od





Appendix 4. Results of the base model for two sub-groups (Significant at 10%\*/ 5%\*\*/1%\*\*\* level). The main differences in significativity with respect to the base model are underlined

The owner (firm) controls 1 warehouse (n=877)

	Core		Metropolitan		Sidelined	
	Est	Std Error	Est	Std Error	Est	Std Error
Intercept	-4.3185***	0.5846	-1.9095***	0.2839	<u>-0.2427</u>	0.1775
jobs	0.0094***	0.00243	<u>0.00306</u>	0.0021	-0.00053	0.00157
van_xl	1.6147***	0.451	0.7998***	0.307	0.1078	0.2839
heavy_xl	0.4032	0.5573	0.6478**	0.2901	-0.104	0.2174
temp_summer	0.1524	0.3962	-0.4547**	0.2163	0.1183	0.156
loc_ori	-0.3586	0.6408	-1.3057***	0.4484	0.655***	0.19
loc_dest	<u>0.5009</u>	0.5017	-0.1496	0.3106	0.6024***	0.1972
leased	<u>1.0011**</u>	0.4136	<u>0.3473</u>	0.2186	-0.034	0.1603
logpark	0.5679	0.4353	0.8491***	0.244	<u>-0.3078*</u>	0.1586
railspur	<u>-1.0896</u>	0.8104	-0.7621**	0.3631	<u>-0.0505</u>	0.2345
jobs*heavy_xl	-0.00921***	0.00328	<u>-0.0017</u>	0.00231	0.000286	0.00183

The owner (firm) controls several warehouses (n=583)

	Core		Metropolitan		Sidelined	
	Est	Std Error	Est	Std Error	Est	Std Error
Intercept	-3.3483***	0.8505	-1.4345***	0.3813	-0.8612***	0.3137
jobs	<u>0.00113</u>	0.00621	0.00832***	0.0026	0.00398	0.00253
van_xl	1.6278***	0.5583	0.8699***	0.3273	-0.3027	0.364
heavy_xl	1.013	0.7683	0.8248**	0.3218	<u>0.4841*</u>	0.2891
temp_summer	<u>-1.0597**</u>	0.5092	-0.8097***	0.2443	-0.0938	0.2148
loc_ori	0.4177	0.835	<u>-0.8305</u>	0.569	0.5798*	0.3156
loc_dest	1.6562***	0.5231	-0.3327	0.2886	1.3119***	0.2149
leased	<u>-0.8856*</u>	0.5163	<u>0.3104</u>	0.2583	<u>-0.3918*</u>	0.2098
logpark	0.6284	0.5933	<u>0.2781</u>	0.2876	0.1416	0.2252
railspur	<u>-13.0161</u>	273.4	-0.5032*	0.3004	-0.4779*	0.2509
jobs*heavy_xl	<u>-0.00218</u>	0.00658	-0.00794***	0.00264	<u>-0.00484*</u>	0.00266

Appendix 5. Results of the extended model (Significant at 10%\*/ 5%\*\*/1%\*\*\* level). The main differences in significativity with respect to the base model are underlined

	Core		Metropolitan		Sidelined	
	Est	Std Error	Est	Std Error	Est	Std Error
Intercept	<u>-3.978</u>	2.5806	-3.1268***	1.1772	<u>5.2087***</u>	0.9743
jobs	0.00574**	0.00288	0.00383**	0.00186	0.000271	0.00138
van_xl	1.1056**	0.4528	0.5559**	0.2655	-0.1046	0.2287
heavy_xl	<u>0.8962*</u>	0.522	0.5156**	0.2474	0.0154	0.1734
temp_summer	-0.4666	0.3862	-0.5418***	0.1901	0.0277	0.1305
loc_ori	0.4781	0.6375	-0.9295**	0.3784	0.6133***	0.1689
loc_dest	<u>0.4331</u>	0.4258	-0.6264***	0.239	0.8529***	0.1473
leased	-0.2939	0.3876	<u>0.0453</u>	0.1918	<u>-0.2914**</u>	0.1307
logpark	0.4642	0.4533	0.3551*	0.2157	<u>-0.2944**</u>	0.1343
railspur	-2.3252***	0.8062	-0.8442***	0.2691	-0.3252*	0.1762
Local_Household_inc	0.000559***	0.000119	0.000354***	0.000056	-0.00017***	0.000045
Local%EmployMNF	-75.4584***	7.4461	-29.0432***	2.1733	-8.5653***	0.9158
jobs*heavy_xl	-0.00625*	0.00328	<u>-0.0031</u>	0.00196	-0.00092	0.00151

Appendix 6. Main facilities of Amazon in France

