City Logistics Policy Toolkit:
A Study of Three Latin American Cities

Final Report

March, 2018

Matthias Winkenbach, Daniel Merchán, Milena Janjevic, Margaret Wilson, Esteban Mascarino,
Xavier Lavenir
Massachusetts Institute of Technology, Cambridge, MA, USA, {mwinkenb@mit.edu, dmerchan@mit.edu, mjanjevi@mit.edu,
mowilson@mit.edu, estmasca@mit.edu, xlavenir@mit.edu}

José A. Hincapié, Isabel Agudelo
LOGYCA / RESEARCH, Bogotá, Colombia, {jhincapie@logyca.org, iagudelo@logyca.org}

Michelle Rodríguez, Andres Regal
Universidad del Pacífico, Lima, Perú, {rodriguez.m@up.edu.pe, a.regalludowieg@alum.up.edu.pe}

Carlos Suárez-Nuñez, Carolina Chávez, Francisco Calero, Ximena Córdova
Universidad San Francisco de Quito, Quito, Ecuador, {cruarez@usfq.edu.ec, norma.chavez@estud.usfq.edu.ec,
francisco.calero@estud.usfq.edu.ec, xcordova@usfq.edu.ec}

The ultimate development objective of this project is to help cities make better urban freight policy choices
to minimize externalities while providing adequate flow of goods within the city. In this report, we present
the final results of three-stage methodology employed to: 1) identify logistics critical areas, 2) derive a
granular characterization of retail and logistics activities in those areas, and 3) identify and assess tailored
urban freight policy choices. We apply this methodology to three case study cities: Bogotá (Colombia),
Lima (Perú), and Quito (Ecuador). The report outlines the different data collection protocols and analytical
methods employed in each stage of the research, and summarizes the main findings.
CONTENTS

1 Introduction ......................................................... 5
  1.1 Motivation .................................................. 5
  1.2 Overall Project Objective ................................. 5
  1.3 Overview of Project Phase I .............................. 6
  1.4 Overview of Project Phase II ............................. 7
  1.5 Overview of Project Phase III ............................ 7
  1.6 City Logistics Policy Toolkit .............................. 9
  1.7 Structure of This Report ................................. 9

2 Phase I: City-Level Data ........................................... 10
  2.1 Data Collection ............................................. 10
    2.1.1 Demographic Data ................................... 10
    2.1.2 Economic Data ....................................... 11
    2.1.3 Road Infrastructure Data ............................ 13
  2.2 Data Processing and Cleansing ........................... 15

3 Phase I: Method for City-Level Cluster Analysis .............. 17

4 Phase I: Results .................................................. 19
  4.1 Single-City Analysis – Bogotá ............................ 19
    4.1.1 Archetype 1: Commercial-Entertainment Zones .... 22
    4.1.2 Archetype 2: Outskirts .............................. 23
    4.1.3 Archetype 3: Residential-Commercial Zones ....... 23
    4.1.4 Archetype 4: Residential-Industrial Zones ......... 24
    4.1.5 Archetype 5: Residential Zones ..................... 25
    4.1.6 Archetype 6: Very Low Density ...................... 26
  4.2 Single-City Analysis – Lima ............................... 27
    4.2.1 Archetype 1: Very Low Density Zones ............... 28
    4.2.2 Archetype 2: Residential Zones ..................... 30
    4.2.3 Archetype 3: Outskirts .............................. 31
    4.2.4 Archetype 4: Commercial-Financial ................. 31
    4.2.5 Archetype 5: Industrial-Residential Zones ......... 31
    4.2.6 Archetype 6: Downtown Residential Zones .......... 33
    4.2.7 Archetype 7: Infrastructure-Driven Zones ........ 33
10.3.2 Analysis results ............................................. 74
10.4 Conclusion and Future Work .................................. 77

11 Engagement with Local Stakeholders and Dissemination of Project Results 78
   11.1 Data collection and analysis workshops ......................... 78
   11.2 Multi-stakeholder analysis workshop ............................ 79
   11.3 Results dissemination workshops ............................... 79

12 Conclusion and Outlook ........................................... 80
1 INTRODUCTION

1.1 Motivation

On average, the global population living in urban areas is expected to keep growing by 65 million per year, according to a study of the UN Department of Social and Economic Affairs. The rapid growth in urban population is particularly pronounced in emerging economies such as Latin America. In 2012, urban population accounted for 79% of the emerging market’s inhabitants, and it has been projected to reach approximately 85% by 2030 (United Nations Population Division 2014). According to projections by McKinsey Global Institute, over the next 20 years only four of the top 25 megacities will be located in the developed world (McKinsey Global Institute 2011).

Increased urbanization directly translates into increased demand for goods and services and their supporting logistics activities. As urban population grows, these logistics operations have to rely on and compete for a limited and already heavily congested transportation infrastructure, particularly the road network and parking spaces. Increased competition for this scarce and hard to expand resource imposes additional layers of complexity and uncertainty to urban freight activities and leads to an increase in the negative externalities of urban goods transport such as congestion, pollutant and green house gas emissions, noise, and other health impairments. Consequently, urban freight is often considered a nuisance from the public perspective, ignoring its essential role in the economic and social viability of urban life. Today, cities often impose regulatory measures on urban freight to mitigate its negative externalities, while lacking a systematic approach to understand the specific freight needs and design appropriate freight policies for urban neighborhoods.

As in most cities around the world, Latin American cities have a wide range of freight regulations. The cities of São Paulo and Mexico City are well known for incentivizing Urban Delivery Vehicles (span., vehículos urbanos de carga, or VUC) and restricting the admissible vehicle weight and length during rush hours. Other cities in the region, like Bogotá, Lima and Quito, have similar strategies. These policies are often driven by physical constraints on roads, as well as traffic management concerns. The latter, however, often ignore the impact of freight restrictions on inner-city access to supply with goods and services. Most cities continue to develop a patchwork of regulations and exceptions in response to pressure from the private sector, as well as political motives. They do not necessarily take an integrated freight system management perspective, which ultimately makes the urban freight system inefficient.

1.2 Overall Project Objective

The primary development objective of this project is to help cities make better urban freight policy choices to minimize externalities while providing adequate flow of goods within the city. Although urban freight vehicles make up a small share of all vehicle traffic, they generate a disproportionate
share of externalities such as: congestion on local streets and highways, infrastructure damage, pollution, greenhouse gases, and noise. This is particularly acute in dense city areas in developing countries with limited or no space for road capacity expansion, and with land uses that have developed organically over time and are potentially incompatible with logistics demands. Cities need to develop systematic solutions to balance the need for efficient urban logistics services and a sustained quality of life in dense urban environments.

Specifically, this project will conduct 3 citywide studies in Bogotá, Quito and Lima to devise a “city logistics policy toolkit” being developed by the MIT Megacity Logistics Lab to make recommendations with respect to urban freight policies and practices based on a systematic, replicable, data-driven approach.

The toolkit allows analyzing a city as a whole, identifying different logistics needs by city areas (at the square-kilometer level) and highlighting critical city areas where the development and deployment of better urban freight policies and practices should be a top priority. Sample data will then be collected in such critical areas to evaluate the potential impact of urban freight policy measures and used to facilitate a discussion between public and private actors.

1.3 Overview of Project Phase I

The first phase of the project aims at developing a city-wide classification (or clustering) of urban areas considering their specific demographic, socio-economic and road-infrastructure properties. Based upon these properties, we seek to derive a limited number of archetypes of city zones that attract and/or generate similar types of logistics flows and requirements. A common example of an archetype zone in cities such as Bogotá, Lima or Quito is the historic center, due to its unique levels commercial activity and type of road infrastructure. These archetypes constitute the baseline to guide the definition of zone-specific logistics policy interventions according to the characteristics of logistics activities prevalent in each zone.

We introduced a data-driven methodology to identify these zone archetypes using publicly available data. A K-means cluster analysis is at the core of this methodology. In this report, we present final results for each of three cities and provided an in-depth description of each archetype. We also discuss the relevance of each archetype for urban freight policy.

Our results suggest 6, 7, and 8 archetypes of city zones for Lima, Bogotá, and Quito, respectively. This classification properly captures the main dis-/similarities in the demographic, economic, and infrastructural properties of the various zones within each city. Particular attention should be paid to archetypes with high intensity of retail activities and higher average population density due to the intensity of freight flows they attract. Given the usually strong correlation between population density and observed traffic challenges, these archetypes should be a priority for logistics policy interventions and the deployment of innovative urban logistics operational practices.
1.4 Overview of Project Phase II

The second project phase entails an in-depth characterization of retail and logistics activities in selected areas where the deployment of better urban freight policies and practices should be a top priority. The characterization also captures the road-network and parking infrastructure available for logistics operations.

Data collection methods to derive this characterization combine field observations of retail and logistics activities with surveys to commercial establishments operators. Furthermore, leveraging a variety of descriptive statistical analysis and data visualization methods, we identify zone-specific characteristics as well as common logistics patterns to inform the design of tailored, yet transferable, urban freight policy interventions.

The intensity of freight operations in each zone is driven by the density and type of retail activity in the zone. Small grocery stores or nanostores (Blanco and Fransoo 2013)), and food service establishments are the predominant retail establishment in most segments and each establishment in these categories generates 2-3 deliveries per day on average. However, the retail composition varies significantly across zones, which ultimately drives the intensity of freight operations in a given zone. Temporal patterns of logistic operations are, however, consistent across zones. In most cases, freight activities are concentrated between 9-11 am and 2-4 pm. As expected, only entertainment districts exhibit a different delivery intensity pattern. In these zones, most delivery operations take place in the afternoon. In general, time-windows for delivery operations are constrained by the hours of operation of retail establishments. These temporal patterns of logistics operations should be carefully taken into considerations to inform, for instance, time-based policy interventions.

1.5 Overview of Project Phase III

For the final project phase, we first propose five policy interventions, along with five complementary private sector operational practices to minimize externalities while providing adequate flow of goods within the city. The proposed policy interventions and operational practices have been based on previous research, literature review and subject matter expertise.

Proposed policy interventions include:

(i) Implement and enforce dedicated un-/loading bays for freight vehicles to reduce traffic disruptiveness from delivery operations.

(ii) Enact time-based access restrictions to zones/roads and incentivize off-hour deliveries to shift delivery operations to off-peak traffic hours.

(iii) Enact restrictions to heavily polluting vehicles, while providing incentives to environmentally friendly vehicle technologies to promote environmentally friendly vehicles
(iv) Create urban transshipment spaces to enable multi-modality.
(v) Deploy technologies to disseminate policies and inform logistics planning to promote information sharing and transparency.

The proposed operational practices include the following:

(i) Seek intra-company and/or inter-company consolidation initiatives to improve asset utilization.
(ii) Implement scheduled / off-hour deliveries to large demand generators to shift delivery operations to desirable time frames.
(iii) Utilize light-freight, environmentally friendly vehicles to gain leadership in sustainable logistics practices.
(iv) Deploy multi-tier networks including satellite facilities to increase proximity to customers.
(v) Deploy alternative delivery modes such as lockers and consolidated drop-off locations to consolidate demand.

We present the findings of a multi-stakeholder analysis to assess the effects and acceptability of these proposed policy interventions and logistics practices. These findings are based on a workshop and interviews conducted in collaboration with policy makers and private sector representatives in Bogotá. Policy interventions need to account for differences in value preferences and operational requirements. Also, different policies will entail different impacts among stakeholders, which makes 'blanket' policy solutions generally unfeasible. For instance, shippers (e.g., consumer packaged goods companies) prioritize practices to improve asset utilization and increase customer proximity, such as deploying multi-tier distribution networks. Logistics service providers and transport operators prioritize actions geared towards sustainable logistics leadership and demand consolidation. Both, shippers and logistics service providers express low ranking for shifting delivery operations to off-peak hours. Government officials exhibit preferences for policies that incentivize environmentally friendly vehicles and promote information transparency. Interestingly, the environmental dimension received higher rankings than the mobility dimension, which may come as a surprise given the major traffic challenges in Bogotá.

Finally, we introduce an traffic simulation tool to quantitatively evaluate the impact of selected policy choices. The simulator effectively captures complex interactions between different system elements, including, but not limited to, flows freight and non-freight vehicles, hourly traffic patterns, temporal variations in intensity of logistics activities, demand uncertainty and spatial distribution of loading and unloading operations. This tool enables a holistic assessment of policy interventions by measuring the impact on congestion, emissions and operational efficiency. The simulator, which is still in prototype phase, is piloted in one of the commercially relevant zones in Quito. Results of this pilot exemplify the potential of this tool.
1.6 City Logistics Policy Toolkit

Throughout the course of this project, we have developed a set of semi-automated tools to clean, process, analyze and visualize data. This set of tools, which we refer to as City Logistics Policy Toolkit, are available in the project site:


Specifically, the site includes:

(i) Python-based scripts to collect, process and analyze city-level data to conduct the cluster analysis for phase I,

(ii) data collection forms and Python-based scripts to clean, process and analyze zone-specific retail and logistics activity data for phase II,

(iii) Excel-based tool to replicate the multi-stakeholder policy assessment workshop for phase III, and

(iv) the Python-based traffic simulator developed to model and quantitatively assess zone-level impacts of proposed policy interventions also developed for phase III

All tools have been properly documented. We have also included samples of input data files required to run the Python-based scripts and replicate the analyses.

1.7 Structure of This Report

The remainder of this report is structured as follows. In Sections 2-5, we present data sources, analysis methods and results of the city-level cluster analysis of project phase I. In Section 6 we present the data collection method for the field study and surveys in selected critical areas. A discussion of key findings follows is presented in Section 7. Proposed policy actions and operational practices are introduced in Section 8. The multi-stakeholder analysis framework and the traffic simulators are described in Sections 9 and 10, respectively. In Section 11 we document different events hosted with local stakeholders in all three cities to gather information and disseminate project results. We conclude this report with a summary of findings and an outlook for future research in Section 12.
2 PHASE I: CITY-LEVEL DATA

In this section, we describe the different datasets used for the first project phase, along with the steps followed to collect, clean and process data. To ensure consistency in data quality and resolution across cities, we rely on publicly and globally available data sources of demographic and road infra-structure information. This data collection approach also ensures the scalability of our methodology to other cities and geographies in the future. The socio-economic information was collected from local sources, and a significant amount of effort had to be invested in finding consistent, reliable and complete local datasets. We observe that local data accessibility continues to be a major challenge for policy research in emerging market cities.

2.1 Data Collection

2.1.1 Demographic Data

LandScan™, a global population database developed by the Oak Ridge National Laboratory (cf., Bright et al. 2015) based on high-resolution satellite imagery, is our source of universally available population data. LandScan provides up-to-date ambient population counts (i.e., average over 24 hours) at a fine-grained spatial resolution close to one square-kilometer. We build our analysis on data from the 2015 LandScan database, the most recent version available. Taking into account the frequency of data updates, global measurement and reporting consistency, spatial resolution, and global data accessibility, LandScan offers significant advantages over any local population census data, which further motivated our choice of this source. To derive population location-specific population density estimates (see, e.g., Figures 1, 2 and 3), we use geo-spatial mapping techniques as described in Section 2.2 to map LandScan population counts to the corresponding square-kilometer segments in each city.

An important consideration in this study relates to the extend of the urban area to be considered. On the one hand, if we strictly limit the considered area to the administrative boundaries of each city, many relevant surrounding zones would have to be excluded. For example, this would be the case for the municipality of Soacha, which falls outside the administrative boundaries of Bogotá, but in practice is considered a vital part of the city. On the other hand, in the immediate surrounding areas of a city there also exist many areas that are not relevant for logistics policy planning due to low population density levels, and consequently low economic activity. The problem of including non-relevant areas in our analysis becomes even more acute if we consider the entire metropolitan area of each city, which is a definition that also varies significantly across cities. For instance, the city of Quito accounts for no more than 10% of the surface area of the Metropolitan District of Quito. To balance these two extreme cases and to enhance the comparability of our three cases study analyses and any potential future analyses to be added for other cities around the world, our
analysis includes all zones within the city and its surroundings that exhibit an ambient population density of at least 1,000 inhabitants per square-kilometer.

![Population density in Bogotá](image)

Figure 1  Population density in Bogotá

### 2.1.2 Economic Data

Our methodology also uses socio-economic variables to derive the archetypes of city zones based upon zone-specific logistics profiles and needs. In particular, we use the number of commercial establishments per industry category as an indicator of the type and intensity of urban logistics activities in a given city segment. This description can be further refined, if appropriate data is available, with additional variables such as the levels of employment per industry category.

To the best of our knowledge, a globally unified and high-resolution database of commercial establishments is not available. This information is generally available from local sources. Nevertheless, we observe limitations in terms of data resolution and accessibility.

In the case of Bogotá, we use neighborhood-level estimates provided by a municipal planning agency, the Secretaría Distrital de Planeación (SDP). The original source of data is the commercial registrations database of the city’s chamber of commerce (Cámara de Comercio de Bogotá (CCB)). In the case of Quito, we use neighborhood-level estimates from the 2010 economic census provided by a municipal research agency, the Instituto de la Ciudad de Quito (ICQ). The original source of this data is the national census bureau (Instituto Nacional de Estadísticas y Censos (INEC)).
In both cases, commercial establishments had been classified according to the United Nations'
International Standard Industrial Classification (ISIC) system (United Nations Statistics Division 2014).

In contrast to Bogotá and Quito, cities in which we use aggregate commercial establishments counts at the neighborhood level and break them down to the square-kilometer level, for Lima we use geo-referenced Point of interest (POI) data provided by Navi Systems EIRL (NS). POI datasets offer high-resolution location information but may present limitations in terms of data completeness: POIs include, to a large extent, retail, food, and service establishments; however information about other industry sectors, such as manufacturing, is usually limited.

We summarize all data sources, types and levels of resolution in Table 1.

<table>
<thead>
<tr>
<th>City</th>
<th>Spatial Resolution</th>
<th>Data Source Type</th>
<th>Entity</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bogotá</td>
<td>Neighborhood</td>
<td>Commercial registrations</td>
<td>SDP, CCB</td>
</tr>
<tr>
<td>Lima</td>
<td>Points of Interest</td>
<td>Geo-coded database</td>
<td>NS</td>
</tr>
<tr>
<td>Quito</td>
<td>Neighborhood</td>
<td>Economic census (2010)</td>
<td>INEC, ICQ</td>
</tr>
</tbody>
</table>

We consolidate all relevant economic activity data for each city segment in a set of primary variables measuring the number of establishments in four categories, based on the corresponding types and intensity of logistics flows: 1) retail and wholesale; 2) food service (i.e. restaurants and bars) and accommodation; 3) manufacturing; and 4) other services. The first two categories include establishments that mostly receive incoming flows of goods (i.e. deliveries). Manufacturing establishments both generate and receive logistics flows. The fourth category, other services, includes a broad array of economic establishments that generate mainly parcel and mail delivery flows (see, Table 2). If data is available, this primary set of variables can be complemented with a similar set of variables measuring the volume of employment for each of these four industry categories. Such data was only available for Quito.

### 2.1.3 Road Infrastructure Data

We use OpenStreetMaps (The OpenStreetMap Foundation 2017) as the primary source of road-network data. For each city segment, we introduce a set of dimensional and topological measures to characterize the road-network infrastructure. Dimensional variables describe the spatial properties of the road-network (cf., Table 3), whereas topological variables inform the road networks’ structure, connectedness and resilience (cf., Table 4) (Boeing 2016).

Combining dimensional and topological data allows for a broader understanding of the configuration of cities (Ratti 2004) (see, e.g., Figure 4). For topological analysis, we use the primal representation of the road network, i.e., we model streets as edges and intersections as nodes.
Table 2  List of socio-economic variables

<table>
<thead>
<tr>
<th>Socio-Economic Variables</th>
<th>Availability</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Primary Variables</strong></td>
<td></td>
</tr>
<tr>
<td>Number of Establishments - Retail and wholesale</td>
<td>Bogotá, Lima, Quito</td>
</tr>
<tr>
<td>Number of Establishments - Food service and accommodation</td>
<td>Bogotá, Lima, Quito</td>
</tr>
<tr>
<td>Number of Establishments - Manufacturing</td>
<td>Bogotá, Quito</td>
</tr>
<tr>
<td>Number of Establishments - Other services</td>
<td>Bogotá, Lima, Quito</td>
</tr>
<tr>
<td><strong>Secondary Variables</strong></td>
<td></td>
</tr>
<tr>
<td>Number of Employees - Retail and wholesale</td>
<td>Quito</td>
</tr>
<tr>
<td>Number of Employees - Food service and accommodation</td>
<td>Quito</td>
</tr>
<tr>
<td>Number of Employees - Manufacturing</td>
<td>Quito</td>
</tr>
<tr>
<td>Number of Employees - Other services</td>
<td>Quito</td>
</tr>
<tr>
<td>Number of Employees - All economic sectors</td>
<td>Quito</td>
</tr>
</tbody>
</table>

Table 3  List of dimensional variables, adapted from Boeing (2016)

<table>
<thead>
<tr>
<th>Dimensional Variables</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intersection density (1/km²)</td>
<td>Number of road intersections per km²</td>
</tr>
<tr>
<td>Road density (1/km²)</td>
<td>Number of road segments per km²</td>
</tr>
<tr>
<td>Total street length (m)</td>
<td>Total length of non-highway and non-primary roads</td>
</tr>
<tr>
<td>Total highway length (m)</td>
<td>Total length of highway roads</td>
</tr>
<tr>
<td>Total primary-road length (m)</td>
<td>Total length of primary roads</td>
</tr>
<tr>
<td>One-way fraction (%)</td>
<td>Fraction of the total street length having directionality constraint (i.e. one-way streets)</td>
</tr>
<tr>
<td>Avg. road-segment length (m)</td>
<td>Mean road-segment length, including streets, primary roads and highways. It is a proxy of block size</td>
</tr>
</tbody>
</table>

Table 4  List of topological variables, adapted from Boeing (2016)

<table>
<thead>
<tr>
<th>Topological Variables</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Node degree</td>
<td>Average of the number of streets incident to a node, over all nodes</td>
</tr>
<tr>
<td>Node connectivity</td>
<td>Expected number of nodes to remove to disconnect a non-adjacent pair of random nodes</td>
</tr>
<tr>
<td>Closeness centrality</td>
<td>Mean of the reciprocal of the sum of the distance from a node to all other nodes in the network, averaged over all nodes</td>
</tr>
<tr>
<td>Betweenness centrality</td>
<td>Fraction of all shortest paths that pass through a node, averaged for all nodes in the network</td>
</tr>
<tr>
<td>Clustering coefficient</td>
<td>Ratio of the existing and possible edges between a node and its neighbors, averaged over all nodes in the segment</td>
</tr>
<tr>
<td>Circuity factor</td>
<td>Avg. ratio of the true network distance to the euclidean distance</td>
</tr>
</tbody>
</table>

To estimate the circuity factor variable listed in Table 4, we use the methodology outlined in Merchán and Winkenbach (2018), in which precise road network distances are obtained from simulating large samples of trips using the Google Distance Matrix web service (Google 2017).
2.2 Data Processing and Cleansing

In preparation for the clustering analysis described in Section 3, all population, socio-economic and road-infrastructure datasets were processed and cleaned to obtain the values of all variables for each square-kilometer city segment. If the data exhibits a resolution that is more fine-grained than the square-kilometer level (e.g., POI data or road-network data) we aggregate all values to the corresponding square-kilometer level. For lower-resolution data, for instance available only on the neighborhood-level, we disaggregate the values proportionally to all the square-kilometer segments or fractions of segments within the corresponding neighborhood.

We use the programing language Python as the main data processing and cleaning environment. Our approach leverages multiple Python libraries, including:

- *Shapely*\(^1\) to manipulate geometric objects,
- *OSMnx*\(^2\) to extract and analyze road-network data from OpenStreetMaps,
- *Seaborn*\(^3\) and *Matplotlib*\(^4\) to visualize data, and
- *Googlemaps*\(^5\), the Python client library for Google Maps service.

After processing all datasets, we remove a small set of city-segments to avoid including non-relevant outliers for the clustering analysis. Removed segments include those with population

---

\(^1\) Shapely GitHub site: github.com/Toblerity/Shapely

\(^2\) OSMnx GitHub site: github.com/gboeing/osmnx

\(^3\) Seaborn GitHub site: github.com/mwaskom/seaborn

\(^4\) Matplotlib GitHub site: github.com/matplotlib

\(^5\) Googlemaps Client GitHub site: github.com/googlemaps/google-maps-services-python
density below 1,000 inhabitants or those with unique road-network properties (e.g. zones where multiple highways and primary roads intersect). Resulting surface areas, total population and avg. population density for each city are summarized in Table 5.

<table>
<thead>
<tr>
<th>City</th>
<th>Area (km²)</th>
<th>Total population (inhab., millions)</th>
<th>Avg. population density (inhab./km²)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bogotá</td>
<td>368</td>
<td>7,4</td>
<td>20,100</td>
</tr>
<tr>
<td>Lima</td>
<td>674</td>
<td>8,3</td>
<td>12,300</td>
</tr>
<tr>
<td>Quito</td>
<td>199</td>
<td>1,4</td>
<td>7,400</td>
</tr>
</tbody>
</table>

All Python codes we develop and employ to clean and process the raw data are available in the accompanying project website.
3 PHASE I: METHOD FOR CITY-LEVEL CLUSTER ANALYSIS

We suggest a multivariate framework that combines principal component analysis (PCA) with \(K\)-means clustering. \(K\)-means clustering is used to efficiently characterize the infrastructural and socio-economic properties of large and complex urban conglomerates in a structured and reproducible way. PCA is used for dimensionality reduction prior to the cluster, i.e., to project higher dimensional data to a lower-dimensional subspace before applying the \(K\)-means clustering algorithm to this subspace (Ding and He 2004). In this section, we provide an overview of these two methods. The stages of our proposed procedure are:

1. partial de-correlation and dimensionality reduction using PCA, and
2. clustering using a \(k\)-means algorithm.

In the first stage, partial de-correlation, principal component analysis (PCA) is used to reduce the number of (independent) variables to consider for a subsequent clustering of the dataset. In data analysis, one often encounters datasets that include a large number of variables. In such situations it is very likely that subsets of these variables are highly correlated with each other (for instance, population density is usually correlated with the number of retail and wholesale establishments in an area). When applying statistical models to such datasets, such as \(K\)-means clustering, highly correlated variables can lead to model overfitting, and the accuracy and reliability of the insights such a model can provide can suffer (see, e.g., Shmueli 2010). One of the key steps in data analysis, therefore, is reducing dimensionality (i.e., the number of independent input variables used by a model) without sacrificing accuracy. To this end, the underlying idea of PCA is to reduce the dimensionality of a data set consisting of a large number of intercorrelated variables by replacing them by a generally smaller set of mutually uncorrelated artificial variables, the so-called principal components (PCs). The PCs of a dataset are linear combinations of the original variables. They are chosen such that the incremental explanatory power of each additional PC on the dataset is maximized. Consequently, the majority of the variation present in the original data set can be explained by a number of PCs that is less than the number of original variables (Jolliffe 2002).

In the second stage, clustering, we divide the entire city into a limited number of archetypes of zones (i.e., clusters), based on the demographic, infrastructural and socio-economic dis-/similarity of city segments. Among the most prominent clustering techniques are hierarchical clustering, \(K\)-means clustering and normal mixtures clustering (cf., MacQueen 1967, Hartigan 1979, Jain and Dubes 1988). While normal mixtures is a probabilistic clustering technique assigning probabilities to every observation denoting their likelihood of being included in a given cluster, the other two techniques deterministically assign each observation to a specific cluster. Here, hierarchical clustering is a combining process starting with every observation being its own cluster then calculating the distances among clusters and combining the two clusters that are closest to each other in a stepwise
process until there is only a single cluster left. The number of desired clusters, $K$, can be chosen after this procedure by cutting the clustering tree at the corresponding position. For $K$-means clustering, the number of desired clusters, $K$, is usually specified in advance. The algorithm then specifies a corresponding number of $K$ seeds within the data that are as distant as possible from each other. Then, in an iterative process, the remaining observations in the dataset are assigned to their closest cluster. Every time an observation is added to a cluster, the centroid of this cluster is recomputed. The process of assigning observations to their closest cluster and recomputing cluster centroids is repeated until the compositions of the $K$ clusters become sufficiently stable.

While hierarchical clustering is most efficient for datasets of limited size, $K$-means clustering is more efficient for very large datasets. In this study, both approaches may be applicable, we however chose to apply $K$-means clustering. Regardless of the clustering approach chosen, the central trade-off at this stage is to determine a number of clusters $K$ that strikes the right balance between having as few and as dissimilar clusters as possible, while ensuring a high degree of homogeneity within these clusters and depicting as much of the real-world structural differences within the city as possible at the same time.

The choice of the ‘right’ number of clusters, $K$, is often ambiguous and does not follow a unique, well-defined rule (Everitt 1979, Hartigan 1985). Among commonly used techniques to determine the optimal number of clusters, $K$, is the use of the $f$ statistic proposed by Pham et al. (2005). In essence, the $f$ statistic measures the distortion (i.e., within-cluster dispersion) in the data compared to a null hypothesis of the data being distributed uniformly. Then, the suggested value of $K$ is the one that minimizes the $f$ statistic. This metric is conceptually similar to the 'gap' statistic introduced in Tibshirani et al. (2001). Nevertheless, in this work we use $f$ as it provides higher discriminatory power and is computationally substantially more efficient (Pham et al. 2005). Other methods have been proposed in the literature such as the ‘elbow’ method, the use of information criteria, silhouette clustering and cross-validation, which however are not discussed here.

In practice, $K$-means clustering is run independently for several different values of $K$, using the ‘right’ number $K$ suggested by the $f$ statistic as an indicative reference or starting point. Subsequently, the level of $K$ that leads to a partition of the dataset that appears the most meaningful to the domain expert is selected.
4 PHASE I: RESULTS

As stated previously, the objective of this analysis is to abstract from the true complexity of the case study cities under consideration, by breaking them into square-kilometer segments that are then used to derive a limited set of city zone archetypes that summarize square kilometers that exhibit similar characteristics in terms of population level, economic activity, as well as infrastructural properties. It is important to note that the zone archetypes are formed by combining square kilometers solely based on their similarity along these three dimensions, not based on their geographic location.

Since an uneven number of variables per dimension (demographic, economic, or infrastructural) might distort the results towards the dimension with the largest number of variables, we adjust the clustering method to account for balance different number of variables by using weights per dimension. In coordination with local experts, we found that slightly biasing the clustering analysis towards economic variables produced clusters with more descriptive socioeconomic and land use characteristics that could be more readily interpreted by those with local knowledge for use in future analyses.

In the following, we will provide a detailed description of the zone archetypes we derived for each of our three case study cities. For each case study, we will characterize each archetype based on its absolute score on each of the demographic, economic, and infrastructural variables we consider for our analysis. The relative scores of each archetype compared to the other archetypes derived for this city are further represented graphically in the form of so-called radar charts. In these representations, the scores of the archetypes have been normalized, with a value of zero corresponding to the lowest score exhibited by any of the archetypes and a value of one corresponding to the highest score achieved for the respective variable by any of the archetypes of that particular city. Moreover, we project our analysis results back onto a map of the respective case study city, both at the city- and at the individual archetype-level, to provide a geo-spatial representation of the results.

4.1 Single-City Analysis – Bogotá

After performing the analysis for Bogotá, the city was divided into 6 archetypes. One of them corresponds to residential areas. Another two are associated to zones with a high concentration of commercial activity. One archetype represents the mixed industrial-residential areas, and one groups the outskirts of the city. The remaining archetype, identified here as ”Very Low Density,” was produced by the clustering algorithm but did not have enough salient socioeconomic or other features to warrant specific characterization or further analysis.
Figure 5 shows the geographical distribution of those six archetypes. Moreover, in Table 6 we provide the absolute numerical score of each archetype for each of the variables included in the cluster analysis for Bogotá. Finally, Figure 6 illustrates the relative scores of all zone archetypes identified for Quito on the clustering variables. In the following, we will describe each archetype for Bogotá individually.
Table 6  Clustering variables summary per archetype for Bogotá

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Archetype 1</td>
<td>26,793</td>
<td>1,728</td>
<td>511.5</td>
<td>391.4</td>
<td>861.2</td>
<td>168.9</td>
<td>1074</td>
</tr>
<tr>
<td>Archetype 2</td>
<td>14,386</td>
<td>150</td>
<td>61.2</td>
<td>51.5</td>
<td>120.0</td>
<td>98.8</td>
<td>469</td>
</tr>
<tr>
<td>Archetype 3</td>
<td>26,055</td>
<td>264</td>
<td>111.6</td>
<td>84.1</td>
<td>234.1</td>
<td>109.3</td>
<td>3,135</td>
</tr>
<tr>
<td>Archetype 4</td>
<td>21,944</td>
<td>435</td>
<td>159.1</td>
<td>155.7</td>
<td>276.5</td>
<td>237.5</td>
<td>624</td>
</tr>
<tr>
<td>Archetype 5</td>
<td>31,843</td>
<td>381</td>
<td>154.1</td>
<td>142.6</td>
<td>284.4</td>
<td>132.8</td>
<td>722</td>
</tr>
<tr>
<td>Archetype 6</td>
<td>8,881</td>
<td>63</td>
<td>26.5</td>
<td>18.5</td>
<td>47.0</td>
<td>40.8</td>
<td>234</td>
</tr>
</tbody>
</table>

Figure 6  Summary of relative archetype scores on clustering variables for Bogotá
4.1.1 Archetype 1: Commercial-Entertainment Zones

This archetype groups the entertainment areas of the city, such as Zona T, where various restaurants, casinos, hotels, and other commercial establishments are located. It also includes Bogotá’s historic center, also known as La Candelaria, a commerce-intense zone. Moreover, these areas also accommodate a considerable number of small, owner-operated manufacturing establishments dedicated to the production of handcrafts and liquors. In fact, the level of the four economic variables we consider is relatively high for this archetype. In particular, the number of both food and general service related commercial establishments exhibits moderate to high scores compared to the other archetypes.

This archetype also has a residential component that is somewhat comparable to the one of Archetype 4 (Residential-Industrial Zones). It exhibits an average population density of 26,793 inhabitants per square kilometer. Nevertheless, the fraction of one-way streets is notably higher in this zone (almost 48%) and the intersection density is lower than Archetype 4, indicating a poorer road network connectivity in the commercial-entertainment zones. See Figure 7 for the geographic location of the areas summarized by this archetype.

Figure 7  Bogotá’s Archetype 1 (red) comprises the city’s commercial and entertainment nodes
4.1.2 Archetype 2: Outskirts

This archetype groups the peripheral areas of Bogotá. It comprises some of the least populated regions, with an average population density of 14,386 inhabitants per square kilometer. In these areas, the level of economic activity is also limited, exhibiting the lowest average value across all archetypes (except for the "very low density" archetype, as noted elsewhere). Similarly, the peripheral nature of these areas leads to low scores on all infrastructure related variables. These regions mainly build on economic activities such as small-scale production. Neighborhoods like Miraflores, Bilbao and Lisboa in the north, and Boita, Ciudad Bolivar and Cazuca correspond to this archetype. See Figure 8 for the geographical location of these zones.

![Figure 8 Bogotá's Archetype 2 (light blue) groups the outskirts of the city](image)

4.1.3 Archetype 3: Residential-Commercial Zones

This archetype includes some highly populated areas of Bogotá, with an average population density of 26,055 inhabitants per square kilometer. In this region, different types of vivid economic activity can be observed. The retail establishments vary in size, from nano-stores to shopping
malls, and are spread in neighborhoods such as Restrepo or San Andresito. Economic activity levels for manufacturing, food and beverage, and other services areas in these zones varies from low to medium. Such activities are mainly developed close to major roadways such as Avenida Norte-Quito-Sur, Avenida Primera de Mayo or Avenida El Dorado. Similarly, the infrastructural variables also exhibit high to mid-range scores for these areas, indicating a good level of accessibility and connectivity. See Figure 9 for the geographic location of the areas summarized by this archetype.

Figure 9  Bogotá’s Archetype 3 (yellow) depicts mixed-use residential and commercial zones

4.1.4 Archetype 4: Residential-Industrial Zones

This archetype depicts Bogotá’s industrial areas including neighborhoods such as Montevideo and parts of Puente Aranda. These area exhibit slightly above average number of manufacturing facilities compared to the other archetypes as well as a relatively high ambient population.

Furthermore, several middle-income residential subareas with an average population density of 21,944 inhabitants per square kilometer are also part of this archetype. However, the relative
density of retail and food establishments in the region is very low. See Figure 10 for the geographic location of the areas summarized by this archetype.

![Figure 10 Bogotá’s Archetype 4 (purple) clusters the industrially intensive regions](image)

4.1.5 Archetype 5: Residential Zones

This archetype groups mainly the residential neighborhoods of Bogotá, such as Villa del Dorado, Sagrado Corazón, and Brasilia. These areas exhibit an average population density of 31,843 inhabitants per square kilometer, the highest among all archetypes. The archetype also includes some commercial areas located within these neighborhoods that are generally placed somewhat further away from larger roadways. In these neighborhoods, we will predominantly find retail establishments such as nano-stores, service companies and small factories. Nevertheless, economic activity across all sectors ranges between low and very low levels compared to other archetypes. With respect to infrastructure, these areas exhibit moderate levels of connectivity with more than 132
intersections per square kilometer and an average road segment length of less than 58 meters. See Figure 11 for the geographic location of the areas summarized by this archetype.

![Figure 11](image.png)

**Figure 11** Bogotá’s Archetype 5 (green) shows the areas of the city which are mainly residential

### 4.1.6 Archetype 6: Very Low Density

This archetype represents a type of output identified by the clustering procedure that groups pixels with very low scores on all clustering variables. The dataset does not allow us to draw meaningful conclusions about the land uses and activity patterns of these areas. However, since these pixels had enough ambient population and other characteristics to pass through the ‘cleaning’ stage of data analysis and warrant inclusion in the clustering stage, they are presented here as output by the clustering algorithm.

Note that there is no radar chart for this archetype included in Figure 6. Radar charts depict relative values of selected descriptive variables across archetypes - that is, each vertex of the radar chart is scaled between 0 and 1, with the minimum value for the category scaled to 0 and the
maximum to 1. Intermediate values are depicted according to their relative positions in this range. Because Archetype 6 has the lowest score across every descriptive variable in the radar charts of Figure 6, it does not produce a meaningful chart (i.e. it plots only a single point) and is omitted from that figure. A similar output is observed in both Lima and Quito.

Figure 12  Bogotá’s Archetype 6 (grey) presents very low density areas without other salient features

4.2 Single-City Analysis – Lima
When applying our analysis methodology, the city of Lima was divided into 7 different archetypes. Three of them were classified as residential areas. One archetype groups downtown areas of intense commercial activity, and one comprises the outskirts of Lima. A final archetype is scattered throughout the city and comprises square kilometers that are mostly characterized by the presence of road
infrastructure in the absence of other salient socioeconomic features. The final archetype, as also seen in the analyses of Quito and Bogotá, is a cluster of very low density areas to which no further analysis will be directed here. Figure 13 provides the geographical distribution of the previously mentioned archetypes. Furthermore, in Table 7 we present, for each archetype, the absolute numerical values associated with each clustering variable considered for the case study of Lima. Finally, Figure 14 illustrates the relative scores of all zone archetypes identified for Lima on the clustering variables. In the following, we will provide a brief description of each individual archetype.

![Figure 13 Geographic distribution of Lima’s archetypes](image)

### 4.2.1 Archetype 1: Very Low Density Zones

This archetype represents a type of output identified by the clustering procedure that groups pixels with very low scores on all clustering variables. The dataset does not allow us to draw meaningful conclusions about the land uses and activity patterns of these areas. However, since
### Table 7  Clustering variables summary per archetype for Lima

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Archetype 1</td>
<td>16,025</td>
<td>22.5</td>
<td>1.09</td>
<td>37.7</td>
<td>47.5</td>
<td>72</td>
</tr>
<tr>
<td>Archetype 2</td>
<td>17,861</td>
<td>52.5</td>
<td>3.64</td>
<td>68.9</td>
<td>171.0</td>
<td>2,379</td>
</tr>
<tr>
<td>Archetype 3</td>
<td>11,977</td>
<td>26.1</td>
<td>0.90</td>
<td>52.1</td>
<td>160.2</td>
<td>174</td>
</tr>
<tr>
<td>Archetype 4</td>
<td>20,481</td>
<td>538.4</td>
<td>84.67</td>
<td>309.1</td>
<td>140.6</td>
<td>1,478</td>
</tr>
<tr>
<td>Archetype 5</td>
<td>13,001</td>
<td>227.8</td>
<td>15.07</td>
<td>233.3</td>
<td>105.4</td>
<td>516</td>
</tr>
<tr>
<td>Archetype 6</td>
<td>26,602</td>
<td>241.5</td>
<td>33.25</td>
<td>168.4</td>
<td>123.1</td>
<td>1,244</td>
</tr>
<tr>
<td>Archetype 7</td>
<td>13,132</td>
<td>91.7</td>
<td>7.94</td>
<td>108.9</td>
<td>135.0</td>
<td>758</td>
</tr>
</tbody>
</table>

**Figure 14  Summary of relative archetype scores on clustering variables for Lima**

these pixels had enough ambient population and other characteristics to pass through the ‘cleaning’
stage of data analysis and warrant inclusion in the clustering stage, they are presented here as output by the clustering algorithm.

Note that there is no radar chart for this archetype included in Figure 14. Radar charts depict relative values of selected descriptive variables across archetypes - that is, each vertex of the radar chart is scaled between 0 and 1, with the minimum value for the category scaled to 0 and the maximum to 1. Intermediate values are depicted according to their relative positions in this range. Because Archetype 61 has the lowest score across every descriptive variable in the radar charts of Figure 14, it does not produce a meaningful chart (i.e. it plots only a single point) and is omitted from that figure. A similar output is observed in both Quito and Bogotá. See Figure 15 for the geographical footprint of this archetype.

![Figure 15 Lima’s Archetype 1 (grey) presents very low-density areas of the city](image)

### 4.2.2 Archetype 2: Residential Zones

Archetype 2 summarizes a kind of residential zone with moderate population density levels (17,861 inhabitants per square kilometer on average). This archetype generally comprises areas with moderate values on all descriptive variables, indicating a moderate degree of economic activity and road connectivity while not being exceptionally dense or industrially or commercially intensive.

See Figure 16 for the geographic coverage of this archetype.
4.2.3 Archetype 3: Outskirts

This archetype gathers the peripheral zones of Lima. It represents the least populated zones of the city (except for Archetype 1), with an average population density of 11,977 inhabitants per square kilometer. Moreover, the level of economic activity is limited, so that the corresponding clustering variables exhibit the lowest scores for this archetype compared with the other archetypes. According to expert knowledge, these areas are mainly located in the eastern side of the city and their emergence and growth has predominantly shaped by recent waves of immigration. See Figure 17 for the geographical location of this archetype.

4.2.4 Archetype 4: Commercial-Financial

This archetype represents the historic center of Lima together with the districts of San Isidro, Gamarra and part of the downtown area. It presents one of the highest values of transient population density, with an average of 20,481 inhabitants per square kilometer. This area also exhibits the highest levels of economic activity, making it a neuralgic center of Lima’s commercial activity.

In terms of infrastructure, the connectivity level of this part of the city is moderate with 140.6 intersections per square kilometer and an average primary classification road length per square kilometer of 1,478 meters. Figure 18 shows the geographic coverage of this archetype.

4.2.5 Archetype 5: Industrial-Residential Zones

This archetype summarizes areas within Lima that are characterized by a mixture of residential and non-residential uses. It presents an intermediate population density level with an average of
Figure 17  Lima’s Archetype 3 (light blue) describes the peripheral areas of the city

Figure 18  Lima’s Archetype 4 (orange) depicts the city’s downtown

13,001 inhabitants per square kilometer as well as a considerable intensity of economic activity, especially manufacturing. Similarly, the infrastructural variables denote a moderate degree of connectivity and accessibility with an average primary road segment length of 516 meters as 105.4 intersections per square kilometer. According to local experts, the archetype matches medium and
high income residential zones with a variety of commercial establishments ranging from shopping malls to drug stores. Figure 19 illustrates the geographic coverage of this archetype.

Figure 19 Lima’s Archetype 5 (purple) presents mixed residential and industrial zones

4.2.6 Archetype 6: Downtown Residential Zones

This sixth archetype collects those residential areas of downtown Lima and zones such as Miraflores. These zones present a very high population density level with, on average, 26,602 inhabitants per square kilometer and moderate levels of economic activity. Regarding infrastructure, this archetype exhibits one of the highest levels of connectivity in Lima, with an intersection density of 123.1 nodes per square kilometer and the highest number of streets per node among any cluster at 3.52. A large fraction (about 88%) of streets in this zone are one-way, a potential hindrance for freight vehicles. See Figure 20 for the geographic coverage of this archetype.

4.2.7 Archetype 7: Infrastructure-Driven Zones

This archetype comprises a few pixels scattered through the city with low to moderate dimensional scores for all but the infrastructural variables. This is an artifact of the clustering process that groups pixels with a high road network density. This archetype scores low on commercial and manufacturing establishments of all types and points of interest (a similar variable derived from publicly available information). Only Archetype 2 (Residential) and Archetype 3 (Outskirts) score lower for these categories. Archetype 7, however, has very high scores for streets per node (3.17) and intersections per square kilometer (135) and moderate values for total primary road length
Figure 20  Lima’s Archetype 6 (yellow) shows one type of the city’s downtown residential zones (758 m) and fraction of one-way streets (0.33). See Figure 21 for the geographic coverage of this archetype.

Figure 21  Lima’s Archetype 7 (red) comprises areas that are mainly characterized by road infrastructure
4.3 Single-City Analysis – Quito

After performing our analysis for Quito, the city was divided in 8 different archetypes. Based on our results, we are able to nicely dissect archetypes of high economic activity based on their industry sector focus. Three of the identified archetypes group the financial and industrial centers of Quito. Another two summarize predominantly residential areas. Two additional archetypes describe the commercial and entertainment districts of the city. Finally, one archetype describes the peripheral region. Figure 22 depicts the geographical distribution of those archetypes across the entire city. In Tables 8 and 9, we provide the absolute numerical value of each archetype for each of the clustering variables considered for Quito. Finally, Figure 23 illustrates the relative scores of all zone archetypes identified for Quito on the clustering variables. In the following, we will describe each archetype for Quito individually.

![Figure 22 Geographic distribution of Quito's archetypes](image)

4.3.1 Archetype 1: Residential Type I Zones

In this archetype, the mean population density is approximately 10,517 inhabitants per square kilometer. Moreover, the level of economic activity in these zones is fairly low. Furthermore, this
Table 8  Required clustering variables summary per archetype for Quito

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Archetype 1</td>
<td>10,517</td>
<td>240</td>
<td>49.2</td>
<td>54.2</td>
<td>130.0</td>
<td>130.8</td>
<td>1,090</td>
</tr>
<tr>
<td>Archetype 2</td>
<td>6,357</td>
<td>90</td>
<td>21.1</td>
<td>23.6</td>
<td>54.9</td>
<td>75.3</td>
<td>206</td>
</tr>
<tr>
<td>Archetype 3</td>
<td>14,204</td>
<td>418</td>
<td>183.5</td>
<td>85.6</td>
<td>727.6</td>
<td>128.1</td>
<td>4,305</td>
</tr>
<tr>
<td>Archetype 4</td>
<td>4,604</td>
<td>69</td>
<td>12.2</td>
<td>15.4</td>
<td>32.6</td>
<td>71.7</td>
<td>275</td>
</tr>
<tr>
<td>Archetype 5</td>
<td>12,828</td>
<td>680</td>
<td>166.3</td>
<td>110.4</td>
<td>350.8</td>
<td>125.7</td>
<td>2,768</td>
</tr>
<tr>
<td>Archetype 6</td>
<td>13,317</td>
<td>3,079</td>
<td>336.1</td>
<td>301.9</td>
<td>751.9</td>
<td>103.0</td>
<td>173</td>
</tr>
<tr>
<td>Archetype 7</td>
<td>8,965</td>
<td>120</td>
<td>33.8</td>
<td>62.9</td>
<td>57.2</td>
<td>74.2</td>
<td>3,025</td>
</tr>
<tr>
<td>Archetype 8</td>
<td>2,951</td>
<td>21</td>
<td>3.5</td>
<td>5.0</td>
<td>10.4</td>
<td>29.5</td>
<td>284</td>
</tr>
</tbody>
</table>

Archetype exhibits the highest score on the number of street intersections and, therefore, a minimal average road length. Given the previously described characteristics, we can conclude that this archetype summarizes residential zones which are mainly located in the southern peripheral suburbs of the city. Guamaní, Turubamba, Quitumbe and La Ecuatoriana are some of the parishes comprised by this archetype. Figure 24 describes the geographic coverage of this archetype.

4.3.2 Archetype 2: Residential Type II Zones

In this second type of predominantly residential area, the population density (6,357 inhabitants per square kilometer) is lower than for Archetype 4, as is the level economic activity. An moderate number of intersections, with 75.3 nodes per square kilometer, suggest that these zones are still relatively well connected. Larger office and industrial structures, including both businesses and institutions such as Universities, can be found in these areas. This residential zone archetype is likely to be populated by medium- and high-income citizens. Parishes such as Concepción and Jipijapa belong to this archetype. Figure 25 depicts the geographical coverage of this archetype.
Figure 23  Summary of relative archetype scores on clustering variables for Quito
4.3.3 Archetype 3: Commercial-Financial District

The zones belonging to this archetype exhibit a considerable level of food service as well as other general service establishments. They are also characterized by a high infrastructure connectivity, which can be inferred from their high intersection density and their correspondingly high road density. Furthermore, employment levels in these areas are among the highest in Quito, with retailers and service companies accounting for the majority of the employed workforce.

Establishments such as banks, hotels, restaurants, and other stores can be found in this region. Besides, office buildings are also important contributors to the land use of this zone archetype. A high level of transit ridership is a direct consequence of the high employment intensity of these areas. Parishes such as Imbabura and Ibarra are included in this archetype. Figure 26 describes the geographic footprint of this archetype.

4.3.4 Archetype 4: Outskirts

This archetype exhibits the lowest scores on almost all the clustering variables considered for Quito. The population density (4,604 inhabitants per square kilometer) as well as the level of economic activity are extremely low. The areas covered by this archetype are also transversed by
Figure 25  Quito’s Archetype 2 (blue) groups a second type of mainly residential zones

relevant peripheral rings, such as Antonio José de Sucre and Simón Bolívar towards the western and eastern city boundaries, respectively. This archetype further yields the highest average road segment length across the all archetypes, which is perfectly correlated with the low level of street intersections. This archetype spans over various parishes in Quito, as depicted in Figure 27.

4.3.5 Archetype 5: Residential-Commercial Zones

The zones belonging to this archetype are characterized by its high population density attaining an average of approximately 12,828 inhabitants per square kilometer. Additionally, the level of economic activity in these areas is moderate when we consider the numbers of various types of commercial establishments as a proxy. This notion is further confirmed by the employment data available to us for Quito, which suggest that the vast majority of those establishments are small, potentially single-owner operated or family owned businesses. Finally, the archetype’s scores for the number of intersections (125.7 nodes per square kilometer) and average total road length (2,768 meters) suggest that the road infrastructure in these areas is highly developed and interconnected.

Several types of stores can be found in these areas covered by this zone archetype. Both population and retail density implies a high level of directed and through traffic in these areas throughout
the day, which may lead to traffic bottlenecks due to the areas’ mix of both one-way and two-way streets. The parishes of Villaflora, Solanda and Chimbacalle are part of this archetype. Figure 28 shows the archetype’s geographic footprint.

4.3.6 Archetype 6: Downtown and Historic Center

This archetype is characterized by the highest score of all archetypes on the number of retail, food, and other service establishments. Moreover, the employment figures for these sectors are also the highest when compared with the other archetypes. It thus represents one of the most commercially relevant regions of the city. This is a common feature of the commercial and entertainment district of Quito, also known as Mariscal Sucre. Establishments such as pubs, restaurants, stores and discos, which are mostly open by evening and night, are located in this geographically highly concentrated region.

The infrastructure of this archetype is characterized mainly by one-way streets. It exhibits one of the highest scores on intersection density (103 intersections per square kilometer). Primary-classification roads are sparse in this archetype, with only 173 meters of primary roadway per square kilometer on average. Finally, the population density in this archetype is only second-highest...
Figure 27  Quito’s Archetype 4 (light blue) groups the peripheral areas of the city

at 13,317 inhabitants per square kilometer. This implies that the amount of people that commute from other regions of Quito to this area must be one of the highest in the whole city. This reinforces the notion that this archetype is also highly relevant in terms of economic activity. Figure 29 illustrates the geographic coverage of locates this archetype.

4.3.7 Archetype 7: Industrial Zones

In this archetype, the ratio of the employment level in manufacturing (767 workers per square kilometer) over the number of corresponding establishments (57 establishments per square kilometer) is one of the highest across the entire city. Moreover, when the infrastructural variables are considered, we observe a moderately high score for the average total primary road length (3,025 meters), as well as a fairly low value for the average intersection density (74.2 intersections per square kilometer). This suggests that block sizes are large enough to accommodate big, labor-intensive commercial and industrial establishments. These values are go hand in hand a rather low number of intersections per square kilometer due to the presence of several main avenues and peripheral ring roads in the northern part of Quito. An average transient population density of approximately 8,965 inhabitants per square kilometer suggests that there are also residential areas
Figure 28  Quito’s Archetype 5 (yellow) entails the mixed residential and commercial areas in the city spread out across this archetype. It comprises parishes such as Comité del Pueblo, Cotocollao and Ponciano. See Figure 30 for the geographic footprint of this archetype.

4.3.8 Archetype 8: Very Low Density Zones

This archetype represents a type of output identified by the clustering procedure that groups pixels with very low scores on all clustering variables. The dataset does not allow us to draw meaningful conclusions about the land uses and activity patterns of these areas. However, since these pixels had enough ambient population and other characteristics to pass through the "cleaning" stage of data analysis and warrant inclusion in the clustering stage, they are presented here as output by the clustering algorithm.

Note that there is no radar chart for this archetype included in Figure 23. Radar charts depict relative values of selected descriptive variables across archetypes - that is, each vertex of the radar chart is scaled between 0 and 1, with the minimum value for the category scaled to 0 and the maximum to 1. Intermediate values are depicted according to their relative positions in this range. Because Archetype 6 has the lowest score across every descriptive variable in the radar charts of Figure 23, it does not produce a meaningful chart (i.e. it plots only a single point) and is omitted.
Figure 29  Quito’s Archetype 6 (brown) depicts the city’s historic center and downtown area

from that figure. A similar output is observed in both Lima and Bogotá. See Figure 31 for the geographical footprint of this archetype.
Figure 30  Quito’s Archetype 7 (purple) locates the city’s industrial zones

5 PHASE I: DISCUSSION

In this section, building on the numerical results and detailed descriptions of the archetypes derived for each city, we discuss how our results compare across case studies, and how differences in data resolution and completeness affect the identification of these archetypes. Further, we discuss the relevance and level of priority of the different archetypes we identified for the definition of urban logistics policies.

Our results suggest 6, 7, and 8 archetypes of city zones for Lima, Bogotá, and Quito, respectively. The selection of the number of archetypes balances statistical criteria and local subject matter expertise. This classification properly captures the main dis-/similarities in the demographic, economic, and infrastructural properties of the various zones within each city.

Overall, our results indicate higher levels of population density for Bogotá. This is observed both at the city level and at the archetype level, where the average population density for comparable archetypes is generally larger in the case of Bogotá (cf., Tables 6, 7, and 8). These results are consistent with specialized population density reports that list Bogotá as the most dense urban area in the Americas (Demographia 2015). In spite of the differences in population density levels among cities, particular attention should be paid to archetypes with higher average population density.
Figure 31  Quito’s Archetype 8 (grey) clusters the low-density areas of the city

due to the intensity of freight flows they attract. Given the usually strong correlation between population density and observed traffic challenges, these archetypes should be a priority for logistics policy interventions and the deployment of innovative urban logistics operational practices.

In all three case studies our analysis resulted in a specific archetype for the zones exhibiting high intensity of retail activities and the elevated population density levels. This archetypes include, for instance, downtown areas, entertainment zones and predominantly commercial neighborhoods. Commercial-residential archetypes also exhibit high levels retail and food-service related activities, as well as generally larger levels of population density compared to all other archetypes. Overall, these three types of zone archetypes represent the majority of last-mile delivery operations in cities and both policy efforts and operational innovation should be geared towards enabling higher mobility for freight vehicles and more efficient and less disruptive loading and unloading operations in these areas.

Large manufacturing and warehousing facilities, the origin of urban logistics flows, are usually concentrated in zones corresponding to industrial and/or outskirts archetypes. For these archetypes, policy measures should focus on accessibility and connectivity between these zones and the rest of the city. Finally, residential zones, in spite of their relatively low levels of economic activity, should
no be disregarded by logistics policy efforts and operational practice innovation: the unbroken global growth in e-commerce volumes and home-deliveries will keep increasing the intensity of last-mile logistics operations in residential areas.

We conclude this discussion by pointing out that we have overcome the data challenges discussed in the first progress report (Winkenbach et al. 2017) by adding gathering additional sources of economic census data for Bogotá and Lima. In general, we have obtained consistent datasets of density of economic activity for each city. Nevertheless, employment data was only available for Quito, yielding a more cohesive clustering. As elaborated in Section 8, one of recommended policy actions to city officials is to increase information accessibility and transparency to better inform urban logistics policies and practice.
6 PHASE II: ZONE-LEVEL DATA

In this section, we describe the data collection steps and analyses conducted for the second project phase. Recall that the objective of the second project phase is to conduct an in-depth assessment of the retail activity and freight operations in critical areas. This assessment should inform the design of policy interventions and urban logistics practices, tailored to the specific needs of each zone. Based upon the cluster analysis results reported Section 4, we select a total 14 zones across all three cities, representing the most relevant archetypes from an urban freight perspective (Table 10). The area observed in each zone covers approximately 1-square kilometer.

<table>
<thead>
<tr>
<th>City</th>
<th>Zone</th>
<th>Archetype</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bogotá</td>
<td>La Candelaria (Centro Histórico)</td>
<td>Commercial-Entertainment</td>
</tr>
<tr>
<td>Bogotá</td>
<td>Zona T</td>
<td>Commercial-Entertainment</td>
</tr>
<tr>
<td>Bogotá</td>
<td>Distrito Financiero</td>
<td>Commercial-Entertainment</td>
</tr>
<tr>
<td>Bogotá</td>
<td>Salitre</td>
<td>Residential-Commercial</td>
</tr>
<tr>
<td>Bogotá</td>
<td>Modelia</td>
<td>Residential-Commercial</td>
</tr>
<tr>
<td>Lima</td>
<td>Centro Histórico</td>
<td>Downtown-Residential</td>
</tr>
<tr>
<td>Lima</td>
<td>Miraflores</td>
<td>Downtown-Residential</td>
</tr>
<tr>
<td>Lima</td>
<td>Jesús María</td>
<td>Downtown-Residential</td>
</tr>
<tr>
<td>Lima</td>
<td>San Isidro</td>
<td>Commercial-Financial</td>
</tr>
<tr>
<td>Lima</td>
<td>Lince</td>
<td>Commercial-Financial</td>
</tr>
<tr>
<td>Lima</td>
<td>La Molina</td>
<td>Industrial-Residential</td>
</tr>
<tr>
<td>Quito</td>
<td>Centro Histórico</td>
<td>Historic Center</td>
</tr>
<tr>
<td>Quito</td>
<td>La Mariscal</td>
<td>Commercial-Entertainment</td>
</tr>
<tr>
<td>Quito</td>
<td>Distrito Financiero</td>
<td>Commercial-Financial</td>
</tr>
</tbody>
</table>

In contrast to project phase I, in which we relied mostly on publicly and globally available data sources, we use internally-developed data collection protocols for this second project phase. These protocols include

i) on-site observations, and

ii) surveys.

This data collection approach ensures consistency and comparability, and also facilitates the scalability of our methodology to other zones and cities. We use the mobile phone based application Fulcrum (Spatial Networks Inc. 2017) to create digital versions of the data collection forms and to expedite the data collection, transmission, and analysis process. A sample of the Fulcrum interface is shown in Figure 32.

Physical versions of the data collection forms are also available on the accompanying project website. Access to the project’s account on Fulcrum can be granted upon request.

In the following, we describe the main two categories of information collected:
6.1 Retail and Logistics Activity

Based upon existing literature and previous data collection experience, an on-site, two-week data collection strategy was designed to gather relevant information about retail activity and freight operations. Four sub-categories of information are identified for field collection: shop inventory, roads and regulations, delivery operations, and disruptions. For the first two categories, the information is collected at the square-kilometer level. For the two remaining categories, data are collected for a relevant street segment, generally one 100-meter block, within the zone.

**Shop inventory:** The shop inventory consists of a geo-referenced collection of all commercial activity in the area, relevant for urban freight purposes. Understanding the inventory of shops, i.e. the composition of the commercial activity of an area across various shop/retail categories, is crucial to understand the freight needs of that area. Retail categories of interest include mainly grocery stores, convenience stores and supermarkets, clothing stores, food-service establishments and pharmacies. For each establishment, we collect data such as store front length, store name, availability of loading area, and geographic coordinates.

**Roads and regulations:** This information identifies the existing road network, parking infrastructure and the corresponding regulations. Details include the specific use of street lanes, number of crosswalks, dimension of sidewalks and availability and dimensions of dedicated loading and unloading areas.

1) retail and logistics activity (collected via on-site observations), and
2) commercial establishment operations (collected via surveys).
Delivery operations: Freight and parcel delivery activities are observed within a street segment for five days. Pickup activities, although less frequent, are also captured. Using the corresponding templates, the following variables are collected: vehicle type, delivery equipment, product type, drop size, vehicle-to-store distance and vehicle-to-store number of trips, number of shops served per stop and duration of delivery.

Disruptions: Information on vehicle and pedestrian disruptions is collected, with particular interest in those caused by freight delivery vehicles. The source of the disruption, its duration, the impact on blocked lanes and number of vehicles affected is also captured.

6.2 Commercial Establishment Operations
To complement the observational data collection of retail and logistics activity, we introduce a survey to further explore the internal operations of retail establishments. In particular, we are interested in better understanding retailer’s replenishment decisions and delivery preferences, as well as suppliers’ preferred delivery modes and time-windows.

We use a stratified sampling approach to ensure statistically significant samples for all retail categories of interest. Between 150 and 200 surveys are being collected per selected zone. The survey takes approximately 15 minutes to complete and includes questions related to:

- General establishment information: including store area, storage area, number of employees, total number of suppliers, availability of parking space.
- Operations: Opening and closing times, peak-demand days and times, preferred delivery time-window, preferred delivery mode.
- Supplier information: product type, order placing method, delivery frequency, delivery day and time-windows, vehicle type.

Due to availability of local resources, we collect survey data only for a subset of zones. To define this subset, we balance two criteria: 1) we ensure having at least one zone per archetype, and 2) we prioritize the more commercially relevant zones within each city. Our final subset of critical zones includes: La Candelaria, Zona T and Distrito Financiero in Bogotá; Miraflores and Centro Histórico in Lima; and all three zones in Quito.
7 PHASE II: ZONE-LEVEL ANALYSIS AND RESULTS

In this section, we analyze retail and logistics activity and commercial establishments operations data in selected urban areas where the deployment of better urban freight policies and practices should be a top priority. While the data collection processes included a large number of variables of interest for urban freight studies in general, our analysis here focuses on comparing and contrasting the most relevant elements of the urban freight system across zones to inform policy choices discussed in Section 8. To complement the discussion presented in this section, in the Appendix to this report (see separate document) we include, for each selected critical zone, a summary of key metrics and of retail density, retail activity, parking and road-network infrastructure and delivery operations.

The density and type of retail activity inform the intensity of freight operations, i.e., the total number of delivery operations per time-period in a given zone. Consider, for instance, the case of Zona T and Distrito Financiero in Bogotá, with levels of retail density of 1,420 and 994 establishments per square kilometer, respectively. As expected, the higher retail density in Zona T will usually translate into a higher number of loading and unloading operations within the zone. Nevertheless, the type of retail also influences the intensity of logistics flows. We observe that in Zona T nearly 57% of the retail activity corresponds to either food service (38%) or clothing stores (19%). In contrast, in Distrito Financiero, the percentage of clothing stores is significantly smaller (2%). Indeed, approximately 55% of the establishments correspond to food-service (49%) and small grocery stores (6%), which generally require more frequent replenishments compared to clothing stores, as suggested by analyses of survey results. Thus, it is important to consider both density and type of retail activity (i.e., shop inventory) to accurately quantify the intensity of delivery operations.

The intensity of freight operations is a critical property of the urban freight systems and should be accurately estimated to yield realistic insights. In addition to the density and retail activity data described above, to quantify the intensity of deliveries we also need to estimate the number delivery operations per establishment. Table 11 collects results of an analysis of delivery observations recorded in 11 square-kilometer districts in 8 Latin American cities between 2015 and 2017 as part of the Megacity Logistics Lab’s KM2 project (Merchán et al. 2015). Data collected in the following districts was used to obtain this result: La Mariscal, Quito; Salitre, Bogotá; Lince, Lima; Zócalo, Mexico City; Miraflores, Lima; Pinheiros, São Paulo; Centro Histórico, Quito; Pocitos, Montevideo; Zona T, Bogotá; La Molina, Lima Modelia, Bogotá. Notice that for this particular analysis we have included data from cities not within the scope of this project. However, given similarities in retail dynamics among these zones and the unavailability of delivery observation records for some of the zones within our scope, we added those additional similar areas to increase
the number of data points available. The standard deviations reported here are also derived from empirical data and represent variation across observed districts, just as the reported mean is the average observed per shop type among all districts. In using this data to model outcomes of various policy measures, i.e., the traffic simulator we introduced in Section 10, we assume this demand to be normally distributed and generate random samples to simulate variability in demand for commercial deliveries.

<table>
<thead>
<tr>
<th>Shop Type</th>
<th>Avg. No. Deliveries Per Day</th>
<th>Std. Dev. Deliveries Per Day</th>
</tr>
</thead>
<tbody>
<tr>
<td>E, Accommodation</td>
<td>2.64</td>
<td>4.45</td>
</tr>
<tr>
<td>F, Food Service and Drinking Places</td>
<td>1.78</td>
<td>2.03</td>
</tr>
<tr>
<td>A, Grocery Stores</td>
<td>2.73</td>
<td>1.69</td>
</tr>
<tr>
<td>O, Other</td>
<td>1.08</td>
<td>0.63</td>
</tr>
<tr>
<td>B, Convenience Stores and Supermarkets</td>
<td>10.88</td>
<td>18.16</td>
</tr>
<tr>
<td>D, Clothing Stores</td>
<td>1.20</td>
<td>1.03</td>
</tr>
<tr>
<td>G, Drugstores</td>
<td>3.31</td>
<td>2.62</td>
</tr>
<tr>
<td>KF, Type F Kiosk</td>
<td>1.00</td>
<td>4.45</td>
</tr>
<tr>
<td>KO, Type O Kiosk</td>
<td>1.00</td>
<td>4.45</td>
</tr>
<tr>
<td>KD, Type D Kiosk</td>
<td>1.00</td>
<td>4.45</td>
</tr>
<tr>
<td>S, School</td>
<td>2.85</td>
<td>4.45</td>
</tr>
</tbody>
</table>

As traffic congestion is usually a concern in densely populated and commerce-intense areas, an analysis of the temporal patterns of delivery operations is necessary, as it could inform, for instance, time-based urban freight policies (i.e., access restrictions). In the majority of cases, delivery operations are particularly concentrated on the time periods between 9 am and 11 am, and 2 pm and 4 pm. Zona T in Bogotá is an exception, as most deliveries take place in the afternoon. This is however expected given that this in an entertainment zone in which food-service establishments operate in the afternoon and evening. Even though our data collection only covered regular business hours, we argue that our results are valid as experience indicates that off-hour deliveries are uncommon in these three cities. We refer the reader to the Appendix (see separate document) for further details.

We explore the heterogeneity of delivery vehicles and availability of parking spaces. A rigid truck is the predominant delivery vehicle across all zones. Still, in zones such as Lince, 90% of the deliveries are executed with two vehicle types, whereas in Centro Financiero or La Mariscal, a wider range of vehicles options is observed. This results are useful, for instance, to better calibrate fleet-related parameters within models to evaluate policy choices, specifically policy interventions related to access restriction based on vehicle size and deployment of parking infrastructure.
Figure 33  Summary of shop weekday operating hours and observations of deliveries in multiple cities, across all shop types

An additional part of our analysis has focused on gaining insights from retailer surveys conducted at hundreds of commercial establishments across the study area. We present the charts in Figure 33 comparing reported store operating times and observed delivery events in six districts across all three cities in the project scope. This result is then used to inform our selection of a time duration over which to simulate ambient and commercial traffic in Phase III, giving us the flexibility to
simulate delivery traffic for a period of typical intensity, peak intensity, or for a period that captures the largest fraction of total daily delivery traffic depending on the condition we wish to examine.
8 PHASE III: PROPOSED POLICY ACTIONS

In this section, we introduce a set of public sector recommended policy interventions and private sector operational practices that form the basis for our analyses in project phase III. These set of policies and practices have been based on previous research (e.g., Winkenbach et al. 2016a,b, Blanco 2015, Merchán and Blanco 2015, Merchán et al. 2015, Blanco and Fransoo 2013, Merchán and Blanco 2016); literature review (Dablanc 2009, 2011, Holguín-Veras et al. 2014, Jaller et al. 2016), subject matter expertise and results from this study.

Proposed policies actions and industry practices

We propose the following sets of recommendations for policy-makers (see Table 12) and private sector decision-makers (see Table 13). These mutually reinforcing, complementary recommendations have been devised based on local infrastructure conditions, investment requirements and local industry dynamics.

Table 12  Recommended policy interventions

<table>
<thead>
<tr>
<th>Code</th>
<th>Objective</th>
<th>Specific Action</th>
</tr>
</thead>
<tbody>
<tr>
<td>G1</td>
<td>Reduce traffic disruptiveness from delivery operations</td>
<td>Implement and enforce dedicated un-/loading bays for freight vehicles</td>
</tr>
<tr>
<td>G2</td>
<td>Shift delivery operations to off-peak hours</td>
<td>Enact time-based access restrictions to zones/roads and incentivize off-hour deliveries</td>
</tr>
<tr>
<td>G3</td>
<td>Promote environmentally friendly vehicles</td>
<td>Enact restrictions to heavily polluting vehicles, while providing incentives to environmentally friendly vehicle technologies</td>
</tr>
<tr>
<td>G4</td>
<td>Enable multi-modality</td>
<td>Create urban transshipment spaces</td>
</tr>
<tr>
<td>G5</td>
<td>Promote information sharing and transparency</td>
<td>Deploy technologies to disseminate policies and inform logistics planning, such as web-based logistics information portal or traffic management systems</td>
</tr>
</tbody>
</table>

Building on this set of recommended policy interventions and private sector operational practices, in Section 9 we introduce a multi-stakeholder analysis for assessing expected impacts and overall acceptability of these recommendations. Furthermore, in Section 10 we develop a traffic micro-simulator to quantitatively assess the impact on operational efficiency, network congestion and on environmental performance of selected policy interventions.
Table 13  Recommended industry practices

<table>
<thead>
<tr>
<th>Code</th>
<th>Objective</th>
<th>Specific Action</th>
</tr>
</thead>
<tbody>
<tr>
<td>P1</td>
<td>Improve asset utilization</td>
<td>Seek intra-company and/or inter-company consolidation initiatives</td>
</tr>
<tr>
<td>P2</td>
<td>Shift delivery operations to desirable time frames</td>
<td>Scheduled / off-hour deliveries to large demand generators</td>
</tr>
<tr>
<td>P3</td>
<td>Gain leadership in sustainable logistics practices</td>
<td>Utilize light-freight, environmentally friendly vehicles</td>
</tr>
<tr>
<td>P4</td>
<td>Increase proximity to customers</td>
<td>Deploy multi-tier networks including satellite facilities</td>
</tr>
<tr>
<td>P5</td>
<td>Consolidate demand</td>
<td>Deploy alternative delivery modes such as lockers and consolidated drop-off locations</td>
</tr>
</tbody>
</table>

9 PHASE III: MULTI-STAKEHOLDER ANALYSIS OF PROPOSED ACTIONS

9.1 Introduction

Inclusion of all relevant stakeholder groups is a major enabler of success of urban logistics policies (Holguín-Veras et al. 2014, Stathopoulos et al. 2012, Lindholm and Browne 2013). It allows to share knowledge and better predict impacts of different policies, develop policies that take into account trade-offs between different stakeholders’ interests and improve stakeholder commitment through a participatory planning process.

With this in mind, we develop a holistic decision-making framework for assessing the effects and the acceptability of different urban freight policies in a multi-stakeholder environment. The decision-making framework is applied to the specific city area within Bogota, the Candelaria. This zone was chosen given its high intensity of freight-related activities (see Section 7). It is characterized as belonging to the Archetype 1 (i.e., commercial-entertainment zone) (see Section 4).

9.2 Research Questions

Urban freight logistics system is composed out of several stakeholder groups with potentially different behaviors and interests along the dimensions of the triple bottom line: economic, social and environmental (Balm et al. 2014, ?). It is generally acknowledged that public and private stakeholders tend to have different perspectives on urban freight issues. Traditionally, public stakeholders wish to minimize traffic congestion, noise, air pollution and safety issues while maintaining the economic development of the city and increasing employment opportunities (Thompson and Taniguchi 2008). Private stakeholders value delivery efficiency, reliability, security and cost effectiveness. Nevertheless, although different stakeholders put different emphasis on the different dimensions of the triple bottom line, this perspective offers a shared vocabulary and a common framework for the evaluation of the urban freight logistics policies. This leads us to the first research question:
**RQ1** How do different stakeholders frame and rank the attributes of the decision-making process along the dimensions of the triple bottom line, i.e., what are the end values (ultimate aims of policies)?

Urban freight logistics policies discussed in the previous section produce effects along three dimensions: economic (e.g., efficiency and costs of operations), environmental (e.g., CO2 emissions) and social (e.g., congestion). They often involve trade-offs between the three dimensions of the triple bottom line. This leads us to the second research question:

**RQ2** How are the means to reach the end values (e.g., policy recommendations and actions) ranked according to different value preferences?

The multi-stakeholder context in which the urban freight policy making takes place introduces additional difficulties when predicting their combined effects. The first complexity arises from the insufficient knowledge about other stakeholder values and preferences, which increases the difficulty of predicting their response to given policies. The second complexity arises from interactions between stakeholder groups that may impact their individual behavioral response. Indeed, the response of one stakeholder group is also influenced by the interactions with other agents. This leads to the third research question:

**RQ3** How do the stakeholders’ preferences change, when they learn about the other stakeholders’ preferences?

### 9.3 Research Model

In order to answer the research questions, we have use a multi-method approach combining literature review, website content analysis, a workshop with an Analytical Hierarchy Process (AHP) procedure and in-depth interviews after the workshop. Figure 34 shows the methodology that was followed for the study.

In order to answer the first research question, we have performed a literature review on the values along the triple bottom line dimensions and a content analysis of websites of a selected sample of key decision makers in Colombia. These values were then ranked during a workshop\(^6\) that included an AHP procedure. In order to answer the second research question, we have ranked the proposed policies (see Section 8) during the workshop that included an AHP procedure. In order to answer the third research question, we have performed in-depth interviews organized after the workshop. We will now discuss each research method and present the results.

---

\(^6\) The workshop was held on November 30, 2017 in Bogotá hosted by LOGyCA.
9.3.1 Content Analysis

Content analysis of websites was used to validated and fine-tune the values along the triple bottom line dimensions as found in the literature. Corporate websites are a key tool for communication, notably for private organizations communicating about their corporate social responsibility (Capriotti and Moreno 2007, Tang et al. 2015). Communicating about a certain value suggests the value is relatively important for the respective stakeholder (Knoppen et al. 2006). At the same time, websites have to meet space restrictions - in contrast to corporate reports - and consequently, are expected to devote more space to values that are relatively more important.

Entity sample. For public entities we aimed to include entities at both municipality and neighborhood levels, just as specific functions (e.g., the transport department of Bogota). For private companies, we aimed to cover different shippers (including food and beverage industries) and logistics service providers, as those would be invited to the workshop, complemented with other sectors to have an overall representation. We aimed to include bigger companies or industry leaders because it is reasonable to assume (a) they will drive the industry standards; (b) they will have the highest impact; (c) they will be involved in the workshop. Companies should be involved in last mile logistics, which may include Business to business (B2B) and Business to consumer (B2C). As a result, 53 organizations’ web sites were content coded (43 private and 10 public).

Codebook. The framework for coding was based on the preceding literature review. It provides clues to coders as to what may describe a certain value. As such, it can include facets that are not impacted by urban freight logistics but that are relevant to capture the relative preferences of
a decision maker. Consequently, the codebook is wider than subsequent operationalization of the triple bottom line perspective for AHP.

In order to increase the validity of our procedure, each dimension of the triple bottom line was defined with a same width (i.e., six values), avoiding overweighting of a particular dimension. The coders do not judge the magnitude of implementation of a certain value (i.e., the degree of actual action) but rather the degree of presence (i.e., the degree of communication). In order to report the degree of presence, we chose to work with a scale from 1 to 4 (1=absent; 2= low presence; 3= intermediate presence; 4= high presence). The selected response scale allows a good tradeoff between the necessity of (a) recording evidence in a way that helps to answer the research question (i.e., the more categories, the better), and (b) facilitate reaching inter-coder agreement (i.e., the less categories, the better).

We developed specific guidelines for coding prescribing the exact time spend in the different coding steps of a website (10 minutes for an initial overall look without coding and after that 20 minutes for coding, collecting pieces of evidence and additional comments before assigning the score 1-4). Coders received instructions on how to scroll through the website: without accessing downloadable reports (given key data should be on the primary pages of a website), and by limiting the coding of financial tables by the companies’ (given that financial reporting does not reflect the importance of the economic bottom line for the company). A pilot test with five coders coding three same website was performed. As a result, additional guidelines on the use of the scale were included (e.g., if a value is part of the mission statement, the score is 4; and, the more clicks to arrive at a piece of evidence, the lower the score). Each website was initially coded by two persons individually. After that, results were compared and in case of contrasting findings, a consensus code was selected. The coders rotated to create multiple coding teams and thus align overall the way of coding.

Findings. First, the use of corporate websites for communicating on triple bottom line values is limited in Colombia. Consequently, we have evaluated the overall score per dimension for the whole sample (n=43 private companies +10 public entities) but also for a filtered sample, where we excluded companies whose scores were not bigger than two on the scale from one to four (n=13 private companies and 3 public entities). Both total and filtered samples point to the same results (see figure 35): (a) government entities communicate more on the social dimension, compared to the environmental and economic dimensions; and (b) private companies communicate more on social and environmental dimensions, compared to the economic dimension. A more in-depth analysis of results reveals the key variables that capture the triple-bottom line for private and public stakeholders. For private companies, these are: (1) revenues and cost of non-compliance for the economic dimension, (2) health and safety and access to education for the social dimensions and (3)
reduced resource consumptions and recycling for the environmental dimension. For public entities, these are: (1) investment in non-human assets and service for the economic dimension, (2) human rights and access to education for the social dimension and (3) reduced resource consumption and recycling for the environmental dimension. When contrasting the results of the coding analysis with values that are potentially impacted by urban freight logistics, we see that they overlap for the economic dimension, but not for the social and the environmental one (i.e., mobility/congestion is a social item for AHP whereas emissions are the environmental item). This may suggest that urban freight is yet not on a top priority of the public entities, confirming the necessity of action research that actively involves stakeholders and shapes their agendas.

![Figure 35: The degree of communication on different values by different stakeholders](image)

**9.3.2 Analytical Hierarchy Process (AHP) Workshop**

*Choice of the method.* AHP focusses on the choice stage of the decision-making process, i.e., it aims to improve deliberation and judgment (Dyer and Forman 1992). It was chosen as the most appropriate method for our study for several reasons. First, it does not require numerical guesses about the outcome of a certain action which is convenient in our setting where information about numerical outcomes is lacking. Secondly, AHP gains in quality of psychometric properties with increased decision complexity. More precisely, when compared to holistic procedures, AHP shows (a) increased reliability (i.e., temporal stability or the minimization of random error), when reliability is measured through the Root mean square (RMS) distance measure (Saaty 1977); and (b) increased criterion validity (i.e., the ability to predict an external phenomenon (Morera and Budescu 1998). Moreover, AHP in contrast to other multi-attribute decision approaches, allows to calculate the obtained consistency (Altuzarra et al. 2007).
### Table 14  List of dimensions and values

<table>
<thead>
<tr>
<th>Dimensions</th>
<th>Values</th>
<th>Clarification</th>
</tr>
</thead>
<tbody>
<tr>
<td>Economic</td>
<td>Increase Profitability</td>
<td>From the companies’ point of view, either by increasing revenue and/or cutting costs.</td>
</tr>
<tr>
<td></td>
<td>Increase Service</td>
<td>Customers experience regarding the delivery process (e.g., on time delivery, delivery frequency, etc.).</td>
</tr>
<tr>
<td>Social</td>
<td>Increase Mobility</td>
<td>Related to fast and efficient citizens’ transportation.</td>
</tr>
<tr>
<td></td>
<td>Increase Public Safety and Occupational Health</td>
<td>Applies to both citizens in general and companies’ employees, excluding the effects caused by particle emissions (e.g., traffic accidents, ergonomic requirements, etc.).</td>
</tr>
<tr>
<td>Environmental</td>
<td>Reduce Greenhouse Gas Emissions</td>
<td>Reduction of CO2 or CO2 equivalent gases that cause Global Warming Effect.</td>
</tr>
<tr>
<td></td>
<td>Reduce Particle Emissions</td>
<td>Reduction of particle emission that affect public health directly or indirectly.</td>
</tr>
</tbody>
</table>

*Adjusting AHP to the workshop setting.* In AHP, pairwise comparisons are made on a 1-9 scale where the nine levels can be labeled with numbers (numerical mode) or with reference phrases (verbal mode). Consequently, AHP does not require units in the comparison. The verbal mode is preferred by many people, but shows lower consistency given that AHP assumes equal distances between two adjacent levels whereas respondents may perceive non-equal distances between two successive verbal levels (Huizingh and Vrolijk 1997). Thus, the verbal mode should only be used when great attention is paid to achieve equivalent interpretations of the different reference phrases. Consequently, in our AHP process, we have used numerical mode. Given that Saaty only presents verbal descriptors for alternative actions and not to rank the values, we have adapted the scale to reflect the degree of importance rather than the degree of impact.

*Identification of the stakeholder values.* When defining values for each dimension of the triple bottom line, we have strived for equal number of values per dimension. Indeed, dimensions with larger number of underlying values tend to receive more weight than when they are less detailed (Ishizaka and Labib 2011). The list of 18 items that was defined in the previous step (coding analysis) was narrowed down to 6 (see Table 1). Finally, this list was reduced to three main values that are impacted by logistics (figure 36).

*Workshop setting.* The aim of the workshop was to attract the interest and cooperation of local participants, encourage careful information processing, facilitate examination of tough trade-offs across multiple value dimensions, and stimulating guided discussion and learning among the
different decision makers (see Gregory and Wellman (2001)) \(^7\). The workshop took a full day (8 hours) and had 19 participants distributed around 3 participant groups representing the three stakeholder groups of urban logistics: shippers, Logistics Service Providers (LSPs) and government or public authorities. Given our willingness to include actors that are active in urban logistics, the shipper group included companies that are in charge of their distribution activities (in opposition to companies outsourcing these activities to a third party). Shippers group was mostly composed of consumer package goods companies and in particular those from the beverage sector. Table 15 provides the composition of different groups.

Our research is action research, which is defined by being (a) research in action rather than research about action; (b) participative; (c) concurrent with action; (d) a sequence of events and an approach to problem solving. This method is deemed very appropriate for operations management research (Coughlan and Coghlan 2002). The principal threat of action research is lack of impartiality of the researcher. Therefore, we involved multiple researchers for key roles in the project (two interviewers per interview, two moderators per table, three researchers to interpret the emerging findings). We also aligned methods by developing and pilot testing guides for moderators and observers. In that regard, we consciously and deliberately enacted the action research cycles, reporting assumptions, inferences, and viewpoints. The sharing and updating of assumptions across all involved researchers, workshop participants and interviewees was thus an important continuous task during the project. Assumptions are intimately tied to the specific empirical context (the Candelaria cluster within Bogota). They are especially important when moving from the general aim of a logistics policy to the specific action of a policy. When doing so, it proved to be important to emphasize the assumption “cultural changes can be made”, in order to not disregard a policy because of foreseen implementation challenges.

\(^7\) The Excel-based tool developed for this workshop is available in the accompanying project website.
Observations made during the workshop. Several observations were made during the workshop. For the public-sector group, most debate was around the values mobility and profitability that finally ranked as second and third. Discussion was evidence-based with entities gaining in leadership when contributing with facts and figures. For the LSPs, most debate happened around the values of profitability and mobility that ended up far apart in priority (first and third in ranking). For the shippers table, most debate happened around the values profitability and emissions that ended up ranked as first and second. Some correlations between values were established - for example, electric vehicles reduce both emissions and costs of fuel therefore impacting both the values of emissions and profitability.

Table 15  Composition of the participant groups during the workshop

<table>
<thead>
<tr>
<th>Stakeholder group</th>
<th>Participants</th>
<th>Researchers</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Governent</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Secretaría Distrital de Movilidad, Logistics coordinator</td>
<td></td>
</tr>
<tr>
<td></td>
<td>ANDI, Logistics and technical coordinator</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Departamento Nacional de Planeación, Consultant</td>
<td>2 moderators</td>
</tr>
<tr>
<td></td>
<td>Universidad Católica, Researcher</td>
<td>1 observer</td>
</tr>
<tr>
<td></td>
<td>Universidad de La Sabana, Researcher</td>
<td></td>
</tr>
<tr>
<td></td>
<td>ANDI, Technical and commercial coordinator</td>
<td></td>
</tr>
<tr>
<td><strong>Shippers</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Solistica, Logistics Manager</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Postobon, Chief of Distribution</td>
<td>2 moderators</td>
</tr>
<tr>
<td></td>
<td>Postobon, Profesional, VP. of Logistics</td>
<td>1 observer</td>
</tr>
<tr>
<td></td>
<td>Postobon, Warehouse and distribution coordinator</td>
<td></td>
</tr>
<tr>
<td><strong>Transport sector</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Exxe Logistica, Logistics coordinator</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Transportempo, Operations manager</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Transportempo, Manager</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Argos, Logistics manager</td>
<td>2 moderators</td>
</tr>
<tr>
<td></td>
<td>Transportempo – Rentingcolombia, Coordinator</td>
<td>1 observer</td>
</tr>
<tr>
<td></td>
<td>Casa Luker, Logistics manager</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Universidad Santo Tomas, Researcher</td>
<td></td>
</tr>
</tbody>
</table>
Table 16 Value ranking matrix for the three stakeholder groups

<table>
<thead>
<tr>
<th></th>
<th>Profitability</th>
<th>Emissions</th>
<th>Mobility</th>
</tr>
</thead>
<tbody>
<tr>
<td>Shippers</td>
<td>41.3%</td>
<td>32.7%</td>
<td>26.0%</td>
</tr>
<tr>
<td>LSPs</td>
<td>69.1%</td>
<td>21.8%</td>
<td>9.1%</td>
</tr>
<tr>
<td>Public sector</td>
<td>13.6%</td>
<td>62.5%</td>
<td>23.8%</td>
</tr>
</tbody>
</table>

Table 17 Policy ranking matrix for shippers

<table>
<thead>
<tr>
<th>Policy</th>
<th>Profitability</th>
<th>Emissions</th>
<th>Mobility</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>P1 Improve asset utilization</td>
<td>33.5%</td>
<td>18.4%</td>
<td>38.2%</td>
<td>29.8%</td>
</tr>
<tr>
<td>P4 Increase proximity to customers</td>
<td>41.5%</td>
<td>27.4%</td>
<td>12.6%</td>
<td>29.4%</td>
</tr>
<tr>
<td>P3 Gain leadership in sustainable logistics practices</td>
<td>7.0%</td>
<td>38.4%</td>
<td>6.3%</td>
<td>17.1%</td>
</tr>
<tr>
<td>P2 Shift delivery operations to desirable time frames</td>
<td>12.1%</td>
<td>9.5%</td>
<td>34.6%</td>
<td>17.1%</td>
</tr>
<tr>
<td>P5 Consolidate demand</td>
<td>6.0%</td>
<td>6.3%</td>
<td>8.2%</td>
<td>6.7%</td>
</tr>
</tbody>
</table>

Results of the workshop. Table 16 shows the value ranking matrix for the three stakeholder groups. It reveals a clear difference between the public sector that values more the environmental and social dimension and the private sector who puts the highest emphasis on the economic dimension. LSPs put a higher emphasis on the profitability than shippers. One possible reason is that this given the highly competitive nature of the logistics market, companies operating in it face higher cost pressure than shippers and therefore put a higher emphasis on the profitability. We can also see that the environmental dimension has consistently received higher rankings than the social dimension, in all three stakeholder categories, which may come as a surprise given the mobility problems in Bogota.

Tables 17 and 18 show the policy ranking matrices for the shippers and the LSPs respectively. The analysis of results reveals several important differences in a way that policies are assessed according to the stakeholder group.

The transport sector has expressed strong preferences for policies P3 and P5 (i.e., leadership in sustainable logistics practices through the use of light-freight, environmentally friendly vehicles and demand consolidation in delivery points). Shippers (mainly composed out of beverage companies) have performed a low ranking for these two policies. This result demonstrated the
importance of the market segment and the commodity type on the behavioral response of urban freight logistics stakeholders. Indeed, different commodities have different operational requirements - the beverages sector is characterized by high-volume and low-value products, rendering the use of light environmentally-friendly vehicles or local delivery points inappropriate from both a practical and an economic standpoint. On the contrary, the LSPs face more defragmented delivery operations and may gain significant efficiency gains from the local demand consolidation as shown of their profitability assessment of that specific policy. Furthermore, a commodity type also impacts the nature of the business relations between different stakeholders (Holguín-Veras, 2010). Consequently, the LSPs may place higher emphasis on implementing sustainable logistics practices as a part of their value proposition to their clients as demonstrated in their assessment of this policy on the profitability.

Shippers have provided the highest ranking to the policies P1 and P4 (i.e., improved asset utilization through consolidation and deployment of multi-tier networks including satellite facilities). This result is largely driven by the strong expected impact of these policies on the profitability value that has received the highest ranking from shippers. The commodity type can again be seen as a factor driving the importance of these policies. Indeed, high asset utilization is required to reach profitable operations in a high-volume and low-value product setting. Additionally, the use of multi-tier networks is only relevant in operational settings with a sufficient overall level of demand, as in the beverages sector.

Finally, both the shippers and the LSPs have expressed a rather low ranking for shifting the delivery operations to desirable time frames. Indeed, although this policy produces a very high impact with regards to the mobility for both private stakeholder groups, it produces a low final ranking given the relatively low importance of that specific value. The ranking of this policy is

<table>
<thead>
<tr>
<th>Table 18</th>
<th>Policy ranking matrix for LSPs</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Profitability</td>
</tr>
<tr>
<td>P3</td>
<td>Gain leadership in sustainable logistics practices</td>
</tr>
<tr>
<td>P5</td>
<td>Consolidate demand</td>
</tr>
<tr>
<td>P4</td>
<td>Increase proximity to customers</td>
</tr>
<tr>
<td>P1</td>
<td>Improve asset utilization</td>
</tr>
<tr>
<td>P2</td>
<td>Shift delivery operations to desirable time frames</td>
</tr>
</tbody>
</table>
Table 19 shows the policy ranking matrix for the public sector. The policy G4 (i.e. promoting the environmentally friendly vehicles) is a policy that has achieved the highest ranking, given its high anticipated impact on the emissions that is highly valued by this stakeholder group. The second highest-ranked policy is G5 (i.e., promoting information sharing and transparency), that is expected to have a considerably high impact on all values and in particular the profitability.

<table>
<thead>
<tr>
<th>Policy</th>
<th>Profitability</th>
<th>Emissions</th>
<th>Mobility</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>G3  Promote environmentally friendly vehicles</td>
<td>6.9%</td>
<td>57.0%</td>
<td>3.4%</td>
<td>37.4%</td>
</tr>
<tr>
<td>G5  Promote information sharing and trans-</td>
<td>60.4%</td>
<td>24.8%</td>
<td>39.9%</td>
<td>33.3%</td>
</tr>
<tr>
<td>parency</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>G2  Shift delivery operations to off-peak</td>
<td>5.2%</td>
<td>4.9%</td>
<td>36.0%</td>
<td>12.4%</td>
</tr>
<tr>
<td>hours</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>G4  Enable multi-modality</td>
<td>16.7%</td>
<td>8.1%</td>
<td>5.0%</td>
<td>8.5%</td>
</tr>
<tr>
<td>G1  Reduce traffic disruptiveness from</td>
<td>10.8%</td>
<td>5.1%</td>
<td>15.7%</td>
<td>8.4%</td>
</tr>
<tr>
<td>delivery operations</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

However, lower for LSPs. Indeed, the feasibility of the off-hour deliveries generally requires extended opening hours at receivers’ locations and is therefore more suited for deliveries towards large freight generators. Its impact on profitability for LSPs who are often delivering towards a larger number of receivers is very low. For shippers from the beverage industry presented at the table and delivering to large generators such as supermarkets, this policy might be slightly more profitable.

9.3.3 Post-workshop Interviews

In order to answer RQ3 and contrast preferences of different stakeholders, we organized interviews with participants in the days after the workshop. During these interviews, we assessed general impression about the workshop content and dynamics, just as the potential shift in preferences, given insight on other stakeholders’ preferences. Appendix 5 provides a summary of interview data. These interviews reveal two important aspects.

The first aspect relates to the general lack of shared awareness - indeed, the majority of interviewees have stated that they were to some extent surprised by the preferences of the other stakeholder groups. Given that proper knowledge of other stakeholder group’s behaviors and preferences is crucial for predicting their behavioral response and tailoring effective urban freight policies, this result confirms the necessity of implementing stakeholder engagement actions that aim in sharing information and promoting mutual understanding.
The second aspect relates to the shift in stakeholder priorities once that they learned about the priorities of the other stakeholder groups. We have observed two rationales for the shift in priorities. On the one hand, the shift in priorities was a consequence of the availability of new information exchanged between the public and the private sector. For example, when learning that government ranks high the policy G4 (i.e., promoting environmentally friendly vehicles), and that they will focus on incentives rather than penalties, some shippers’ ranking of the policy P3 (i.e., gain leadership in sustainable logistics practices) has increased. Similarly, for LSPs, the newly acquired knowledge about the potential incentives has reinforced their ranking of the P3. On the government side, when learning that shippers have provide a low ranking for the P3, one participant highlighted the necessity of discussing these efforts further with the private sector and possibly readjusting their policies in view of the newly acquired information. Also, two government representatives highlighted that the high ranking of the policy P5 from the LSPs (i.e. demand consolidation) side would push them to reevaluate the importance of the policy G1 (i.e., reducing traffic disruptiveness from delivery operations).

On the other hand, shifts in priorities were motivated by foreseen synergies among actors within the private sectors, which would result in easier implementation of the policies. For example, when LSPs learned that shippers have a strong preference for moving closer to the customer (P5), their ranking of that policy increased given the potential synergies that can be created. Similarly, when learning the row ranking of the policy P2 (i.e., shift delivery operations to desirable time frames) by the shippers, their ranking of the policy decreased even more. Indeed, given that urban freight logistics policies require system-wide changes, companies are aware of the necessity of joining efforts to achieve those changes. For example, retailers’ practices of refusing off-hour deliveries can only be impacted if a large number of logistics service providers joins in promoting this new measure.

9.4 Findings and Discussion

This study focuses on a decision-making process along the triple bottom line in a multiple stakeholder setting. Content analysis reveals that stakeholders do not emphasize values impacted by urban freight logistics when communicating through their website. More precisely, the values ‘mobility’ (social dimension) and ‘emissions reduction’ (environmental dimension) do not appear in a significant manner on the websites.

The comparison of the results of the content analysis and the workshop reveals a certain discrepancy, showing that values that are used in overall communications are not necessarily the same as those used for decision making.

The study also confirmed the necessity of focusing the scope of the urban logistics policies.
First of all, logistics policies should have a clear geographic focus on a specific area in the city. Indeed, leadership in environmental practices is especially important in zones where operations are visible to the customers (such as the Candelaria), leading to creation of competitive advantage and a correlation between the environmental and economic dimension. Furthermore, if a zone (such as the Candelaria) does not represent a lot of business for a company, companies are more open to experiment (positively value) novel logistics policies.

Secondly, logistics policies should have a clear focus on a specific stakeholder group and a specific sector. For example, leadership in environmental practices is especially important for logistics service providers and arguably less in other sectors. Also, policies that aim in consolidating the demand are not applicable to beverages sector because of: (a) cash payments upon delivery; (b) heavy load; (c) return logistics. Retail, on the other hand, where (a) payments occur before delivery, might be more open to this policy, if (b) the load is not too heavy so the customer can carry from drop-off point to customer’s site.

Finally, logistics policies should also target specific companies within a sector. Indeed, not all companies can aim to differentiate through sustainable practices (i.e., making the business case for environmentally-friendly vehicles is easier for companies that will attract more customers through a green positioning).

We were able to generate iterations of preference structures, where stakeholders became aware of other stakeholders’ potential actions. We were able to highlight two mechanisms for the shifts in preferences. One the one hand, better information exchange between the public and the private sector allowed to reduce policy incoherence. On the one hand, information exchange between private sector actors enables to achieve more synergies between policies pursued by companies. This is a vital step towards actual implementation of policies.

The study also confirms the importance of dissociating values and the means of achieving those values as already noted by Keeney (1996) - indeed, some policies are highly rated in terms of impact that they can achieve on certain values (e.g., off-hour deliveries have a high impact on the mobility) but are focused on aspects that are not considered important by the decision-makers.

With regards to the contribution to AHP methodology, our research demonstrated the importance of the stating and continuous reminding of assumptions behind each policy and action. For example, in our case, assumptions were made in order to avoid participants disregarding certain policies because of the anticipated difficulty in implementation (e.g., assumption that retailers are willing to receive goods after opening hours) or the anticipated cost. It was required to repeat these assumptions all throughout the workshop.
10 PHASE III: A TRAFFIC SIMULATOR TO QUANTITATIVELY ASSESS URBAN FREIGHT POLICIES

10.1 Overview

In this section, we introduce a traffic simulation tool to evaluate the impact of proposed urban freight policy interventions (Table 12). Simulations are an effective way of evaluating policies as they can capture relations that may be too complicated to capture in an analytical model (Daamen et al. 2017). A microscopic simulator is a specific class of traffic simulator which models the behavior of individual vehicles and pedestrians and how they interact with each other and the transportation infrastructure. This type of simulator is particularly useful in evaluating the impact of the urban freight policies as we can model the interaction between freight vehicles and the general traffic throughout key parts of a city where the policy is implemented. This allows us to understand the impact of this policy on the transportation system as a whole by measuring the impact on congestion, emissions and other key logistic metrics.

We select policy G1, implementation of un-/loading delivery bays to illustrate the potential of this tool. In order to evaluate this policy, a traffic simulator was extended by implementing a model on top of it which allowed us to simulate logistical systems within a road network under the stated policy. The next sections will describe the developed model in more detail, the inputs required to run a simulation, the key performance indices measured by the tool and any future planned developments. We will conclude this section by presenting some preliminary results obtained by the tool when analyzing the impact of the policy in the historic region of downtown Quito, Ecuador. It is important to note that these results are preliminary as the tool has been developed for the project and is still in prototype phase.

10.2 Simulation Model and Tool

In order to assess the impact of the policy using a simulator, a model was developed and implemented on top of an existing traffic simulator to extend the tool further. This model allows us to implement freight vehicles, dedicated delivery bays and customers into the system to simulate how the fleet of freight vehicles operating in a particular region, under the specified policy, would serve the customers, and measure the impact on the system. Specifically, we are interested in measuring the following outputs:

- Total pollutants emitted by the freight vehicles
- Total distance traveled by the freight vehicles
- Total distance walked by the delivery personnel operating the freight vehicles
- Service time of the freight vehicles to deliver goods
10.2.1 Data

In order to run a scenario on the simulation tool to assess the impact of the dedicated bays policy in a particular area, the following variables need to be specified:

Geographical region. A file detailing the transportation network we wish to measure the policy impact on needs to be specified. This network should contain all the information needed to realistically model the scenario such as one-way streets, speed limits, traffic lights or truck restrictions on certain roads.

Time window to simulate. The time window where we wish to simulate the freight operations should be specified.

Traffic intensity. The traffic occurring in the time window and region we wish to simulate needs to be provided. This allows the simulator to simulate the interaction between the freight operations and general traffic.

Freight operations intensity. The intensity of the freight operations within the simulated time window and region needs to be provided in the form of individual customer orders distributed across the network. A second file detailing the intensity of the logistic operations throughout the time window should be specified so that the simulator models how the number of freight vehicles changes with time.

Properties of freight vehicles. The properties of the freight vehicles need to be specified. These properties include the logistic properties such as a vehicle’s capacity, number of personnel on board or the time needed to make a delivery. The physical properties of the vehicles such as maximum speed, size or acceleration should also be specified.

Dedicated freight un-/loading bays. The locations of all dedicated un-/loading bays for freight operations should be specified.

10.2.2 Simulation Environment

The Simulation of Urban MObilility (SUMO) traffic simulator was chosen as the traffic simulator which we will leverage to evaluate the policy (Krajzewicz et al. 2012). SUMO is an open source microscopic traffic simulation package which allows users to analyze complex traffic systems. It is actively being developed by the Institute of Transportation Systems at the German Aerospace Center. It was chosen as the simulator as it is a microscopic in nature and open source. This allowed us to extend the simulator further with the model we developed to include additional functionalities such as logistic systems and dedicated delivery bays. Furthermore, SUMO has its own graphic user interface, as shown in Figure 37, so that the user can observe how the system develops with time and how the vehicles interact with one another.
The figures below show two views of a SUMO network loaded with vehicles – the left shows an overview of the network, where the size of each vehicle has been emphasized for illustration purposes, and the right view shows vehicles at an intersection with a traffic light.

10.2.3 Model Dynamics and Limitations

In order to simulate the freight operations under the dedicated delivery bay policy, the following features had to be included in our model:
Logistic freight vehicles and delivery personnel. An overview of the design of the logistic operations is shown in Figure 40. Each freight vehicle has a specified number of delivery personnel on board who make the deliveries to the customers by exiting the truck and walking to the customer location. The time taken to unload the goods from the freight vehicle and transfer them to the customer can each be respectively specified by the user.

Each freight vehicle is part of a fleet of vehicles belonging to a certain distribution company. Each distribution company can therefore have multiple freight vehicles operating in the region and each vehicle can have different properties such as capacity or speed. On a top level, there are multiple different distribution companies operating in the regions making deliveries to the customers.

Establishments receiving goods. The customers receiving goods from the freight vehicles were simply modeled as coordinates on the road network. The size of the goods being received by the customer can be specified which influences the time taken to transfer the delivery from the freight vehicle to the customer.

Dedicated delivery bays. The dedicated un-/loading bays were modeled as simple parking spots large enough to fit a single freight vehicle of any size. Only freight vehicles can park in these spots – other general vehicles were restricted from doing so.

Delivery routing problem. The final main feature implemented in the model was to determine how the fleet of freight vehicles should serve the customers optimally. This was done separately for each distribution company which operates a fleet of freight vehicles in the region. Firstly, clusters of customers were formed based on the distance of customers to the closest un-/loading bay. Secondly
a capacitated vehicle routing was solved to determine in which order these cluster of customers should be served. Finally, when the freight vehicles stops to serve a cluster of customers, a second capacitated vehicle routing problem is solved to determine the optimal order in which the delivery personnel should serve the customers.

Figure 41 below shows several of the implemented features in the SUMO simulator, namely a freight vehicle, a delivery personnel, a customer delivery and a dedicated un-/loading bay.

![Figure 41 Model features developed in the SUMO traffic simulator](image_url)

**Limitations.** When developing the model, we made several assumptions about the dynamics of the logistic freight vehicles. Firstly, we assumed that when making deliveries to customers, freight vehicles can only ever stop at the dedicated delivery bays. We restricted these freight vehicles from stopping in the middle of the road or parking on on-street parking when making deliveries. This assumption was made as it would have been too complicated, for the first iteration of this tool, to model the dynamics of how freight vehicles park when serving customers. One consequence of this is that if there are no free spots at any un-/loading bay within a specified distance of the customer location, the freight vehicle will idle around the network until one becomes available.

Secondly, as we have not assigned any time windows to individual deliveries, it is assumed that as long as the deliveries are made within the total simulated time window (e.g. Monday from 08:00-14:00), the customer orders can be served in any order.

**Future development.** As mentioned earlier, the tool developed to evaluate this policy is still in prototype phase and there are many planned future extensions. Most notably:

(i) Removing the restriction where the freight vehicles can only park at the dedicated un-/loading delivery bays. This would allow us to model more realistic freight vehicle behavior as they will likely not wait for a dedicated spot to become available if there isn’t one – they would rather look for on-street parking or double park.
(ii) Extending the type of delivery units operating in the system such as bicycle couriers
(iii) Improving the calibration of operational parameters such as delivery time, un-/loading times
     and number of customers served per vehicle.
(iv) Developing an extended set of demand and traffic scenarios to evaluate the robustness of the
     system under operational conditions.
(v) Extending the simulation model to assess other policy interventions (see Table 12).

10.3 Application to Case Study
A preliminary analysis was done, using the developed tool, on the historic downtown region of
Quito, Ecuador, to assess the impact of dedicated delivery bays on the region. Specifically, we
evaluated the impact of 20, 38, 58, 76, 96, 115, 134 and 154 delivery bays in the downtown region.
This range for the number of bays is defined based on a previous exploratory study conducted
by the authors of this report in the historic center of Santiago, Chile, zone which shares many
infrastructure and retail activity characteristics with downtown Quito.

10.3.1 Parameter settings
The following parameters were used as inputs to the model:

Geographical region. The transportation system of the historic downtown (Centro Histórico)
region of Quito, Ecuador, was generated by extracting data from OpenStreetMaps (OSM) (The
OpenStreetMap Foundation 2017), a free provider of geospatial data, as shown in Figure 42. The
boundary coordinates of this region have respective latitude and longitude coordinates of (-0.224,
-78.517) and (-0.216, -78.507). Recall that downtown Quito is one of the selected zones (see Table
10) for in-depth policy analysis discussed in 6.

Figure 42 Quito downtown region in OSM
Time frame to simulate. The time frame chosen to simulate the freight operations was a Monday morning period from 08:00 to 14:00. Based on the field observations and survey data in this zone discussed in Sections 6 and 7 (see Figure 33), this time-window concentrates 36.1% of of the freight activity in commercial areas.

Traffic intensity. The traffic intensity within the specified time window, in the downtown region, was estimated by calculating the average speed across all roads in the region and increasing the flow of vehicles into the system until a similar speed profile was obtained. The average speeds of the network were obtained using the Google Matrix API (Google 2017).

Freight operations intensity. The delivery intensity was based on the 08:00-14:00 time window specified above. Specifically, a demand model, based on the field observations and survey analyzed in Section 7 (see Table 11), is used to generate a spatial distribution of demand across the downtown area. The demand was also split by the type of customer (e.g. small grocery store, clothing store, pharmacy, etc.).

Properties of freight vehicles. Two types of vehicles were used to serve the demand in downtown Quito: a small truck (20 units of capacity) and medium truck (40 units of capacity). For this simulation, the capacity of the truck is determined based on the average number of establishments served. The proportion of each truck making the deliveries varied with the type of customer. These parameters have been defined based on subject matter expertise and field observations.

Dedicated freight un-/loading bays. The simulation was run for the following number of dedicated delivery bays: 20, 38, 58, 76, 96, 115, 134 and 154. While the circumstances that make a location ideal for loading bays will depend on some difficult-to-model factors such as neighborhood character or infrastructural characteristics not appearing in our simulated transportation network, a chief driver of where to locate these loading bays should be their proximity to areas of intense commercial activity. To achieve a reasonably efficient allocation of these parking locations throughout the district, we use a K-means clustering algorithm to group the commercial demand into \( p \) clusters, where \( p \) represents the number of loading zones to model in the square-kilometer district. The demand-weighted centroid of each cluster becomes a loading zone, which is then moved to the nearest point that lies on a street and provided to the simulated freight vehicles as a parking location. We then modify \( p \), generating several different scenarios under which we simulate traffic conditions for different levels of accommodation of freight traffic.

10.3.2 Analysis results

The multiple scenarios with 20, 38, 58, 76, 96, 115, 134 and 154 respective delivery bays were then evaluated using the simulator under the conditions outlined above. Each simulation was run
from 08:00 AM to 14:00 PM and was ended abruptly at the end of this time window, regardless of whether the freight vehicles had completed their journeys.

Figure 43 shows the proportion of deliveries which were completed by the fleet of freight vehicles within the time window versus the number of delivery bays in the system.

The proportion of completed deliveries increases non-linearly with the number of dedicated delivery bays in the system. Indeed, there is an initial rapid increase in completed deliveries which then begins to level off as the number of dedicated delivery bays in the system increases further. The reason the proportion of completed deliveries is relatively low, with a maximum value of 46% when there are 154 bays in the system, is because the freight vehicles, in the scenario, are relatively large with capacities of 20 and 40 units. Since, the trucks are filled to capacity, there are 95 total freight vehicles in the system in the simulated period. Therefore, they cannot complete their assigned deliveries in the simulated time window. If the capacities of the supply trucks were to be reduced, we would see more trucks in the system and higher proportion of completed deliveries.

Figure 44 shows the total \( CO_2 \) emitted by the fleet of freight vehicles versus the number of bays in the system and Figure 45 shows the emitted \( CO_2 \) per completed delivery. The total \( CO_2 \) emitted by the freight operations reduces non-linearly with the number of delivery bays primarily because as more delivery bays are made available vehicles will have to drive less to find an available delivery bay. There is an initial rapid decrease in emissions and it levels off as the number of delivery bays increases. The same is true for the \( CO_2 \) emitted per completed delivery with an even more pronounced effect - the \( CO_2 \) emitted per completed delivery levels off at a faster rate since more deliveries are completed as we increase the number of delivery bays. The ‘elbow’ in Figure 45 is an important piece of information to assess infrastructure capacity requirements. In this example, we
observe marginal improvements in $CO_2$ for a capacity level of more than 80 delivery bays, a result that should be contrasted with the financial investment required to deploy the additional capacity.

We are also interested in the total distance traveled by the freight vehicles in the system as well as the distance traveled by the delivery personnel operating the freight vehicles. Figure 46 shows the total distance traveled by the freight vehicles. The distance traveled decreases non-linearly with the number of delivery bays. It may be expected that the distance traveled increases with the number of delivery bays in the system since the freight vehicles are less restricted to where they must park. However, one assumption of our model is that the freight vehicles can only make deliveries to customers when parked at a dedicated delivery bay. Thus, the trucks spend considerable time traveling around the network looking for available delivery bays. This effect is more significant when there are fewer dedicated delivery bays in the system.

![Figure 44](image1.png)  
**Figure 44**  
Total $CO_2$ emitted by freight vehicles

![Figure 45](image2.png)  
**Figure 45**  
$CO_2$ emitted by freight vehicles per completed delivery
In the current version of the simulator, we can only measure the distance traveled by pedestrians which have completed all of their assigned deliveries. However, in the simulated scenario, no pedestrians complete their assigned deliveries and thus it was not possible to measure the distance traveled by the delivery personnel.

10.4 Conclusion and Future Work
We introduce a prototype version of a traffic simulation tool to evaluate the impact of proposed urban freight policy interventions. This tool has been designed to understand the impact of selected urban freight policies on the transportation system as a whole by measuring the impact on congestion, emissions and other key logistic metrics. A case study in the most commercially relevant zone in Quito illustrates the potential of this tool. We expect to continue developing this simulator to address current limitations, including scalability and transferability.
11 ENGAGEMENT WITH LOCAL STAKEHOLDERS AND DISSEMINATION OF PROJECT RESULTS

In this section we describe the activities we conducted over the course of the project to engage with local research partners and stakeholders in each city. In total, we conducted five workshops over the course of the project:

- Two workshop data collection and analysis workshops
- One multi-stakeholder analysis workshop
- Two results dissemination workshops

11.1 Data collection and analysis workshops

The main purpose of these workshops was to design and align data collection and analysis strategies with local research teams. The first workshops took place on March 28, 2017 in Bogotá, hosted by LOGyCA. The primary objective of this workshops was to introduce the data collection methodology to all local teams and devise data collection plans.

A second data collection and analysis workshop was hosted at the MIT campus in Cambridge, MA, on August 21-25, 2017. The aim of this week-long meeting was to review data collection progress, devise strategies to overcome data completeness challenges and align on data analysis approaches.

Figure 47 Workshop at MIT
11.2 Multi-stakeholder analysis workshop

As detailed in Section 9, we hosted a multi-stakeholder workshop to gather information on decision makers' preferences and perceived impacts of proposed urban freight policy interventions and operational practices. The workshop was hosted in LOGyCA on November 30, 2017. Seventeen decision makers representing private companies, public sector agencies and research institutions participated in the workshop.

11.3 Results dissemination workshops

We hosted two final workshops to present project results and stimulate urban freight policy discussion with local stakeholders in Lima and Quito (Figure 48). These workshops were hosted by Universidad San Francisco in Quito and Universidad del Pacifico in Lima on January 29 and 30, 2018, respectively. More than 70 participants attended these two events, which also had extensive media coverage (see, for instance, a news piece in this link). A summary video of the Lima workshop is available here.

In these final workshops also collected anecdotal evidence on the acceptability, feasibility and impact of the proposed policy interventions and operational practices. In general, participants agreed on the need to device mutually reinforcing policy interventions and operational practices. Furthermore, to implement these recommendations, existing mobility and land use regulations will have to be revised and strengthened.

Finally, participants highlighted the importance of this study, mainly because it raises awareness on this increasingly relevant topic and proposes policy interventions and operational practices that are contextually relevant. Future research efforts should continue to closely engage local stakeholders through workshops and task-forces to strengthen the process of designing sustainable urban freight policies.
12 CONCLUSION AND OUTLOOK

We report 3 citywide studies in Bogotá, Quito and Lima to devise a “city logistics policy toolkit” to make recommendations with respect to urban freight policies and practices based on a systematic, replicable, data-driven approach. The toolkit allows analyzing a city as a whole, identifying different logistics needs by city areas (at the square-kilometer level) and highlighting critical city areas where
the development and deployment of better urban freight policies and practices should be a top priority. The toolkit includes several semi-automated software tools to clean, process, analyze and visualize data, which we have made available in the project website: https://github.mit.edu/pages/MIT-MLL/WB_2018_Policy_Toolkit/

Our results for the first project phase propose 6-8 archetypes of city zones for logistics policy planning. This classification properly captures the main dis-/similarities in the demographic, economic, and infrastructural properties of the various zones within each city. Particular attention should be paid to archetypes with high intensity of retail activities and higher average population density due to the intensity of freight flows they attract.

The second project phase entails an in-depth characterization of retail and logistics activities within these priority zones. We identify zone-specific characteristics as well as common logistics patterns to inform the design of tailored, yet transferable, urban freight policy interventions. Small grocery stores and food service establishments are the predominant retail establishment in most zones. However, the retail composition varies significantly across zones, which ultimately drives the intensity of freight operations. Temporal patterns of logistic operations are, however, consistent across zones and typically concentrated in morning, 9-11 am, and afternoon, 2-4 pm, time-windows. These time-windows also reflect the implicit operational constraints on logistics activities imposed by the hours of operation of retail establishments.

Building on the results for phase I and II, we propose five policy interventions, along with five complementary private sector operational practices to minimize externalities while providing adequate flow of goods within the city. Proposed policy interventions range from initiatives geared towards providing infrastructure of parking and transshipment of goods, to actions designed to incentivize the use of environmentally friendly vehicles and information sharing. The proposed operational practices for the private sector complement and reinforce proposed policies.

We also report the findings of a multi-stakeholder analysis to assess the effects and acceptability of these proposed policy interventions and logistics practices. We argue that policy interventions need to account for differences in value preferences and operational requirements. While shippers prioritize practices to improve asset utilization and increase customer proximity, logistics operators prioritize actions geared towards sustainable logistics leadership and demand consolidation. Both, shippers and logistics operators express find unfeasible to shift delivery operations to off-peak hours due to numerous challenges. Government officials exhibit preferences for policies that incentivize environmentally friendly vehicles and promote information transparency.

Finally, we introduce an traffic simulation tool to quantitatively evaluate the impact of selected policy choices. The simulator effectively captures complex interactions between different system
elements and enables a holistic assessment of policy interventions by measuring the impact on congestion, emissions and operational efficiency.

Future research directions include extending current tools to yield more comprehensive assessments of selected policies. Specifically, we expect to continue developing the traffic simulation tool to address current limitations and increase its transferability to other policies and local contexts. We also plan on developing data-driven visual interfaces for policy design that facilitate interaction among multiple stakeholders. Furthermore, the fast expansion of online retailing, particularly in emerging markets, will bring new urban freight challenges that will need to be explored by means of rigorous analytical frameworks.
LIST OF ACRONYMS

CCB Cámara de Comercio de Bogotá ................................................................. 11

ICQ Instituto de la Ciudad de Quito ............................................................... 11

INEC Instituto Nacional de Estadísticas y Censos ......................................... 11

ISIC International Standard Industrial Classification ..................................... 13

NS Navi Systems EIRL .................................................................................... 13

PC principal component .................................................................................. 17

PCA principal component analysis .................................................................. 17

POI Point of interest ....................................................................................... 13

SDP Secretaría Distrital de Planeación .............................................................. 11

AHP Analytical Hierarchy Process ................................................................. 56

B2B Business to business ............................................................................... 57

B2C Business to consumer ............................................................................. 57

RMS Root means square ............................................................................... 59

LSP Logistics Service Provider ..................................................................... 61

SUMO Simulation of Urban MObility .............................................................. 69

OSM OpenStreetMaps .................................................................................... 73
REFERENCES


Shmueli NRBPC Galit; Patel (2010) Data Mining for Business Intelligence (Wiley), URL http://dx.doi.org/10.1017/CBO9781107415324.004.


Winkenbach M, Roset A, Spinler S (2016b) Strategic Redesign of Urban Mail and Parcel Networks at La Poste. *Interfaces* 46(5).