CrossMark

Commodity flow estimation for a metropolitan scale freight modeling system: supplier selection considering distribution channel using an error component logit mixture model

Takanori Sakai[1] · B. K. Bhavathrathan[1] · André Alho[1] · Tetsuro Hyodo[2] ·
Moshe Ben-Akiva[3]

Abstract

Freight forecasting models have been significantly improved in recent years, especially in the field of goods vehicle behavior modeling. On the other hand, the improvements to commodity flow modeling, which provide inputs for goods vehicle simulations, were limited. Contributing to this component in urban freight modeling systems, we propose an error component logit mixture model for matching a receiver to a supplier that considers two-layers in supplier selection: distribution channels and specific suppliers. The distribution channel is an important element in freight modeling, as the type of distribution channel is relevant to various aspects of shipments and vehicle trips. The model is estimated using the data from the Tokyo Metropolitan Freight Survey. We demonstrate how typical establishment survey data (i.e. establishment and outbound shipment records) can be used to develop the model. The model captures the correlation structure of potential suppliers defined by business function and provides insights on the differences in the supplier choice by distribution channel. The reproducibility tests confirm the validity of the proposed approach, which is currently integrated into a metropolitan-scale agent-based freight modeling system, for practical use.

Keywords Freight demand model · Urban freight · Commodity flow · Error component logit mixture · Disaggregate

✉ Takanori Sakai
  takanori@smart.mit.edu

1   Singapore-MIT Alliance for Research and Technology, 1 CREATE Way, 09-02, CREATE Tower,
    Singapore 138602, Singapore

2   Department of Logistics and Information Engineering, Tokyo University of Marine Science
    and Technology, Tokyo, Japan

3   Intelligent Transportation Systems Lab, Massachusetts Institute of Technology, Cambridge, USA

Introduction

Urban freight has become a globally important policy subject. Efficient operations of urban freight systems are critical in the metropolitan regions that become larger and denser. Furthermore, the evolution of logistics and supply chain practices has rapidly transformed the way freight operations take place. In this context, it is of utmost importance to develop a comprehensive and flexible urban freight simulator that can analyze the impacts of relevant policies such as land use changes, vehicle restrictions, the use of urban freight consolidation centers, the designation of loading/unloading spaces, crowdsourced deliveries, and other "smart" measures for freight operations.

Several urban freight modeling systems have been proposed in the literature (e.g. Boerkamps et al. 2000; Fischer et al. 2005; Wisetjindawat et al. 2005, 2006; Hunt and Stefan 2007; Nuzzolo and Comi 2014; Moeckel and Donnelly 2016; Alho et al. 2017) and, overall, freight modeling has been advanced significantly during the last few decades. However, the models for the commodity flows within a metropolitan area are underdeveloped (Comi et al. 2012). This is attributable to the heterogeneity of urban freight agents and the commodities handled as well as the difficulty of obtaining data, particularly disaggregate data. For urban freight, it is especially important to consider the nuances of sourcing behaviors by receiver type, such as offices, stores, restaurants, logistics facilities, factories, cultural and educational facilities, and residential facilities, which are relevant to commodity flows. Due to the data limitations, it is quite common to use aggregate economic statistics such as input–output tables to estimate commodity flows in freight traffic simulations (Ben-Akiva and de Jong 2013). On the other hand, a lot of efforts and resources were spent on collecting disaggregate data until this day. Several metropolitan-scale establishment surveys, which are also called commodity flow surveys, were conducted in recent years or are currently ongoing for informing urban freight models (Hunt et al. 2006; Alho and e Silva 2015; Toilier et al. 2016; Cheah et al. 2017; Oka et al. 2018). Exploring urban freight modeling approaches that make full use of those survey data is one of the important research topics.

One of the research gaps in urban commodity flow modeling is the consideration of distribution channels. Without the consideration of distribution channels, the model, for example, cannot differentiate direct and indirect shipments that go through logistics facilities, i.e. warehouses, distribution centers, truck terminals, and other intermediate facilities. While many freight models that deal with commodity flows were proposed in the past, the use of the facilities is considered mostly in the context of inter-city commodity flows (Huber et al. 2015); only a small number of the models explicitly consider the use of logistics facilities in the urban commodity flow estimation. Although an entire supply chain does not fit within a metropolitan region in many cases, the spatial distribution of urban logistics facilities is suspected to influence freight traffic flows (Dablanc and Rakotonarivo 2010; Sakai et al. 2015a). Furthermore, without considering the distribution channels, one cannot analyze the impact of the changes in urban structure. If some factories that have mainly been serving local markets move out, the local demands should be served by other suppliers, which could be other factories in the area or the logistics facilities that receive goods from outside. This paper aims to fill this research gap by proposing and testing a new approach that reproduces aggregated logistics chain structures within a metropolitan area through the estimation of the commodity flows. Here, we define a "logistics chain" as the connection between the locations of production and consumption, consisting of means of transportation, mainly goods vehicles in urban freight, and the intermediate locations

that goods go through. Our approach is particularly appealing as it requires only typical establishment survey data, i.e. establishment and outbound shipment records.

For the model specification, we assume that commodity flows are demand-driven; a receiver, directly or indirectly through the cost minimization principle, determines a distribution channel (defined by the supplier function type, i.e. factory, logistics facility, office/store), and a supplier which has a specific location. While this is a strong assumption, it makes a simple and robust model specification possible. We present a framework for the supplier choice model and develop an error components logit mixture model that considers the complex correlation structure among alternative suppliers. We use the data from the Tokyo Metropolitan Freight Survey (TMFS) conducted in 2013 to estimate the model. The focus of this paper is the modeling approach, but the estimated parameters in the models using Tokyo's data also provide interesting insights on the supplier choice mechanism. The proposed framework is integrated into a comprehensive metropolitan-scale agent-based modeling system, named "SimMobility" (Adnan et al. 2016). The SimMobility accommodates various decisions and behavioral aspects related to urban freight traffic, ranging from those based on long-term "strategic" visions, e.g. business characteristics and overnight parking locations for goods vehicles, to those relevant to daily logistics operations, e.g. vehicle touring, route choice, and delivery/pick-up parking (Alho et al. 2017). The proposed approach is also applicable to the commodity flow estimation in other existing urban freight modeling systems.

The rest of the paper consists of the following: in second section, the recent developments of freight models are reviewed, focusing on disaggregate commodity flow estimations and the consideration of distribution channels in the models; in third section, a proposed model structure and the calibration data are presented; in fourth section, the model specifications and the estimation process are detailed; in fifth section, the estimated models and the elasticity effects are discussed; in sixth section, tests for the model reproducibility are presented; and finally, in seventh section, the conclusions of this research are provided.

Literature review

In the last two decades, a number of freight forecasting models were proposed, considering the behaviors of individual agents and/or the decision-making processes in logistics operations (Chow et al. 2010; Comi et al. 2012; De Jong et al. 2013). These models are used for testing various logistics practices and policies, for which traditional aggregate-level freight models are not suited. Chow et al. (2010) provide a review of the two emerging groups of freight forecast models, "logistics models" and "vehicle touring models", and De Jong et al. (2013) summarize the recent model development (2004 and after) mainly focusing on European models. There are various behavioral aspects that are considered in freight models, such as freight/trip production and consumption, commodity flow, mode choice, shipment frequency and size, vehicle scheduling and routing, and vehicle parking. The commodity flow estimation, the focus of this research, is one of the most essential components in a forecast model regardless of its scale, i.e. regional, national or urban.

A large number of approaches for the commodity flow estimation were proposed in the past, most of which are based on aggregate data. As disaggregate (i.e. establishment level) commodity flow data are often not available, the zone-to-zone level commodity flow estimation is still widely applied. Even recently proposed urban freight modelling systems, such as Nuzzolo and Comi (2014) and Moeckel and Donnelly (2016), treat commodity

flows at an aggregate level. However, the aggregate models are passible of aggregation biases. Furthermore, the aggregate commodity flows have limitations for their use in a microsimulation of logistics behaviors (Ben-Akiva and de Jong 2013). Ben-Akiva and de Jong propose an Aggregate-Disaggregate-Aggregate (ADA) freight model system that specifies the procedure to disaggregate zone-to-zone flows to establishment-to-establishment flows. The ADA freight model system was implemented in the Swedish national freight transport model (SAMGODS) (Abate et al. 2018). Zhao et al. (2015) propose another approach to disaggregate zone-to-zone flows to firm group-to-firm group flows through a linear programing (LP) problem. They solve the market system optimum of commodity flows with the constraints on zone-to-zone flows, obtained from Freight Analysis Framework Version 3 (FAF3) data, for California, U.S.

Regarding fully disaggregate approaches, Boerkamps et al. (2000) propose the conceptual structure of Goodtrip model, one of the earliest urban freight microsimulation modeling systems. In the Goodtrip model, the commodity flows are estimated at an actor-to-actor level. Activity type (such as "consumers, supermarkets, stores, offices, distribution centers of retailers, and producers") is determined for taking the differences in flows by shipper and receiver types (type of "linkage") into account. However, the choice of the linkage type is not considered and the use of intermediate facilities is not modelled. An urban freight microsimulation model proposed by Wisetjindawat et al. (2005) includes "purchasing decision module" that consists of three components: distribution channel choice, (supplier) location choice, and supplier choice, all proposed to be multinomial logit models, for estimating commodity flows. In "distribution channel choice" model, the distribution channel is defined by supplier industry type (e.g. warehousing industry and manufactures), but not by function type (e.g. warehouse and factory). Although the work of Wisetjindawat et al. is pioneering, they do not fully describe the model estimation method and the implications from the estimation. Roorda et al. (2010) propose another conceptual framework for an agent-based logistics microsimulation. The proposed "commodity contract formulation" process is corresponding to the commodity flow estimation. In the model, a customer chooses a supplier from a choice set that can fulfill the requirements on order size and frequency. Similarly, Liedtke (2009) proposes "sourcing module" in which suppliers are chosen based on production size, commodity usefulness, and distance, for INTERLOG, a national level freight simulator. Both Roorda et al. and Liedtke do not explicitly consider distribution channels and logistics chain structures. There are also a couple of novel practices of the commodity flow estimation in the disaggregate freight modeling, taking deterministic modeling approaches. Gliebe et al. (2015) propose the use of a Bayesian game-based approach for connecting buyers and sellers. The proposed procurement market game (PMG) buyer–seller matching model was developed as a part of the meso-scale freight forecasting model for the Chicago Region (RSG 2015). Furthermore, Livshits et al. (2017) propose an agent-based computational economics (ACE) approach for the market-clearing process that matches buyers and suppliers in an agent-based behavioral freight model for the Arizona Sun Corridor Megaregion.

Despite the recent developments of disaggregate models for the commodity flow estimation, there is room for improving the consideration of the choice of distribution channel, particularly for urban freight modeling systems. In the existing urban freight modeling systems, the function type of an agent is underrepresented and often associated with industry type. However, it is fair to hypothesize that factories, warehouses and offices have quite different behavioral characteristics even if all those belong to manufacturing industry. Consideration of logistics facilities (such as warehouses and distribution centers) is especially important in simulating commodity flows. Huber et al. (2015) review more than

a hundred freight models and find that only a small number of models consider the use of logistics facilities. Proposing one of such models, Davydenko and Tavasszy (2013) apply a unique approach to address the use of intermediate facilities, by extending the SMILE model lineage. The SMILE model is an aggregate, national level freight model developed by Tavasszy et al. (1998). The proposed logistics chain model, which has a multinomial logit structure, chooses intermediate locations for Production-Consumption (P–C) flows, and estimates commodity flows (i.e. shipment ODs). The above-mentioned model of RSG (2015) for the Chicago Region also considers the decisions of indirect/direct shipments (in their "distribution channel model") and, for an indirect shipment, a local warehouse in the study area is assigned as the first or last transfer stop; however, the warehouse location is simply based on the random selection. The similar structure with the RSG's distribution channel model is proposed for the Arizona Sun Corridor Megaregion, although the model was not estimated due to the data limitation (Livshits et al. 2017). Focusing on the logistics chains that go through intermediate facilities within a metropolitan area, Sakai et al. (2017) propose a logistics chain model, which pairs freight demand generating locations with logistics facilities, which become the origins or destinations of shipments. They estimate the models by commodity type and the type of demand locations. The estimated models are used for estimating indirect commodity flows. Alternatively, rule-based and/or optimization methods that consider the use of warehouses can be applied to estimate commodity flows (e.g. Friedrich 2010). However, the required data for such models is not compatible with the extension of the approach to replicate the entire urban commodity flows in a metropolitan area.

In this paper, we improve and expand the framework of Wisetjindawat et al. (2005) and Sakai et al. (2017) for the estimation of the commodity flows within a metropolitan area, covering various distribution channel types. We contribute to the state-of-the-art urban freight models by proposing a method to model the multi-level decision of supplier function choice (i.e. distribution channel) and supplier choice. The model is particularly innovative by considering the distribution channels based on business function for the urban commodity distribution and aggregately reproducing the parts of logistics chains that are within the study area. It should be noted that our approach does not replicate the entire logistics chains that go beyond the scale of a metropolitan area. In fact, the data for the entire logistics chains are extremely difficult to obtain through a survey. However, the proposed approach in this paper just requires the shipment and supplier records, both of which are usually covered in a standard establishment survey for urban freight.

Model structure and calibration data

Analytical settings

In this paper, we define a shipment as a quantity of goods shipped together from one location to another without transshipment. A shipment based on this definition is also called a "leg". Logistics chains that connect producers and consumers consist of legs as shown in an illustrative logistics chain (Fig. 1). We use "daily attraction" (DA) as the unit of the data points for the analysis. A DA is defined as a daily freight attraction associated with a receiver facility that has specific function and location, which should be supplied by one of potential suppliers. In the simulation context, DAs, or freight attraction at another time scale, should be estimated by a freight attraction model. We
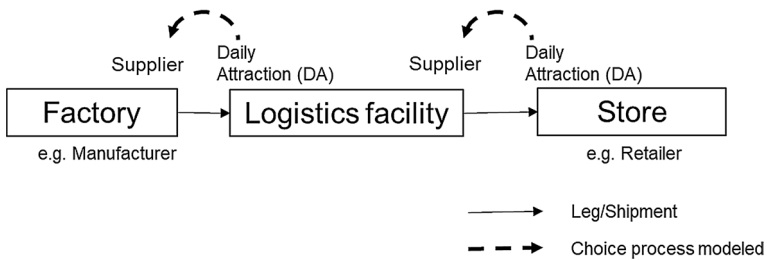
Fig. 1 Illustrative logistics chain consisting of two legs

consider five receiver function types: office/store/restaurant (OSR^R), logistics facility (LF^R), factory (FC^R), cultural and educational facility (CEF^R), and residential facility (RF^R). The superscript "R" indicates that those are receiver establishments. We define a supplier as an establishment, categorized into one of three function types: office/store (OS^S), logistics facility (LF^S), or factory (FC^S). As with receiver functions, the superscript "S" indicates that those are suppliers. Each supplier is associated with a geographical location and every supplier handling a required commodity type is a potential origin of a shipment for a receiver demanding goods.

The aim of the proposed model is to select a supplier, which is associated with one of three function types, for a DA at a receiver, who also has its own function type, for estimating intra-metropolitan commodity flows. In other words, the model estimates the shipment OD (or commodity flows) along with the "leg type", which is defined by the functions of the upper and lower ends of the leg. Given the receiver function type, the "leg type" is equivalent to "distribution channel" (Boerkamps et al. 2000; Wisetjindawat et al. 2005). Leg type is important to replicate freight traffic in an urban freight simulation model. For example, shipment size, vehicle type used, and delivery tour pattern, are all tied to the leg type; a shipment from a factory to a logistics facility and that from a logistics facility to a store can be different in these aspects and, if any, such differences must be considered.

The model is designed so that a receiver, to whom a DA is associated, selects a supplier. It does not necessarily reflect the reality in which the origin of the shipment is determined in the course of the complex interactions between both receivers and suppliers. However, this is a reasonable premise applied in previous freight modeling efforts (Wisetjindawat et al. 2005, 2006; Liedtke 2009; Roorda et al. 2010). This is further justified as, in general, the need for goods movement is associated with the receiver (i.e. demand) side and the receiver directly/indirectly has an influence on the location of the shipment origin through the effort of minimizing the total cost.

The connections between the receivers and suppliers established by the model are to be used for generating shipments. A carrier operation model, which is included in Sim-Mobility (Alho et al. 2017) (and, also, in other disaggregate urban freight simulators), takes these shipments as inputs and generates goods vehicle tours and trips, which are, in turn, the inputs for a traffic simulation. The type of the distribution channel can be considered at each of the steps that follows the commodity flow estimation. Regarding the framework of the SimMobility, the feedback structure is considered so that the outputs of the traffic simulation, for example travel time, are returned to the other "upstream" simulation components that include the commodity flow estimation.

Data

We use the data of the 2013 TMFS. The TMFS is an urban freight establishment survey conducted by the Transport Planning Commission of the Tokyo Metropolitan Region (TPCTMR) for the Tokyo Metropolitan Area (TMA), currently the most populous metropolitan area in the world. An area of 23 km^2 was covered and the responses were collected from 43,131 establishments (a response rate of 31.6%) in various industry sectors such as manufacturing, wholesale, transportation, service, retail and restaurant. The TMFS collected both establishment information and shipment records through a paper-based survey. The establishment information includes location (address), industry and function types, commodity types handled, employment size, floor area, and the year of establishment. The shipment records were collected for a typical day. Commodity type, weight, and the information of origins and destinations including their locations at the municipality level (315 municipalities in the survey area) and function types are covered. Each data point (i.e. establishment) has an expansion factor, which is calculated by the TPCTMR based on industry type, employment size, and geographic location, for taking extraction rates into account. The average expansion factor is 4.61. We use these expansion factors to reproduce the population of supplier establishments, which are considered as alternatives in the choice model, and also to expand the shipment records for parameter estimations. It should be noted that, in the shipment records in the TMFS data, "offices and stores" with distribution functions, such as those of wholesalers, are labeled as "offices and stores", as with those without distribution functions.

We estimate the models for nine commodity types shown in Table 1. This paper will focus the discussion on one commodity type ("household/light manufacturing goods") for demonstrating the modeling approach and interpretation. It covers sundry items, clothing and textiles, products of paper, rubber, wood, and furniture. The numbers of the associated shipments and supplier establishments are summarized in Table 2. It should be noted that these shipments do not include external shipments, i.e. the shipments sent to or received from outside of the study area. As indicated by Sakai et al. (2017), the supplier choice mechanisms for intra-metropolitan and inter-regional shipments are quite different. For considering the external shipments in a simulation, one must define the external productions and consumptions and connect these external locations with the establishments in the study area using another set of models, which can be estimated with the similar framework proposed in this paper. Most of the urban freight movements

Table 1 Commodity type categories

No.	Commodity type
1	Agricultural products
2	Food products
3	Household/light manufacturing
4	Wood and paper products
5	Minerals, ore, stone, cement, ceramics or glass
6	Metals or articles of metal
7	Machinery, appliances, and mechanical parts
8	Chemicals, rubber or plastics
9	Mixed goods and parcel

Table 2 Number of shipments and supplier establishments (after expansion)

Receiver	Household/light manufacturing goods				All commodity types			
	Supplier				Supplier			
	Office/store (OS^S)	Logistics facility (LF^S)	Factory (FC^S)	Other^a	Office/store (OS^S)	Logistics facility (LF^S)	Factory (FC^S)	Other^a
No. of shipment								
Office/store/restaurant (OSR^R)	24,274	27,212	18,275	347	91,165	169,177	77,902	9668
Logistics facility (LF^R)	3909	8174	5662	319	18,857	88,704	36,461	1842
Factory (FC^R)	6540	9229	12,448	121	36,425	45,448	95,271	1922
Cultural and educational facility (CEF^R)	19,978	3302	3039	124	41,842	18,492	10,465	563
Residential facility (RF^R)	4232	8234	3040	8	16,352	54,550	16,809	6868
Other^b	281	50	301	17	9559	9923	15,617	4765
No. of supplier establishment	5198	2626	5328	298	102,666	16,783	54,156	5668

^{a}"Other" are not considered in the model estimation as the number of the associated shipments is small; ^{b}"Other" are not considered as most of them are construction sites, which we exclude from the scope of analysis

in the TMA are by goods vehicles; therefore, the models in the present research are for the shipments transported by goods vehicles.

Model specification and estimation

Specification

The proposed structure of the supplier choice consists of two levels and we apply the logit mixture model with error components (Fig. 2) that allows us to capture the complex correlation structure among alternatives. The upper and lower levels are defined by supplier function types and individual suppliers, respectively. This structure is selected as the suppliers within the same function category (i.e. OS^S, LF^S or FC^S) are expected to correlate with each other. Also, the cross-nested structure is considered as the suppliers from two different function categories may correlate. For example, in some case, a commodity is highly likely supplied to a receiver by either of a OS^S or a LF^S but not by a FC^S. In Fig. 2, "downstream" (DWS) is the nest that includes both OS^S and LF^S and "upstream" (UPS) is that including LF^S and FC^S. This arrangement mimics the three-level nested logit model in which, the choice from UPS and DWS is at the first level, the one from OS^S, LF^S and FC^S is at the second level, and the individual supplier choice is at the third level. The proposed explanatory variables for the utility of a supplier for a receiver (Table 3) are the travel time between a supplier and a receiver, the freight production (i.e. the total weight of outbound shipments) of a supplier, and the weight of the shipment (i.e. DA). We decided to use travel time, not travel distance, because the travel time can capture traffic conditions, although these two variables are highly correlated to each other. Overall, we assume a supplier that is located closer and has a larger capacity than others should have a higher potential to be selected by a receiver. In addition, depending on the size of demand, a supplier with a specific function type would be more likely to be chosen. In general, a larger demand is more likely to be catered by a logistics facility or a factory than an office/store.
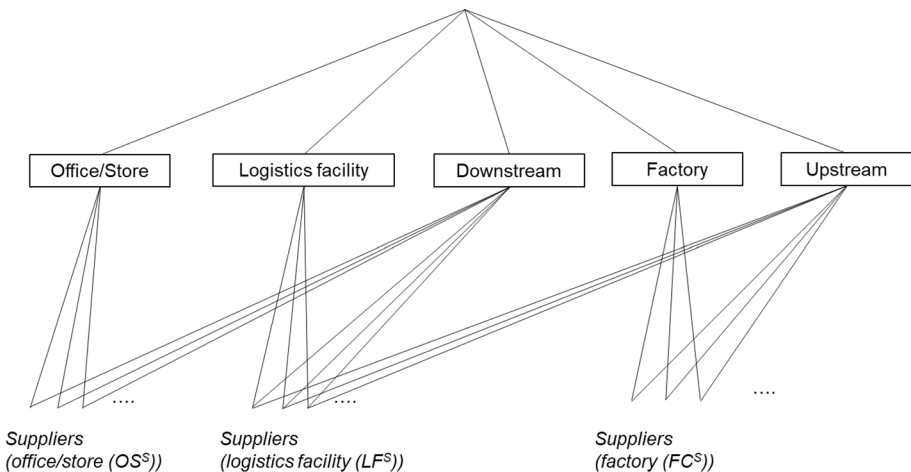

Fig. 2 Nesting structure for supplier choice

Table 3 Explanatory variables

Variable	Description
Travel time	Estimated travel time for the shortest path. Daily average speeds, which are estimated based on traffic counts conducted in 2010 and the Bureau of Public Roads (BPR) formula, are used for computing travel time for each link
Freight production	Total weight of the outbound shipments of a supplier per day obtained from the 2013 TMFS
Weight	Weight of the shipment of a typical day, i.e. daily attraction (DA), obtained from the 2013 TMFS

Each model is specific to a combination of receiver function *rf* (OSR, LF, FC, CEF, or RF) and commodity type *c*. For the simplicity of notification, *rf* and *c* are omitted in the following model specification.

The utility of a supplier with supplier function *sf* (*sf* = 1: office/store; *sf* = 2: logistics facility; *sf* = 3: factory), S^{sf}, for a daily attraction DA is:

$$U_{DA,S^{sf}} = V_{DA,S^{sf}} + M_{DA,S^{sf}} + \varepsilon_{DA,S^{sf}} \tag{1}$$

where $V_{DA,S^{sf}}$: Deterministic component of the utility. $M_{DA,S^{sf}}$: Random component that captures the correlation structure among alternatives. $\varepsilon_{DA,S^{sf}}$: Identically and independently Gumbel distributed random component (normalized: scale = 1, location = 0).

The deterministic component is given by the following equation:

$$
\begin{aligned}
V_{DA,S^{sf}} = & \left(\beta_{TIME}^{OS} \cdot Dum_{OS\,S^{sf}} + \beta_{TIME}^{LF} \cdot Dum_{LF\,S^{sf}} + \beta_{TIME}^{FC} \cdot Dum_{FC\,S^{sf}} \right) \cdot lnTIME_{DA,S^{sf}} \\
& + \left(\beta_{FP}^{OS} \cdot Dum_{OS\,S^{sf}} + \beta_{FP}^{LF} \cdot Dum_{LF\,S^{sf}} + \beta_{FP}^{FC} \cdot Dum_{FC\,S^{sf}} \right) \cdot lnFP_{S^{sf}} \\
& + \left(\beta_{Const}^{LF} + \beta_{Weight}^{LF} \cdot lnWEIGHT_{DA} \right) \cdot Dum_{LF\,S^{sf}} \\
& + \left(\beta_{Const}^{FC} + \beta_{Weight}^{FC} \cdot lnWEIGHT_{DA} \right) \cdot Dum_{FC\,S^{sf}}
\end{aligned}
$$

$$\tag{2}$$

where $lnTIME_{DA,S^{sf}}$: The natural log of travel time between supplier S^{sf} and daily attraction *DA*. $lnFP_{S^{sf}}$: The natural log of freight production (i.e. total weight of outbound shipments) by supplier S^{sf}. $lnWEIGHT_{DA}$: The natural log of the weight of daily attraction *DA*. $Dum_{OS\,S^{sf}}$: Dummy variable. 1 if *sf* = 1 (OS); 0 otherwise. $Dum_{LF\,S^{sf}}$: Dummy variable. 1 if *sf* = 2 (LF); 0 otherwise. $Dum_{FC\,S^{sf}}$: Dummy variable. 1 if *sf* = 3 (FC); 0 otherwise. β_{TIME}^{OS}, β_{TIME}^{LF}, β_{TIME}^{FC}, β_{FP}^{OS}, β_{FP}^{LF}, β_{FP}^{FC}, β_{Const}^{LF}, β_{Const}^{FC}, β_{Weight}^{LF}, β_{Weight}^{FC}: Parameters to be estimated.

The continuous variables are log-transformed as it significantly improves the significance of the variables and the fit of the models. We do not directly include the costs that are incurred upon the formulation of the receiver-supplier relationship as the explanatory variables. However, the deterministic component of the utility shown above could be regarded as consisting of the terms that represent transportation cost, which depends on the "impedance" between a supplier and a receiver, the scale of a supplier, and the relative purchasing price of commodity. We assume the followings: (i) the average transportation cost increases monotonically with travel time and the weight of goods; (ii) the scale is represented by freight production (i.e. total weight of outbound shipments); (iii) for a certain commodity type, the average price of the commodity monotonically

increases with the weight of goods and differs by the function type of suppliers. For example, (2) can be rewritten for LF^S as:

$$V_{DA,S^{sf}=2} = \ln\left(TIME_{DA,S^{sf}=2}^{\beta_{TIME}^{LF}} \cdot WEIGHT_{DA}^{\beta_{Weight_1}^{LF}}\right) + \beta_{FP}^{LF} \cdot lnFP_{S^{sf}=2} + \ln\left(\tilde{\beta}_{Const}^{LF} \cdot WEIGHT_{DA}^{\beta_{Weight_2}^{LF}}\right) \quad (3)$$

where

$$\beta_{Weight}^{LF} = \beta_{Weight_1}^{LF} + \beta_{Weight_2}^{LF}$$

and;

$$\beta_{Const}^{LF} = ln\left(\tilde{\beta}_{Const}^{LF}\right)$$

The terms in (3) represent the effects of transportation cost, the scale of a supplier, and the relative purchasing price (against OS^S), respectively. Note that $\beta_{Weight_1}^{LF}$ and $\beta_{Weight_2}^{LF}$, i.e. the coefficients relevant to the interaction and main effects of the weight of DA, cannot be estimated separately. The heterogeneity within the suppliers that are the same in terms of the three explanatory variables (Table 3) is captured as a part of the random component.

The identification of the parameters to be estimated within $M_{DA,S^{sf}}$ is required to obtain definite estimations. We followed the procedure proposed by Walker et al. (2007), ensuring that the Order Condition, the Rank Condition, and the Equality Condition are satisfied, and determined the two nests, FC and UPS, for which the covariance, σ^{FC} and σ^{UPS}, are fixed as 0. Thus, $M_{DA,S^{sf}}$ is given by the following equation:

$$M_{DA,S^{sf}} = \begin{cases} \sigma^{OS}\eta_{DA}^{OS} + \sigma^{DWS} \cdot \eta_{DA}^{DWS} & if\ sf = 1 \\ \sigma^{LF}\eta_{DA}^{LF} + \sigma^{DWS} \cdot \eta_{DA}^{DWS} & if\ sf = 2 \\ 0 & if\ sf = 3 \end{cases} \quad (4)$$

where $\eta_{DA}^{OS}, \eta_{DA}^{LF}, \eta_{DA}^{DWS}$: Random terms following the normal distribution with zero mean and one standard deviation. $\sigma^{OS}, \sigma^{LF}, \sigma^{DWS}$: Parameters to be estimated.

Given some values for $\eta_{DA}^{OS}, \eta_{DA}^{LF}$ and η_{DA}^{DWS}, the conditional probability for $S^{sf\prime}$ to be chosen by a receiver for DA is:

$$P_{DA}\left(S^{sf\prime}|\eta_{DA}^{OS},\eta_{DA}^{LF},\eta_{DA}^{DWS}\right) =$$

$$\exp\left(V_{DA,S^{sf\prime}} + M_{DA,S^{sf\prime}}\left(\eta_{DA}^{OS},\eta_{DA}^{LF},\eta_{DA}^{DWS}\right)\right) \Bigg/ \sum_{S^{sf}} \exp\left(V_{DA,S^{sf}} + M_{DA,S^{sf}}\left(\eta_{DA}^{OS},\eta_{DA}^{LF},\eta_{DA}^{DWS}\right)\right) \quad (5)$$

Unconditional probability is given by the following equation, using the multivariate density function of ηs, $f()$:

$$P_{DA}\left(S^{sf\prime}\right) = \int_{\eta_{DA}^{OS}} \int_{\eta_{DA}^{LF}} \int_{\eta_{DA}^{DWS}} \left(P_{DA}(S^{sf\prime}|\eta_{DA}^{OS},\eta_{DA}^{LF},\eta_{DA}^{DWS}) \times f\left(\eta_{DA}^{OS},\eta_{DA}^{LF},\eta_{DA}^{DWS}\right)\right) d\eta_{DA}^{DWS} d\eta_{DA}^{LF} d\eta_{DA}^{OS} \quad (6)$$

Estimation

The computational burden becomes an issue in estimating an error components logit mixture model with a large choice set. For example, we consider 13,152 potential suppliers that

handle household/light manufacturing goods as the alternatives. Different approaches to select the choice set have been tested and discussed in the past research (Lemp and Kockelman 2012; Guevara and Ben-Akiva 2013). Guevara and Ben-Akiva compare three methods for sampling alternatives, i.e. population shares method, 1_0 method and Naïve method, and show that Naïve method, i.e. the random selection from all of the potential alternatives, is suitable and practical for logit mixture models. Furthermore, the draws of the random terms, ηs, must be specified. After testing various combinations of choice set size and the size of Halton draws, we decided to use 50 Naïve choices of alternatives with 1000 Halton draws as the estimated parameters are stable under such setting. We use PythonBiogeme (Bierlaire 2016) for the model estimation based on the maximum simulated likelihood method.

Estimation results and discussion

Models for household/light manufacturing goods

Table 4 shows the estimation results of five models for different receiver types. The adjusted Rho-square is the highest for CEF^R, 0.651, and the lowest for LF^R, 0.225. Most explanatory variables are significant for a 95% confidence interval, with the exception of Dum_{FC} in FC^R model, ln FP_{LF} in CEF^R model, and Dum_{LF} and ln $Weight \times Dum_{FC}$ in RF^R model. We first focus on the covariance parameters, as we will later discuss the effects of variables based on elasticities.

Compared with the covariance parameters in the other models, $Sigma_{DWS}$ in the models for CEF^R and RF^R are very high, 6.03 and 14.3, respectively. This indicates there are two uncorrelated categories identified for supplier function, i.e. non-factory and factory, while the suppliers in OS^S and LF^S (i.e. within non-factory category) are highly correlated. What is common between the two receiver types is that they are end-consumers of goods, not intermediate facilities or producers. On the other hand, the models for LF^R and FC^R shows the significant $Sigma_{OS}$ and $Sigma_{LF}$ (2.46 and 1.13, and 0.83 and 0.91, respectively) but not $Sigma_{DWS}$. This indicates that the correlations exist mainly within each supplier function category for these facility groups. Lastly, all three covariance parameters are significant in the model for OSR^R, corroborating the complex correlation structure within and across supplier function types. The estimated models for other commodity types are shown in the "Appendix".

Elasticities

To compare the magnitude of the effects of the variables, we compute elasticity effects for the pairs of receiver and supplier function types. The average elasticities are defined by the following equations:

For travel time/freight production:

$$\frac{1}{N_{S^{sf}}} \frac{1}{N_{DA}} \sum_{DA} \sum_{S^{sf}} \left(\frac{\partial P_{DA,S^{sf}} / P_{DA,S^{sf}}}{\partial x_{DA,S^{sf}} / x_{DA,S^{sf}}} \right) \tag{7}$$

For weight:

Table 4 Estimated error components logit mixture models (household/light manufacturing goods)

| | Model by receiver function | | | | | | | | | |
| | Office/store/restaurant (OSR^R) | | Logistics facility (LF^R) | | Factory (FC^R) | | Cultural and educational Facility (CEF^R) | | Residential facility (RF^R) | |
	Coef.	t-value	Coef.	t-value	Coef.	t-value	Coef.	t-value	Coef.	t-value
Explanatory variable										
ln Time$_{OS}$	−1.94	−163.5	−1.49	−44.1	−1.52	−75.8	−3.50	−161.6	−3.20	−87.6
ln Time$_{LF}$	−1.70	−149.4	−1.16	−62.4	−2.17	−94.7	−2.61	−82.9	−2.81	−97.9
ln Time$_{FC}$	−2.07	−172.1	−1.19	−58.4	−1.62	−122.5	−2.43	−60.7	−2.48	−62.0
ln FP$_{OS}$	0.408	128.1	0.418	58.2	0.452	78.6	0.427	87.2	0.575	61.3
ln FP$_{LF}$	0.232	72.0	0.281	47.7	0.399	55.9	0.001	0.1	0.176	24.7
ln FP$_{FC}$	0.391	108.2	0.302	49.8	0.416	99.0	0.584	51.6	0.625	53.9
Dum$_{OC}$	0	–	0	–	0	–	0	–	0	–
Dum$_{LF}$	−2.27	−19.6	−1.46	−6.2	3.63	20.5	−7.06	−30.6	−0.386	−1.6
Dum$_{FC}$	−1.35	−11.3	−1.45	−5.9	0.137	0.9	−27.4	−30.1	−15.2	−15.4
ln Weight×Dum$_{OS}$	0	–	0	–	0	–	0	–	0	–
ln Weight×Dum$_{LF}$	0.769	52.5	0.374	15.3	0.218	17.0	1.050	47.4	0.429	22.1
ln Weight×Dum$_{FC}$	0.679	45.5	0.265	10.1	0.283	22.9	3.670	25.1	−0.121	−1.6
Covariance parameter										
Sigma$_{OS}$	2.15	35.4	2.46	11.2	0.831	5.5	0.003	0.0	0.805	6.8
Sigma$_{LF}$	1.46	28.2	1.13	7.1	0.911	9.4	0.009	0.1	0.000	0.0
Sigma$_{DWS}$	1.07	14.9	0.0713	0.1	0.006	0.0	6.03	22.3	14.3	13.2
L(0)	−252,033		−60,798		−99,516		−102,539		−56,308	
L($\hat{\beta}$)	−171,708		−47,122		−64,836		−35,750		−28,013	
ρ^2	0.319		0.225		0.348		0.651		0.502	
$\bar{\rho}^2$	0.319		0.225		0.348		0.651		0.502	
Sample size:	14,172		3701		6230		5323		2575	

$$\frac{1}{N_{DA}} \sum_{DA} \frac{\partial \left(\sum_{S^{sf}} P_{DA,S^{sf}} \right) \Big/ \sum_{S^{sf}} P_{DA,S^{sf}}}{\partial x_{DA} \Big/ x_{DA}} \qquad (8)$$

where $P_{DA,S^{sf}}$: Probability for a supplier S^{sf} to be selected for a daily attraction DA. $x_{DA,S^{sf}}$, x_{DA}: An explanatory variable (Travel Time, Freight Production, or Weight) without log transformation. $N_{S^{sf}}$: Number of alternatives (potential suppliers). N_{DA}: Number of observations (i.e. DA).

We compute the elasticities with respect to travel time/freight production for all receiver and supplier pairs and then calculate their average. As for weight, a DA specific variable, we first compute the elasticities of the probability of selecting a supplier function group for each DA and then calculate the average. For taking the error structures into account, we use Monte Carlo method. The random values are generated for η_{DA}^{OS}, η_{DA}^{LF}, η_{DA}^{DWS} and the average elasticities (defined above) are computed 1,000 times. The means of the simulated average elasticities are calculated.

Household/light manufacturing goods

The results are shown in Table 5. Travel time has a very strong effect on supplier choice, regardless supplier–receiver function pair. Especially, the effect is strong for CEF^R and RF^R, which indicates that suppliers that are closely located are highly likely to be selected for the needs of end-consumers. On the other hand, the effects are relatively weak for LF^R, with the elasticities ranging between -1.47 and -1.10. The effect is the smallest when the supplier is also a logistics facility (LF^S), which is due to the fact that logistics facilities often function as intermediate points for long logistics chains. Another interesting finding is that, for FC^R, the travel time from LF^S is more important than from the other supplier types (i.e. OS^S and FC^S), indicating logistics facilities should be located close to factories (FC^R) for being selected as suppliers.

The scale of production is also a contributing factor to be selected as a supplier. Production has no effect only for one receiver-supplier functions pair, CEF^R–LF^S. In general, the elasticities associated with LF^S are small. For the shipments from intermediate facilities, the volume of goods handled has only limited effects.

Lastly, the weight of DA relates to the decisions on supplier function. As naturally expected, a larger demand is less likely to be supplied by OS^S. The choice of LF^S or FC^S for a heavier demand depends on the receiver function; as for FC^R and CEF^R, a heavier

Table 5 Average elasticity effects of variables (household/light manufacturing goods)

Receiver	Travel time				Freight production				Weight		
	Supplier										
	OS^S	LF^S	FC^S	ALL^S	OS^S	LF^S	FC^S	ALL^S	OS^S	LF^S	FC^S
OSR^R	-1.90	-1.64	-2.04	-1.91	0.40	0.22	0.39	0.36	-1.78	0.79	0.49
LF^R	-1.47	-1.10	-1.17	-1.28	0.41	0.27	0.30	0.34	-1.54	0.54	-0.07
FC^R	-1.50	-2.10	-1.58	-1.66	0.45	0.39	0.41	0.42	-1.04	0.03	0.34
CEF^R	-3.38	-2.57	-2.41	-2.83	0.41	0.00	0.58	0.40	-3.25	-0.06	7.88
RF^R	-3.13	-2.70	-2.45	-2.77	0.56	0.17	0.62	0.51	-0.57	0.61	-0.90

demand is more likely to be supplied by a FC^S, while as for OSR^R, LF^R and RF^R, a LF^S is more likely to handle such demand.

Commodities other than household/light manufacturing goods

The average elasticities of other commodities are computed as shown in Table 6. Some features are common across commodity types. For example, excluding *mixed goods and parcel*, the effects of travel time tend to be low for LF^R (between -1.61 and -0.93 for LF^R-ALL^S), while high for RF^R (between -3.37 and -1.60). The travel time is least influential when the LF^R receive from FC^S (between -1.06 and -0.66). There are also some differences on the elasticity effect of travel time by commodity type; by receiver type, the commodity type that shows the smallest absolute elasticity effect (with ALL^S) is different. On the other hand, the travel time is critical for *mixed goods and parcel*, having much larger absolute elasticity effects than other commodities in many cases.

As for freight production, like household/light manufacturing goods, the elasticity effects tend to be smaller when the suppliers are LF^S, although in some cases, such as FC^R–LF^S and LF^R–LF^S pairs for *food products* and *mixed goods and parcel*, the effects are large. For *agricultural product* received by OSR^R, LF^R, FC^R, and CEF^R, the freight production is most sensitive when the suppliers are OS^S, showing the elasticity effects of 0.50, 0.51, 0.48 and 0.58, respectively. For *wood and paper products*, the freight production has large effects for the pairs of OSR^R–FC^S (0.53), LF^R–OS^S (0.67), LF^R–FC^S (0.52), and FC^R–FC^S (0.44).

Lastly, a smaller DA is more likely to be catered by OS^S and a larger DA by LF^S. There are some cases where this characteristic does not apply; for instance, when a OSR^R receives *minerals, ore, stone, cement, ceramics or glass*, a larger demand is less likely to be handled by LF^S.

Observed versus estimated OD flows

We compare the estimated OD flows using the estimated models, against the observed OD flows. The purpose of this exercise is to test the reproducibility of the model at the OD level. Note that the process of comparing the simulation results against the data we used for the model estimation is different from the process of model validation, which requires a validation data set that is not used for the model estimation. For computing the OD flows, the TMA is divided into 18 zones. The OD flow is compared for each of 324 (18×18 zones) OD pairs. The Monte Carlo method is used for the OD flow estimation; the simulation for supplier choices is repeated 20 times and the averages are computed. We calculate R^2 as the measure of the reproducibility; R^2 is defined as:

$$R^2 = 1 - \sum_{i,j} \left(y_{i,j} - \widehat{y_{i,j}} \right)^2 \Big/ \sum_{i,j} \left(y_{i,j} - \bar{y} \right)^2 \tag{9}$$

where $y_{i,j}$: Observed number of shipments for an OD pair i, j. $\widehat{y_{i,j}}$: Estimated number of shipments for an OD pair i, j. $\bar{y}$: Average observed number of shipments for all OD pairs.

Figures 3 and 4 show the comparison of each OD flow for all shipments and for the shipments from each receiver function type. Figure 3 considers the shipments of household/light manufacturing goods only and Fig. 4 covers all commodity types.

Table 6 Average elasticity effects of variables (other than household/light manufacturing goods)

	Travel time				Freight production				Weight		
	Supplier										
	OS^S	LF^S	FC^S	ALL^S	OS^S	LF^S	FC^S	ALL^S	OS^S	LF^S	FC^S
OSR^R											
Agricultural products	−2.33	−2.33	−2.47	−2.37	0.50	0.07	0.38	0.34	−0.22	−0.37	0.90
Food products	−2.78	−2.43	−2.04	−2.34	0.27	0.30	0.31	0.30	−5.37	1.56	−3.13
Wood and paper products	−2.24	−2.33	−1.43	−1.88	0.26	0.21	0.53	0.38	0.76	2.79	−6.10
Minerals, ore, stone, cement, ceramics or glass	−1.16	−1.34	−1.56	−1.40	0.37	0.15	0.08	0.18	−1.46	−5.08	4.99
Metals or articles of metal	−2.41	−1.70	−1.02	−1.44	0.32	0.05	0.09	0.14	1.77	1.20	−8.10
Machinery, appliances, and mechanical parts	−1.64	−2.24	−1.29	−1.54	0.27	0.22	0.09	0.17	−0.76	0.43	−0.46
Chemicals, rubber or plastics	−1.66	−1.43	−0.79	−1.18	–	0.13	0.53	0.29	−1.05	0.30	1.08
Mixed goods and parcel	−1.75	−2.96	−3.44	−2.61	0.03	0.40	–	0.25	−8.07	0.26	−2.19
LF^R											
Agricultural products	−0.68	−2.34	−0.84	−1.22	0.51	0.23	0.25	0.36	4.53	−0.64	3.00
Food products	−2.27	−1.19	−0.82	−1.28	0.37	0.52	0.32	0.40	−2.67	0.68	−0.83
Wood and paper products	−2.11	−2.13	−1.04	−1.61	0.67	0.29	0.52	0.51	−1.83	0.83	−0.63
Minerals, ore, stone, cement, ceramics or glass	−1.55	−2.25	−1.06	−1.41	0.26	0.20	0.21	0.22	−2.60	−1.19	1.66
Metals or articles of metal	−1.44	−1.46	−0.66	−0.93	0.53	0.24	0.35	0.38	−1.53	0.22	0.20
Machinery, appliances, and mechanical parts	−1.02	−0.92	−0.99	−0.99	0.37	0.15	0.32	0.32	−0.73	0.67	−0.36
Chemicals, rubber or plastics	−1.48	−1.88	−0.84	−1.25	0.03	0.09	0.27	0.16	−1.57	−0.39	0.90
Mixed goods and parcel	−3.09	−2.32	−2.11	−2.55	0.36	0.53	–	0.44	−4.32	0.15	−1.52
FC^R											
Agricultural products	−1.73	−1.60	−1.33	−1.59	0.48	0.29	0.18	0.35	−0.70	0.53	−1.27
Food products	−1.23	−1.60	−1.64	−1.53	0.55	0.75	0.40	0.54	1.85	1.26	−1.03
Wood and paper products	−2.08	−2.17	−1.21	−1.68	0.29	0.09	0.44	0.32	−6.26	−3.17	1.44
Minerals, ore, stone, cement, ceramics or glass	−1.89	−2.18	−1.28	−1.61	–	0.44	0.20	0.17	−2.13	0.11	0.32
Metals or articles of metal	−2.54	−1.56	−1.34	−1.65	0.30	0.17	0.29	0.28	−1.09	1.29	−0.13

Table 6 (continued)

	Travel time				Freight production				Weight		
	Supplier										
	OS^S	LF^S	FC^S	ALL^S	OS^S	LF^S	FC^S	ALL^S	OS^S	LF^S	FC^S
Machinery, appliances, and mechanical parts	−1.86	−1.60	−1.45	−1.61	0.24	0.30	0.17	0.21	−3.24	0.69	0.32
Chemicals, rubber or plastics	−2.32	−2.17	−1.06	−1.66	0.23	0.28	0.27	0.26	−4.66	1.40	1.07
Mixed goods and parcel	−5.76	−5.36	−4.65	−5.44	0.45	0.78	0.33	0.64	2.03	−0.29	−0.96
CEF^R											
Agricultural products	−1.38	−0.11	−3.35	−1.55	0.58	–	–	0.26	3.60	3.81	−3.06
Food products	−2.29	−1.43	−3.16	−2.40	0.38	0.09	0.45	0.32	2.34	−2.60	4.44
Wood and paper products	−1.12	−1.57	−1.79	−1.54	0.10	–	0.02	0.04	−1.72	2.16	1.45
Metals or articles of metal	−1.05	−2.28	−2.04	−1.84	–	–	–	–	38.30	26.24	−11.26
Machinery, appliances, and mechanical parts	−2.25	−2.04	−0.74	−1.44	0.19	–	–	0.07	−0.05	−0.14	1.94
Chemicals, rubber or plastics	−2.12	−3.33	−2.98	−2.79	0.27	0.10	0.51	0.36	−1.06	2.50	0.99
RF^R											
Agricultural products	–	−3.05	–	−1.60	–	0.05	0.28	0.16	–	1.01	−15.15
Food products	−4.07	−2.67	−3.31	−3.30	–	0.16	0.46	0.25	−3.09	2.23	−3.15
Wood and paper products	−2.73	−2.32	−2.20	−2.39	0.50	0.08	0.21	0.27	−0.20	0.15	0.20
Minerals, ore, stone, cement, ceramics or glass	−2.23	−1.88	−2.54	−2.34	0.21	0.04	0.04	0.10	−1.92	0.45	0.17
Metals or articles of metal	−2.65	−1.93	−1.65	−1.92	0.21	0.18	0.35	0.30	−1.80	3.56	−1.11
Machinery, appliances, and mechanical parts	−3.63	−3.64	−1.60	−2.58	–	–	–	–	−0.63	0.17	0.63
Chemicals, rubber or plastics	−3.84	−3.31	−3.12	−3.37	–	0.42	0.52	0.35	−5.47	1.39	−2.31
Mixed goods and parcel	−0.35	−4.33	–	−2.96	–	0.19	–	0.12	−70.70	0.00	–

[a]For some receiver type and commodity type pairs, the models were not estimated as the numbers of shipments for those pairs are negligible. Also, for the same reason, only two supplier types are considered in some estimated models. [b]Average elasticity effects are not available for some cases. Also, some estimated coefficients which show the opposite signs from the expected ones were fixed to zero and the models were re-estimated
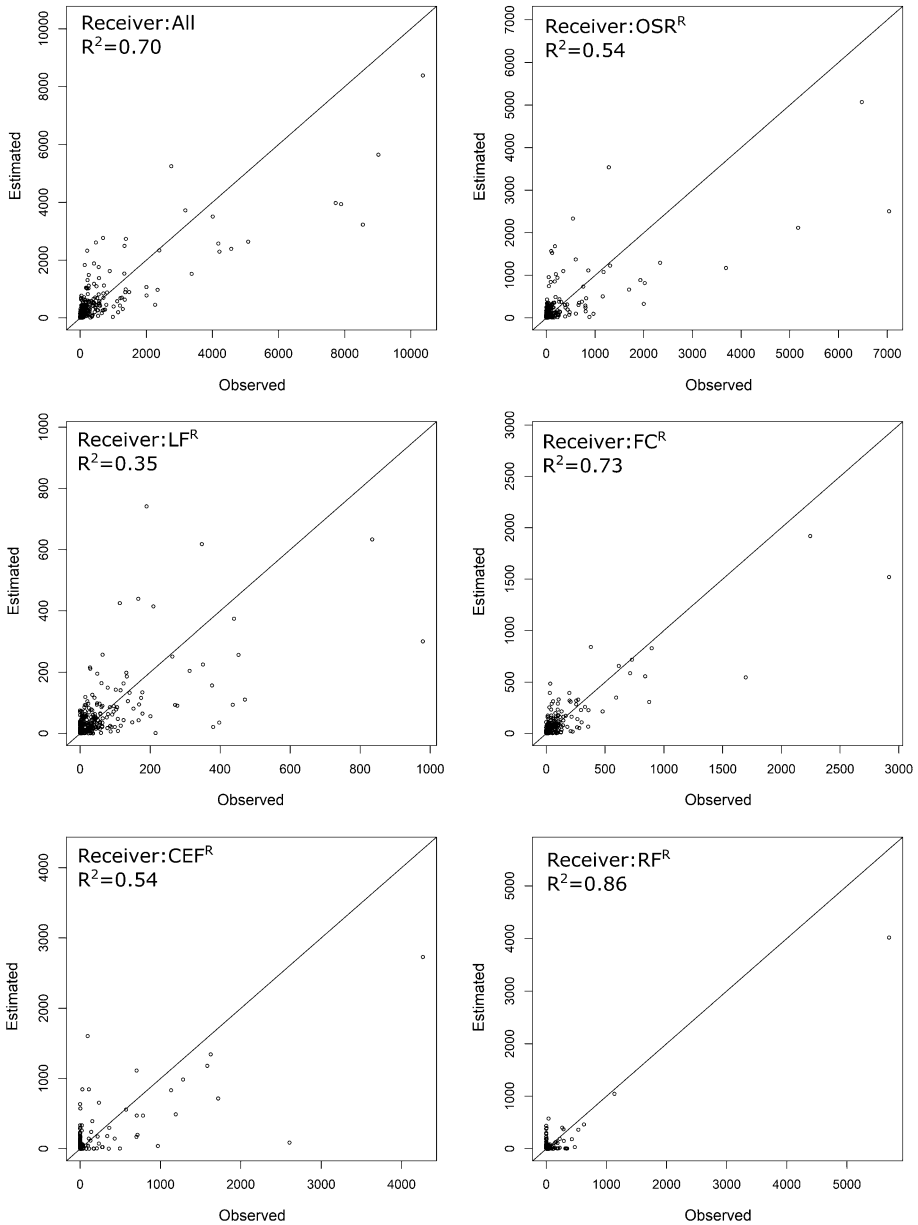
Fig. 3 Comparison of OD flows: observed versus. estimated (household/light manufacturing goods)

Figure 3 shows the high overall reproducibility of the model for household/light manufacturing goods with R^2 of 0.70 (top-left). On the other hand, the reproducibility for the shipments that LF^{R}s receive is low, showing R^2 of 0.35 (middle-left). Also, the shipments to OSR^{R}s show relatively low R^2 (0.54, top-right). Both LF^{R}s and OSR^{R}s, especially LF^{R}s, are the potential intermediate locations of logistics chains. The result highlights the
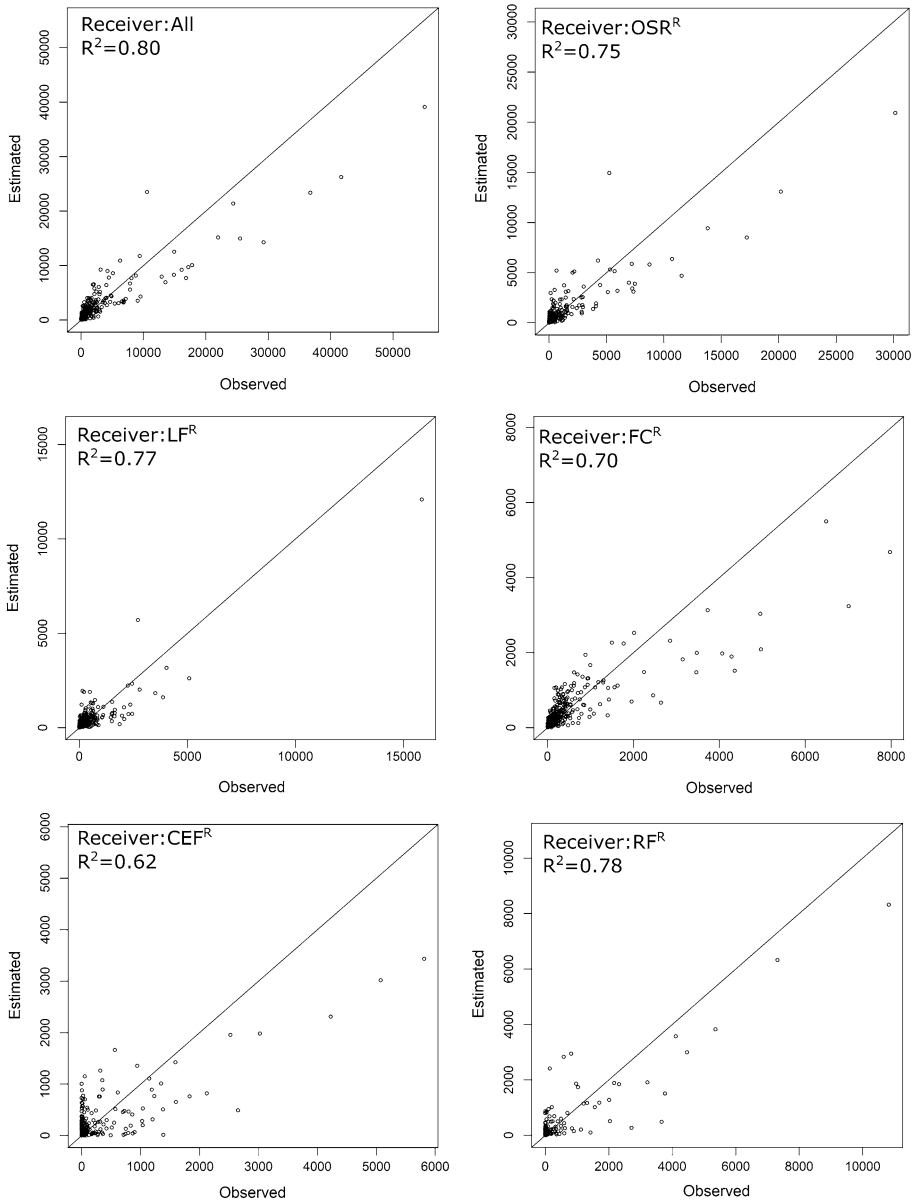
Fig. 4 Comparison of OD flows: observed versus estimated (all commodity types)

difficulty of estimating the flows of the shipments sent to those facilities for household/ light manufacturing goods. In contrast, the shipments to FC^Rs and RF^Rs, which are likely to be the consumers of goods received (in case of FC^Rs, goods are "consumed" to produce other goods), show relatively high R^2s (0.73, middle-right, and 0.86, bottom-right, respectively), indicating this type of the shipment flows can be well-captured with the proposed model framework. The reason for the relatively low R^2 for the shipments to CEF^Rs, which

is also 0.54 (bottom-left), is unclear. As the underestimation for the two OD pairs which are the largest in terms of the observed shipment counts greatly lowers the R^2, there may be some large-scale suppliers for CEF^Rs (e.g. schools) and the model may not properly capture the decisions of choosing them. Figure 4 shows the even higher reproducibility of the model when all commodity types are considered (R^2 is 0.80; top-left). The models for LF^R also achieve much higher R^2 (0.77, middle-left) than the case only for household/light manufacturing goods; the errors for different commodity types are canceled out when all commodities are aggregated.

Additionally, we calculate R^2 in terms of weight, instead of shipment counts, for the shipments of all commodity and receiver types. The computed weight-based R^2 is 0.46, which is lower than the R^2 based on the number of shipments. The weight distribution of the intra-metropolitan shipments is highly skewed. The analysis using the 2013 TMFS indicates that about 5% of the heaviest shipments account for 80% of the total shipment weight. Such high skewness makes it challenging to reproduce the observed OD flows in terms of weight using a disaggregate model. It is worth noting that, if the data is segmented into small groups (e.g. the groups by OD pair), the highly skewed distribution of shipment weight population may impair the representativeness of the sample sets for the groups, even with a large-scale survey data. For example, Sakai et al. (2015b) show that the highly skewed distribution of establishment-level freight generations causes a large error in the estimation of the average freight generation rate based on sample data.

Conclusions

Despite the fact that the type of distribution channel is relevant to the decisions on various aspects in logistics operations, the selection of distribution channel was not considered in most of the existing urban commodity flow models. Leveraging establishment and shipment records, which become increasingly more available for cities around the world, we propose a framework of the error component logit mixture model that considers the choices at two levels, the supplier function choice and the supplier choice. Given the function of the receiver, the supplier function choice is equivalent to the selection of distribution channel. We tested the model framework using the data obtained from the 2013 TMFS, a large-scale establishment survey.

The estimated model allows us to analyze the relationships among the distribution channels. The covariance parameters indicate that the correlations among the distribution channels are different by the function type of a receiver. For example, non-factory suppliers are highly correlated when a receiver is an end-consumer. Furthermore, the estimated model provides valuable insights about the relationship between the choice of the supplier/supplier function and the variables, i.e. distance (travel time), freight production, and the weight of demand by the receiver function and the commodity type. For instance, regarding household/light manufacturing goods, when a logistics facility supplies goods to a factory, the distance to the demand is crucial for an establishment to be a competitive supplier. The model also captures the fact that a large-size demand is more likely met by a supply from logistics facilities and factories. It also indicates that travel time is valued more highly for *mixed goods and parcel* than the other commodity types. Lastly, R^2 obtained from the comparison of the simulated versus observed shipment counts motivates us to further use the proposed modeling framework in urban freight modeling systems, although it requires

caution to interpret the estimated weight-based flows if the demand weight distribution is highly skewed.

The 2013 TMFS data we use has some limitation in differentiating the establishments that have only office functions from those serving both as office and distribution facilities. The business function information is valuable for urban freight modeling and we encourage to collect accurate records for it in future surveys. The receiver/supplier functions are also relevant to modeling other aspects of urban freight including carrier selection and goods vehicle tours. The development of new modeling/simulation approaches to deal with such elements is a future work. The integration of the simulations for the different decision-making processes in urban freight is especially important because it allows the feedbacks on the key decision parameters, especially those related to costs. For example, the accurate estimation of travel time through traffic simulations is important for a realistic simulation of *mixed goods and parcel* flows and a policy test focusing on this commodity type.

Author contributions TS Research Design, Literature Review, Model Specification, Analysis, and Manuscript Writing. BKB Model Specification and Analysis. AA Model Specification, Manuscript Writing and Editing. T Hyodo: Data Preparation and Analysis. MBA Research Design, Model Specification and Analysis.

Compliance with ethical standards

Conflict of interest On behalf of all authors, the corresponding author states that there is no conflict of interest.

Appendix: Error component logit mixture models

Tables 7, 8, 9, 10 show the estimated models for commodity types other than household/light manufacturing goods.

Table 7 Error component logit mixture models (agricultural products/food products)

Explanatory variable[a]	Agricultural products					Food products				
	OSR^R	LF^R	FC^R	CEF^R	RF^R,b	OSR^R	LF^R	FC^R	CEF^R	RF^R
$\ln \text{Time}_{OS}$	−2.37 (−148.1)	−0.679 (−18.6)	−1.74 (−16.9)	−1.40 (−21.8)	—	−2.81 (−149.3)	−2.29 (−56.2)	−1.24 (−23.8)	−2.34 (−77.6)	−4.21 (−40.5)
$\ln \text{Time}_{LF}$	−2.40 (−189.9)	−2.49 (−192.9)	−1.67 (−36.9)	−0.111 (−1.1)	−3.16 (−92.2)	−2.53 (−322.5)	−1.23 (−99.6)	−1.63 (−58.5)	−1.49 (−88.8)	−2.76 (−65.9)
$\ln \text{Time}_{FC}$	−2.51 (−129.0)	−0.838 (−20.2)	−1.36 (−21.9)	−3.51 (−37.5)	—	−2.06 (−215.6)	−0.830 (−64.3)	−1.69 (−93.7)	−3.18 (−69.3)	−3.32 (−25.5)
$\ln \text{FP}_{OS}$	0.503 (110.9)	0.511 (38.2)	0.483 (15.1)	0.586 (25.7)	—	0.272 (98.9)	0.376 (47.3)	0.556 (31.8)	0.384 (57.2)	—
$\ln \text{FP}_{LF}$	0.069 (22.4)	0.246 (66.4)	0.305 (19.0)	—	0.051 (6.6)	0.312 (149)	0.541 (106.6)	0.759 (53.1)	0.093 (28.2)	0.165 (16.9)
$\ln \text{FP}_{FC}$	0.391 (63.2)	0.253 (19.5)	0.182 (9.3)	—	0.281 (12.6)	0.318 (118.4)	0.327 (88.6)	0.410 (74.7)	0.457 (41.8)	0.463 (13.9)
Dum_{LF}	4.66 (39.7)	24.6 (51.4)	1.57 (2.3)	−4.37 (−4.9)	—	−5.33 (−37.1)	−9.41 (−38.2)	1.19 (2.5)	3.65 (10.9)	−15.0 (−19)
Dum_{FC}	−0.042 (−0.3)	6.00 (14.2)	1.75 (2.3)	25.8 (31.3)	−13.5 (−23.8)	−6.05 (−39.3)	−8.49 (−34.9)	9.83 (22.3)	0.543 (1.3)	−16.8 (−14.8)
$\ln \text{Weight} \times \text{Dum}_{LF}$	−0.039 (−4.2)	−1.19 (−27.4)	0.168 (4.6)	0.056 (0.5)	—	1.57 (64.7)	0.600 (26.4)	−0.126 (−4.7)	−1.51 (−26.4)	1.23 (21.6)
$\ln \text{Weight} \times \text{Dum}_{FC}$	0.297 (19.0)	−0.352 (−16.4)	−0.079 (−2.0)	−1.78 (−14.0)	−4.01 (−19.2)	0.507 (17.5)	0.329 (13.6)	−0.619 (−20.3)	0.642 (8.3)	−0.014 (−0.1)
Covariance parameter										
Sigma_{OS}	0.923 (13.7)	0.000 (0.0)	0.453 (0.7)	0.000 (0.0)	0.000 (0.0)	3.95 (54.2)	2.69 (19.1)	1.32 (5.7)	4.37 (22.7)	0.487 (1.4)
Sigma_{LF}	0.000	1.77	0.056	0.001	2.29	0.829	0.000	0.051	3.39	0.000

Table 7 (continued)

	Agricultural products					Food products				
	OSR^R	LF^R	FC^R	CEF^R	RF^{R,b}	OSR^R	LF^R	FC^R	CEF^R	RF^R
	(0.0)	(12.1)	(0.0)	(0.0)	(11.4)	(18.0)	(0.0)	(0.3)	(23.3)	(0.0)
Sigma$_{DWS}$	2.390	0.037	0.000	0.000	0.000	3.52	1.32	1.17	1.31	6.57
	(35.5)	(0.1)	(0.0)	(0.0)	(0.0)	(53.8)	(13.7)	(7.5)	(6.9)	(10.7)
L (0)	−186,139	−136,771	−8620	−6792	−22,671	−508,188	−145,328	−48,978	−72,655	−19,958
L($\hat{\beta}$)	−117,201	−64,546	−5995	−3534	−10,907	−302,334	−107,906	−32,249	−49,094	−7300
ρ^2	0.370	0.528	0.305	0.480	0.519	0.405	0.258	0.342	0.324	0.634
$\bar{\rho}^2$	0.370	0.528	0.303	0.478	0.519	0.405	0.257	0.341	0.324	0.634
N.	6277	4358	494	522	1050	28,779	10,294	4028	4890	1164

t values are shown in the parentheses

[a]Variables were not considered if they show the opposite signs from the expected (i.e. lnTime must be negative and lnFP must be positive). [b]All alternatives are either LF^S or FC^S

Table 8 Error component logit mixture models (wood and paper products/minerals, ore, stone, cement, ceramics or glass)

	Wood and paper products					Minerals, ore, stone, cement, ceramics or glass[b]			
	OSR^R	LF^R	FC^R	CEF^R	RF^R	OSR^R	LF^R	FC^R	RF^R
Explanatory variable[a]									
ln $Time_{OS}$	−2.28	−2.14	−2.10	−1.18	−2.82	−1.19	−1.57	−1.91	−2.26
	(−100.8)	(−48.3)	(−45.5)	(−63.0)	(−34.4)	(−22.2)	(−20.2)	(−40.9)	(−14.4)
ln $Time_{LF}$	−2.41	−2.23	−2.20	−1.58	−2.41	−1.40	−2.36	−2.22	−1.97
	(−121.2)	(−81.0)	(−48.8)	(−26.6)	(−42.1)	(−21.4)	(−34.0)	(−37.1)	(−18.3)
ln $Time_{FC}$	−1.45	−1.05	−1.25	−1.80	−2.21	−1.58	−1.08	−1.31	−2.59
	(−70.6)	(−34.2)	(−73.8)	(−38.9)	(−25.6)	(−28.8)	(−27.6)	(−52.3)	(−27.8)
ln FP_{OS}	0.265	0.675	0.296	0.106	0.517	0.375	0.263	–	0.210
	(46.8)	(37.9)	(28.6)	(24.4)	(20.1)	(22.0)	(13.8)	–	(6.6)
ln FP_{LF}	0.215	0.308	0.089	–	0.086	0.159	0.211	0.448	0.037
	(41.3)	(36.4)	(8.1)	–	(6.0)	(8.8)	(11.0)	(16.6)	(1.8)
ln FP_{FC}	0.541	0.525	0.454	0.019	0.214	0.077	0.212	0.206	0.045
	(79.4)	(44.2)	(73.4)	(1.8)	(6.5)	(6.3)	(18.2)	(28.5)	(3.0)
Dum_{LF}	−0.478	2.33	−0.726	−4.76	0.693	7.33	6.11	−4.31	−0.081
	(−2.7)	(7.0)	(−1.7)	(−7.6)	(1.3)	(9.7)	(9.6)	(−8.6)	(−0.1)
Dum_{FC}	−6.05	−9.64	−10.0	−2.06	−3.96	−3.32	−6.59	−7.39	2.11
	(−25.2)	(−14.6)	(−20.8)	(−4.8)	(−5.0)	(−1.3)	(−9.7)	(−18.8)	(1.7)
ln Weight×Dum_{LF}	0.597	0.459	0.526	2.03	0.064	−0.487	0.194	0.310	0.349
	(43.6)	(15.5)	(12.6)	(15.9)	(1.6)	(−7.8)	(4.3)	(15.8)	(4.2)
ln Weight×Dum_{FC}	−2.02	0.208	1.31	1.66	0.074	0.868	0.586	0.339	0.308
	(−8.7)	(3.6)	(17.2)	(14.2)	(1.3)	(3.5)	(9.5)	(15.0)	(2.9)
Covariance parameter									
$Sigma_{OS}$	0.000	1.82	2.85	2.55	0.961	0.008	1.27	0.081	1.620
	(0.0)	(11.1)	(17.0)	(10.4)	(3.4)	(0.0)	(1.8)	(0.1)	(2.5)
$Sigma_{LF}$	0.000	0.215	2.40	0.000	0.000	1.46	0.000	0.004	0.096
	(0.0)	(0.3)	(11.8)	(0.0)	(0.0)	(4.0)	(0.0)	(0.0)	(0.1)
$Sigma_{DWS}$	13.1	6.86	5.31	0.000	0.000	4.68	2.73	1.58	5.430
	(10.3)	(9.8)	(17.5)	(0.0)	(0.0)	(3.9)	(7.4)	(6.5)	(5.3)
L (0)	−105,370	−38,928	−45,172	−32,676	−9561	−8320	−10,910	−19,256	−4401
L($\hat{\beta}$)	−70,078	−25,451	−32,203	−23,323	−4087	−6455	−7924	−14,462	−2758
ρ^2	0.335	0.346	0.287	0.286	0.573	0.224	0.274	0.249	0.373
$\bar{\rho}^2$	0.335	0.346	0.287	0.286	0.571	0.223	0.273	0.248	0.370
N.	5297	2087	2832	1056	601	590	781	1508	308

t values are shown in the parentheses

[a]Variables were not considered if they show the opposite signs from the expected (i.e. lnTime must be negative and lnFP must be positive). [b]The model for CEF^R were not estimated as such shipments are limited

Table 9 Error component logit mixture models (metals or articles of metal/machinery, appliances, and mechanical parts)

	Metals or articles of metal					Machinery, appliances, and mechanical parts				
	OSR^R	LF^R	FC^R	CEF^R	RF^R	OSR^R	LF^R	FC^R	CEF^R	RF^R
Explanatory variable[a]										
ln Time$_{OS}$	−2.50	−1.46	−2.58	−1.06	−2.73	−1.67	−1.04	−1.88	−2.33	−3.69
	(−59.0)	(−21.9)	(−90.2)	(−4.2)	(−29.5)	(−90.7)	(−45.6)	(−85.7)	(−87.0)	(−26.7)
ln Time$_{LF}$	−1.83	−1.59	−1.61	−2.30	−2.01	−2.44	−0.969	−1.67	−2.17	−4.05
	(−43.7)	(−40.1)	(−88.1)	(−5.4)	(−17.4)	(−104.3)	(−39.1)	(−81.1)	(−66.4)	(−32.7)
ln Time$_{FC}$	−1.03	−0.665	−1.37	−2.10	−1.67	−1.30	−1.00	−1.48	−0.742	−1.60
	(−26.7)	(−25.4)	(−155.4)	(−28.7)	(−28.2)	(−45.8)	(−43.1)	(−143.6)	(−6.5)	(−9.6)
ln FP$_{OS}$	0.335	0.536	0.304	–	0.213	0.273	0.379	0.238	0.200	–
	(33.7)	(29.3)	(50.8)		(9.0)	(50.7)	(52.8)	(48.4)	(28.2)	
ln FP$_{LF}$	0.049	0.263	0.176	–	0.190	0.240	0.162	0.315	–	–
	(6.1)	(23.8)	(28.5)		(6.2)	(43.7)	(22.5)	(52.7)		
ln FP$_{FC}$	0.092	0.353	0.292	–	0.351	0.092	0.327	0.170	–	–
	(10.5)	(45.6)	(106.5)		(17.2)	(14.7)	(46.9)	(61.6)		
Dum$_{LF}$	−1.59	5.72	−5.84	30.5	−10.4	5.10	0.079	−2.81	0.792	4.07
	(−3.8)	(11.2)	(−27.5)	(4.4)	(−9.7)	(30.9)	(0.3)	(−14.3)	(3.0)	(3.9)
Dum$_{FC}$	−19.0	−2.52	−6.95	134	−9.47	−8.06	−1.04	−2.45	−65.9	−21.0
	(−10.0)	(−5.3)	(−39.8)	(1.1)	(−13.1)	(−12.7)	(−4.6)	(−14.9)	(−1.6)	(−9.8)
ln Weight×Dum$_{LF}$	−0.135	0.303	0.464	−3.68	0.973	0.327	0.280	0.953	−0.054	0.368
	(−4.2)	(5.8)	(29.6)	(−3.9)	(14.6)	(25.7)	(24.4)	(24.3)	(−3.7)	(5.9)
ln Weight×Dum$_{FC}$	−2.34	0.300	0.187	−15.3	0.125	0.081	0.073	0.864	1.19	0.575
	(−7.3)	(5.9)	(12.9)	(−1.0)	(2.8)	(2.0)	(5.8)	(22.2)	(1.2)	(3.5)
Covariance parameter										
Sigma$_{OS}$	6.41	4.15	3.18	0.587	0.009	0.447	0.000	4.45	0.000	2.40
	(12.1)	(8.8)	(30.3)	(0.6)	(0.0)	(2.5)	(0.0)	(22.6)	(0.0)	(11.0)
Sigma$_{LF}$	7.09	2.21	0.000	1.12	0.000	0.106	0.017	0.518	0.061	0.071

Table 9 (continued)

	Metals or articles of metal					Machinery, appliances, and mechanical parts				
	OSR^R	LF^R	FC^R	CEF^R	RF^R	OSR^R	LF^R	FC^R	CEF^R	RF^R
	(10.4)	(11.9)	(0.0)	(0.7)	(0.0)	(0.4)	(0.1)	(4.6)	(0.1)	(0.2)
Sigma$_{DWS}$	24.3	0.000	0.000	47.9	0.159	6.870	1.230	0.910	29.0	4.05
	(7.0)	(0.0)	(0.0)	(0.9)	(0.2)	(12.9)	(7.5)	(9.1)	(1.3)	(4.5)
L (0)	−28,141	−27,230	−142,554	−2443	−4520	−77,391	−37,533	−128,925	−30,123	−7533
L($\hat{\beta}$)	−19,513	−20,309	−108,038	−1830	−2719	−47,407	−30,103	−98,303	−16,267	−2050
ρ^2	0.307	0.254	0.242	0.251	0.398	0.387	0.198	0.238	0.460	0.728
$\bar{\rho}^2$	0.306	0.254	0.242	0.247	0.396	0.387	0.198	0.237	0.460	0.727
N.	1809	1960	9871	135	304	4064	2812	9246	888	364

t values are shown in the parentheses

aVariables were not considered if they show the opposite signs from the expected (i.e. lnTime must be negative and lnFP must be positive)

Table 10 Error component logit mixture models (chemicals, rubber or plastics/mixed goods and parcel)

	Chemicals, rubber or plastics					Mixed goods and parcel[b]			
	OSR^R	LF^R	FC^R	CEF^R	RF^R	OSR^R	LF^R	FC^R	$RF^{R\ c}$
Explanatory variable[a]									
ln Time$_{OS}$	−1.71	−1.50	−2.36	−2.21	−3.87	−1.76	−3.10	−5.80	−0.355
	(−58.1)	(−54.2)	(−96.1)	(−52.7)	(−96.2)	(−23.4)	(−15.8)	(−21.4)	(−1.1)
ln Time$_{LF}$	−1.47	−1.97	−2.24	−3.36	−3.52	−3.05	−2.40	−5.52	−4.47
	(−45.3)	(−70.6)	(−88.6)	(−20.4)	(−169.4)	(−55.2)	(−91.8)	(−48.1)	(−45.3)
ln Time$_{FC}$	−0.801	−0.857	−1.08	−3.02	−3.15	−3.45	−2.11	−4.66	–
	(−23.1)	(−38.3)	(−76.3)	(−43.1)	(−144.8)	(−5.3)	(−7.2)	(−14.2)	–
ln FP$_{OS}$	–	0.034	0.233	0.276	–	0.028	0.363	0.453	–
	–	(3.8)	(32.8)	(21.3)	–	(1.8)	(12.5)	(9.3)	–
ln FP$_{LF}$	0.137	0.097	0.285	0.104	0.444	0.409	0.550	0.801	0.191
	(14.3)	(14.9)	(37.6)	(3.1)	(69.0)	(18.8)	(49.2)	(25.1)	(6.9)
ln FP$_{FC}$	0.533	0.273	0.274	0.519	0.530	–	–	0.333	–
	(48.4)	(41.5)	(69.4)	(28.9)	(80.0)	–	–	(3.9)	–
Dum$_{LF}$	−4.65	3.08	−4.20	1.24	−7.89	0.391	−7.30	−3.05	4.73
	(−11.0)	(12.3)	(−15.4)	(1.3)	(−26.9)	(0.7)	(−6.8)	(−2.0)	(0.0)
Dum$_{FC}$	−15.1	−8.53	−11.7	−0.932	−10.3	4.24	−5.24	−3.69	–
	(−29.1)	(−26.6)	(−47.0)	(−1.8)	(−36.2)	(1.3)	(−2.2)	(−1.4)	–
ln Weight×Dum$_{LF}$	0.312	0.244	1.29	0.993	2.09	1.54	0.735	−0.44	14.4
	(8.8)	(12.2)	(25.4)	(8.0)	(43.8)	(21.4)	(7.9)	(−4.7)	(0.0)
ln Weight×Dum$_{FC}$	0.493	0.510	1.22	0.571	0.962	1.09	0.461	−0.567	–
	(10.7)	(17.0)	(24.3)	(7.9)	(21.4)	(4.3)	(3.1)	(−3.9)	–
Covariance parameter									
Sigma$_{OS}$	0.000	0.720	6.28	3.26	3.97	0.000	2.64	1.99	0.000
	(0.0)	(2.7)	(26.6)	(11.5)	(35.5)	(0.0)	(5.1)	(7.1)	(0.0)
Sigma$_{LF}$	2.56	0.018	0.293	2.60	0.000	0.000	0.000	0.000	0.001
	(7.5)	(0.1)	(0.5)	(5.4)	(0.0)	(0.0)	(0.0)	(0.0)	(0.0)
Sigma$_{DWS}$	2.69	1.31	2.37	0.000	3.14	2.210	0.000	0.000	0.000
	(8.1)	(7.2)	(16.4)	(0.0)	(48.0)	(2.0)	(0.0)	(0.0)	(0.0)
L (0)	−32,445	−43,800	−102,465	−13,686	−179,258	−7657	−18,932	−11,997	−8378
L($\hat{\beta}$)	−25,399	−32,683	−78,467	−8370	−71,354	−3736	−8940	−1792	−985
ρ^2	0.217	0.254	0.234	0.388	0.602	0.512	0.528	0.851	0.882
$\bar{\rho}^2$	0.217	0.254	0.234	0.387	0.602	0.510	0.527	0.850	0.882
N.	2340	3163	7753	1028	10,600	602	1471	726	966

t values are shown in the parentheses

[a]Variables were not considered if they show the opposite signs from the expected (i.e. lnTime must be negative and lnFP must be positive). [b] The model for CEF^R were not estimated as such shipments are limited. [c] All alternatives are either OS^S or LF^S

References

Abate, M., Vierth, I., Karlsson, R., de Jong, G., Baak, J.: A disaggregate stochastic freight transport model for Sweden. Transportation. (2018). https://doi.org/10.1007/s11116-018-9856-9

Adnan, M., Pereira, F. C., Azevedo, C.M.L., Basak, K., Lovric, M., Feliu, S.R., Yi, Z., Joseph, F., Christopher, Z. Ben-Akiva, M.: SimMobility: a multi-scale integrated agent-based simulation platform. In: 95th Annual Meeting of the Transportation Research Board, Washington, DC (2016)

Alho, A., Bhavathrathan, B.K., Stinson, M., Gopalakrishnan, R., Le, D.T., Ben-Akiva, M.: A multi-scale agent-based modelling framework for urban freight distribution. Trans. Res. Proc. **27**, 188–196 (2017)

Alho, A., e Silva, J.D.A.: Lisbon's establishment-based freight survey: revealing retail establishments' characteristics, goods ordering and delivery processes. Eur. Transp. Res. Rev. **7**(2), 16 (2015)

Ben-Akiva, M., de Jong, G.: The aggregate-disaggregate-aggregate (ADA) freight model system. In: Ben-Akiva, M., Meersman, H., Van de Voorde, E. (eds.) Freight Transport Modelling, pp. 69–90. Emerald Group Publishing Limited, Bingley (2013)

Bierlaire, M.: PythonBiogeme: a short introduction. Report TRANSP-OR 160706, Series on Biogeme. Transport and Mobility Laboratory, School of Architecture, Civil and Environmental Engineering, Ecole Polytechnique Fédérale de Lausanne, Switzerland (2016)

Boerkamps, J., van Binsbergen, A., Bovy, P.: Modeling behavioral aspects of urban freight movement in supply chains. Transp. Res. Rec. **1725**, 17–25 (2000)

Cheah, L., Zhao, F., Stinson, M., Lu, F., Ding-Mastera, J., Marzano, V., Ben-Akiva, M. Next-generation commodity flow survey: a pilot in Singapore. In: 10th International Conference on City Logistics, Phuket, Thailand (2017)

Chow, J.Y., Yang, C.H., Regan, A.C.: State-of-the art of freight forecast modeling: lessons learned and the road ahead. Transportation **37**(6), 1011–1030 (2010)

Comi, A., Delle Site, P., Filippi, F., Nuzzolo, A.: Urban freight transport demand modelling: a state of the art. Eur. Transp. (2012). http://hdl.handle.net/10077/6123

Dablanc, L., Rakotonarivo, D.: The impacts of logistics sprawl: how does the location of parcel transport terminals affect the energy efficiency of goods' movements in Paris and what can we do about it? Proc. Soc. Behav. Sci. **2**(3), 6087–6096 (2010)

Davydenko, I., Tavasszy, L.: Estimation of warehouse throughput in freight transport demand model for the Netherlands. Transp. Res. Rec. **2379**, 9–17 (2013)

De Jong, G., Vierth, I., Tavasszy, L., Ben-Akiva, M.: Recent developments in national and international freight transport models within Europe. Transportation **40**(2), 347–371 (2013)

Fischer, M., Outwater, M., Cheng, L., Ahanotu, D., Calix, R.: Innovative framework for modeling freight transportation in Los Angeles County, California. Transp. Res. Rec. **1906**, 105–112 (2005)

Friedrich, H.: Simulation of logistics in food retailing for freight transportation analysis. Karlsruher Institut für Technoligie (KIT), Karlsruhe (2010)

Gliebe, J., Smith, C., Wies, K., Heither, C., Doyle, J., Outwater, M.: An agent-based gaming approach to simulating the evolution of commodity flows. In: Urban Freight and Behavior Change (URBE) Conference, Rome (2015)

Guevara, C.A., Ben-Akiva, M.E.: Sampling of alternatives in logit mixture models. Transp. Res. Part B Methodol. **58**, 185–198 (2013)

Huber, S., Klauenberg, J., Thaller, C.: Consideration of transport logistics hubs in freight transport demand models. Eur. Transp. Res. Rev. **7**(4), 32 (2015)

Hunt, J., Stefan, K., Brownlee, A.: Establishment-based survey of urban commercial vehicle movements in Alberta, Canada: survey design, implementation, and results. Transp. Res. Rec. **1957**, 75–83 (2006)

Hunt, J.D., Stefan, K.J.: Tour-based microsimulation of urban commercial movements. Transp. Res. Part B Methodol. **41**(9), 981–1013 (2007)

Lemp, J.D., Kockelman, K.M.: Strategic sampling for large choice sets in estimation and application. Transp. Res. Part A Policy Pract. **46**(3), 602–613 (2012)

Liedtke, G.: Principles of micro-behavior commodity transport modeling. Transp. Res. Part E Logist. Transp. Rev. **45**(5), 795–809 (2009)

Livshits, V., You, D., Zhu, H., Jeon, K., Vallabhaneni, L., Camargo, P., Pourabdollahi, Z.: Mega-regional multi-modal agent-based behavioral freight model. Final report. Strategic Highway Research Program 2 (SHRP2 C20 IAP Grant). http://www.azmag.gov/Portals/0/Documents/MagContent/TRANS_2017-02-13_SHRP2-TRANS_2017-06-06-C20-MAG-Next-Generation-Freight-Demand-Model-Update.pdf. (2018)

Moeckel, R., Donnelly, R.: A model for national freight flows, distribution centers, empty trucks and urban truck movements. Transp. Plan. Technol. **39**(7), 693–711 (2016)

Nuzzolo, A., Comi, A.: Urban freight demand forecasting: a mixed quantity/delivery/vehicle-based model. Transp. Res. Part E Logist. Transp. Rev. **65**, 84–98 (2014)

Oka, H., Hagino, Y., Kenmochi, T., Tani, R., Nishi, R., Endo, K., Fukuda, D.: Predicting travel pattern changes of freight trucks in the Tokyo Metropolitan area based on the latest large-scale urban freight survey and route choice modeling. Transp. Res. Part E Logist. Transp. Rev. (2018). https://doi.org/10.1016/j.tre.2017.12.011

Resource Systems Group, Inc.: User's guide and model documentation: agent-based supply chain modeling tool. The report for Chicago Metropolitan Agency for Planning. http://www.cmap.illinois.gov/data/transportation/modeling (2015)

Roorda, M.J., Cavalcante, R., McCabe, S., Kwan, H.: A conceptual framework for agent-based modelling of logistics services. Transp. Res. Part E Logist. Transp. Rev. **46**(1), 18–31 (2010)

Sakai, T., Kawamura, K., Hyodo, T.: Locational dynamics of logistics facilities: evidence from Tokyo. J. Transp. Geogr. **46**, 10–19 (2015)

Sakai, T., Kawamura, K., Hyodo, T.: Urban freight survey sampling: challenges and strategies. In: 94th Annual Meeting of the Transportation Research Board, Washington DC (2015b)

Sakai, T., Kawamura, K., Hyodo, T.: Logistics chain modeling for urban freight: pairing truck trip ends with logistics facilities. Transp. Res. Rec. **2609**, 55–66 (2017)

Tavasszy, L.A., Smeenk, B., Ruijgrok, C.J.: A DSS for modelling logistic chains in freight transport policy analysis. Int. Trans. Oper. Res. **5**(6), 447–459 (1998)

Toilier, F., Serouge, M., Routhier, J.L., Patier, D., Gardrat, M.: How can urban goods movements be surveyed in a megacity? The case of the Paris region. Transp. Res. Proc. **12**, 570–583 (2016)

Walker, J.L., Ben-Akiva, M., Bolduc, D.: Identification of parameters in normal error component logit-mixture (NECLM) models. J. Appl. Econom. **22**(6), 1095–1125 (2007)

Wisetjindawat, W., Sano, K., Matsumoto, S.: Supply chain simulation for modeling the interactions in freight movement. J. East. Asia Soc. Transp. Stud. **6**, 2991–3004 (2005)

Wisetjindawat, W., Sano, K., Matsumoto, S.: Commodity distribution model incorporating spatial interactions for urban freight movement. Transp. Res. Rec. **1966**, 41–50 (2006)

Zhao, M., Chow, J.Y., Ritchie, S.G.: An inventory-based simulation model for annual-to-daily temporal freight assignment. Transp. Res. Part E Logist. Transp. Rev. **79**, 83–101 (2015)

Takanori Sakai is a Postdoctoral Associate in the Future Urban Mobility IRG at the Singapore-MIT Alliance for Research and Technology (SMART). He received a Ph.D. degree from the Department of Urban Planning and Policy at the University of Illinois at Chicago. His research mainly focuses on urban freight transportation planning and modeling and transportation geography. His papers were selected as the recipients of the Best Paper Award by the Urban Freight Transportation Committee at Transportation Research Board Annual Meetings in 2017 and 2018.

B. K. Bhavathrathan is an Assistant Professor in Civil Engineering at the Indian Institute of Technology Palakkad. Prior to this, he served as a Postdoctoral Associate in the Future Urban Mobility IRG at the Singapore-MIT Alliance for Research and Technology (SMART). He received a Ph.D. degree from the Department of Civil Engineering, Indian Institute of Technology Bombay. His primary research interests are on urban road networks, especially pertaining to their service deterioration.

André Alho is a Senior Postdoctoral Associate in the Future Urban Mobility IRG at the Singapore-MIT Alliance for Research and Technology (SMART). He received a Ph.D. degree in Transportation Systems at the Instituto Superior Técnico, University of Lisbon, Portugal. He is passionate about fundamental and applied research in Urban Freight / City Logistics, Supply Chains and Transportation in general. André's work has led to new insights in the field of Urban Freight, which have been published in peer-reviewed journals and presented at multiple international conferences on this topic.

Tetsuro Hyodo is a Professor at Tokyo University of Marine Science and Technology. He received a Ph.D. degree from the Department of Civil Engineering, Tokyo Institute of Technology, in 1989. His research topics cover almost all transportation modes and his major interest is discrete choice modeling. He serves as a member of the Japanese National Transportation Council. He received the Best Overall Presentation Award at Transportation Research Board SHRP2 in 2010.

Moshe Ben-Akiva is the Edmund K. Turner Professor of Civil and Environmental Engineering at the Massachusetts Institute of Technology (MIT), and Director of the MIT Intelligent Transportation Systems (ITS) Lab. He holds a Ph.D degree in Transportation Systems from MIT and 4 honorary degrees and coauthored the textbook Discrete Choice Analysis.