*sustainability*

MDPI

*Article*

# Investigation of Freight Agents' Interaction Considering Partner Selection and Joint Decision Making

**Dapeng Zhang and Xiaokun (Cara) Wang ***

Department of Civil and Environmental Engineering, Rensselaer Polytechnic Institute, Troy, NY 12180, USA; zhangdapeng@live.com

* Correspondence: wangx18@rpi.edu; Tel.: +1-518-276-2098

Received: 18 March 2020; Accepted: 28 April 2020; Published: 1 May 2020

check for updates

**Abstract:** Freight transportation plays an increasingly important role in sustainable development. However, freight travel demand has not been understood comprehensively, due to its unique features: freight activities are the result of collaboration among freight agents. It distinguishes freight transportation from passenger transportation, in which travel decisions are made mostly by individuals. Specifically, two processes in the collaboration can be observed: partner selection and joint decision making. Using the supplier-customer collaboration as an example, partner selection is a process for suppliers and customers to evaluate their potential partners and select the best one. Joint decision making allows suppliers and customers to seek common interests and make compromises. As a traditional travel demand model cannot model the two processes effectively, this research develops an innovative econometric model, spatial matching model, to bridge the gap. The proposed model is specified based on freight agents' behavioral, estimated by Bayesian MCMC methods, and demonstrated by numerical examples. The proposed model and estimation methods can recover the coefficient values in the econometric models, and establish the relationship between the influential factors and the observed matching behavior. The analysis improves the understanding of freight travel demand in a behavioral-consistent manner and enriches the body of freight demand modeling literature.

**Keywords:** joint decision making; freight transportation; partner selection; Bayesian MCMC; econometric modeling

## 1. Introduction

An efficient freight system can transport an extraordinarily large number of products for daily life, stimulates demand for goods, and employs millions of people. While bringing tremendous benefits, freight transportation can consume a good number of natural resources and incur negative impact on the environment and sustainability. Hence, establishing an efficient freight system is important, which often begins with thorough transportation planning. Traditional planning tools model travel demand at the zonal level, assuming each zone can represent all individual decision-makers within the zone. However, the resulting aggregate zonal model cannot capture the individual heterogeneity of freight agents. In other words, existing planning models, which miss individual behavior, are inconsistent with what freight agents behave in the real world, leading to poor freight planning. Recently developed models, mostly variants of discrete outcome models, analyze freight demand at the disaggregate level. These models have been able to capture some of the individual heterogeneity but have not yet been sufficient.

Individual heterogeneity in freight transportation can be characterized by the collaboration between freight agents. Travel demand of passenger transportation is usually perceived as a result of an individual's demand for work, shopping, and entertainment. Most individual's travel decisions are independent to the behavior of other travelers. Freight activities, on the other hand, have to be conducted by multiple agents. For example, dry bulk commodities are delivered based on agreements of large manufacturers, logistic companies, and retail stores. Small package delivery enabled by e-commerce are scheduled by sellers, parcel consolidators, parcel carriers, and receivers. The involvement of multiple agents can be seen in most freight deliveries. Holguin-Veras et al. [1] summarized that many freight activities were results of freight agents' interactions and understanding these interactions is important for both academia and industry in understanding transport efficiency, pollutions, and sustainability.

The interaction of freight agents can be understood by two processes: partner selection and joint decision making. Using the supplier-customer interaction as an example, partner selection is a process for the supplier (or the customer) to evaluate the counterparty's characteristics, such as locations, industry sectors, and sizes. In particular, the supplier begins with the assessment of all customers in the market and based on this assessment, selects the customer that best fulfils the supplier's desire. Sometimes, when the supplier's most favorable customer dislikes the supplier, the supplier would move on to the next best customer until the desired customer also likes the supplier. In the joint decision-making process, a freight decision is made based on common interests as well as compromises between suppliers and customers. Without considering partners, each agent's decision is made by themselves. Such a decision may or may not be accepted by the partner and consequently, they must compromise. Note that the partner selection and joint decision-making processes are executed at the same time: the joint freight decisions cannot be made without establishing a partnership and, conversely, a partnership cannot be established without a successful joint decision making.

The observed partnership in a market and joint decision outcomes can enable the development of an innovative econometric model to explain the collaboration behavior. As the two discussed processes are executed simultaneously, a structural equation model can be employed with one equation modeling the partner selection process and the other modeling the joint decision-making process. For the partner selection, the observed matching relationship (e.g., which two agents in a market are matched and which two agents are not matched) implies a series of inequality conditions of the pairwise utility (e.g., the preference of matching). Generally, the utility of an unmatched pair should be smaller than matched pairs with certain requirements. Based on the utility conditions, the effect of factors determining the pairwise utility can be identified, thus revealing factors' effects on agents' attractiveness. In the joint decision making process, discrete outcome models can be estimated with independent variables of characteristics of each agent (e.g., supplier's size and customer's industry sector) and joint factors (e.g., distance between two agents). Such two equations can be connected by specifying a correlation in the error term, characterizing the sample selection feature: the joint decisions can only be made by matched agents.

Such an innovative modeling framework, the spatial matching model, is developed upon a two-sided matching model [2], which analyzes firm's initial public offering (IPO) considering the partnership of firms and investment banks. This paper proposes an extension to the model that can further improve the understandings of freight behavior: (1) the spatial interaction between freight agents can be characterized; (2) the matching structure is many-to-many, in comparison to the one-to-many matching structure in the literature; and (3) the joint decision-making outcome is ordinal (e.g., shipping frequency), in comparison to the binary outcome of firm's IPO.

A numerical example, investigating trip frequency considering supplier-customer collaboration, is followed to illustrate the proposed model's applicability. The numerical example estimates model's parameters and interprets the result with experiment data. Note that real data, mostly unavailable now to researchers, will likely be available in future studies because of continuously advancement in data collection and dissemination. The best data that can take advantage of the proposed model is operational data of freight companies. For example, the ideal dataset contains information about the upstream or downstream partners regarding operation, locations, and delivery trip frequency. These data have been usually recorded by freight companies without additional surveying efforts. Not capable of analyzing the data today is mostly due to data dissemination: companies often treat such data confidential or are not enticed to share broadly. In the future, the proposed model has a great potential to be adopted by practitioners as the proposed model is ready for use.

Although a numerical example cannot not fully showcase the proposed model, it can demonstrate the potential advantages in understanding freight trip generation and distribution. Traditional four-step models treat them as two separate steps: trip frequency is not related with origin-destination matching. These assumptions are acceptable for most passenger trips: passengers' demand can be met at many places in nearby zones. However, the demand of freight agents can be satisfied by only a few partners who may be located far away. In other words, trip generation and distribution steps in freight transportation are interrelated at a disaggregate level: trip generation cannot be observed until a proper supplier-customer matching is established. The proposed spatial matching model can capture the interaction and provide a more behavioral-consistent freight modeling alternative.

This paper is organized as the following. The next section reviews the literature of related studies. Then, the methodology is introduced and followed by a validation using experiments. Finally, the result of numerical examples is discussed and followed by conclusions.

## 2. Literature Review

Travel demand models use collected data to explain past travel demand and forecast future travel demand. Results can drive decisions in network planning and highway engineering. The most widely-used passenger model in practice is the sequential four-step model but freight travel demand often violates this model's assumptions. This research develops an innovative econometric model to capture the unique nature of freight travel demand.

### 2.1. Freight Demand Models

The sequential four-step model [3] has been widely adopted in the transportation practice for both freight and passenger. A typical model follows the steps of trip generation, trip distribution, mode choice, and route assignment. From a larger perspective, the model can be viewed in two stages [4]. The first stage is to use the characteristics of travelers and land use features to measure the travel demand. The second stage loads the demand onto the transportation road network. An accurate measurement of travel demand at the first stage serves as a foundation of the second stage, enabling thorough transportation planning. However, the first stage is valid in aggregate analyses, but fails to perform in disaggregate situations.

Aggregate analyses are typically conducted using input-output models and gravity models. Pioneered by Wassily Leontief [5], the input-output models are used to analyze the interdependencies among industries in an economy. The basic form of an input-output model consists of a system of equations, in which each one formulates the distribution of an industry's product over the market [6]. When the investigated industries are located at different geographic regions, the model can analyze the freight demand across the regions, leading to the regional input-output model. The regional input-output model analytical framework has been employed in many empirical analyses. For example, Robison and Miller [7] used input-output models to study the timber economy of the West-central Idaho Highlands and found that input-output models were efficient techniques in cross-region trade analyses. Hewings et al. [8] employed an input-output framework to investigate the interdependence of the inner-city communities and suburbs of Chicago metropolitan area. Jackson et al. [9] designed a flow matrix based on the input-output analytical framework to investigate the interdependence of the 51 states of the U.S. A series of papers [10–12] discussed the world input-output models, and Duchin et al. [13] applied such a model, covering 189 countries around the world.

Gravity models characterize the travel demand between large economics as being stronger than that between small ones, and nearby economics attract each other more than faraway ones. Recent research of gravity models mainly focus on resistance terms [14], zero trade flows [15,16], and distance measurements [17,18]. As an important factor in gravity models, distance is specified by different methods. The most commonly used method is the distance between the centers of the investigated regions. Based on the characteristics of regions, centers are usually capitals, the largest cities, or the centroids. Adjacency is also considered in some of the literature, which is due to the considerations of freight costs and political costs. Other subtle factors, such as trade cost, market access, economic geography, and language similarities are also perceived to capture the concept of distance in gravity models.

However, these aggregate analyses assume that travelers can be represented by homogeneous geographic regions, usually census units based on geography, economics, and administrative divisions. For example, freight flow is usually conceptualized by the characteristics of the origin and destination counties, rather than the disaggregate corporations and individuals who ship and receive the cargoes. Such aggregation allows lower data requirements and computational burden, but is unable to capture the heterogeneity of individual travelers.

The development of discrete outcome models [19] enables studies of travel demand at the disaggregate level. However, constrained by the behavioral framework, these models have to assume that travel decisions are made by one individual. For freight travel demand, either the supplier or the customer, but not both, determines the frequency. The characteristics of the customers (or suppliers) are sometimes used as exogenous variables to help explain the supplier's (or customers') behavior. Later, group decision models, mainly focusing on intra-household collaboration, were investigated to recognize the fact that multiple agents may jointly make decisions. Srinivasan and Bhat [20] investigated intra-household activity travel patterns by examining interactions among household members. Zhang et al. [21] investigated household discrete choice behavior considering heterogeneous group decision making mechanisms. Marcucci et al. [22] analyzed joint decision making in three-member households. However, the matching relationship of household members is predetermined and not impacted by the joint decision-making process.

The importance of freight agents' joint decision making at a disaggregate level has been discussed in the literature, although sound quantitative methodologies have not yet been developed. A series of freight delivery time studies [1,23,24] raised concerns about freight agents' interactions. Freight carriers prefer the night time due to smooth traffic and lower costs, but receivers prefer the day time because no additional labor is needed to receive cargos in business hours. The delivery time is reasonably assumed to involve a partner selection and a joint decision-making process. In addition, Holguin-Veras et al. [1] argues that many freight activities are results of the freight agents' interactions, and that these interactions determine the supply chain's response to freight policies. For example, delivery rates [25], sizes, and frequency are impacted jointly by suppliers, carriers, and receivers. Disregarding interactions among agents may prevent the research community from fully understanding the decision mechanism, leading to misleading assessments of policy effects on each individual decision maker and, consequently, poor predictive power. However, no sound quantitative methodologies have been proposed and applied to solve freight agents' interaction problems. Hence, the proposed econometric model enriches the literature of freight modeling.

### 2.2. Matching Models

Collaboration in supply chains has drawn researchers' increasing attention since the boom of information technology. Collaboration encourages all players of supply chains to engage in planning, forecasting, replenishment, and sharing. Benefits include cost reduction, sales improvement, and accurate forecast. Cao and Zhang [26] claimed that collaborative advantage is a strategic benefit gained over competitors in the marketplace through supply chain partnering and partner enabled knowledge creation. They summarized that collaborative advantages as five dimensions of process efficiency, offering flexibility, business synergy, quality, and innovation. Ramanathan and Gunasekaran [27] analyzed the long-term collaboration decisions considering the execution of current collaboration. In terms of the modeling technique, simulation methods are popular. For example, Hudnurkar and Rathod [28] used a beer game simulation technique to investigate the impact of collaboration in vender managed inventory and planning, forecasting, and replenishment. Results substantiate the improvement by collaborative techniques in inventory management, service level, and overall supply chain cost. Dorigatti et al. [29] proposed a systematic and reliable framework of agent-based simulation to support the collaborative interaction analysis in supply chains. Econometric modeling is scarce. Nyaga and Whipple [30] used a structure equation model to examine the satisfaction and performance of collaboration from both suppliers' and customers' perceptions. They found that information sharing has a greater influence on suppliers in comparison to buyers, because joint effort provides the supplier with greater access to the buyers, enables suppliers to share concerns, seek relationship benefits, and safeguard their dedicated investments. However, to the best knowledge of the authors, there is no econometric study analyzing the matching behavior in supply chains.

The limitations of existing aggregate and disaggregate analyses enable the development of an innovative analytical framework to characterize freight travel demand. Based on the nature of freight activity behavior (e.g., partner selection and joint decision making), matching models in the existing literature may serve as a foundation to developing the innovative methodology in this paper. The matching model has a relatively short history. In 2010, Dale Mortensen earned the Nobel Prize in Economics for the analysis of markets with search frictions, which explicitly explains the process of partner selection. Matching model frameworks can solve problems that are raised by multiple decision makers. When the matching relationship data is observed, matching models in econometrics can analyze factors in forming such matching. Relationship data include which firms do business with other firms, which men are married with which women, and which bidders won which auction items, among other data involving agent collaboration. The basic economic idea is that individuals would like to match with the most attractive partners. Economists seek influential factors that determine the observed matching and estimate the parameters of these factors.

An important matching literature is Sorensen [1], which uses a two-sided matching model to explain firms' IPO with the bank-firm matching. It first uses a latent variable equation to explain the bank-firm partner selection and then uses a binary outcome model to formulate firms' IPO. The first equation considers matching utility of all possible pairs, while the second equation only considers the IPO of matched pairs, leading to a sample selection process. A Bayesian MCMC approach is employed to estimate the parameters of factors in determining all parameters in the two equations. This modeling framework takes each decision maker position to analyze behavioral-consistently the selection of partners. Such a method works as groundwork for this paper, which relaxes the outcome being ordinal. In addition, spatial effect will be emphasized and the matching structure extends from one-to-many (e.g., firms can only get investment from one bank) to many-to-many (e.g., each supplier can trade with multiple customers, and each customer can trade with multiple shippers) relationship. Note that the spatial effect is specified as a linear-in-parameter term in the proposed model, as opposed to autocorrelation terms in spatial econometrics.

The econometric matching model has also been discussed in a limited number of empirical studies. Chen [31] specifies the utility equations for each side of the partner to analyze the premium of bank loans. This study is an extension of Sorensen [1] in which the utility of paired partners is investigated instead of the utility of decision makers, respectively. However, this model may suffer from identification problems if extended to ordered outcomes. To the research team's best knowledge, no other studies have conducted the research in a similar sample correction manner. Other studies have analyzed the matching process from the market's perspectives, using different estimation methods, or without considering the mutual selection process. For example, Zhang and Wang [32] use the matching model to analyze the interaction between airlines and airports. Choo and Siow [33] investigated the stable matching relationship from the market's perspective, rather than each decision maker's perspective. Similar marriage analyses can be also found in Siow [34]. Hitsch et al. [35] also analyze a marriage dataset, but the methodology does not consider the mutual preference by sorting pairwise utility. Fox [36] and Levine [37] use maximum score estimators to identify parameters in the matching model. Ma et al. [38] use a variational inequality model to investigate rider-sharing surge pricing at the origin-destination with a constraint of ride matching.

In summary, this research develops an innovative spatial matching model to analyze freight travel demand in a behavior-consistent manner. Grounded by the existing matching model, this research incorporates the spatial relationship, the many-to-many matching structure, and ordered joint decision outcomes at the disaggregate level.

## 3. Methodologies

The proposed methodology, which is potentially widely applicable on data consolidated from freight companies, is a two-equation econometric modeling that can be estimated by a Bayesian MCMC method.

### 3.1. Model Specification

The unique feature of freight demand is the agent's collaboration and two processes can be observed: partner selection and joint decision making. Two equations can be specified to conceptualize the two processes, respectively. For partner selection, the matching equation uses the conditions of pairwise utility to identify the parameters. The joint decision-making process is characterized by a standard ordered probit model with the freight delivery frequency as the dependent variable. In addition, a variance-covariance matrix is specified to capture the involved sample selection feature: joint decisions can be only made between matched agents.

The partner selection process considers a market consisting of two sides (e.g., suppliers and customers), although agents of any side may also interact with other participants that do not belong to the two sides (e.g., carriers). Each agent of one side is assumed to have full information about agents in the other side and look for the best partners. The pairwise utility of all pairs (matched and unmatched) $u_{ij}$ can be specified as

$$u_{ij} = \alpha_d d_{ij} + \alpha'_w w_{ij} + \eta_{ij} \tag{1}$$

where $d_{ij}$ is the distance between supplier $i$ and customer $j$, capturing their spatial relationship. The spatial relationship can be defined by Euclidean distance, network distance, travel time, and other reasonable measurements. This linear specification of the distance variable is without loss of generality: the distance can be also in a nonlinear term that can be transformed to a linear function. The term $w_{ij}$ contains other influential characteristics of both sides, which can be the characteristics of each side (e.g., supplier's size and/or population density at the customer location) and their joint factors (e.g., cooperation history in years of supplier $i$ and customer $j$). The $\alpha\{\alpha = [\alpha_d; \alpha_w]\}$ are the parameters to be estimated. The error term $\eta_{ij}$ contains unobserved effects determining the pairwise utility and is assumed to follow a normal distribution.

The pairwise utility $u_{ij}$ cannot be observed directly from the data. Its absolute magnitude is not of interest, but its relative magnitude is able to characterize the matching relationship. Specifically, inequality conditions can be employed to describe the relative magnitude of pairwise utility, which outstands the matching model from traditional econometric models. The inequality conditions are introduced for unmatched pairs and matched pairs, respectively.

For an unmatched pair,

$$u_{ij} < \overline{u_{ij}} = \max\left[\min_{i' \in \tau(j)} u_{i'j}, \min_{j' \in \tau(i)} u_{ij'}\right] \tag{2}$$

The term $\overline{u_{ij}}$ can be perceived as the opportunity cost for supplier $i$ and customer $j$ to deviate from the existing pair and form a new match together. The valuation of the worst supplier in customer's current pairs is $\min_{i' \in \tau(j)} u_{i'j}$, and the worst customer in the supplier's current pairs is $\min_{j' \in \tau(i)} u_{ij'}$. When $u_{ij}$ exceeds $\overline{u_{ij}}$, the supplier $i$ and customer $j$ would both prefer to form a new match together. When $u_{ij}$ is smaller than $\overline{u_{ij}}$, both of the supplier $i$ and customer $j$ would stay with their current matches.

The $u_{ij}$ of matched pairs is constrained by the following conditions:

$$u_{ij} > \underline{u_{ij}} = \max\left[\max_{i' \in S(j)} u_{i'j}, \max_{j' \in S(i)} u_{ij'}\right] \tag{3}$$

where $S(i) = \left\{ j \in J : u_{ij} > \min_{i' \in \tau(j)} u_{i'j} \right\}$ and $S(j) = \left\{ i \in I : u_{ij} > \min_{j' \in \tau(i)} u_{ij'} \right\}$.

The term $\underline{u_{ij}}$ can be perceived as the opportunity cost for supplier $i$ and customer $j$ to stay with their existing pairs. The sets $S(i)$ and $S(j)$ are the feasible deviations from the existing pairs for supplier $i$ and customer $j$, respectively. The feasible set of supplier $i$ is the customers that the supplier prefers to, compared with the supplier's current partners. When $u_{ij}$ is greater than $\underline{u_{ij}}$, the agents prefer to stay with their current partners.

The delivery frequency $y_{ij}$, in an ordered outcome, is the result of joint decisions made by supplier $i$ and customer $j$. The function takes the form of

$$\begin{aligned} y_{ij}^* &= \beta_d d_{ij} + \beta_x x_{ij} + \varepsilon_{ij} \\ y_{ij} &= C \text{ if } \mu_{C-1} \leq y_{ij}^* < \mu_C \end{aligned} \tag{4}$$

where $d_{ij}$ is the distance between supplier $i$ and customer $j$ and $x_{ij}$ contains influential characteristics of agent $i$ and agent $j$ on the joint decision. The term $\beta\{\beta = [\beta_d; \beta_x]\}$ are the parameters to be estimated. The $C$ is the ordinal outcome and divided by threshold $\mu_C$. The error term $\varepsilon_{ij}$ is assumed to follow a normal distribution, leading to an ordered probit model. Spatial relationship indicator, such as distance, can also be added as an explanatory variable in $x_{ij}$, capturing the effect of spatial relationship on the delivery frequency decisions.

The matching with different partners is likely to result in different shipping frequency. Thus, the matching equation should connect with the probit equation to capture the simultaneous effect between the partner selection process and the joint decision-making process. A variance-covariance matrix is assumed to capture the correlation of the two equations: $\varepsilon_{ij}$ and $\eta_{ij}$ follow a multivariate normal distribution

$$\begin{pmatrix} \varepsilon_{ij} \\ \eta_{ij} \end{pmatrix} \sim N\left(0, \begin{bmatrix} 1+\delta^2 & \delta \\ \delta & 1 \end{bmatrix}\right) \tag{5}$$

The variance $\varepsilon_{ij}$ and $\eta_{ij}$ are assumed to be $1+\delta^2$ and 1 due to identification concerns and without loss of generality. The covariance is defined as $\delta$, indicating the correlation between the matching equation and the ordered probit equation. Only positive covariance is considered in this research due to the research context: the error terms in the two equations capture the unobserved effects on matching and delivery frequency, respectively. If the supplier and the customer would prefer to cooperate with each other, the delivery frequency between them would be expected higher, leading to a positive covariance.

### *3.2. Estimation*

The observed matching relationship is the key to identify and estimate the parameters in the spatial matching model. When a supplier has matched with enough customers, other customers cannot match with this supplier. Therefore, interaction exists among all possible pairs: the partner selection of each pair relies on the decisions of all other pairs. As a result, the error terms of all pairs should be integrated simultaneously to evaluate the likelihood function. The most commonly used estimation methods (e.g., the maximum likelihood estimation) are good at dealing with low-dimensional likelihood functions. As the dimensionality can easily go to thousands, maximum likelihood methods become infeasible for the proposed spatial matching model. As an alternative, the Bayesian MCMC estimation can avoid the evaluation of high-dimensional integrals by simulating the distributions of parameters and thus, is used in this research.

The Bayesian MCMC estimation seeks to simulate the posterior distributions of unknown parameters using the assumed prior distributions and the likelihood function. Normal distributions are assumed as the prior distributions for most parameters in this research because they are conjugate distributions, which ease the tractability in the estimation. Large variances (e.g., 10,000) are used for the prior distributions to construct uninformative priors. With such prior distributions, posterior distributions can rely more on the observed data than using informative priors.

The posterior distribution of $u_{ij}$ depends on whether individual $i$ and $j$ are matched with each other. If they are matched,

$$u_{ij} \sim N\left(w_{ij}\alpha + \left(y^*_{ij} - x_{ij}\beta\right)\delta/\left(1+\delta^2\right), 1/\left(1+\delta^2\right)\right) \tag{6}$$

and truncated below at $\underline{u_{ij}}$. Otherwise,

$$u_{ij} \sim N\left(w_{ij}\alpha + \left(y^*_{ij} - x_{ij}\beta\right)\delta/\left(1+\delta^2\right), 1/\left(1+\delta^2\right)\right) \tag{7}$$

and truncated above at $\overline{u_{ij}}$.

The latent outcome of ordered probit equation $y^*_{ij}$ can follow a normal posterior distribution

$$y^*_{ij} \sim N\left(x_{ij}\beta + \left(u_{ij} - w'_{ij}\alpha\right)\delta, 1\right) \tag{8}$$

truncated according to the ordinal outcome and the values of thresholds. For example, if $y_{ij} = 2$, then $y^*_{ij}$ is truncated as $(\mu_1, \mu_2)$.

The thresholds of ordered outcome $\mu_c$ can follow a uniform posterior distribution

$$\mu_c \sim U\left(\max_{y_{ij}=c}\left(y^*_{ij}\right), \min_{y_{ij}=c+1}\left(y^*_{ij}\right)\right) \tag{9}$$

The coefficient βis assumed to follow a normal prior

$$\beta \sim N\left(\beta_0, \sum_{\beta}\right) \tag{10}$$

The conditional posterior distribution of $\beta$ is a multivariate normal distribution of $N(-M_\beta^{-1}N_\beta, -M_\beta^{-1})$ where

$$M_\beta = \sum_{\beta}^{-1} + \sum_{ij\in\mu_m} x_{ij}x'_{ij} \tag{11}$$

And

$$N_\beta = -\sum_{\beta}^{-1}\beta_0 - \sum_{ij\in\mu_m} x_{ij}\left(y^*_{ij} - u_{ij}\delta + w'_{ij}\alpha\delta\right) \tag{12}$$

The term $\alpha$ is assumed to follow a normal prior

$$\alpha \sim N\left(\alpha_0, \sum_{\alpha}\right) \tag{13}$$

The conditional posterior distribution of $\alpha$ is a multivariate normal distribution of $N(-M_\alpha^{-1}N_\alpha, -M_\alpha^{-1})$ where

$$M_\alpha = \sum_{\alpha}^{-1} + \sum_{ij\in M_m} w_{ij}w'_{ij} + \sum_{ij\in\mu_m} \delta^2 w_{ij}w'_{ij} \tag{14}$$

And

$$N_\alpha = -\sum_{\alpha}^{-1} \alpha_0 + \sum_{ij \in M_m} -w_{ij} u_{ij} + \sum_{ij \in \mu_m} \delta w_{ij} \left( y_{ij}^* - x_{ij}\beta - u_{ij}\delta \right) \tag{15}$$

The term $\delta$ is assumed to follow a normal prior

$$\delta \sim N\left( \delta_0, \sum_{\delta} \right) \tag{16}$$

The conditional posterior distribution of $\delta$ is a multivariate normal distribution of $N(-M_\delta^{-1} N_\delta, -M_\delta^{-1})$ where

$$M_\delta = \sum_{\delta}^{-1} + \sum_{ij \in M_m} \left( u_{ij} - w_{ij}\alpha \right)^2 \tag{17}$$

And

$$N_\delta = -\sum_{\delta}^{-1} \delta_0 + \sum_{ij \in \mu_m} \left( y_{ij}^* - x_{ij}\beta \right)\left( u_{ij} - w_{ij}' \alpha \right) \tag{18}$$

The parameters are simulated based on the posterior distributions given above. The parameters are updated sequentially until the convergence is reached. The estimation is conducted in MATLAB.

### 3.3. Desired Data Structure

Unlike traditional econometric models where the modeled data is a set of observations for each independent individual, this matching model needs data that contain the relationship information of matching relationships between individuals. Using the freight supplier-consumer example, let $i(i = 1 \ldots I)$ denote the one side of the market (e.g., suppliers) and $j(j = 1 \ldots J)$ denote the other side of the market (e.g., customers). The desired relationship information of supplier $i$ is the observed matching with a list of $j$, denoted as $\tau(i)$. For a market consisting of three suppliers and three customers, assume supplier 1 is matched with customer 2 and 3. Then, $\tau(1) = 2, 3$. The proposed model has to know $\tau(i)$ and $\tau(j)$ for any agents in the market.

Apart from the matching relationship, the model also desires influential factors that determine pairwise utility and the observed ordinal outcome of the joint decision making. The requirements of these data do not differ much from the traditional econometric models. For example, an influential factor that determines the supplier-customer pairwise utility is the customer's debt. Building the matching model needs the debt data of all customers in the market. The observed ordinal outcome can be the delivery frequency. Then, building the matching model needs the observed delivery frequency of matched suppliers and customers.

As disaggregate data is often associated with safety, privacy, and confidentiality concerns, obtaining the desired data for the proposed model may require additional efforts. For example, the data can be obtained by consolidating operation records from freight suppliers and customers, which contain their business partners and the delivery frequency. Due to the limitation of research time and resource, this paper does not analyze empirical data. However, a numerical example is used to demonstrate the proposed model by assuming the desired data can be obtained.

## 4. Model Validation

Experiments are conducted to test the parameters' recovery capability, which is measured by root-mean-square deviation (RMSE) statistics. The model is validated if values of the estimated parameters are close to the pre-defined values of parameters that generate the experiment data.

Two matching structures are investigated in the experiment. The first structure has an unbalanced number of agents on each side (five suppliers and 500 customers) and a predetermined number of partners for each side: Each supplier matches with two customers and each customer matches with

200 customers. The second structure has an equal number of agents on each side (50 suppliers and 50 customers) but instead of a predetermined number of matched partners, a constraint (each agent can have 20 to 30 partners) is given. The specific number of partners for each agent is data-driven, given that the utility conditions define the matching relationships and thus, the observed partners.

For both matching structures, the covariance term $\delta$ is evaluated at 0.2, 0.5, and 0.8, respectively. Parameters $\alpha$, $\beta$, and $\mu$ are investigated at the same values in all experiments. The independent variables in both equations are randomly generated by uniform distributions of (0, 3). With the pre-defined values of parameters, the pairwise utility of matching and the ordered outcome utility are calculated. Then, the observed matches and delivery frequency are determined by the utility conditions and pre-defined thresholds.

Each experiment updates the posterior distributions of all parameters for 100,000 iterations. The first 50,000 iterations are removed from the final results as burn-in and the mean and standard deviation of the parameters are calculated based on the final 50,000 iterations. The result of experiments is reported in Table 1.

**Table 1.** Experiment results for two matching structures.

| Parameter | True Value | Average of Means from Posterior Distributions Matching Structure 1 Unbalanced & Predetermined | | | Matching Structure 2 Balanced & Constrained | | |
|---|---|---|---|---|---|---|---|
| $\delta$ | True Value | 0.2 | 0.5 | 0.8 | 0.2 | 0.5 | 0.8 |
| | Estimated Value | 0.253 | 0.524 | 0.895 | 0.289 | 0.577 | 0.799 |
| $\alpha_d$ | −0.6 | −0.624 | −0.597 | −0.532 | −0.579 | −0.549 | −0.613 |
| $\alpha_{w1}$ | −0.3 | −0.320 | −0.360 | −0.348 | −0.255 | −0.283 | −0.310 |
| $\alpha_{w2}$ | 0.9 | 0.820 | 1.039 | 1.050 | 1.092 | 1.334 | 0.926 |
| $\beta_d$ | 0.9 | 0.982 | 0.855 | 0.972 | 0.948 | 0.857 | 0.875 |
| $\beta_{x1}$ | −0.6 | −0.656 | −0.566 | −0.656 | −0.567 | −0.664 | −0.674 |
| $\beta_{x2}$ | −0.3 | −0.296 | −0.295 | −0.338 | −0.323 | −0.293 | −0.297 |
| $\mu_1$ | −0.3 | −0.168 | −0.312 | −0.343 | −0.116 | −0.425 | −0.428 |
| $\mu_2$ | 0.3 | 0.392 | 0.318 | 0.329 | 0.468 | 0.174 | 0.173 |
| **Average RMSE** | | 0.024 | 0.018 | 0.025 | 0.037 | 0.054 | 0.022 |

The experiment results show good parameters' recovery capability in general, successfully validating the proposed spatial matching model. Larger average RMSE of the matching structure 2 indicates greater deviations from true parameter values. The reason may be that the number of partners for each agent is not predetermined, resulting in a more flexible matching relationship.

## 5. Numerical Example

A numerical example, experiment data of the matching structure 1 with $\delta$ equals to 0.2, is used to further demonstrate the proposed spatial matching model. Although this study does not perform real-data analysis, the numerical example is still able to show the performance of the model and feed into interpretation discussion. Although the result does not show any practical policy implication, planners could adopt this interpretation method when the appropriate data is available.

Table 2 shows the estimation results of the spatial matching model and a standard ordered probit model without considering the matching relationship.

**Table 2.** Estimation result of a numerical example using the spatial matching model.

| | Spatial Matching Model | | | | Standard Ordered Probit Model | | |
|---|---|---|---|---|---|---|---|
| **Variable** | **Coef.** | **Standard Deviation** | **t-stat** | ***P*-value** | **Coef.** | **t-stat** | ***P*-value** |
| **Matching Equation** | | | | | | | |
| Distance | −0.624 | 0.040 | −15.60 | 0.000 | | | |
| Supplier's History | 0.820 | 0.094 | 8.72 | 0.000 | | | |
| Customer's Debt | −0.320 | 0.042 | −7.62 | 0.000 | | | |
| **Ordered Probit Equation** | | | | | | | |
| Distance | 0.296 | 0.047 | 6.30 | 0.000 | 0.507 | 10.3 | 0.000 |
| Supplier's Size | −0.982 | 0.051 | −19.25 | 0.000 | −0.905 | −17.43 | 0.000 |
| Population Density at that Customer's Location | 0.656 | 0.065 | 10.09 | 0.000 | 0.61 | 14.27 | 0.000 |
| cut1 | −0.168 | | | | 0.458 | | |
| cut2 | 0.392 | | | | 1.005 | | |
| **Error Term** | | | | | | | |
| | 0.253 | 0.126 | 2.01 | 0.053 | | | |
| **Goodness-of-Fit** | | | | | | | |
| Log-likelihood | −2543 | | 0.000 | | −746 | 0.000 | |
| Log-likelihood at null | −2755 | | | | −788 | | |

### 5.1. Matching Equation

This numerical example only includes three explanatory variables in the matching equation while empirical research can analyze as many important explanatory variables as possible. The value of the estimated coefficient would reveal the explanatory variable's impact on the pairwise utility. A positive coefficient indicates that the variable improves the matching utility and vice versa. Based on the estimation result, freight companies can adjust the current characteristics in order to improve their attractiveness to the other side of the market. On the other hand, policy makers can understand the agent's partner selection mechanisms, and make proper interventions to reshape the matching relationship.

Specifically, the distance between the supplier and the customer is −0.624, implying closer agents are more likely to be cooperative with each other. Deliveries between closer agents are associated with lower transportation cost and thus, higher profit. This distance captures the spatial relationship of freight agents, which is one of the important focuses of this research. Traditional planning analytical framework defines the distance between two regions that represent decision makers within the region, which roughly approximates the spatial relationship. The distance investigated in this research can improve the accuracy of spatial relationship measurement by executing disaggregate-level analyses.

The supplier's history has a coefficient of 0.820, indicating that experienced suppliers can attract good customers. The estimated coefficient of customer's debt is −0.320. Low debt means that the customer can pay off the commodity in time, which is an important consideration for suppliers.

### 5.2. Ordered Probit Equation

The ordered probit model can be understood by the marginal effect, which is reported in Table 3. Except for the distance, the result of the supplier's size and population density at the customer's location in the spatial matching model is similar to the result of the standard ordered probit model. The marginal effect can be interpreted using the supplier's size as an example. If the supplier's size increases by one unit, the probability of observed low frequency would increase by 33.1%, the probability of moderate frequency would increase by 4.0%, and the probability of high frequency would decrease 37.1%.

**Table 3.** Marginal effects of the ordered probit equations in the spatial matching model and the standard ordered probit model.

| | Spatial Matching Model | | | Standard Ordered Probit Model | | |
|---|---|---|---|---|---|---|
| | **Low Frequency** | **Moderate Frequency** | **High Frequency** | **Low Frequency** | **Moderate Frequency** | **High Frequency** |
| Distance | −0.102 | −0.012 | 0.114 | −0.181 | −0.020 | 0.202 |
| Supplier's Size | 0.331 | 0.040 | −0.371 | 0.324 | 0.036 | −0.360 |
| Population Density at that Customer's Location | −0.235 | −0.028 | 0.263 | −0.219 | −0.024 | 0.243 |

The interpretation of distance is more complicated because it is the variable in both the matching equation and the ordered probit equation. The marginal effects differ significantly in both models due to this reason. The difference implies that missing the matching relationship would result in an erroneous interpretation of spatial relationship and consequently insufficient understanding of freight travel demand. Note that the marginal effect is obtained by increasing distance of all pairs by one unit and calculating the change in the results ordinal outcome.

### 5.3. Error Term

In this simulation study, the covariance term turns out to be significantly different from zero. The error term of the matching equation contains the unobserved factors determining the matching utility. In this freight demand context, the factor may be the supplier-customer cooperation history: a long history is associated with a high possibility to form a matching relationship. The error term of the ordered probit equation may include highway density at the supplier or customer's location and other factors determining delivery frequency. The positive covariance implies that the cooperation history and highway density are positively correlated. Furthermore, long history not only increases matching utility, but also increases the probability of high delivery frequency and conversely, a dense highway network not only increases delivery frequency, but also increases the probability to form a match.

The significant covariance also confirms the existence of the sample selection bias. Suppose that one seeks to estimate the ordered probit equation. If the matching relationship is not considered, the sample in the ordered probit equation consists of all possible pairs of suppliers and customers. The expectation of the ordered outcome utility is

$$E\left(y_{ij}^{*}\middle|x_{ij}\right) = \beta' x_{ij} \tag{19}$$

In practice, delivery frequency outcomes can be only observed for matched pairs. In other words, the delivery frequency outcome is conditional on the matching outcome. Therefore, when the matching is considered, the expectation of ordered outcome utility is written as

$$E\left(y_{ij}^{*}\middle|x_{ij}, sample\ selection\ rule\right) = E\left(y_{ij}^{*}\middle|x_{ij}, U\right) = \beta' x_{ij} + f(\delta) \tag{20}$$

where U is the utility conditions described by Equations (2) and (3). As U is a complicated conditionality, a closed-form mathematical expression cannot be derived for the conditional expectation. However, the conditional expectation is a function of the covariance term $\delta$, according to the bivariate normal distribution features (Hamedani & Tata, 1975). Therefore, without considering the matching, the delivery frequency outcome is biased in estimation and consequently, the estimation of coefficients is biased.

The correlation of the two equations is an evidence that trip generation and trip distribution are interrelated in this simulation analysis. Such a finding and modeling technique enrich the existing literature in jointly modeling the two steps.

## 6. Conclusions

This paper develops an econometric model, the spatial matching model, to characterize freight travel demand. The unique feature of freight travel demand to passenger travel demand is the collaboration between multiple agents. Due to this unique feature, some of the traditional four-step model assumptions no longer hold, such as the independence of trip generation and distribution. The proposed model can capture partner selection and joint decision making, which sheds lights on the freight modeling literature.

The observed partnership of freight agents can infer the conditions of pairwise utility for all possible agent pairs. Based on the utility conditions, the parameters determining the pairwise utility can be identified. Along with the observed joint decision outcomes, a two-equation structural model can formulate the partner selection and the joint decision-making processes. Each equation can contain the spatial relationship as independent variables, which gives the name of spatial matching models.

The proposed model is validated by experiments and estimated by numerical examples. Analyses demonstrate that the proposed model can understand the freight travel demand in a behavioral-consistent manner. This research contributes to the literature of freight transportation, which becomes increasingly important but is still under-investigated.

The proposed model is a starting effort in analyzing the matching and joint decision-making behavior. Some of the challenges are to be studied in future work: (1) due to data availability issue, the paper used a numerical example to demonstrate the proposed model. The reason for not having proper empirical data is that many freight companies treat these data as highly confidential. However, such data do exist, and the next step is to have it properly studied. In other words, researchers in freight companies can leverage the proposed model into their work now, or academia and freight companies can collaborate in the future. (2) The specification of both equations may introduce extra error terms for each side. More terms can be added to the equation and the corresponding estimation techniques can be developed.

**Author Contributions:** D.Z. conducted the research, developed the model and drafted the manuscript; X.(C.)W. conceptualized the research and revised the manuscript. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research received no external funding.

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

1. Holguín-Veras, J.; Aros-Vera, F.; Browne, M. Agent interactions and the response of supply chains to pricing and incentives. *Econ. Transp.* **2015**, *4*, 147–155. [CrossRef]
2. Sørensen, M. How smart is smart money? A two-sided matching model of Venture Capital. *J. Financ.* **2007**, *62*, 2725–2762. [CrossRef]
3. Manheim, M.L. *Fundamentals of Transportation Systems Analysis*; MIT Press: Cambridge, MA, USA, 1979.
4. McNally, M.G. The Four-Step Model. In *Handbook of Transport Modelling*; Emerald Publishing: Bingley, UK, 2008; pp. 35–53. [CrossRef]
5. Leontief, W. Quantitative input and output relations in the economic systems of the United States. *Rev. Econ. Stat.* **1936**, *18*, 105–125. [CrossRef]
6. Miller, E.; Blair, P. *Input-Output Analysis: Foundations and Extensions*; Cambridge University Press: New York, NY, USA, 2009.
7. Robison, M.H.; Miller, J.R. Cross-Hauling and Nonsurvey Input—Output Models: Some Lessons from Small-Area Timber Economies. *Environ. Plan. A* **1988**, *20*, 1523–1530. [CrossRef]
8. Hewings, G.; Okuyama, Y.; Sonis, M. Economic Interdependence within the Chicago Metropolitan Area: A Miyazawa Analysis. *J. Reg. Sci.* **2001**, *41*, 195–217. [CrossRef]
9. Jackson, R.; Schwarm, W.; Okuyama, Y.; Islam, S. A method for constructing commodity by industry flow matrices. *Ann. Reg. Sci.* **2006**, *40*, 909–920. [CrossRef]

10. Leontief, W. Sructure of the world economy: Outline of a simple input-output formulation. *Am. Econ. Rev.* **1974**, *64*, 823–834. [CrossRef]
11. Nyhus, D.; Birth, P. The Trade Model of a Dynamic World Input-Output Forecasting System. 1975. Available online: http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.252.7938 (accessed on 30 April 2020).
12. Dietzenbacher, E.; Los, B.; Stehere, R.; Timmer, M. The construction of world input–output tables in the WIOD project. *Econ. Syst. Res.* **2013**, *25*, 71–98. [CrossRef]
13. Duchin, F.; Lange, G.; Thonstad, K.; Idenburg, A.; Cropper, M. The future of the environment: Ecological economics and technological change. *J. Econ. Lit.* **1996**, *34*, 818–819. [CrossRef]
14. Rose, A.; Van Wincoop, E. National Money as a Barrier to International Trade: The Real Case for Currency Union. *Am. Econ. Rev.* **2001**, *91*, 386–390. [CrossRef]
15. Santos Silva, J.; Tenreyro, S. The Log of Gravity. *Rev. Econ. Stat.* **2006**, *88*, 641–658. [CrossRef]
16. Helpman, E.; Melitz, M.; Rubinstain, Y. Estimating trade flows: Trading partners and trading volumes. *Q. J. Econ.* **2008**, *123*, 441–487. [CrossRef]
17. Limao, N.; Venables, A. Infrastructure, geographical disadvantage, transport costs, and trade. *World Bank Econ. Rev.* **1999**, *15*, 451–479. [CrossRef]
18. Disdier, A.; Head, K. The Puzzling Persistence of the Distance Effect on Bilateral Trade. *Rev. Econ. Stat.* **2008**, *90*, 37–48. [CrossRef]
19. McFadden, D. Conditional logit analysis of qualitative choice behavior. In *Frontiers in Economics*; Zarembka: New York, NY, USA, 1971; pp. 105–142.
20. Srinivasan, S.; Bhat, C. Modeling household interactions in daily in-home and out-of-home maintenance activity participation. *Transportation* **2005**, *32*, 523–544. [CrossRef]
21. Zhang, J.; Kuwano, M.; Lee, B.; Fujiwara, A. Modeling household discrete choice behavior incorporating heterogeneous group decision-making mechanisms. *Transp. Res. Part B Methodol.* **2009**, *43*, 230–250. [CrossRef]
22. Marcucci, E.; Stathopoulos, A.; Rotaris, L.; Danielis, R. Comparing Single and Joint Preferences: A Choice Experiment on Residential Location in Three-Member Households. *Environ. Plan. A* **2011**, *43*, 1209–1225. [CrossRef]
23. Holguín-Veras, J.; Silas, M.; Polimeni, J.; Cruz, B. An investigation on the effectiveness of joint receiver–carrier policies to increase truck traffic in the off-peak hours. *Netw. Spat. Econ.* **2008**, *8*, 327–354. [CrossRef]
24. Holguín-Veras, J.; Xu, N.; de Jong, G.; Maurer, H. An Experimental Economics Investigation of Shipper-carrier Interactions in the Choice of Mode and Shipment Size in Freight Transport. *Netw. Spat. Econ.* **2011**, *11*, 509–532. [CrossRef]
25. Zhang, D.; Wang, X.; Holguin-Veras, J.; Zou, W. Investigation of carriers' ability to transfer toll increases: An empirical analysis of freight agents' relative market power. *Transportation* **2019**, *46*, 2291–2308. [CrossRef]
26. Cao, M.; Zhang, Q. Supply chain collaborative advantage: A firm's perspective. *Int. J. Prod. Econ.* **2010**, *128*, 358–367. [CrossRef]
27. Ramanathan, U.; Gunasekaran, A. Supply chain collaboration: Impact of success in long-term partnerships. *Int. J. Prod. Econ.* **2014**, *147*, 252–259. [CrossRef]
28. Hudnurkar, M.; Rathod, U. Collaborative supply chain: Insights from simulation. *Int. J. Syst. Assur. Eng. Manag.* **2012**, *3*, 122–144. [CrossRef]
29. Dorigatti, M.; Guarnaschelli, A.; Chiotti, O.; Salomone, H. A service-oriented framework for agent-based simulations of collaborative supply chains. *Comput. Ind.* **2016**, *83*, 92–107. [CrossRef]
30. Nyaga, G.; Whipple, J. Examining supply chain relationships: Do buyer and supplier perspectives on collaborative relationships differ? *J. Oper. Manag.* **2010**, *28*, 101–114. [CrossRef]
31. Chen, J. Estimation of the Loan Spread Equation with Endogenous Bank-Firm Matching. *Struct. Econom. Models* **2013**, *31*, 251–289. [CrossRef]
32. Zhang, D.; Wang, X. Understanding Many-to-Many Matching Relationship and Its Correlation with Joint Response. *Transp. Res. Part B Methodol.* **2018**, *108*, 249–260. [CrossRef]
33. Choo, E.; Siow, A. Who Marries Whom and Why. *J. Political Econ.* **2006**, *114*, 175–201. [CrossRef]
34. Siow, A. How does the marriage market clear? An empirical framework. *Can. J. Econ.* **2008**, *41*, 1121–1155. [CrossRef]
35. Hitsch, G.; Hortacsu, A.; Ariely, D. Matching and sorting in online dating. *Am. Econ. Assoc.* **2013**, *100*, 130–163. [CrossRef]

36. Fox, J. Estimating Matching Games with Transfers. 2016. Available online: https://www.econstor.eu/bitstream/10419/149760/1/856188158.pdf (accessed on 30 April 2020).
37. Levine, A. Licensing and Scale Economies in the Biotechnology Pharmaceutical Industry. 2009. Available online: http://apps.olin.wustl.edu/workingpapers/pdf/2010-11-004.pdf (accessed on 30 April 2020).
38. Ma, J.; Xu, M.; Meng, Q.; Cheng, L. Ridesharing user equilibrium problem under OD-based surge pricing strategy. *Transp. Res. Part B Methodol.* **2020**, *134*, 1–24. [CrossRef]