FREIGHT MOBILITY RESEARCH INSTITUTE
College of Engineering & Computer Science
Florida Atlantic University

Project ID: Y3R8-19

INCORPORATING FREIGHT IN REGIONAL LAND USE PLANNING MODELS

Final Report

by

Principal Investigator: Sabya Mishra, Ph.D., P.E.
The University of Memphis, Email: smishra3@memphis.edu

Mihalis M. Golias, Ph.D.
Ali Riahi Samani
The University of Memphis

for
Freight Mobility Research Institute (FMRI)
777 Glades Rd.
Florida Atlantic University
College Park, MD 20742

January 2022

ACKNOWLEDGEMENTS

This project was funded by the Freight Mobility Research Institute (FMRI), one of the twenty TIER University Transportation Centers that were selected in this nationwide competition, by the Office of the Assistant Secretary for Research and Technology (OST-R), U.S. Department of Transportation (US DOT).

DISCLAIMER

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

EXCUTIVE SUMMARY

Research on firm location choice has traditionally received less attention compared to residential location choices. This study focuses on modeling the location choice of smaller economic units (establishments) within the framework of the North American Industrial Classification System (NAICS) sectors. It seeks to uncover critical insights into the interaction between land use and transportation networks, addressing several key objectives:
This study aims to address this gap by focusing on the location choice of smaller economic units, known as establishments, within the context of the North American Industrial Classification System (NAICS) sectors. The primary objectives of this research are as follows:
The study employs a discrete choice model to capture establishment location preferences, utilizing data from the state of Tennessee, USA. The results reveal that spatial location determinants can be categorized into four distinct categories: accessibility, neighborhood characteristics, office profile, and the presence of other activities. Among these, agglomeration, land value, office size, square footage, and surrounding land use conditions emerge as the most crucial location determinants. These findings are valuable for transportation planners, providing insights into the intricate connections between establishment locations, demographic conditions, and transportation networks.
Amid the dynamic landscape of economic activity, establishment relocations play a pivotal role in shaping travel patterns and land-use conditions. This research addresses the critical issue of establishments' relocation behavior, particularly in light of recent global events such as the COVID-19 pandemic. The study presents a modeling approach to assess and compare establishment relocations before, during, and after the pandemic. The analysis comprises two key steps: the relocation decision and the relocation action. Data from the state of Tennessee, USA, is leveraged, and a Random Forest classification approach is employed to model both phases.

Key findings from this research are as follows:

- In the post-COVID era, the importance of establishment characteristics in relocation decisions diminished by half, with relocations primarily driven by office profile and accessibility.
- Accessibility remains a significant factor, although its attractiveness reduced by 20.9% in the post-COVID analysis.
- These insights enrich our understanding of establishment relocation behavior and offer valuable information for urban and transportation planners. They provide a nuanced perspective on the impacts of the COVID-19 pandemic, informing policy development and enhancing travel behavior modeling. In summary, this project encapsulates two critical research domains—firm location choice and establishment relocation behavior—providing a holistic understanding of the complex interactions between spatial determinants, industry sectors, and external factors, offering actionable guidance for urban and transportation planning.

This page is intentionally left blank

1.0 WHAT INFLUENCES THE LOCATION CHOICE OF ESTABLISHMENTS? AN ANALYSIS CONSIDERING ESTABLISHMENT TYPES AND ACTIVITIES INTERACTIONS

1.1 INTRODUCTION

transportation planners to understand the interaction between transportation networks and the socioeconomic condition of a region. Although the location choice of businesses indicates the job opportunities and directly affects the travel patterns of workers, decisions made by businesses on where to locate are usually given less consideration than the residential location (Balbontin and Hensher, 2021). Moreover, most studies in the literature modeled the determinant of firms' location choices, while evaluating the location determinants of smaller economic units, referred as to establishments has received less attention (Chin, 2020). An establishment is a district economic unit that produces goods or services at a single physical location, while a firm is a legal entity that consists of one or more establishments under common ownership (Buczkowska, 2017). Due to this structural difference, the decision-making of establishment on location choice would vary compared to firms' location choices. In another word, since firms might consist of multiple establishments, to maximize the benefit, they will consider criteria for decision-making that maximize the benefit of the group, and not necessarily each individual. While modeling establishment location choice will provide this opportunity to evaluate how establishments can maximize their benefits individually. This study aims to explore the determinants of business location choices by targeting establishments.

Moreover, it is important for transportation and urban planners to assess which group of establishments are interacting more with the transportation network and demographic conditions of the neighborhoods (Paleti et al., 2021). Modeling the location choice of establishments provides more in-depth information regarding the decisions of businesses, and capable planners to understand the differences between the location choice of different activities (e.g., farming, warehouses, retail sectors, etc.) (Chin, 2020; Sharma and Mishra, 2022). However, in literature, empirical evaluation of how the location determinant would vary over establishment types is not well-addressed. This might be partly attributed to data scarcity on establishments' physical attributes and detailed information (Kang, 2020). While there is a substantial amount of research devoted to identifying industry-specific location factors, little is known about the influence that establishment type has on the assessment of location criteria (Kimelberg and Williams, 2013a). Moreover, to develop transportation policies and network improvement, it is important to know which type of establishments would interact more with the transportation network, and changes in transportation conditions would have stronger effects on which type of activities. Besides, to understand the decision-making of establishments, it is crucial to understand how they interact which each other, and whether the presence of one establishment would attract or repel other activities (Balbontin and Hensher, 2019). To answer these questions, the current study aims to investigate the location choice of establishments of different types, evaluate how the location determinants vary among different activities, assess the interactions between establishments, and

estimate the importance of accessibility for different establishment types. The rest of this study is organized as follows: section two presents the literature review, literature gaps, and research objective are discussed. Then the methodology applied to develop location choice models and the data collection procedure are provided in section three and four. The model development results are provided in section five, and the location determinants of each NACIS sector are discussed. Finally, the conclusion section presents a summary of the study and avenues for future research.

1.2 LITERATURE REVIEW

Early studies in the field of business location choice indicated the positive and negative factors in the location choice of firms (Pellenbarg et al., 2002). Balbontin and Hensher (2019) provided an overview of the main business location determinants and characterized them into three main categories: accessibility, office profile such as rent, office size, and business profile such as agglomeration. Among all business location choice determinants, transportation planners are more interested in the influence of accessibility. Willigers and Van Wee (2011) showed that the presence of a high-speed train service significantly improves the attractiveness of a location for offices in the Netherlands. Weterings and Knoben (2013) found that a closer distance to a train station has a positive influence on businesses' location choices. Jiang et al. (2018) showed that electronic information manufacturing firms tend to choose areas closer to transportation infrastructure, and the effect of airport accessibility is significant. Moreover, studies in the USA mostly focused on the importance of the accessibility to interstate and highways (Kang, 2020; Yuan, 2021).

Another important location determinant is the surrounding geographical environment (Malecki, 2009). Studies showed that the proximity to knowledge sources and local absorptive capacity is the main location determinant for knowledge-based start-ups (Baptista and Mendonça, 2010), and generally, the availability of appropriate labor and the population density have significant effects on the location choices of business (Holl and Mariotti, 2018). Also, assessing the distribution of warehouses in Los Angeles highlighted the association between minority neighborhoods and warehouse locations (Yuan, 2021). Regarding the interaction between businesses' location choices, to the best of our knowledge, no study has evaluated the effect of this variable. Most studies focused on the effect of agglomeration while the results are varied. Several studies found a positive effect of agglomeration on business relocations, suggesting that businesses are more likely to relocate to areas with more agglomeration or levels of specialization (Ye et al., 2019). However, another study suggested that the presence of the same activity reduces the probability of starting a firm (Backman and Karlsson, 2017).

Few studies have addressed the businesses' location choices at the establishment level. Chin (2020) evaluated the location choice of new establishments by focusing on the relationship between the uniqueness of the certain region and spatially bounded characteristics while the results confirmed the importance of economic, demographic, and geographic conditions at the neighborhood level. Kang (2020) investigated warehousing decentralization by comparing the location choices of warehouses built in 1980 with ones established after 2000 in Los Angeles. Chen et al. (2021) investigated the changes in the spatial distribution of new electronic information manufacturing establishments in China. In a recent study, Ahmed et al. (2022) addressed establishments' intra-firm and inter-firm location choices; where the results showed that establishments from the same firm rather locating farther from one another, while still choosing to co-locate with other establishments from the same industry. Also, van der List (2022) developed a mode of location choice for new establishments in Germany, considering taxes, labor markets, and spillovers. Hawkins and Nurul Habib (2022) developed an establishment location choice model at the individual level and found that professional service establishments tend to locate near passenger rail stations, while industrial establishments tend to locate near major highways.

In literature, studies rarely compare the location determinants of different activities. Kimelberg and Williams (2013) compared the most important location factors for three different industries, office, manufacturing, and retail. Using the data collected from surveys, they found that office respondents are significantly more likely to assign higher ratings to quality-of-life factors,

such as crime rates, amenities, housing, and schools. Sakai et al. (2020) investigated the location factors for logistics facilities considering activity categories. Results showed the importance of accessibility for the group of facilities that serve retail shops and end-consumers industries. Ahmed et al. (2022) compared the location choices of wholesale and retail trade where the results showed that establishments in the wholesale industry tend to locate in lower population density areas partially due to their larger land footprints while retail establishments that sell everyday goods such as grocery stores tend to locate in high population density areas.

1.2.1 Literature Gaps, Objectives, and Contributions

This study aims to address four literature gaps. Although the literature on firms' location choices is rich, analyzing the decision behind the smaller economic unit referred as to establishments has received less attention. Hence, the first objective of this study is to model the location choice of a business at the establishment level (Chin, 2020). Moreover, the literature fails to provide a comprehensive comparison of how the location determinant of business varies across different activities. To address this gap, the second objective of this study is to compare the variation between the location determinants of different activities. In this regard, twenty different activities, categorized by North American Industry Classification System (NAICS) are selected as the classification criteria, and multiple discrete choice models will be applied to these twenty categories to evaluate how the determinants of location choice would vary across NAICS sectors. Also, it is crucially important for transportation planners to understand which type of establishments would interact with the transportation network. Therefore, the third objective of the study is to compare the magnitude of the effect of accessibility between different NAICS sectors and evaluates the difference in the Willingness to Pay (WTP) for better accessibility. In addition, the literature fails to show the effect of interactions between different activities in their allocation choice. Hence, the fourth objective of this research is to evaluate how the presence of one activity would affect the location choice of other activities.

To sum up, this study contributes to the literature by, *first*, modeling the location choice of businesses by targeting establishments of different types; *second*, comparing the location determinants of different activities considering the NAICS sectors; *third*, investigating the interaction between activities by assessing how the presence of one activity would affect the location choice of other activities, and *fourth*, evaluating the importance of accessibility by measuring the WTP of different activities for better accessibility.

1.3 METHODOLOGY

In this study, we use discrete choice modeling to indicate the location determinants of establishments in different NAICS sectors. NAICS categorizes establishments into 20 categories, therefore 20 discrete choice models are developed in this study considering the location attributes (e.g., accessibility, land value, population density, etc.) and individual attributes (e.g., employment, business growth, etc.) as independent variables. Discrete choice models assume that the establishment e_t (where $e_t \epsilon E_t$ and E_t is total establishments of type t) selects the parcel i among a choice set of G sites where $i \in G$, and G is the total number of parcels. The selection of a site can be defined by an unobservable utility function $U_{e_t,i}$, such that, parcel i will be selected over parcel j ($j \in G$) if/only if $U_{e_t,i} > U_{e_t,j}$ ($i \neq j$). The utility ($U_{e_t,i}$) can be formulated as follows:

$$U_{e_t,i} = \beta_i + \sum_n \beta'_{i,n} \times x_n + \sum_m \beta''_{i,m} \times w_m + \varepsilon_i \quad (1)$$

Where β_i is the constant term, x_n is a vector of location i (alternative) attributes, W_m is a vector of the attributes of the establishment e_t, $\beta'_{i,n}$ and $\beta''_{i,m}$ are vectors of the parameters to be estimated using maximum likelihood, ε_i is the error term, and $U_{e_t,i}$ is assumed to be linear. Hence, the probability of selecting an alternative (location/parcel) i by the establishment e_t (which is the general form of the MNL model) can be estimated as follows:

$$P_{e_t,i} = \frac{exp(U_{e_t,i})}{\sum_j exp(U_{e_t,j})} \quad (2)$$

However, when the number of alternatives is large (in this study each parcel in the state of Tennessee can be an alternative), it would be computationally difficult to estimate the model. In addition, it also increases the likelihood that the independence of irrelevant alternatives (IIA) is violated because the unobserved attributes of locations in the same neighborhood are likely to be similar (McFadden, 1977). To overcome this problem (Manski, 1977) proposed a discrete-choice modeling framework incorporating probabilistic choice sets. In this approach, the first step formulates a subset of choice alternatives (C) from the universal choice set (G). This step is referred to as sampling in some studies (Rashidi et al., 2012). The first step can be done by using criteria for selecting the choice set (referred to as labeling) or by random (Ben-Akiva et al., 1985). The actual choice alternatives (C) are unknown; we only observe the chosen alternative (j). With the IIA assumption, parameters can be consistently estimated using only a subset (C) of the alternatives from the universal choice set (G) (McFadden, 1977). In this study, we followed the random sampling approach which is more common in the literature firm location choice, and since it reduces the chance of violating IIA (Kang, 2020). Hence, for each of the chosen alternatives, we randomly selected four alternatives to formulate a choice set (a choice set of 5). To select the choice set size, a trial sample was selected considering 10% of the entire data set, and different models were developed considering choice sets of 2 to 50. It was observed that the beta estimates (coefficients) stabilized for choice sets of 5 and more. Hence, in this study, models are developed considering 5 choice sets (an already selected parcel and four alternatives). Considering the output of the first state, in the second step, conditional on the formulated random choice set (C), an actual

choice of an alternative i is estimated which is the probability that an establishment selects a choice at i is $P_{e_t}(i|C)$. The general model is formulated as follows:

$$P_{e_t}(i) = \sum_{C \in G} P_{e_t}(i|C) P_{e_t}(C) \tag{3}$$

1.4 CASE STUDY AND DATA

As a case study, this study evaluates the location choice of establishments in the state of Tennessee, USA. The population of Tennessee, which comprises 95 counties, was 6,975,218 in 2021, with 4,368,040 of those people working in the state's 315,709 establishments. **Figure 1** illustrates the distribution of different types of establishments in the State of Tennessee in 2021. As **Figure 1** shows, the density of establishments is significantly more in four major cities in the state, Memphis, Nashville, Knoxville, and Chattanooga.

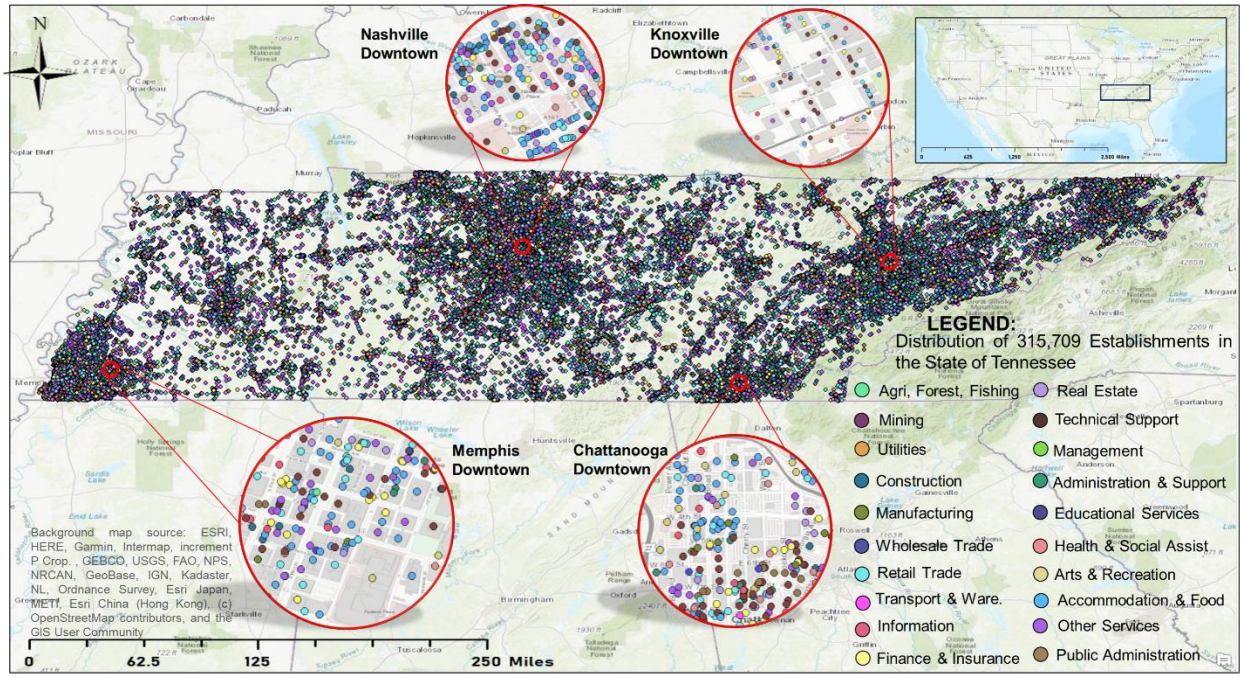


Figure 1: Spatial distribution of different types of establishments in the State of Tennessee in 2021.

In addition to **Figure 1**, the frequency of establishments in each category is presented in **Figure 2**. As this figure shows, health care and social assistance, retail trade, and other services are the top three types of establishments with the highest frequency, and mining, management of companies and enterprises, and utilities are the three categories of establishments with the least frequency in the state of Tennessee.

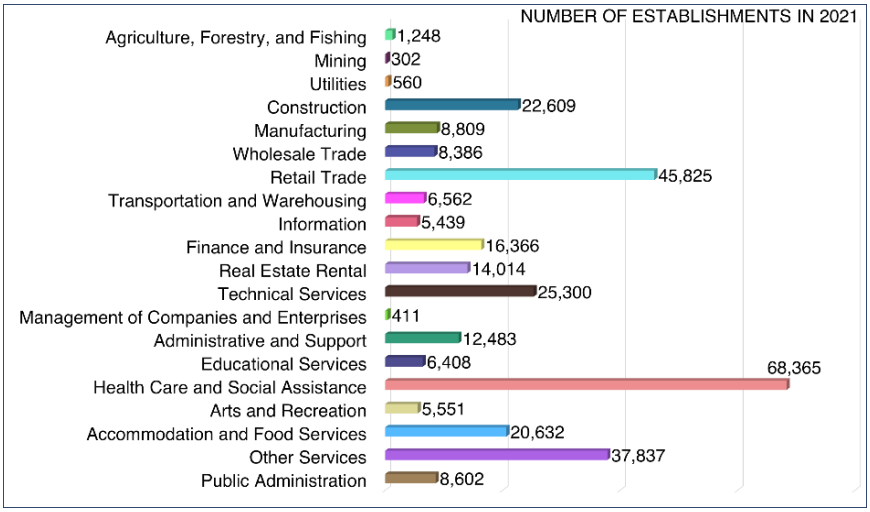


Figure 2: The frequency of establishments of different types in the state of Tennessee in 2021.

To gain a comprehensive understanding of all determinants of establishments' location choices, the following four sources for data were collected as follows:

Establishment's information: Detailed information regarding the establishments in the state of Tennessee is collected from the InfoUSA data set. InfoUSA provides detailed information for companies from local shops to global enterprises. Establishment information is collected from 2018 to 2021providing a panel data set containing details such as NAICS, Standard Industry Code, owner, address, location, office profile, business profile, credit history, business value, employment, headquarters, and franchise information.

Census data: demographic and socio-economic conditions of the surrounding neighborhood of establishments were collected through the US Census Bureau's 2021 American Community Survey (ACS) summary files that provide estimates of population and housing characteristics from 2017 through 2021. Total population, total employment and unemployment, poverty rate, the population of different ages, education and income groups, individual average income, the number of vacant houses, and house price are collected at the block group level. The state of Tennessee has 4,125 block groups, and the data collected are spatially joint to establishments.

Parcel Data: Parcel data refers to a combination of both spatial and nonspatial attribute files, presenting land ownership in a local jurisdiction. Generally, working with parcel data is challenging since the content, currency, structure, and coverage of parcel data sets vary significantly across jurisdictions and regions. These differences create a challenge to obtain a standardized data set (Samani et al., 2022). However, the state of Tennessee provides cleaned and standardized parcel data which is available through the Tennessee Comptroller of Treasury website. Information regarding the land value, building information, land area, and the land use condition of the surrounding neighborhood (i.e., residential, industrial, agricultural, and vacant/developable area) are collected from parcel data.

Transportation network: The transportation network is used to calculate the accessibility of each establishment. In this regard, the distance to the closest interstates entrance, urban highways entrance, all highways (urban and rural), and major arterials are calculated, considering the free flow travel time and the shortest path.

After preparing a cleaned data set, multicollinearity between independent variables is checked to finalize the models' explanatory variables. Independent variables are grouped into four categories: office profile, accessibility, neighborhood characteristics, and the presence of different NAICS sectors. **Table 1** describes the independent variables in detail.

1.5 RESULTS

The results of developing 20 discrete choice models (parameter estimates) are presented in the form of charts in **Figs. 3 and 4.** These figures demonstrate 40 charts, each representing a single explanatory variable, where the y-axis presents the establishment's types (NAICS sectors), and the x-axis presents the value of estimated coefficients for that specific explanatory variable. Presenting the models' results in this form provides a clear view of the difference between the effect of each location determinant (explanatory variables) on the location choices of different establishment types (NAICS sectors). In addition to **Figs. 3 and 4**, details of all developed models are provided in the form of tables, presenting variables' coefficients and *t*-value, in **Appendix B**. Parameter selection (eliminating/keeping variables) is followed considering the *t*-value and the improvement in the goodness of fit measures, AIC and R-squared.

1.5.1 Models' Parameter Estimates

1.5.1.1 Office Profile

The office profile contains three variables: land value, square feet, and office size. As **Figure 3** shows, *land value* showed negative impacts on the location choice of all types of establishments, where the construction has the highest magnitude. This shows that establishments generally tend to select locations with lower land value. However, the effect of land value was not significant for location choices of establishments related to management, utilities, and mining, which can be justifiable due to the type of activities an establishment related to management, mining, or utilities, which mostly depends on the availability of resources. The office's *square feet* showed mixed effects on the location choices. Square feet showed significant positive effects on the location choices of establishments related to public administration, wholesale trade, financial and insurance, information, transportation and warehouse, and agriculture, forestry, and fishing, showing that these types of establishments rather larger places, as many of these establishments require large storages. However, establishments related to real estate, management, health and social assistance, and construction prefer a location with smaller square feet. The *number of offices* showed negative signs for the most type of establishments. Most activities preferred to select locations with a smaller number of offices. Other services, accommodation and food, and real

Table 1: Explanatory variables description.

Variable	Type	Unit	Min	Max	Mean
Office Profile					
Land value	Cont.*	Million $	0.001	133	1.2
Square feet	Cont.	$(Feet)^2$	0	2.8×10^6	11,226
Office size	Cont.	$(Feet)^2$	0	115	10.2
Accessibility					
Interstates[1]	Cont.	Feet	17.57	87.2×10^4	17.5×10^3
Urban Highways[1]	Cont.	Feet	7.89	46.6×10^4	8.59×10^3
All Highway[1]	Cont.	Feet	5.16	46.6×10^4	7.23×10^3
Major Arterials[1]	Cont.	Feet	3.15	15.9×10^4	0.76×10^3
Neighborhood Characteristics					
Population	Cont.	Pop /$(mile)^2$	0	13,370	1,891.9
Unemployment	Cont.	Pop /$(mile)^2$	0	750	50.9
Large Households[2]	Cont.	Pop /$(mile)^2$	0	730	34.9
Highly Educated Population[3]	Cont.	Pop /$(mile)^2$	0	718	32.1
High-Income HH[4]	Cont.	HH /$(mile)^2$	0	3,909	164.2
Poverty Ratio < 1	Cont.	N/A	0	3,302	227.1
Pop < 18 Years Old	Cont.	Pop /$(mile)^2$	0	5,126	458.5
Pop > 65 Years Old	Cont.	Pop /$(mile)^2$	0	2,826	281.2
Commercial Area[5]	Cont.	Percentage	0	0.98	0.158
Industrial Area[5]	Cont.	Percentage	0	0.69	0.051
Agricultural Area[5]	Cont.	Percentage	0	0.989	0.198
Metropolitan[6]	Cat.**	N/A	0	1	N/A
CSA[7]	Cat.	N/A	0	1	N/A
Interaction between Establishments					
Presence of NAICS[8]	Cat.	N/A	0	1	N/A

* Continuous ** Categorical

[1] Travel distance (ft) to the closest entrance is considered

[2] Households with 5 or more members in block group

[3] The number of people with a graduate degree or more in the block group

[4] Households with an annual income of $100,000 or more in the block group

[5] The percentage of parcels with commercial, industrial, and agricultural land-use at the block group

[6] If the location is in a metropolitan area

[7] CSAs are areas where at least 15% of the population from one community will commute to another community for employment or commerce

[8] Presence of each NAICS within a 1-mile distance, 20 binary variables each presenting presence if one sector

estate showed the largest magnitude. However, establishments related to public administration, health and social assistance, and technical services prefer locations with more professional offices, which is justifiable since these types of activities usually have a high number of employees and prefer to have places with more offices.

1.5.1.2 Accessibility

Four variables represent accessibility in this study: distances to the interstate entrance, urban highways, all highways (urban and rural), and major arterials. As **Figure 3** shows, *distance to interstates* showed significant and negative signs in modeling the location choice of establishments related to management, health and social assistance, constructions, other services, retail trade, and mining. These types of establishments tend to select locations close to interstates. These results might be affected by the type of case study as interstates are not stretched all over the state of Tennessee. *Distance to urban highways* showed significant effects on the location choices of wholesale trade, arts and recreation, management, other services, mining, transportation

and warehouses, and health and social assistance. Moreover, the coefficients of *distance to all highways* were significant for utilities, public administration, administrative and support, wholesale trade, construction, and accommodation and food. These results showed the importance of accessibility to highways for establishments that have heavy truck traffic (e.g., wholesale trade, transportation and warehousing, and mining). The last accessibility measure is *the distance to major arterials*, where decisions made by establishments of different types showed more correlation with accessibility. Distance to major arterials showed significant effects on the location choices of wholesale trade, technical services, transportation and warehouse, administrative and support, public admin, manufacturing, construction, retail trade, health and social assistance, and accommodation and food. Compared to other measures of accessibility, distance to major arterials showed the largest magnitudes in modeling establishments' location choices, and among all types of establishments, wholesale trade had the largest coefficient magnitude for distance to major arterials. Generally, it can be inferred that establishments that interact with their customers directly, value the accessibility to major arterials more than other types of activities. Among all types of establishments, only location choice models of construction, wholesale trade, and health and social assistance showed significant coefficients for three variables related to accessibility.

1.5.1.3 Neighborhood Attributes

Various variables related to neighborhood attributes were tested in this study to provide a comprehensive insight into the correlation between the location choice of establishments and the surrounding environment. In addition to the variables provided in **Table 1**, many variables (e.g., employment, population average income, gender, and ethnicity) were eliminated due to multicollinearity or insignificance coefficients. *Population density* is the first variable showing a significant coefficient in the developed models while it had mixed effects across different NAICS sectors. However, as **Figure 3** shows, the magnitudes of its effect are low, and wholesale trade shows the highest negative magnitude. These results are justifiable as establishments related to wholesale trade tend to locate in lower-density areas as they need better access to highways, large square feet, and low land prices. On the other hand, educational services and food and accommodation showed the largest positive coefficient, showing the high interaction of these types of establishments with the neighborhood population.

The *density of unemployment* is the next variable that showed significant effects in modeling establishments related to mining, educational services, health and social assistance, arts and recreation, and other services. Except for health and social assistance, increases in the density of unemployment increase the chance of selecting a location. As **Figure3** shows, establishments related to mining are usually located in areas with high unemployment rates, this large magnitude, shows the interaction between the establishment lotion choice and neighborhood conditions clearly, whereas activities related to mining are usually located in areas with low welfare index.

The *density of high-income households* and the population with a *poverty ratio < 1* are the two tested variables related to the financial condition of the neighborhood. Establishments related to utilities and management will be attracted to the neighborhoods with a higher density of households with high incomes. Moreover, the density of high-income households has negative impacts on the location choice of agriculture. The *poverty ratio < 1* showed significant coefficients for modeling the location choices of establishments related to public administration, other services, accommodation and food, health and social assistance, educational service, administrative and support, technical services, retail trade, construction, and mining. However, expect the coefficient for modeling the location choices of mining, the poverty ratio showed a low magnitude. The large magnitude of the effect of the poverty ratio on mining, emphasizes the fact that establishments

related to mining are located in areas with low welfare indexes. In addition, the *density of large households* was tested which showed significant effects on the location choices of the establishments related to health and social assistance, educational services, and administrative and support, while the magnitude of the effect was low. Also, the *density of highly educated population* was tested which showed significant effects on the location choices of establishments that require access to the skilled and educated population such as health and social assistance, educational services, and technical services, while similar to the density of large households, the magnitudes of the coefficients were low.

The *density of population under 18* and *over 65 years old* were tested to investigate the effect of population age on the location choices of establishments. The population under 18 showed significant coefficients in modeling location choices of mining, educational services, and accommodation and food, where the attractiveness of a location for mining related establishments reduces with increases in the density of the population under 18 while establishments related to educational services would rather areas with a high density of population under 18. This point also can be inferred that educational services attract households with students to be located near them. In addition, the density of the population over 65 showed a mixed effect on the location choices of establishments while mining and management showed the largest negative magnitudes, and the coefficient of the population over 65 was positive for health and social assistance and finance and insurance, two significantly important activities for this age group.

Among all variables related to neighborhood attributes, variables representing the land use conditions showed the largest magnitude on the location choices. The percentage of *commercial areas* showed significant coefficients in most types of establishments, but technical services, finance and insurance, information, utilities, and mining. Also, the effect of commercial areas percentage was positive only for accommodation and food and health and social assistance, and the increases in the percentage of commercial areas reduce the attractiveness of a location for other types of establishments. The percentage of *industrial areas* showed significant negative effects in modeling the location choice of utilities, health and social assistance, other services, and retail trade. These results emphasize the required atmosphere for these types of establishments. For instance, it is understandable that establishments related to health care and social assistance would rather not be close to industrial areas. In contrast, increases in the percentage of industrial areas raise the attractiveness of a location for establishments related to Wholesale trade, transportation and warehouse, and manufacturing. The percentage of *agricultural areas* showed significant positive effects on the location choice of agriculture, transportation and warehouse, wholesale trade, manufacturing, accommodation and food, and retail trade.

Finally, the effect of the type of area was assessed by adding *metropolitan* and *CSA* variables to the models. Establishments related to accommodation and food, arts and recreation, health and social assistance, real estate, retail trade, and construction rather locating in a metropolitan area as the coefficients of the binary variable for metropolitan showed significant positive value. Also, establishments related to other services, technical services, transportation and warehouses, and wholesale trade tend to select a location that is not in a metropolitan area. In addition, the *CSA* area showed significant negative effects on the location choice of establishments related to mining, public administration, health and social assistance, retail trade, and manufacturing, showing that these types of activities usually tend to locate in areas with high transit from other location (mostly far from the downtown). However, positive coefficients were observed in molding the action choices of establishments related to accommodation and food, administrative and support, and technical services, showing that areas with high commute rates

from other neighborhoods showed a positive effect on the location choice of these types of establishments.

1.5.1.4 Interaction between NAICS

The interactions between the location choice of establishments are modeled such that the effect of the presence of a type of establishment is assessed on the location choice of other establishments. **Figure 4** provides the parameter estimated for the coefficients of the interactions between different types of establishments. As this figure shows, the presence of activities showed significant effects on the location choice of most types of establishments. The interesting point is the large positive magnitude of the presence of similar activities on the location choices of establishments of different types. For instance, in modeling the location choice of establishment related to agriculture, forest, and fishing, although the presence of other activities such as manufacturing, retail trade, information, real estate, technical services, educational services, health care and social assistance, arts and recreation, and other services are significant, the magnitude of the effect of the presence of same activities (agriculture, forest, and fishing) is significantly larger than the coefficients of the presence of other types of establishments. In addition, the presence of some types of establishments would repel other establishments to select a parcel close to them. Establishments related to agriculture, mining, and utility are the best examples. On the other hand, in some cases the presence of one type of activity would attract others, for instance, the presence of management would attract establishments related to administration and support.

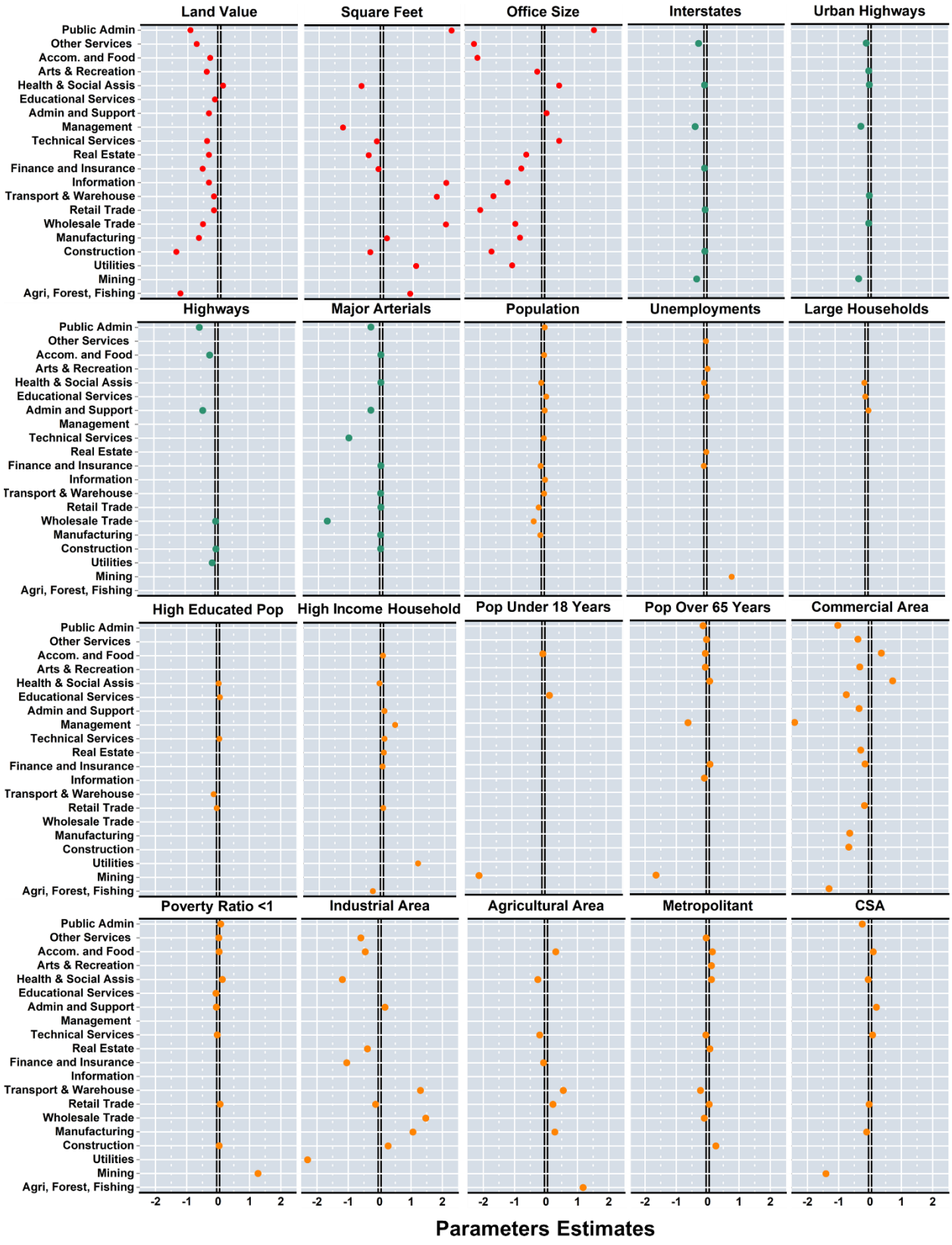


Figure 3: Parameter estimates of the developed model for accessibility, office profile, and neighborhood attributes.

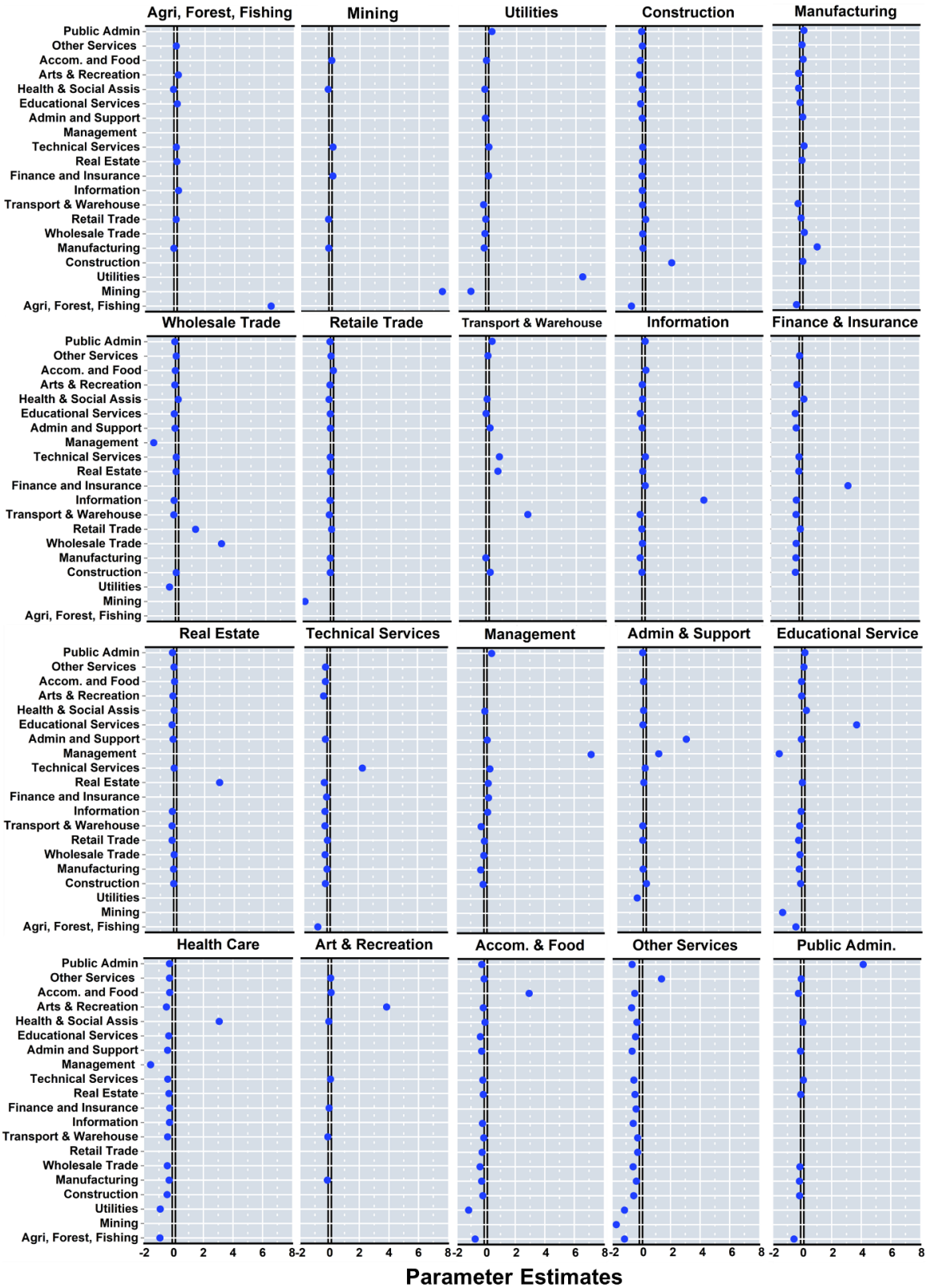


Figure 4: Parameter estimates of the developed model for interactions between establishments.

1.5.2 Elasticity analysis

Applying discrete choice models provides information regarding the significant determinant but cannot show the magnitude of the effect of each determinant in the location choice process. Therefore, we estimate the elasticity for each significant variable. Elasticities are generally calculated to measure the magnitude of a specific variable's impact on outcome probabilities (Samani et al., 2022). Elasticity is estimated from the partial derivative for each observation n as follows:

$$E_{x_{ki}}^{P(i)} = \frac{\partial\, P(i)}{\partial x_{ki}} \times \frac{x_{ki}}{P(i)} \quad (4)$$

where $P(i)$ is the probability of outcome i and x_{ki} indicates the value of variable k for outcome i. By taking the partial derivative, Equation (4) becomes as follows:

$$E_{x_{ki}}^{P(i)} = [1 - P(i)]\, \beta_{ki}\, x_{ki} \quad (5)$$

Where β_{ki} is the coefficient of variable k for outcome i. Elasticity estimated from Equation (5) is only convenient for continuous variables and is not valid for indicator variables. Since our independent variables are mixed of continuous and categorical variables a pseudo-elasticity needs to be calculated to estimate an approximate elasticity of categorical variables. The pseudo-elasticity can be defined as:

$$E_{x_{ki}}^{\lambda_i} = \frac{\exp(\Delta\, \beta_i\, x_i) \sum_{\forall I} \exp(\beta_{kI}\, x_{kI})}{\exp(\Delta\, \beta_i\, x_i) \sum_{\forall I} \exp(\beta_{kI}\, x_{kI}) + \sum_{\forall I \neq I_n} \exp(\beta_{kI}\, x_{kI})} - 1 \quad (6)$$

Where $x_{kI}\ is$ the value of variable k for outcome i, λ_i is the expected frequency for observation i; β_i is a vector of estimable parameters; x_i is a vector of explanatory parameters; I_n indicates the set of alternate outcomes with x_k in the function that determines the outcome, and I is the set of all possible outcomes. Elasticity provided in Eq. 6, is known as direct elasticities because they accurately capture the impact that a change in a variable controlling the chance of an alternate outcome, outcome i, has on the likelihood that outcome i will be selected (Washington et al., 2020). The results of the elasticity analysis are provided in **Figure 5** This figure provides the elasticity analysis result in the form of a heat map such that, the positive effects are indicated in blue color, and negative the effects are presented in red color. In the following subsections, the elasticity analysis of the significant variables is provided. However, for brevity, we focus more on variables that showed an elasticity larger than ±5%.

1.5.2.1 Office Profile

Elasticity analysis for *land value* showed that the largest magnitude belongs to construction, such that 1% increase in the land price, would reduce the probability of selecting a location by construction-related establishments by up to 28%. The importance of land price for construction is understandable due to the fact that it would directly affect the revenue of the establishment. The second sensitive activity to the land price was agriculture which showed -14% elasticity. Manufacturing and wholesale trade are the third and fourth most sensitive activities to the land price, where the elasticity shows -12% and -9.5% respectively. The largest effect of office *square feet* was observed in public administration, such that a 1% increase in the office square feet would increase the probability of selecting a location up to 75%. Then establishments related to information, wholesale trade, and transportation and warehousing showed the largest elasticity (42%, 37%, and 30%). While management with -24% showed the largest negative effect. Generally, establishments that required large space showed high elasticity to land price and square feet. As **Figure 5** shows, other services and real estate with elasticities of -39% and -36%

respectively had the largest negative sensitivity to *the office size*, and public administration had the largest positive elasticity (29%), which shows the required condition for this type of activity.

1.5.2.2 Accessibility

As **Figure 5** shows, establishments related to management showed the largest sensitivity to *distance to the interstate* (-9.41%). After management, health and social assistance with -6.05%, and construction with -4.72% had the highest elasticity. Establishments related to wholesale trade showed the highest sensitivity to the *distance to urban highways*, such that 1% increase in the distance to urban highways reduces the chance of selecting a location by a wholesale trade business up to 7.61%. These results show the specific condition of the wholesale trades where two factors, being in an urban area and having access to highways (due to high traffic of trucks), come to play an important role. Arts and recreation with -5.81% and management with -5.32% had the second and third largest sensitivity to urban highways, respectively. Elasticity analysis showed that only the decisions of establishments related to utilities will be affected by more than 1% with changes in *the distance to all highways*. Finally, *distance to major arterials* showed a -33.5% effect on the location choice of wholesale trade, and the location selected by technical services showed -18.4% affected by 1% increase in the distance to major arterials. Generally, the wholesale trade shows high interaction with transportation networks which will be discussed further.

1.5.2.3 Neighborhood Attributes

Elasticity analysis showed that mining-related establishments have the highest sensitivity to the *density of unemployment* (6.2%). *High-income households* showed strong effects on the location of establishments related to utilities and management whereas the elasticity analysis showed 20% and 8.3% respectively, showing these types of establishments would rather be located in well-established and high-profile neighborhoods. Also, a *low poverty ratio* increases the chance of selecting a location by mining related establishments by 9.4%. Moreover, increases in the *population under 18* reduce the chance of selecting a location for establishments related to mining by 15% and rise the chance of the presence of educational services by 3.4%. In addition, a location with a large *population of over 65* years old has a 12% lower chance to be selected for establishments related to mining and management.

Among all variables related to neighborhood attributes, the surrounding land use conditions showed the largest elasticity in the location choice of the establishments. Elasticity analyses for models of establishments related to management showed that 1% increase in the percentage of *commercial areas* would reduce the chance of selecting a location by up to 55%. After management, agriculture, public administration, constriction, and manufacturing showed the highest elasticity (-20%, -18%, -12%, and -12% respectively). These results are in line with the preferences of these types of establishments on selecting a location with lower prices and larger square feet, as both factors are not usually available in an area with high commercial land use. In contrast, increases in *commercial areas* increase the likelihood of selecting a location by health and social assistance by up to 13%. As **Figure 5** shows, establishments related to utilities showed the largest sensitivity (-41%) and health and social assistance and other services respectively showed -21% and -11% to the percentage of *industrial areas*. This is in line with the nature of these types of establishments, which requires to be far from industrial areas. In contrast, 1% increase in the industrial areas increases the location choice of Wholesale trade, transportation and warehouse, and manufacturing by 26%, 22%, and 19% respectively. These high elasticities can be interpreted as, first, these types of establishments can be categorized into industrial establishments, therefore they tend to be located in the same environment, second, they tend to be close to other

industries to reduce their logistics costs. Elasticity analysis showed that the magnitudes of the effects of *agricultural areas* on location choices of agriculture, transportation and warehouse, wholesale trade, manufacturing, accommodation and food, and retail trade, are 18%, 9.5%, and 8.8% respectively. Finally, if a location is in a metropolitan area, it will have a higher chance to

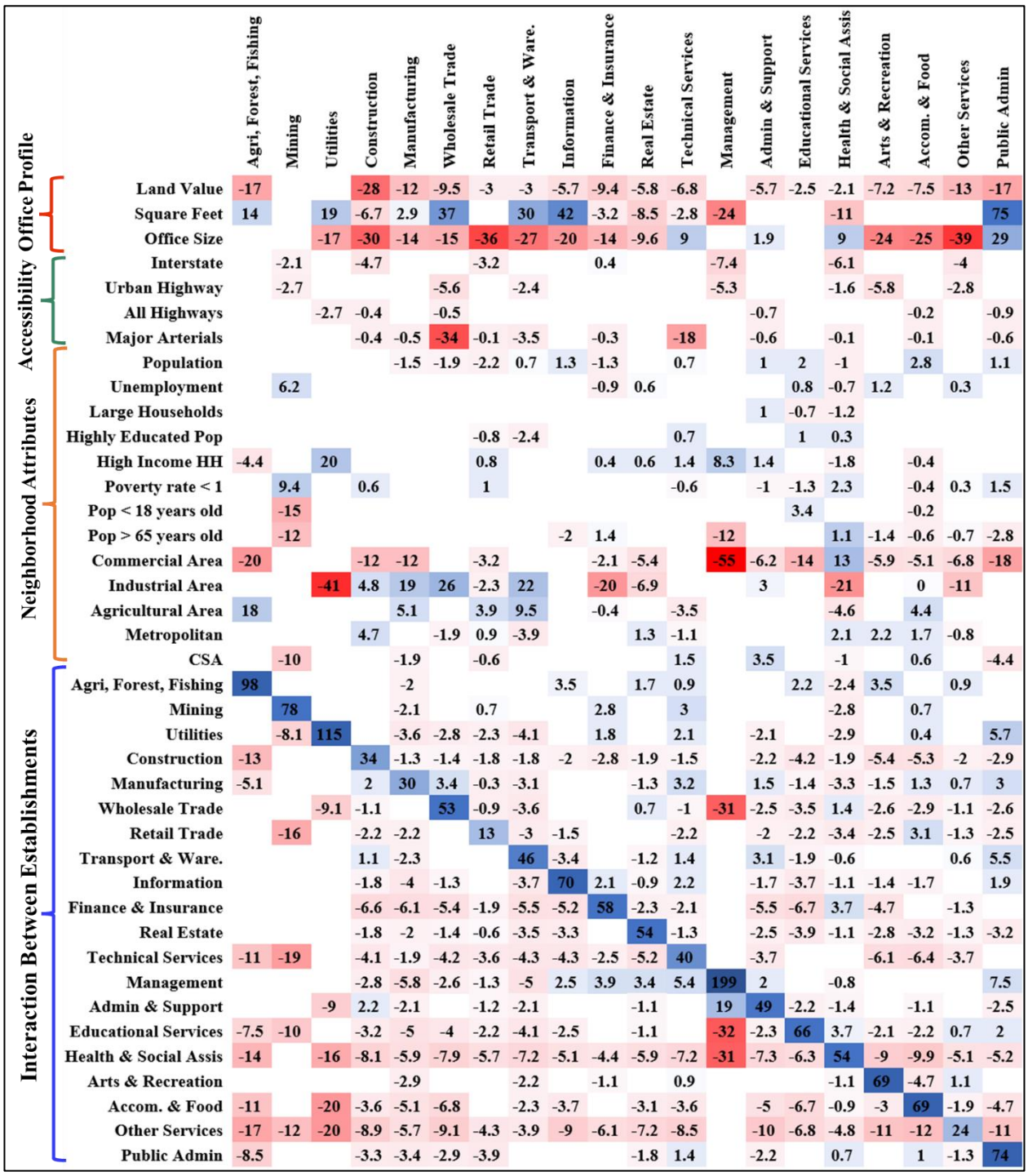


Figure 5: The results of elasticity analysis for all developed models.

be selected by establishments related to construction where the elasticity analysis shows 5.2%, which is understandable as most construction establishments are located in urban areas.

1.5.2.4 Interaction between NAICS

As **Figure 5** shows, the presence of similar activity has a large significant positive effect on the location choices of all types of establishments, such that, the presence of similar activity would increase the chance of selecting a location for establishments related to management by 199%. After management, establishments related to utilities (115%), agriculture, forestry, fishing (98%), mining (78%), and information (70.1%) showed the largest sensitivity to the presence of similar activities. Retail trade showed the lowest sensitivity to the presence of similar activity where the elasticity was 13%. **Figure 5** highlights the significant negative effects of the presence of educational services (32%), health, and social assistance (31%), and wholesale trade (31%) on the location choice of establishments related to management. However, the presence of establishments related to administration and support increases the probability of selecting a location for management by 19%. The interactions between agriculture, forest, and fishing with other services (17%) and construction (13%) were significantly negative. The presence of technical services and retail trade would reduce the probability of selecting a location by mining respectively up to 19% and 16%. The presence of other services, accommodation and food, and health and social assistance would affect the location choice of establishments related to the utility by -20%, -20%, and -16% respectively. The interactions between establishments related to construction, manufacturing, retail trade, transportation and warehousing, information, finance and insurance, real estate, and technical services with other establishments were relatively weak as they all show less than 10% effects. Lastly, the presence of an establishment in the category of other services reduces the probability of selecting a location by admins and support (-10%), arts and recreation (-11%), accommodation and food (-12%), and public administration (-11%).

1.5.3 Willingness To Pay

Estimating the willingness to pay (WTP) is one of the most important behavioral post-analyses of choice studies. WTP reveals how much a decision-maker (here an establishment) is willing to pay for an improvement in another attribute. In this study, WTP is incorporated to evaluate the importance of accessibility to the transportation network and the presence of different activities for different types of establishments and is calculated considering the land value. Considering **Eq. 1**, the WTP for attribute m can be estimated using **Eq. 7** (Breidert et al., 2006):

$$WTP_{m,i} = \frac{\partial U_{e_t i} \big/ \partial\, x_{m,i}}{U_{e_t i} \big/ \partial\, x_{LV,i}} = \frac{\beta_{m,i}}{\beta_{LV,i}} \tag{7}$$

where $x_{LV,ni}$ refers to the land value attribute and $\beta_{LV,ni}$ indicates the coefficients of the land value attribute. One of the main goals of this study is to understand which types of establishments tend to pay more for better accessibility to transportation networks. In this regard, the WTP of different types of establishments for better accessibility to interstates' entrances, urban highways' entrances, all highways, and major arterials, were calculated for establishments that showed significant coefficients for both land value and accessibility attributes. **Figure 6** presents the results of calculating WTP for better accessibility to the transportation network for different types of establishments.

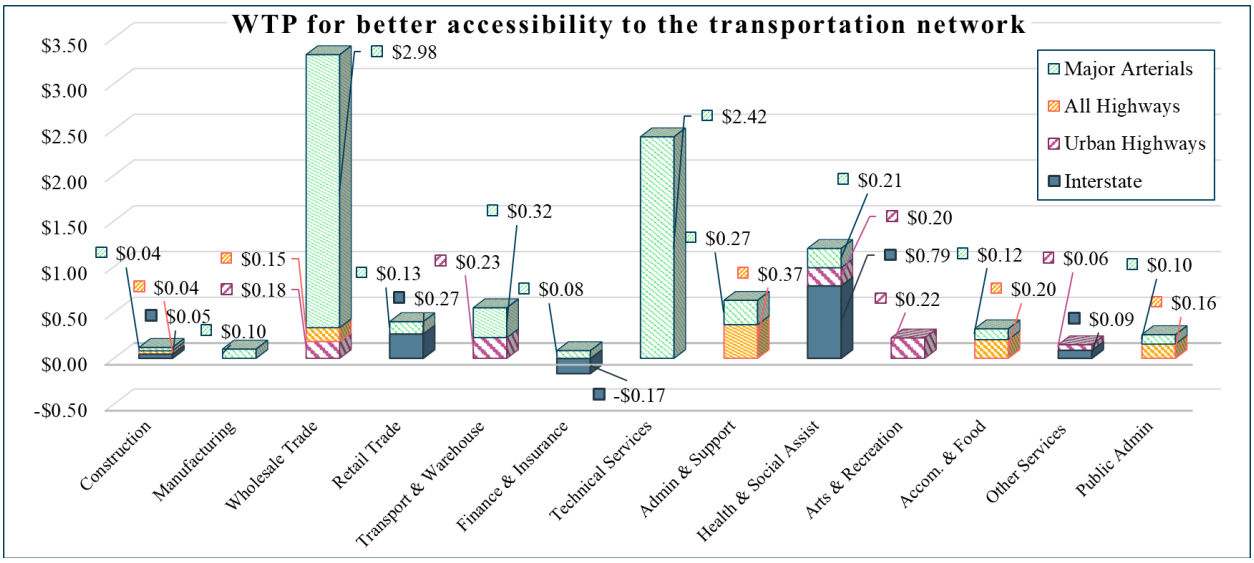


Figure 6: Willingness to Pay (WTP) for better accessibility of different establishment types.

As **Figure 6** shows, the WTPs of the wholesale trade and technical services section for better accessibility to major arterials are relatively high; such that, respectively, they are willing to pay $2.97 and $2.42 for a location that is one foot closer to a major arterial. Moreover, WTP for access to a major arterial is the most repeated in **Figure 6**, showing that it is very important for all types of establishments to have better accessibility to major arterials, especially for types of activities that are in touch with their customers directly, such as wholesale trade and technical services. Wholesale trade is the only activity that showed positive WTP for better accessibility to both urban and all highways. Establishments related to health and social assistance showed the highest WTP for better accessibility to interstate entrances ($0.79), showing how accessibility to freeways and highspeed corridors is important for this type of establishment. Also, since establishments related to health care and social assistance are distributed usually in big cities (please see Appendix A), shorter distances to interstates provide better access to health care related facilities for smaller cities and suburban neighborhoods. Retail trade has the second highest WTP for interstate access, where WTP shows $0.27 for a location one foot closer to an interstate entrance. Establishments related to administration and support showed the highest WTP for a location closer to all highways ($0.37 for each foot). Establishments related to finance and insurance showed the only negative WTP for better accessibility ($-0.17). As **Figure 6** shows, in this type of establishment, a closer distance to interstates reduces the attractiveness of a location for selection. This result is understandable as finance and insurance usually do not have heavy truck traffic, and on the other hand, other results showed that they tend to be in areas with high elderly population density and residential areas, which usually are not close to interstates. Moreover, establishments related to wholesale trade, transportation and warehousing, health and social assistance, arts and recreation, and other services, show WTPs of less than $0.25 for better access to urban highways. To sum up, considering WTPs for all types of accessibility, wholesale trade, technical services, and health care and social assistance showed the highest total WTP to have better access to the transportation network.

In addition to estimating WTP for better accessibility, WTP is estimated for the presence of different types of activities, in the form of a heat map in **Figure 7**. In this figure, in addition to the value of WTP, green and red colors are assigned to positive and negative WTPs correspondingly. In order to make **Figures 7** and **8** comparable, the WTPs for the effect of the presence of different types of activities are calculated for each foot, hence a unit change (from mile to foot) is applied to **Eq. 7**. **Figure 7.** should be interpreted as, the amount of money each type of establishments (columns) is willing to pay to be one foot closer to establishment types listed in the rows. For instance, an establishment related to agriculture, forestry, and fishing is willing to pay $98.8 to get one foot closer to an establishment with the same type and the WTP for a location reduces by $13 for each foot getting closer to an establishment related to construction. Similar to the results of conducting elasticity analysis, **Figure 7** emphasizes the strong effect of the presence of similar activity in the neighborhood. The highest WTP belongs to educational services where they are willing to pay $488 for each foot getting closer to a place where the same activity exists. The second and third highest positive WTPs belong to health and social assistance and wholesale

	Agri, Forest & Fishing	Construction	Manufacturing	Wholesale Trade	Retail Trade	Transport & Warehouse	Information	Finance & Insurance	Real Estate	Technical Services	Admin & Support	Educational Services	Health & Social Assist	Arts & Recreation	Accom. & Food	Other Services	Public Admin
Agri, Forest, Fishing	$98.8		-$3.2		$4.5		$11.5		$5.2	$2.4		$16.2	-$22.1	$9.0		$1.3	
Mining			-$3.5		-$14.5			$5.7		$8.3			-$23.7		$5.9		
Utilities			-$5.8	-$12.9	-$11.1	-$26.7		$3.2		$5.8	-$6.9		-$25.3		-$3.3		$6.6
Construction	-$13.0	$22.7	-$2.0	-$7.7	$13.4	-$11.1	-$6.3	-$4.6	-$5.7	-$4.4	-$6.9	-$31.1	-$17.4	-$14.2	-$15.7	-$2.9	-$3.3
Manufacturing	-$5.0	$1.2	$31	$17.2	-$2.2	-$24.5			-$2.3	$8.7	$5.2	-$12.2	-$30.0	-$5.7	$7.8	$1.1	
Wholesale Trade		-$0.7		$259	$139	-$24.5	-$11.5		$4.0	-$2.9	-$8.0	-$25.7	$12.6	-$7.1	-$7.8	-$1.6	-$3.1
Retail Trade		-$1.4	-$3.5		$83.6	-$20.1	-$7.5			-$5.8	-$6.3	-$16.2	-$31.6	-$6.6	$5.9	-$1.8	-$2.9
Transport & Ware.		$2.2	-$3.5			$303	$4.6		-$4.0	$3.9	$9.8	-$13.5	-$4.7				$6.4
Information		-$1.2	-$6.4	-$6.0	-$12.3	-$24.5	$233	$4.3	-$2.9	$6.3	-$5.2	-$28.4	-$9.5	-$3.8	$10.4		$2.3
Finance & Insurance		-$4.3	-$9.6	-$26.7	-$3.3	-$35.7	-$17.2	$114	-$7.5	-$5.8	-$17.8	-$50.1	$33.1	-$12.3		-$2.1	
Real Estate		-$1.2	-$3.2	-$6.0	-$23.4	-$23.4	-$10.9		$172	-$3.4	-$8.0	-$28.4	-$11.0	-$7.1	-$2.6	-$2.1	-$3.7
Technical Services	-$11.1	-$2.7	-$3.2	-$20.7	-$8.9	-$27.9	-$14.3	-$5.0	-$16.6	$110	-$12.6			-$16.1	-$14.4	-$5.5	
Management		-$1.9	-$9.3	-$9.5	-$7.8	-$32.3	$8.6	$7.5	$10.9	$14.6	$6.9		-$6.3				$8.6
Admin & Support		$1.4	-$3.2		-$14.5	-$14.5			-$3.4		$161	-$16.2	-$12.6		-$5.9		-$2.9
Educational Services	-$7.6	-$2.2	-$7.9	-$18.9	-$35.7	-$26.7	-$8.6		-$3.4		-$7.5	$488	$33.1	-$5.2	-$7.8	$1.1	$2.3
Health & Social Assist	-$14.4	-$5.4	-$9.3	-$37.9		-$46.8	-$16.6	-$10.0	-$18.9	-$19.9	-$24.1	-$46.0	$484	-$23.7	-$18.3	-$7.6	-$6.0
Arts & Recreation			-$4.7			-$14.5		-$1.4		$2.4			-$9.5	$181	$5.9	$1.6	
Accom. & Food	-$11.0	-$2.4	-$8.2	-$33.6	-$27.9	-$15.6	-$12.6		-$9.8	-$9.7	-$16.1	-$50.1	-$7.9	-$8.0	$192	-$2.9	-$5.6
Other Services	-$17.0	-$5.9	-$9.0	-$45.6	-$24.5	-$24.5	-$29.8	-$11.8	-$23.0	-$23.3	-$34.4	-$50.1	-$42.6	-$29.8	-$27.4	$36.1	-$12.4
Public Admin	-$8.6	-$2.2	-$5.5	-$13.8					-$5.7	$3.9	-$6.9		$6.3		-$17.0	-$1.8	$85

Figure 7: Willingness to Pay (WTP) for the presence of different types of activities.

trade where these two establishment types are willing to pay $484 and $259 to be located next to an establishment with the same type. In addition to the importance of agglomeration, **Figure 7** highlights the importance of the presence of educational services on the WTP of establishments related to health and social assistance (WTP is $33.1). On the other hand, the presence of some activities would reduce the WTP of an establishment to be located close to them. The largest negative WTP belongs to educational services, such that one foot closer to establishments related to finance and insurance, accommodation and food, and other services reduces the WTPs by $50.1. Moreover, the presence of establishments related to health and social assistance reduces the WTP of establishments related to wholesale trade by $46.8.

1.6 IMPLICATIONS FOR RESEARCH AND PRACTICE

The results of modeling the location choices of different types of establishments highlighted the importance of the presence of the same activity in selecting a location, which can be related to agglomeration. The evaluation of the influence of agglomeration on location/relocation choices of establishments is well-addressed in the literature and the results of this study are in line with Guimaraes et al. (2003), Gabe and Bell (2004), De Bok and Van Oort (2011), Lee and Hwang (2016), Wu et al. (2019) and Ye et al. (2019) who showed the positive effect of the presence of similar activities in selecting a location. For establishments that required a specific source, i.e., agriculture, forestry, and fishing (NAICS 11) and mining (NAICS 21), the presence of more than one establishment of the same type is obvious due to the need for a specific source. This study showed that the presence of the same activity showed the most positive parameter in attracting an establishment to select a location for all establishment types, except public administration. In addition, among all types of establishments, retail trade showed the lowest sensitivity to the presence of similar activity, showing less interest in competition in this type of establishment compared to others. In literature, Backman and Karlsson (2017) stated that the presence of the same business will reduce the likelihood of new firms' location choices.

Moreover, the presence of other services (NAICS 81) and health care and social assistance (NAICS 62) in a location would significantly reduce the interest of all other types of establishments to select a parcel close to them. In addition, the interactions between establishments related to agriculture, forestry, and fishing (NAICS 11), mining (NAICS 21), and utility (NAICS 22) were significantly negative which shows the nature of these activities and can be inferred that these types of establishments would rather be far from other activities. The presence of logistic facilities, which is categorized under NAICS 48-49, transportation and warehousing, showed positive effects on the location choices of establishments related to construction (NAICS 23), technical services, administrative and support (NAICS 56), other services (NAICS 81), and public administration (NAICS 92), which compared to literature, we expected to observe a more significant effect. For instance, Sakai et al. (2020) incorporated the relationship between logistic providers, industrial logistics, and distributors, and showed the significant effects of closer distance to logistic facilities on the location choices of firms.

The neighborhood attributes, which contains a combination of demographic and land use condition, had significant effects on the decision made by establishments. The percentage of commercial, industrial, and agricultural areas in the block group, which is showing what type of land use is more prevalent in the neighborhood, showed large significant effects on the location choices compared to other neighborhood-related determinants. This is emphasizing the importance of policy-makers decisions in assigning a specific land use to a neighborhood and designing the growth plan in attracting or repelling activities. In other words, the dominant land use in an area would affect the location choice of establishments of all types. In literature, the effect of land use conditions was evaluated in terms of land use diversity (Limtanakool et al., 2006) or the degree of land use (Bodenmann, 2004) while the results were highly dependent on the case study.

Moreover, interpreting the results of the developed model showed the independency between demographic conditions and location choice of establishments. In the models of the location choices of establishments related to mining (NAICS 21), the selection of locations with high unemployment rates and high poverty rates, and low populations under 18 and over 65 was significant. Also, interpreting the effects of population age shows the correlation between establishments related to educational services (NAICS 61) and the population under 18, and the

connection between the population over 65 and establishments related to health care and social assistance (NAICS 62) and finance and insurance (NAICS 52).

The results of current studies supported the results of the study conducted by (Bodenmann and Axhausen, 2012 and Hensher et al. (2017) and who stated that the population with graduate degrees has a positive effect on firms' location choice. We showed that establishments related to educational services (NAICS 61), technical services (NAICS 54), and health care and social assistance (NAICS 62) are located close to areas with a high educated population. Also, the results of this study are in line with the findings of Chin (2020), who showed the importance of economic, demographic, and geographic conditions at the neighborhood level. Moreover, the results of this study showed significant interactions between the neighborhood environment and establishments' location choices which suggests considering establishment decision-making on the integrated land use transport models. A good example could be the interaction between the neighborhood's properties and the location choice of establishments related to mining (NAICS 22). In recent years, (Hensher et al., 2019) proposed an integrated land-use transport model that incorporates the simultaneous locations of firms and jobs. The results of the current study can be incorporated into the land use transport model development.

A surprising result of this study was the little effect of accessibility on the location choice compared to the effects of office profile and neighborhood conditions. The distance to major arterials was the most significant accessibility measure, and the accessibility of interstate and highways did not show a very large effect on the decision made by establishments. The results of evaluating the effect of distance to interstates are in line with Gabe and Bell (2004) who showed the negative effect of distance to interstates on the number of businesses investing per location at the municipality level. Accessibility measures were only among the top three important location determinants for establishments related to wholesale trade (NAICS 42) and technical services (NAICS 54). One possible reason behind this small effect can be related to the study area since interstates and highways are not passing through all counties and cities in the state of Tennessee.

To sum up, **Table 2** provides the top three most important positive and negative location determinants for each type of establishment, which provide a general view of the difference between the location determinants of different activities. As this table shows, the presence of the same activity, office size, land value, land use conditions, and presence of establishments related to NAICSs 62 and 81 are the most repeated location determinants. **Table 2** shows that the presence of similar activity is the most important positive parameter in establishments' location choices, except for public administration (NAICS 92), where the square feet is the first positive parameter. Also, this table emphasizes the importance of land-use conditions, as the percentage of agricultural, commercial, and industrial areas appear as the top three important parameters in the majority of activity types. The land value is also one of the most repeated negative parameters in the location choices of different types of establishments. Distance to transportation network only appears in three types of establishments. The distance to major arterials is the most important negative parameter in the location choices of establishments related to wholesale trade (NAICS 42) and technical services (NAICS 54), and the distance to interstates entrance is the third most important variable in the location choices of health care and social assistance. Locating in a metropolitan area appears only for establishments related to construction (NAICS 23) and is the second most positive parameter. In addition, location in a CSA area is the third positive factor in the location choices of establishments related to real estate (NAICS 53) and administrative and support (NAICS 56). **Table 2** can also help transportation planners by providing important variables required for developing an integrated land-use transportation model. The location

determinants provided in this table can be further incorporated into modeling and estimating the number of job opportunities created, as suggested by Hensher et al. (2019).

To help the transportation planner to understand the importance of accessibility for establishments, this study evaluated the WTP for better accessibility. Wholesale trade, technical

Table 2: Top three strongest positive and negative location determinants for each type of establishment.

Establishment	Positive	Negative	Establishment	Positive	Negative
Agriculture, Forestry, Fishing (NAICS 11)	1. Similar activity* 2. Agricultural area 3. Square feet	1. Commercial area 2. Land value 3. NAICS 81	**Real Estate (NAICS 53)**	1. Similar activity 2. NAICS 51 3. CSA	1. Office size 2. Square feet 3. NAICS 81
Mining (NAICS 21)	1. Similar activity 2. Poverty rate 3. Unemployment	1. NAICS 54 2. NAICS 44-45 3. Population < 18	**Technical Services (NAICS 54)**	1. Similar activity 2. Office size 3. NAICS 31-33	1. Major arterials 2. NAICS 81 3. Land value
Utilities (NAICS 22)	1. Similar activity 2. High-income HH 3. Square feet	1. Industrial area 2. NAICS 81 3. NAICS 72	**Management of Companies (NAICS 55)**	1. Similar activity 2. NAICS 56 3. Office size	1. Commercial area 2. Square feet 3. Population > 65
Construction (NAICS 23)	1. Similar activity 2. Metropolitan 3. Industrial area	1. Office size 2. Land value 3. Commercial area	**Administrative & Support (NAICS 56)**	1. Similar activity 2. Presence of 72 3. CSA	1. NAICS 81 2. NAICS 62 3. Land value
Manufacturing (NAICS 31-33)	1. Similar activity 2. Industrial area 3. Agricultural area	1. Office size 2. Land value 3. Commercial area	**Educational Services (NAICS 61)**	1. Similar activity 2. Population < 18 3. Total population	1. Commercial area 2. NAICS 81 3. NAICS 51
Wholesale Trade (NAICS 42)	1. Similar activity 2. Square feet 3. Industrial area	1. Major arterials 2. Office size 3. Land value	**Health & Social Assistance (NAICS 62)**	1. Similar activity 2. Commercial area 3. Office size	1. Industrial area 2. Square feet 3. Interstates
Retail Trade (NAICS 44-45)	1. Similar activity 2. Agricultural area 3. Commercial rea	1. Office size 2. NAICS 62 3. Land value	**Arts & Recreation (NAICS 71)**	1. Similar activity 2. NAICS 11 3. Metropolitan	1. Office size 2. NAICS 81 3. Land value
Transport & Warehousing (NAICS 48-49)	1. Similar activity 2. Square feet 3. Industrial area	1. Office size 2. NAICS 62 3. Land value	**Accommodation & Food (NAICS 72)**	1. Similar activity 2. Agricultural area 3. Population	1. Office size 2. NAICS 81 3. Land value
Information (NAICS 51)	1. Similar activity 2. Square feet 3. NAICS 11	1. Office size 2. NAICS 81 3. Land value	**Other Services (NAICS 81)**	1. Similar activity 2. Unemployment 3. NAICS 71	1. Office size 2. Land value 3. Commercial area
Finance & Insurance (NAICS 52)	1. Similar activity 2. NAICS 21 3. NAICS 51	1. Agricultural area 2. Office size 3. Land value	**Public Administration (NAICS 92)**	1. Square feet 2. Similar activity 3. Office size	1. Commercial area 2. Land value 3. NAICS 81

* Presence of establishment with the same NAICS code

services, and health and social assistance are the most important establishments categories that transportation planners should focus on as the WTP of these types of establishments for better accessibility to major arterials and the interstate was much more than other types of establishments. Moreover, in general, the importance of major arterials over other types of roads, especially for establishments that are directly in touch with their customers, needs to be considered in their decision-making and budget assignments.

1.7 CONCLUSION

This part of study aimed to understand and evaluate the location choice of establishments, the smallest economic unit, to assess how the location determinants would vary across different establishment types. A discrete choice model was applied to model the location choice, where first, the choice sets (alternatives) are modeled, and then the actual choice of each establishment is modeled, using a multinomial logit model. Using the data collected for the state of Tennessee, models were developed to indicate the significant parameters in the location choice of establishments.

Then elasticity analysis was conducted to evaluate the magnitude of each significant parameter. The location determinants of establishments are classified into four categories, office profile, accessibility, neighborhood attributes, and interaction between establishments, and showed that the location determinants of establishments vary across different NAICS sectors. Elasticity analysis showed that the presence of the same activity, land value, office size, square feet, and land use conditions are the most important and most repeated location determinants of establishments of different types. Moreover, the presence of establishments related to other services (NAICS 81) and health and social assistance (NAICS 62) had consistent negative effects on the location choice of others as the results showed that other types of establishments do not prefer to select a location close to these two types of establishments.

Among accessibility variables, the distance to major arterials showed a significant effect on the location choices. To indicate the importance of accessibility for establishments, willingness to pay was calculated for different types of activities. Results showed that establishments tend to pay more for better accessibility to major arterials, specifically establishments related to wholesale trade (NAICS 42) and technical services (NACIS 54). Moreover, better accessibility to interstates was significant for establishments related to health and social assistance. The importance of major arterials over other types of roads needs to be considered in transportation planners' decision-making and budget assignment as this research showed and compared the significant correlation between the distance to major arterials and other types of roads.

This study is conducted under some limitations. Due to data availability, to model establishment location choices this study had to assume that establishments make their decision individually and independently. While an establishment can be a member of a firm or a franchise. Therefore, in the real world, the decision-making process of this establishment is not occurring independently and depends on the firms' policies and strategies.

Future studies can involve other variables such as crime rate or other office conditions. Due to the scale and the condition of the study area, this research did not consider the effect of accessibility to public transportation, hence future studies can incorporate the accessibility to public transport due to the important role it plays in residential and business location choice. The application of other modeling approaches e.g., latent class models, hybrid models, and/or multilevel models can reveal more details regarding the preference of establishments in location choice. The current study modeled establishments' location choices considering 2-digit NAICS sector classification. Considering the possible significant heterogeneity among establishments with the same 2-digit NAICS sectors, future studies can investigate modeling establishments' location choices at a finer level (e.g., 3-digit NAICS code). This study investigated the effect of the presence of different types of activities within one mile on the location choice of establishments. Two points should be mentioned here, first, 1 mile was arbitrary, and we decided

to go with one unit of the distance, second, binary variables were considered for the presence of each type of activity to specifically target the effect of the presence of different types of establishments, and regardless of the number of establishments. Therefore, further investigation can be applied to provide more insight into the interaction between different types of establishments, considering their numbers, types, and logistic policies and approaches. In this study, we incorporated 4 parcels as the choice alternatives for each establishment (creating choice sets of 5), future studies can conduct robustness analysis for selecting the optimal number of choice sets, but doing so in our research was out of the scope. Finally, since this research considered a statewide area as the case study, applying the model to a bigger (national level) or smaller (county or city) study area might lead to different results. Also, the type of variables can be changed, e.g., one important factor in encouraging or discouraging establishments is the national business policy or tax policy which needs to be considered in analyzing the location choices of establishments on a larger scale.

2.0 EVALUATING RELOCATION BEHAVIOR OF ESTABLISHMENTS: EVIDENCE FOR THE SHORT-TERM EFFECTS OF COVID-19

2.1 INTRODUCTION

Establishment relocations, in addition to redistribution of economic activities, affect the travel pattern in short-term and transportation network and land-use conditions in long-term. Studies showed that even a short-distance, intra-city company relocation can disrupt employees' daily routines and reshape their own and other people's mobility (Ding et al., 2017; Rau et al., 2019). Hence, assessing the relocation behavior of businesses, why establishments relocate, and what are the spatial relocation determinant, are important questions for travel behavior studies and land-use modeling. Due to rapid economic development, the world is witnessing an ever-increasing competition of businesses to gain more profits which are associated with relocation in some cases. In addition, many variables such as, less revenue than expected, changes in land-use conditions, governmental policies, environmental conditions, can lead to establishments relocations (Nilsen et al., 2020). Considering variety of effective parameters, including the attitudes and preferences, and since the decision-making mechanism are typically not observed, modeling establishments' relocation behavior is challenging (Balbontin and Hensher, 2021; Paleti et al., 2020). This study aims to propose a modeling approach to effectively assess the relocation behavior of establishments.

Generally, studies on establishments' location and relocation choices aimed to assess the effective parameters (determinants) in their decision-making. These decisions are typically modeled using discrete choice models, such as random utility maximization which the decision-maker is assumed to choose the alternative that provides the highest utility (Haque et al., 2021). Within the class of discrete choice models, the multinomial logit (MNL), and its generalizations (e.g., nested logit, cross nested logit, etc.) are commonly used to analyze business location/relocation choices (Paleti et al., 2020). However, the MNL model has several limitations because it assumes that the probability of each alternative is independent of the features of the rest of the alternatives (Pineda-Jaramillo, 2019). Furthermore, MNL fails to capture complex non-linear relationships between dependent and explanatory variables. To overcome these challenges, the implication of an ensemble machine learning approach, Random Forest (RF), will be explored to model the effect of establishments' characteristics (i.e., size, value, age, type, business changes, etc.), office profile (e.g., land value, office size, and square feet), environments conditions (e.g., population density, agglomeration, rent, and, land-use mix), and accessibility (i.e., distance to interstates, highways, and major arterials) on establishments' relocation behavior.

The world has been struggling with Covid-19 for the last few years, and the pandemic has had irrefutable effects on every part of society. Besides, changes in the supply chain, working from home, e-shopping, virtual meetings, and other forms of online activities are more common which have affected establishments' policies and approaches. Although researchers have been investigating the effect of Covid-19 in different areas, evaluating the responses of establishments

to these changes during and after the pandemic is not addressed yet, while understanding these responses is important for travel behavior analysis and land-use modeling. Hence, this study addresses this concern by comparing the relocation behavior of establishments, before, during, and after the pandemic and capturing the changes in spatial relocation changes in post-Covid conditions.

To sum up, this study contributes to the literature by proposing a modeling approach for assessing the relocation behavior of establishments and investigating the changes in the decision-making of establishments before, during, and after the pandemic by developing different models and applying statistical analysis. The rest of this study is organized as follows: section two provides the literature review of business relocation studies and modeling approaches; sections three and four discuss the methodology, case study, and data collection procedure; section five illustrates the results of the model calibration, performances, and results; and finally, sections six and seven provide discussion, conclusion, and avenues for the future studies.

2.2 LITERATURE REVIEW

The literature suggests that there is a clear distinction between factors affecting the choice of the "initial" and "readjusted" location (Balbontin and Hensher, 2019). First-time locating firms have significantly different preferences than those of relocating firms (Elgar et al., 2015). The literature on business relocation decisions suggests that key factors can be separated into two categories: push factors and pull factors (Pellenbarg et al., 2002). The push factors refer to situations that motivate the company to move; and the pull factors refer to parameters of the new location, which attract a company. An early study showed that two types of motives are associated with relocation decisions: (i) the need to reduce operating costs; and (ii) the access to an appropriate workforce and opportunities for collaboration (Kimelberg and Williams, 2013b). Rossi (2019)discussed factors that attract firms to move to Ticino, Switzerland, and found that institutional environments, good accessibility, and availability of suitable production inputs are the most important factors. Also, another study addressed the relationship between the migration tendency of firms and a set of relocation factors, where the results showed positive significant association between transportation supply and firm-level relocation decisions (Targa et al., 2006). Based on the data collected from commercial registers of three Swiss Cantons, distance to the previous location, local taxes, and cantonal business development strongly influence relocating firms' destination choices. A study on relocation of the electronic and information industry in China showed that labor costs, market scale, land rent, transport cost, and preferential policy are important relocation determinants (Y. Jiang et al., 2018). The assessment of firms' relocation in Norway showed that relocation decision is influenced by a firm's internal and external characteristics, and firms seem to have different preferences regarding what makes a location attractive or not (Nilsen et al., 2020).

To model business relocation decisions, various methods are applied, while the majority of studies applied random utility maximization models. A nested logit model was applied to analyze the decisions of manufacturing establishments in South Korea on whether and where to relocate (Yi, 2018). In another study, an MNL model was developed to model firms' relocation choices in Seoul metropolitan areas. Results showed that the network distance for relocation is an important factor for manufacturing firms (An and Wan, 2018). Linear regression was applied to estimate the number of firm relocations in the Netherlands (Risselada et al., 2013). Cox regression with time-varying covariate was incorporated to estimate separate models for different relocation distances (Weterings and Knoben, 2013). An econometric model based on a large set of data on firms in China, considering global, regional, and local factors in developing the industrial relocation model. where results showed that the factors that affect the choice of location are largely determined by the specific capacity of the company (Zhu et al., 2017). To the best of our knowledge, although machine learning methods are incorporated to model residential decisions, such as location (Scheuer et al., 2021), relocation (Xue and Yao, 2022), mode choice (Cheng et al., 2019), and travel pattern (Alkhereibi et al., 2021), these methods have not been applied to model establishments location/relocation decisions.

2.3 METHODOLOGY

A relocation behavior of an establishment consists of two parts: 1) whether to move or not; and 2) where to move. In this study, the first part models the relocation of an establishment considering the information of the year that led to the relocation decision. In other words, if an establishment is relocated in 2020, the information collected about this establishment in 2019 will be modeled to provide details on why this establishment is decided to move. Thus, the dependent variable of this part is a binary variable that is equal to one if the establishment is relocated and otherwise it is zero. In this study, the modeling of the first part is referred to as "relocation decision" models and the effective variables are referred as to "push factors". In modeling the second part, the spatial relocation choice of an establishment is modeled considering the information of both current (relocated to) and previous (relocated from) locations. This part aims to model the effective variables in location selections of relocated establishments and is titled "relocation action" in this study. The dependent variable is a binary variable that is equal to one for the current location and equal to zero for the previous location. Also, the variables incorporated in this step are referred as to "pull factors".

In this study to model relocation decisions and relocation actions, a machine learning approach, Random Forest (RF), is incorporated. RF is an ensemble learning approach that combines individual decision trees through a bagging process leading to outputs with lower variance (Breiman, 2001). The bagging is the process of building individual models (decision trees) on the bootstrapped samples formulated by randomly drawing data points with replacement. Then each model is trained on bootstrapped subsets individually and is combined using an averaging process. While some observations will appear more than once in the bootstrapped samples, others will be left out in sampling known as out-of-bag (OOB) observations. This is very helpful since it can be treated as a test sample for checking the error rate of individual models known as out-of-sample error. The OOB error is calculated by dividing the number of false predictions for all models by the total quantity of data in the OOB dataset:

$$Error_{OOB} = \left(1 - \frac{1}{N}\sum_{i \in OOB} \delta_i\right) \times 100 \quad (1)$$

$$\delta_i = \begin{cases} 1 & if\ predicton\ is\ correct \\ 0 & otherwise \end{cases} \quad (2)$$

Where $Error_{OOB}$ is the prediction error rate using the OOB sample; δ_i is the correctness indicator variable; N is the number of observations in the OOB sample (Cheng et al., 2019). Before model development, three parameters of RF need to be calibrated for model specification: the total number of trees n (forest size), the number of splitting variables m, and the maximum tree depth $max\ depth$. After the calibration, 10-fold cross-validation will be applied to validate the final models' accuracy. Since this study aims to evaluate the effect of Covid-19 on establishments' relocation behavior, six RF models will be developed to model relocation decisions and actions of establishments in three periods of time: before the pandemic (2019-20), during the pandemic (2020-21), and after the pandemic (2021-22).

Finally, to provide a comprehensive insight into establishments' relocation actions, the changes between spatial relocation choices over different years will be assessed using statistical analyses. In this regard, relocation distances and the differences between the attributes of current and previous locations will be compared between before, during, and after the pandemic, to investigate if there are meaningful changes between the establishment relocation actions of

different years. A general view of the different sections of the methodology applied in this study to evaluate the relocation behavior of establishments is provided in **Figure 8.**

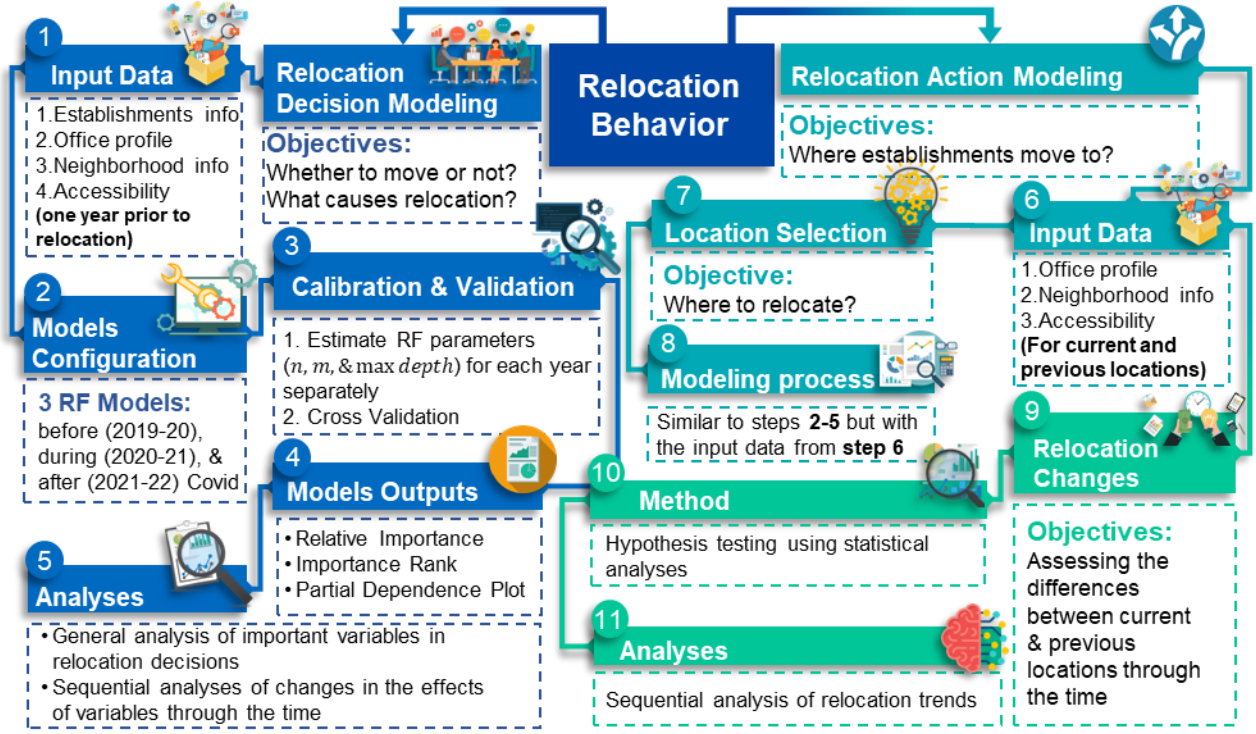


Figure 8: The flowchart of the methodology.

2.4 CASE STUDY AND DATA

This study assesses the relocation behavior of establishments using the state of Tennessee, USA as a case study. The state of Tennessee has 95 counties, and in 2021, hosted 6,975,218 population, 4,368,040 employment, and 315,709 establishments. **Figure 9** shows the relocation rates (within the state) and the rate of new business from 2013 to 2021. As this figure shows, during the pandemic, the number of new establishments decreased significantly, and after the pandemic, the number of new and relocated establishments increased significantly compared to before the pandemic.

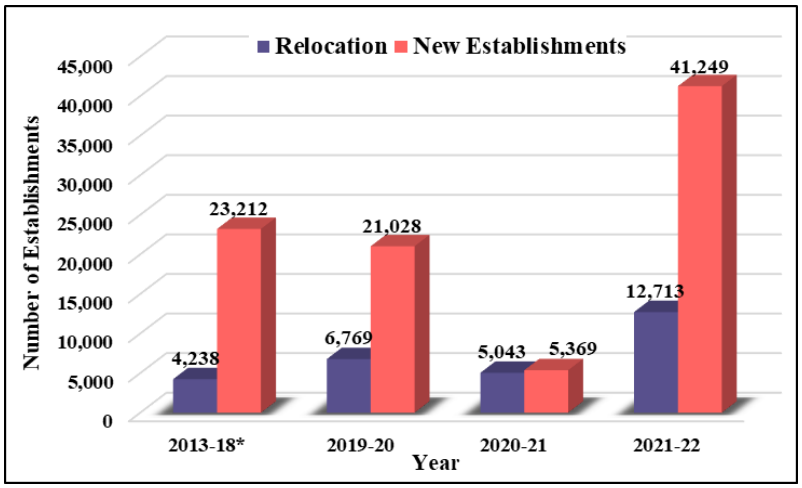


Figure 9: Within the state relocation and new business in the state of Tennessee (*The information provided for 2013-2018 is the average value per year).

Also, the distribution of establishments relocations across different activities, based on North American Industry Classification System (NAICS), is provided for three periods in **Figure 10**. As this figure shows, in all provided periods, the retail trade sector (NAICS 44-45) had the highest rate of relocations. After that, establishments related to other services (NAICS 81) and health care and social assistance (NAICS 62) showed the highest relocation rates. Management of companies and enterprises (NAICS 55) showed the lowest relocation rate in the state of Tennessee. Moreover, establishments that utilize natural resources, such as agriculture, forestry, fishing, and hunting (NAICS 11), mining (NAICS 21), and utilities (NAICS 22) showed low relocation rates.

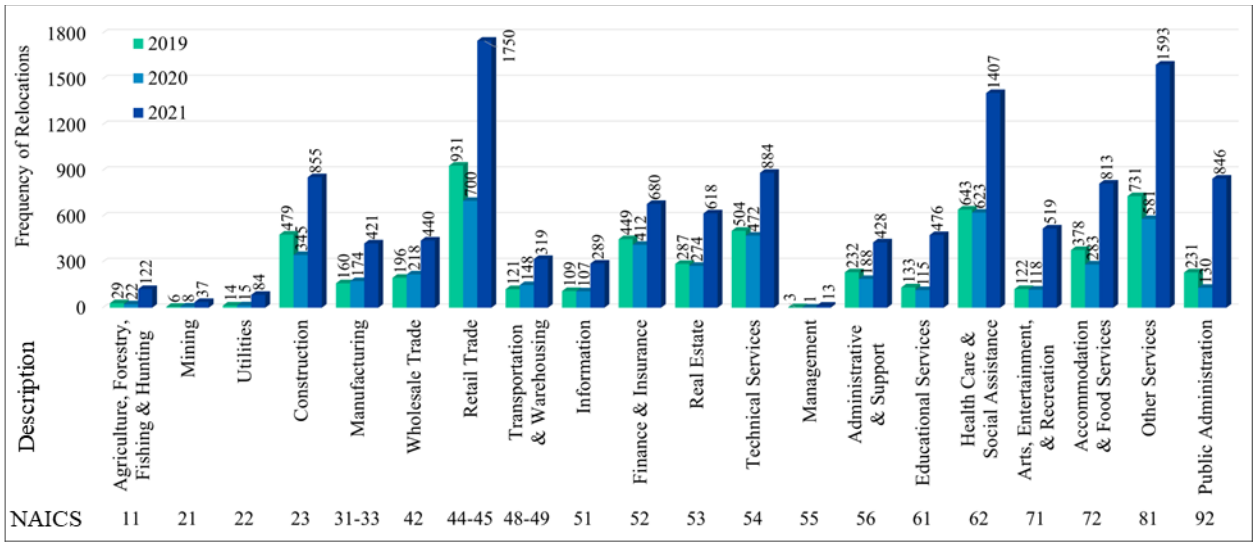


Figure 10: The frequency of relocations across NAICS sectors from 2019 to 2021.

A comprehensive data collection procedure was followed to accumulate possible relocation determinants from four different sources. The establishment's detailed information is collected from the InfoUSA data set (available at http://www.infousa.com). InfoUSA provides detailed information for companies from local shops to global enterprises. Establishments' information such as type, owner, address, business profile, credit history, business value, number of employments, size, and franchise status is collected from 2018 to 2022. In addition, by considering the spatial locations of establishments, the numbers of the same establishments (establishments with exactly same types) and similar establishments (establishments with same 2-digit NAICS code) within 1 mile are estimated for each establishment to capture the effect of agglomeration. Census data was used to collect the population density and the median rent price of the surrounding neighborhood at the block group level.

Moreover, the land value, square feet, and land-use diversity of the surrounding area were collected from parcel data. Parcel data refers to a combination of both spatial and nonspatial attribute files, presenting land ownership in a local jurisdiction (Ali R. Samani et al., 2022). The land-use diversity of the surrounding area was collected by calculating the land-use mix using the following equation (Cao and Ermagun, 2017):

$$land\ use\ mix = 1 - \left(\frac{\left|\frac{Res}{Tot} - \frac{1}{5}\right| + \left|\frac{Ind}{Tot} - \frac{1}{5}\right| + \left|\frac{Com}{Tot} - \frac{1}{5}\right| + \left|\frac{Agr}{Tot} - \frac{1}{5}\right| + \left|\frac{Pub}{Tot} - \frac{1}{5}\right|}{8/5} \right) \quad (3)$$

$Land\ use\ mix$ shows the diversity of land-use conditions within one mile for each establishment and ranges from 0 (dominated by one land-use) to 1 (balanced distribution of various land-uses). In Equation 3, Res refers to the number of residential parcels; Ind, Com, Agr indicate the number of industrial, commercial, and agricultural parcels respectively, and Pub denotes parcels with public access; and Tot is the sum of all five types of land-uses within 1 mile.

Finally, the accessibility of each establishment to the transportation network was estimated by calculating the distance to the closest interstate entrance, urban and rural highways entrance, and major arterials. To sum up, from the data collected, 20 explanatory variables are selected to model relocation behavior, and these variables are provided in **Table 3**. As this table shows, the explanatory variables are categorized into four categories: office profile, neighborhood conditions, accessibility, and establishments' characteristics.

Table 3: List of explanatory variables used to model relocation decisions and behaviors of establishments

Variable Name	Type	Min	Max	Mean	Description
Office profile					
Square feet	Int.	0	99×10^7	256,772	Square feet (sq)
Value/square feet	Dbl.	0	188,507	48	Land value per square feet ($/sq)
Neighborhood condition					
Population Density	Int.	0	208,817	7,298	Population density at the block group level
Rent	Int.	99	3,500	965	Median rent price at the block group level ($)
Land-use mix	Dbl.	0	0.909	0.19	Equation (3)
Same Establishment	Int.	0	295	40	The number of same establishments within 1 mile
Similar Establishment	Int.	0	550	133	The number of establishments with the same 2-digit NAICS within 1 mile
Metropolitan	Cat.	0	1	N/A	Not in the CBSA (0), or metropolitan (1)
Accessibility					
Dist. to interstate	Dbl.	0.01	86.82	7.89	Distance to interstates entrances (mile)
Dist. to highways	Dbl.	0.01	37.76	2.44	Distance to urban and rural highways (mile)
Dist. to major arterials	Dbl.	0.01	27.17	1.31	Distance to major arterials (mile)
Establishment's characteristics					
Employment	Int.	1	35,000	11	The number of employees in the establishment
Age	Int.	1	313	16	Establishment's age
Credit score	Int.	0	99	82	Value of the establishment
Franchise	Cat.	0	1	N/A	Franchise (1) and otherwise (0)
Business condition	Cat.	-1	1	N/A	Shrinking (-1), constant (0), or growing (1)
Female owner	Cat.	0	0	N/A	Female owner (1) and otherwise (0)
High income	Cat.	0	1	N/A	High-income business (1) and otherwise (0)
Establishment Size	Cat.	1	3	N/A	Small (1), medium (2), and big (3)
Work from home	Cat.	0	1	N/A	Working from home (1) and otherwise (0)

Cat. = Categorical, Int. = Integer, and Dbl. = Double

2.5 RESULTS

This section provides the results of modeling establishments' relocation decisions and relocation actions. In the following subsections, first, results of calibrating and validating the RF models are provided. Then, the results of developing RF models for assessing relocation decisions and relocation actions will be presented. Developing RF provides variables' Relative Importance (RI) and Partial Dependence Plots (PDP). Permutation-based importance measure is used to find the mean of accuracy decrease on the OOB when there is a random shuffle of the variable. The variables with the largest average decrease in accuracy are considered the most important. The PDP shows the graphical depiction of the marginal effect of an independent variable on the dependent variable, accounting for the average influences of all other variables in the model (Greenwell, 2017). Since random forest does not follow a hypothesis of a predetermined relationship (e.g., linear, log-linear, and nth power), PDP can show whether the relationship between the target and a feature is linear, monotonic, or more complex. And finally at the end of this section, the results of conducting statistical analysis are provided to assess if there are significant changes in the relocation actions of different years.

2.5.1 Model Calibration and Validation

Three parameters of RF: the number of trees (n), the number of random features selected for each tree (m), and the maximum depth of trees ($max\ depth$), are needed to be calibrated. The calibration procedure of RF contains two steps. In the first step, the optimum value of n is estimated; and in the second step, $max\ depth$ is predicted considering different combinations of m values. **Figure 11** provides a general view of the effect of these two steps on the error rate of relocation action models. **Figure 11a** presents the effect of changing n on the error rate of the relocation action models of three periods, where default settings are considered as $m = 6$ and max $depth > 4{,}500$ (fully grown). As **Figure 11a** shows, after a certain n, the error rate does not change considerably. Hence, $n = 300$ is selected for the relocation decision and $n = 200$ is estimated for relocation actions models. Moreover, **Figure 11b** presents the comparison results of evaluating $max\ depth$ with different combinations of m, considering the corresponding calibrated n (the output of the first step $n = 200$) for relocation action and in 2019-20 (for brevity, the calibration procedure of $max\ depth$ is demonstrated only for one of six models). Results suggest considering $m = 12$ for both models and $max\ depth$ equal to 3,100 and 1,300 for the relocation decision and relocation action models, respectively. After calibrating the required parameters, the final OOB errors of relocation decision models are 18.02%, 18.97%, and 13.28% respectively for before, during, and after the pandemic. Also, the models of relocation action showed low error rates for the same periods (16.33%, 18.17%, and 14.53.% respectively). To validate models' accuracy, 10-fold cross-validation was applied to the developed models. All models showed high accuracy in predicting relocation decisions and actions such that relocation decision showed 81.12%, 80.09%, and 86.36%, and the relocation action showed 82.86%, 80.93%, and 84.75% accuracy in 10-fold cross-validation, respectively for before, during, and after the pandemic.

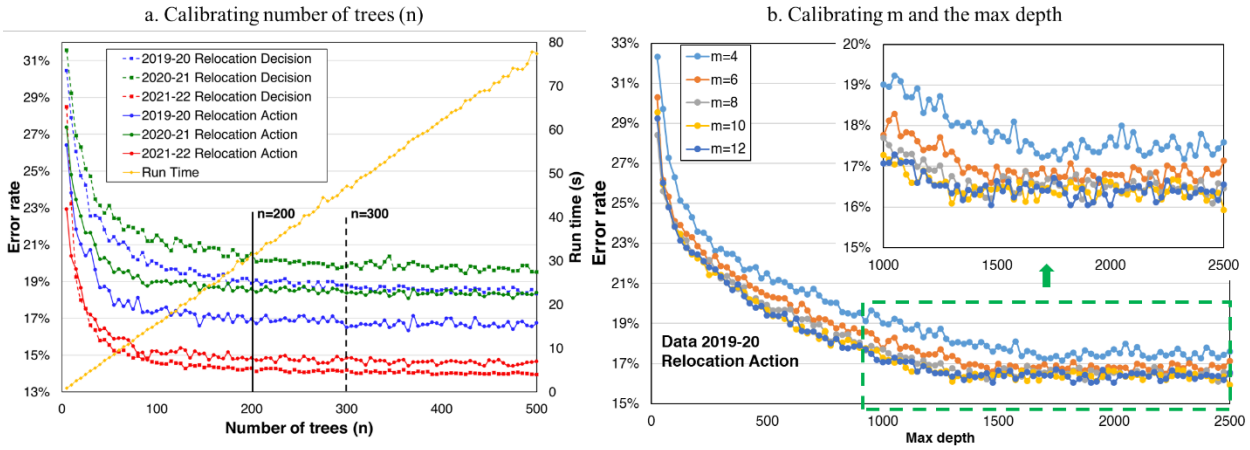


Figure 11: Results of calibrating RF: a. the number of trees (n) and b. the number of random features selected for each tree (m) and the max depth of trees (max depth) for relocation action in 2019-20.

2.5.2 Relocation Decision

The estimated results of relative importance and ranks of relocation decision determinants extracted from developed RF models for each year are provided in **Table 4**. In this table, explanatory variables are categorized into four categories, office profile, neighborhood condition, accessibility, and establishment characteristics. The percentages of RI of variables in these categories are summed to provide a general view of the importance of each category in the decisions of establishments on whether to relocate or not. Also, PDPs are provided for each variable in **Figure 12**. In these figures, the y-axis shows the relocation probability, and the x-axis shows the changes in the explanatory variable.

In general, similarities between the RI of variables before and during the pandemic are high, while the result of modeling post-pandemic is changed significantly. Results showed that before the pandemic neighborhood condition (33.4%), office profile (26.3%), establishment's characteristics (23.6%), and accessibility (16.8%) respectively are the most important groups in the relocation decision. During the pandemic, the establishment's characteristics (25.1%) are the second most important, and office profile (22.0%) and accessibility (20.7%) respectively are the third and fourth important groups. However, the RIs of groups have changed significantly after the pandemic such that, office profile (31.0%) is the most important group, and neighborhood condition (29.9%), accessibility (26.0%), and establishment's characteristics (13.4%) are respectively second, third, and fourth. Interestingly, the importance of accessibility after the pandemic is increased significantly compared to before and during the pandemic. In addition, as **Table 4** shows the relocation decision is less affected by the establishment's characteristics after the pandemic, and the importance of establishment condition is reduced by 10.2%. which is the highest percentage change among all four categories.

Table 4: The Relative Importance (RI) and rank of explanatory variables for relocation decision.

Variable	2019-2020		2020-2021		2021-2022	
	Rank	**RI**	**Rank**	**RI**	**Rank**	**RI**
Office profile						
Square feet	1	16.6%	1	13.9%	1	23.0%
Value/square feet	3	9.7%	3	8.1%	5	8.0%
Sum		**26.3%**		**22.0%**		**31.0%**
Neighborhood condition						
Population Density	8	6.1%	9	5.8%	4	8.8%
Median Rent	6	6.3%	7	6.5%	6	7.3%
Land-use mix	2	10.1%	2	9.9%	8	6.1%
Same Business	16	1.8%	16	1.9%	16	1.6%
Similar Business	5	7.6%	6	6.5%	9	4.2%
Metropolitan	17	1.5%	17	1.6%	12	1.9%
Sum		**33.4%**		**32.2%**		**29.9%**
Accessibility						
Dist. to interstate	7	6.3%	5	6.5%	2	9.6%
Dist. to highways	14	2.3%	8	6.3%	7	7.0%
Dist. to major arterials	4	8.2%	4	7.9%	3	9.4%
Sum		**16.8%**		**20.7%**		**26.0%**
Establishment's characteristics						
Number of employment	12	2.8%	14	2.8%	14	1.8%
Age of establishment	10	4.8%	10	5.2%	15	1.6%
Credit score	9	5.7%	11	5.2%	11	3.8%
Franchise	18	0.8%	18	0.9%	18	0.4%
Business condition	13	2.5%	13	2.9%	17	0.9%
Female owner	19	0.6%	19	0.6%	19	0.1%
High income	20	0.0%	20	0.3%	20	0.1%
Establishment Size	11	4.2%	12	4.9%	10	2.9%
Work from home	15	2.2%	15	2.3%	13	1.8%
Sum		**23.6%**		**25.1%**		**13.4%**

In all three periods, the square feet of the office is the most important variable in the relocation decision. As **Table 4** shows, the importance of square feet after the pandemic (23.0%) is increased compared to before (16.6%) and during (13.9%) the pandemic. Also, as **Figure 12A** shows, establishments with higher square feet showed a lower tendency to relocate. And, when the office's square feet is less than 20,000 ft^2, the probability of relocation is higher for an establishment with the same square feet in 2021-22 compared to during and before the pandemic. Also, in 2021-22, the relocation probability of establishments with large square feet is significantly lower than before and during the pandemic. **Figure 12B** presents the PDP for value/ square feet, which was the third most important variable in modeling the relocation decision during and before the pandemic. As this figure shows, the probability of relocation increases as the value per square feet increases. Also, as table 2 shows the RI of value/square feet is reduced in 2020-21and 2021-22.

The importance of the population density is increased in post-covid and as **Table 4** shows it has the fourth highest RI in relocation decisions in 2021-22 (8.8%). Also, as **Figure 12C** shows, despite 2019-20, a population density over 10,000 significantly reduces the chance of relocation, in 2021-22, and establishments in less populated areas are more likely to decide to move. The median rent of the neighborhood showed similar RIs in all periods and is ranked 6 before and after the pandemic and 7 during the pandemic. As **Figure 12D** shows, establishments in neighborhoods with median rent between $1,000-1,500 have the highest probability of relocation. The land-use

mix is the second most important variable before and during the pandemic. As **Figure 12E** shows, increases in land-use diversity reduce the tendency to relocate. This is in line with the changes in the effect of population density, as in more dense areas typically the land use diversity is higher. Also, similar to median rent, the PDPs of land use mix for before and after the pandemic shows similar behaviors. The influence of agglomeration on relocation decisions is measured by evaluating the number of similar and same activities within 1 mile. The agglomeration of the same activities did not show high importance as the RIs were less than 2% and the rank was 16 in all three periods. Based on **Figure 12F**, an establishment with the same number of same establishments within 1 mile, is more likely to relocate in 2021-22, compared to before and during the pandemic. However, the agglomeration of similar establishments showed high importance before the pandemic and the importance rank is 5. While, the importance of this variable is reduced significantly after the pandemic, such that the importance rank dropped from 5 in 2019-20 to 9 in 2021-22. Based on **Figure 12G,** increases in the number of similar establishments surge the likelihood of relocations in all periods. Also, after the pandemic, the lower number of similar establishments showed a higher probability of relocation. Locating in a metropolitan area showed low RIs in all three periods (less than 2%). Based on **Figure 12O**, establishments in metropolitan areas show a lower probability to relocate.

Generally, variables related to accessibility mostly are among the most important variables, except the distance to highways which has the 14th rank in 2019-20. Distance to interstate has the second rank in 2021-22 model and showed 9.6% RI, while RIs for before and during the pandemic are 6.3% and 6.5%, respectively. As **Figure 12H** shows, when the distance to interstates is less than 10 miles, establishments with a longer distance to interstates have a higher probability of relocation, however, when the distance to interstates is more than 10 miles, the chance of relocation reduces slightly. Also, considering the constant distance to the interstates, the relocation probability of an establishment is higher after the pandemic. Distance to highways showed the lowest importance among accessibility measures. The highest RI of distance to highways was observed in the model on relocation decisions in 2021-22 (7.0%). Also, as **Figure 12I** shows, the probability of relocation will increase with a longer distance to highways. Distance to major arterials was the most important variable in accessibility measures, such that it showed significant RI in all three years and ranked 4 in models of 2019-20 and 2020-21, and 3 in the model of 2021-22. As **Figure 12J** shows increases in the distance to major arterials will reduce the probability of relocation significantly, and most relocations happen for establishments that are close to major arterials.

Among variables related to the establishment's characteristics, credit score has the highest RIs in all periods As **Figure 12K** shows, establishments with better credit scores tend to relocate less. After the credit score, the age of the establishment showed the highest RIs. Also, based on **Figure 12L**, older establishments tend to relocate less, and relocation is more relevant among establishments aged between 10-15. Similar to other variables related to establishments' characteristics, the RIs of credit score and age reduce after the pandemic. **Figure 12** also demonstrates PDPs for the female owner, business conditions, and establishment size. As **Figure 12M** shows, establishments with female owners showed a lower probability to relocate. **Figure 12N** shows the effect of business conditions on the probability of relocation. As was expected, establishments with a shrinking business condition tend to relocate more than other groups. And finally, **Figure 12P** depictures the changes in the probability of relocation considering the size of the establishments, such that small establishments are more likely to relocate compared to medium and big size companies.

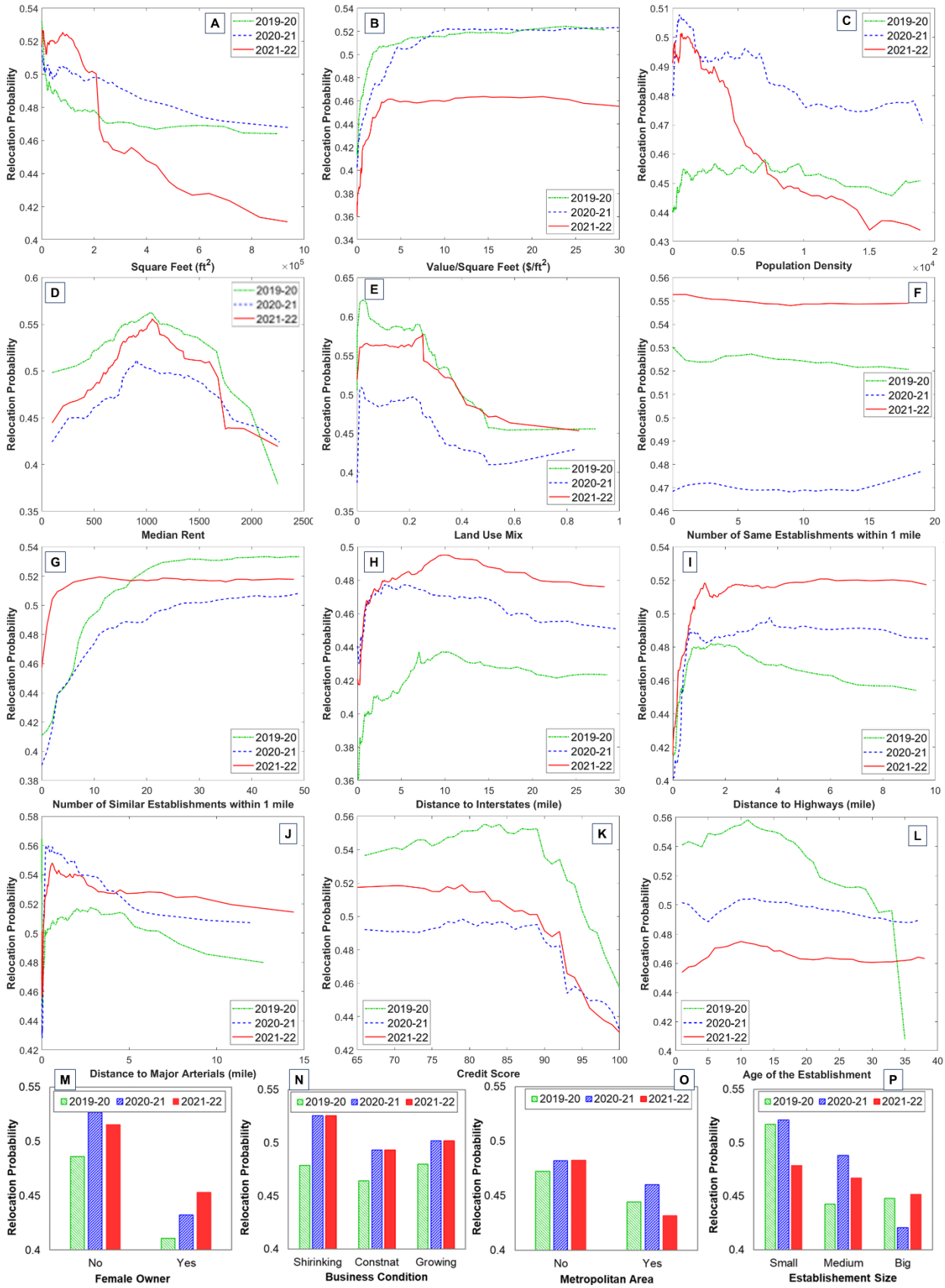


Figure 12: The PDP of relocation decisions and push factors.

2.5.3 Relocation Actions

To indicate the pull factors in relocation behavior, RF models were developed considering office profile, neighborhood condition, and accessibility of the destination (the place an establishment relocate to) as the exoplanetary variables. **Table 5** provides the ranks and RIs of the explanatory variables after developing pull factor models for three periods. Moreover, the PDPs of the relocation choices and pull factors are provided in **Figure 13**, to show how the probability of selecting a spot to relocate to would change corresponding to changes in explanatory variables.

Table 5:The Relative Importance (RI) and ranks of explanatory variables for relocation actions

	2019-2020		2020-2021		2021-2022	
Variable	**Rank**	**RI** *	**Rank**	**RI**	**Rank**	**RI**
Office profile						
Square feet	4	9.6%	1	16.4%	1	19.5%
Value/square feet	8	3.9%	9	3.8%	6	10.6%
Sum		**13.5%**		**20.2%**		**30.1%**
Neighborhood condition						
Population Density	5	8.1%	7	8.9%	8	5.9%
Rent	7	5.6%	6	9.1%	9	4.4%
Land-use mix	9	3.7%	8	6.7%	7	8.4%
Same Business	10	0.7%	10	1.1%	10	0.8%
Similar Business	3	10.7%	5	9.3%	3	13.6%
Sum		**28.8%**		**35.1%**		**33.1%**
Accessibility						
Dist. to interstate	2	24.3%	4	14.1%	4	11.8%
Dist. to highways	1	26.2%	3	15.2%	5	11.1%
Dist. to major arterials	6	7.2%	2	15.4%	2	13.9%
Sum		**57.7%**		**44.7%**		**36.8%**

Generally, as **Table 5** shows, the most important group of factors affecting the spatial relocation choices is the accessibility group, which showed 57.7%, 44.7%, and 36.8 RIs, respectively for before, during, and after the pandemic. After the pandemic, although the RI of the accessibility group still is the largest, it dropped down to 36.8%. This shows that the attractiveness of accessibility for establishments is reduced after the pandemic. In addition, compared to before the pandemic, the importance of variables related to office profile is increased significantly in post-covid conditions, as it is raised from 13.5% to 30.1%. Also, the sum of RIs for variables related to neighborhood conditions is not changed significantly before and after the pandemic, compared to other groups of variables.

The square feet of the office showed the highest RIs during and after the pandemic, 16.4% and 19.5% respectively, which are significantly increased compared to before covid (9.6%). Also, as **Figure 15A** shows, during and after the pandemic, locations with higher square feet have lower attractiveness, while before the pandemic, increases in the square feet of the office enhance the probability of selecting a location. The value/square feet showed low RIs in the models of before and during the pandemic, 3.9% and 3.8% respectively, and ranked 8 and 9 (out of 10). However, the importance of value/square feet increases to 1.6% and ranks as the 6^{th} important factor. As **Figure 15B** shows, relatively, the probability of selecting a location will reduce with increases in value/square feet, specifically in 2021-22 where the changes in the likelihood are more significant.

Regarding the effects of neighborhood conditions on relocation actions, as **Table 5** shows, the importance of population density is reduced in 2021-22 compared to the period before the

pandemic. **Figure 15C** shows that in all three periods, neighborhoods with a population density of fewer than 5,000 people/acres attract establishments to relocate the most. In addition, in models of 2020-21 and 2021-22, increases in the population density reduce the probability of selecting a location, while, in the model of before the pandemic, when the population density is more than 10,000 people/acres the probability of selecting a location increase. **Figure 13C** presents a big difference between the attractiveness of dense areas before and after the pandemic. Based on **Figure 13D**, the PDP of the median rent shows a similar trend for all three periods such that increases in the median rent reduce the probability of selecting a location. The most attractive areas for establishment before and after the pandemic are neighborhoods with a median rent ranging between $500 to $1,000, while $1,000-1,500 shows the highest probability of location choice in 2020-21. Also, the median rent shows a higher RI in modeling the year 2020-21 (9.1%). The effect of land use diversity is presented in **Figure 13E**, where increases in the land-use diversity index enhance the attractiveness of a location. The importance of land use diversity is surged after the pandemic (RI showed 8.4% and ranked 7), while before and during the pandemic the rank were 9 and 8, respectively. The agglomeration of similar establishments was the most important variable related to neighborhood conditions The RIs respectively show 10.7%, 9.3%, and 13.6% for before,

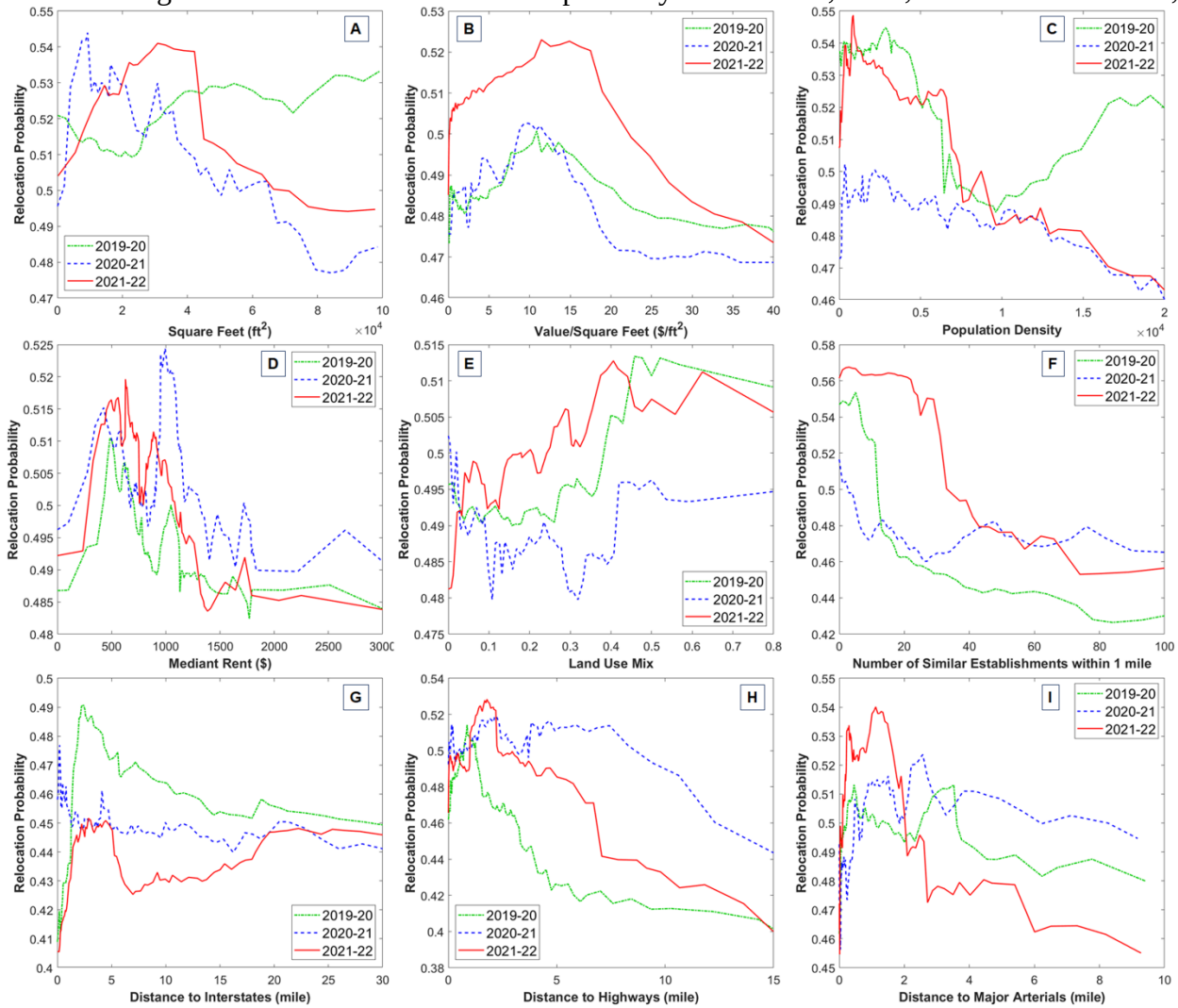


Figure 13: The PDP of location actions and pull factors.

during, and after the pandemic. Also, the number of similar establishments ranked 3 in models of before and after the pandemic. Besides, **Figure 13F** shows that the probability of selecting a location will reduce significantly when the number of similar activities is more than 15 in 2021-19 and 2020-21, and 30 in 2021-22. Finally, the number of the same establishments showed the least important variable and the RIs were less than 1.1%.

As Table 3 shows, the accessibility measures were the most important group of variables. The importance of distance to interstates is reduced over time such that the RI of distance to interstates reduces from 24.3% in 2019-20 to 11.8% in 2021-22. The distance to interstates is the second most important variable in modeling the relocation actions before the pandemic, while the rank of this variable in modeling the after the covid condition is reduced to 4. In addition, as Figure 6G shows, in all three periods, the most attractive locations for relocation are those within less than 5 miles to interstates. However, based on this figure, before and during the pandemic, increases in the distance to interstates would reduce the probability of selecting a location, while the PDP of 2021-22 for locations with more than 10 miles distance to interstates, the probability of selecting a location will increase. The distance to highways is the most important variable in the relocation actions of establishments before the pandemic. Similar to interstates, the importance of distance to highways is reduced over time, where the RI is reduced from 26.2% in 2019-20 to 11.1% in 2021-22. Figure 6H shows the relationship between distance to highways and the probability of relocation, where increases in the distance to highways reduce the probability of selecting a location in all three periods, and a location with less than 5 miles distance to highways showed the highest attractiveness. Finally, the distance to major arterials was the second most important variable in molding the relocation actions during and after the pandemic, such that the RIs are 15.4% and 13.9%, respectively. As Figure 6I shows, increases in the distance to major arterials reduce the probability of selecting a location, while the attractiveness of locations closer to major arterials is increased over time, such that, before the pandemic, a location within 3-4 miles distance to major arterials had the highest probability of selection, and after the pandemic, considering all other variables constant, a 2-mile showed the highest probability of relocation.

2.5.3.1 Analyzing Relocation Changes

The statistical analysis of evaluating if there are meaningful changes between the establishment relocation actions of different years is provided in this subsection. **Table 6** provides the average and standard deviation of differences between the attributes of current and previous locations of relocated establishments in each period. For instance, this table shows that establishments in 2019-20 relocated to a place that on average had 558 ft^2 bigger area while in 2020-21 establishments selected locations to move which on average had 1851 ft^2 smaller area. Since the data of relocation behavior changes were non-parametric, the Kruskal-Wallis test was applied followed by post-hoc analyses with Bonferroni correction to measure if there is a significant difference between the mean of each group (years) and where this difference is happening. Kruskal-Wallis test is a nonparametric statistical method of analysis of variance, that usually is applied when the underlying population distributions are unknown (26). The results of these statistical analyses are tabulated in **Table 7.**

Analyzing the changes in the relocation distance (the distance from previous to current location) shows that, establishments relocated to closer locations in post-covid. As **Table 7** shows, there are significant differences between relocation distances of different years ($p < 2.2\ e-16$), and post-hoc analysis showed that these differences exist between all groups. Moreover, **Figure 14** presents the distribution of relocation distance. As this figure shows, the majority of relocations

occurred within 1 mile of the previous location, in all three periods. However, 75% of relocations occurred within 3.1 miles which is significantly lower than 2020-21 and 2019-20. Moreover, the rate of relocating >20 miles is higher before the pandemic compared to during and after the pandemic.

Table 6: The mean and SD (in parenthesis) of differences between current and previous locations of relocated establishments.

variables	2019-20	2020-21	2021-22
Relocation distance (mile)	5.446 (15.79)	7.102 (24.162)	3.689 (8.975)
Office profile			
Square feet ($\times 10^3$)	558 (1329)	-1851 (5806)	-2966 (6155)
Value/square feet	-14.8 (320.4)	-4.2 (283.1)	1.1 (416.3)
Neighborhood condition			
Population density (pop/acres)	-746.6 (8010)	-392.8 (7009)	-216.8 (6555)
Rent ($)	116.5 (438.3)	123.3 (462.5)	-42.2 (347.4)
Land-use mix	0.009 (0.203)	0.0017 (0.206)	-0.0099 (0.204)
Same business	-1.03 (17.48)	3.04 (62.80)	6.68 (106.05)
Similar business	-7.13 (78.32)	8.33 (136.94)	14.21 (222.47)
Accessibility			
Dist. to interstate (mile)	0.121 (4.185)	0.188 (4.2)	0.233 (2.813)
Dist. to highways (mile)	-0.02 (4.76)	0.121 (2.396)	-0.127 (2.074)
Dist. to major arterials (mile)	0.164 (1.882)	0.133 (1.803)	0.056 (1.454)

Table 7: Results of statistical analyses on relocation changes.

	Kruskal-Wallis test		**Post-hoc analysis**		
Variable	*Chi-Square**	*p-value*	2019 vs 20	2019 vs 21	2020 vs 21
Relocation distance	922.7	**< 2.2 e-16**	**1.6e-08**	**< 2e-16**	**< 2e-16**
Square feet	21.629	**2.01e-05**	**0.017**	**8.6e-05**	0.271
Value/square feet	83.358	**< 2.2 e-16**	0.54	**1.6e-15**	**3.3e-09**
Population density	51.965	**5.2 e-12**	**0.00017**	**6.3e-13**	0.35306
Rent	33.687	**4.84e-08**	0.978	**8.1e-08**	**6.4e-06**
Land-use mix	7.7708	**0.02054**	**0.024**	**0.068**	0.954
Same business	68.722	**1.2 e-15**	**9.7e-05**	**3.3e-16**	**0.014**
Similar business	299.49	**< 2.2 e-16**	**< 2e-16**	**< 2e-16**	**2.8e-07**
Dist. to interstate	9.2187	**0.00995**	**0.048**	**0.017**	0.343
Dist. to highways	24.8	**4.12 e-06**	**0.032**	0.358	**3.5e-07**
Dist. to major arterials	1.9925	0.3693	0.59	0.988	0.72

$^{*}df = 2$

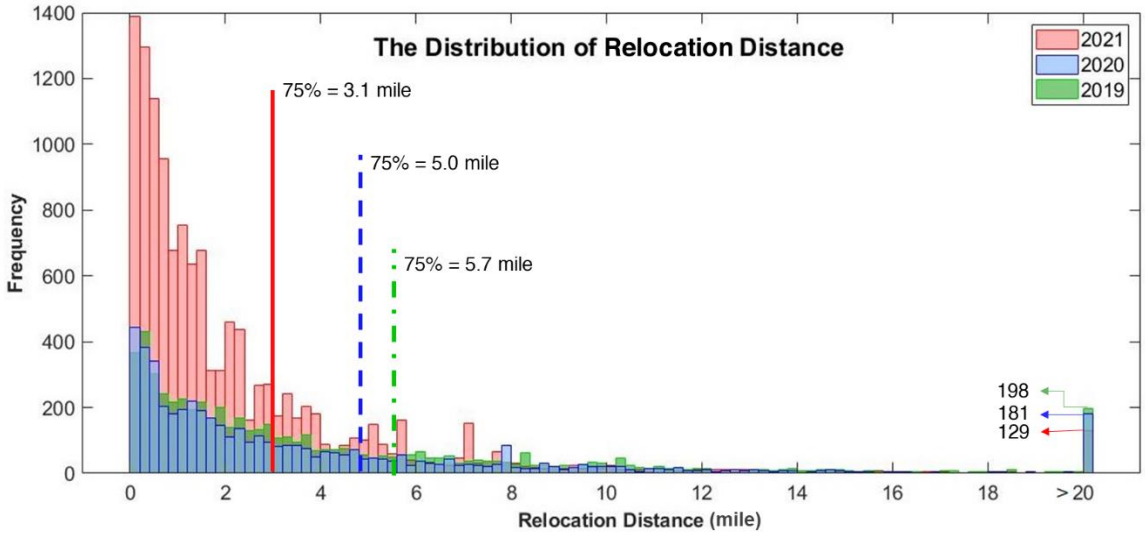


Figure 14: The distribution of relocation distances.

In addition to relocation distance, the Kruskal-Wallis test showed that there are significant differences between the mean of all variables, except the distance to major arterials. Post-hoc results showed that there are significant differences between the means of changes in the square feet of office during and after the pandemic compared to before pandemic. As *Chi-Square* shows, establishments tend to select a smaller location compared to their previous location, during and after the pandemic. The value/square feet showed significant increases after the pandemic, such that establishments relocated to more expensive buildings. Establishments on average tend to relocate from more to less dense areas, as the means of population density are negative in all three periods. Also, the post-hoc test showed a significant difference between the changes in the population density during and after the pandemic with before the pandemic. Changes in the differences in median rent showed a significant decrease after the pandemic, and establishments on average relocated to neighborhoods with lower median rent. Changes in the land-use mix showed significant differences between during and after the pandemic and before the pandemic, such that on average establishments tend to relocate to an area with lower land-use diversity. The post-hoc test showed that there are significant differences between the means of the number of same and similar businesses of different years. As **Table 6** shows, compared to before the pandemic, establishments tend to relocate to a location where the number of same and similar activities is higher than their previous location. Finally, in the accessibility group, results showed that there are significant differences between the means of changes in the distance to the interstates during and after the pandemic compared to before Covid. Establishments tend to relocate to places that are further from interstate entrances. Besides, analysis of changes in the distance to highways shows that there are only significant differences between the relocation behavior of establishments during the pandemic compared to before and after Covid-19.

2.6 DISCUSSION

This study aimed to model the establishments' relocation procedure by investigating push and pull factors and assessing the effect of Covid-19 on establishments' decision-making. Push factors in some studies are discussed as the relocation tendency (Alkay, 2010). RF classification method was applied, which in addition to high accuracy in forecasting the relocation decisions and behaviors, provides the non-linear relationship between dependent and independent variables in the form of PDP. The importance of assessing the relocation behavior of establishments for transportation planners is rooted in the interaction between transportation networks and the location of an establishment. This study well demonstrated the importance of accessorily (closeness to transportation network) in relocation behavior such that accessibility had the highest RI in relocation behavior analysis, and the probability of selecting a location reduces significantly with increases in the distance to the transpiration network which is in line with the results of the study conducted by (Yuan, 2021). Although many companies have changed their policies, and nowadays, working from home is more common, the importance of accessibility is irrefutable. Even though results showed that the importance of accessibility is reduced significantly in post-Covid analysis.

Agglomeration was one the most important variables in both pull and push factor analyses, which was shown in previous studies (Backman and Karlsson, 2017; De Bok and Van Oort, 2011). This study showed that when the number of similar activities within one mile of a location is more than 10 the tendency to relocate increases significantly, and if it is more than 20, establishments are less interested to relocate that location. However, the number of the same establishments did not show a significant effect in both pull and push factor analyses. Also, this study showed that a significant part of relocation decisions depends on establishment characteristics where the results were in line with the finding of studies conducted by (Brouwer et al., 2004; Lampón et al., 2015). However, one of the findings of this study is the reduction in the importance of the establishment's characteristics on relocation decisions, suggesting that in post-Covid conditions environmental conditions trigger the relocation decision. For instance, reductions in the importance of the age of establishment in relocation decisions in the post-Covid period show the magnitude of the effect of covid on establishments' policies and attitudes, and it can be interpreted that even old establishments considered relocation in post-covid conditions.

Besides, statistical analyses were applied to evaluate the changes in the relocation behavior by comparing the difference between the attributes of the current and previous location and the relocation distance in three periods. The results of this section brought up other research questions. As results showed that establishments tend to move to areas with less population density, the decentralization of business and analysis of relocating of behavior concerning distance to CBD could provide important results for urban planners, hence conducting studies similar to research conducted by Kang (2020) on the decentralization of warehouses, are needed. This study only considered relocations that occurred within a state, therefore variables such as access to ports, different tax rules, and other governmental policies which can be effective in the relocation of the establishment were not discussed.

2.7 CONCLUSION

Due to the significant effects of business relocations on residential relocation, travel patterns, and land-use conditions it is crucial for transportation and urban planners to assess and understand the relocation behavior of establishments. This study aimed to investigate the relocation procedure of smaller economic units (establishments) and assess the effect of covid-19 on relocation behaviors. The relocation procedure was divided into two segments, 1) relocation decision (why does an establishment move?); 2) relocation selection (where the establishment relocate to?). Random Forest (RF) classification was applied to model these two sections for three periods, before (2019-2020), during (2020-2021), and after (2021-2022) the pandemic. Results of training and testing the RF models showed that this approach can forecast the relocation decision and relocation selection up to 86.1.2% and 84.8% respectively.

This study categorized the determinants of relocation decisions into four categories, establishment characteristics, office profile, neighborhood condition, and accessibility. Interpreting the results of developing RF models showed that variables related to neighborhood conditions have the highest importance in relocation decisions before and during the pandemic, while office profile showed the highest importance after the pandemic. Moreover, results of modeling selecting a location to relocate showed that after the pandemic the importance of accessibility was significantly reduced compared to before the pandemic selected a location. However, accessibility was the most important location determinant for all three periods. In addition to significant changes in the importance of accessibility, the importance of the office profile increased significantly after the pandemic. Finally, this study conducted statistical analysis to evaluate changes in relocation behavior by comparing the differences between the attributes of previous (relocated from) and current (relocated to) locations in three periods. While results showed significant effects of Covid-19 on the relocation behavior of establishments.

Considering the effect of Covid-19 on different parts of society and the importance of analyzing post-Covid conditions, future studies can assess the post-covid condition by collecting more data, as the responses of the establishments might not be captured in short-term. Besides, due to heterogeneity in the behavior of establishments, the application of other modeling approaches, e.g., latent class, hybrid models, and other machine learning methods can be a good path for future studies. Moreover, future studies can assess the relocation behavior after covid-19 at national and international levels. Finally, capturing the effect of relocation behavior on travel patterns and residential relocation/location choices is another interesting research topic for future studies.

3.0 CONCLUSION

This study has delved into the multifaceted world of establishment location choices and relocation behaviors, shedding light on the intricate interplay of economic activity, spatial determinants, and external influences, particularly the COVID-19 pandemic.

In the first part of the research, we embarked on an exploration of establishment location choices, the smallest economic units, with a specific focus on the North American Industrial Classification System (NAICS) sectors. Leveraging a discrete choice model, we modeled the location preferences of these establishments, followed by a detailed examination of the determinants of location choices. Our findings classified these determinants into four distinct categories: office profile, accessibility, neighborhood attributes, and interactions between establishments. Significantly, the study revealed variations in location determinants across different NAICS sectors. Key determinants encompassed the presence of the same activity, land value, office size, square footage, and land use conditions, with a unique emphasis on the importance of accessibility, particularly in relation to major arterials. Our results also provided valuable insights into establishments' Willingness to Pay for enhanced accessibility, offering practical implications for transportation planners.

In the second part, the research pivoted to the critical realm of establishment relocations, with a particular focus on the influence of the COVID-19 pandemic. The relocation process was dissected into two vital segments: the relocation decision and the relocation selection. Through the application of Random Forest (RF) classification, we developed models for these segments, spanning three periods: before, during, and after the pandemic. The results of these models demonstrated their ability to accurately forecast relocation decisions and selections. Furthermore, the study categorized determinants of relocation decisions into four key categories: establishment characteristics, office profile, neighborhood conditions, and accessibility. The findings unveiled the evolving importance of these determinants, with neighborhood conditions being pivotal before and during the pandemic, and office profile gaining prominence after the pandemic. Accessibility, while consistently significant, experienced shifts in importance across these periods. This research also undertook a robust analysis to assess changes in relocation behavior and confirmed the significant impact of COVID-19 on the decisions and selections of establishments.

As this study unfolds, it naturally unveils avenues for future research. Assessing the post-COVID condition with a more extensive dataset and exploring heterogeneous establishment behavior through alternative modeling approaches could offer deeper insights. Extending the analysis to national and international levels, and considering the ripple effects of relocation behavior on travel patterns and residential choices, are promising research paths. In a dynamic world, the study underscores the need for continuous exploration and understanding of establishment behaviors in the ever-evolving urban and economic landscape.

4.0 REFERENCES

Ahmed, U., Hawkins, J., Roorda, M.J., 2022. Establishment location choice model considering intra-firm interactions. Journal of Transport Geography 102, 103391.

Alkay, E., 2010. What determines the relocation tendency of manufacturing firms?

Alkhereibi, A.H., Tahmasseby, S., Mohammed, S., Muley, D., 2021. Blue collar laborers' travel pattern recognition: Machine learning classifier approach. Transportation research interdisciplinary perspectives 12, 100506.

An, Y., Wan, L., 2018. Modelling industrial firm relocation with impacts of spatial dependence. International Journal of Urban Sciences 22, 80–103.

Backman, M., Karlsson, C., 2017. Location of new firms: Influence of commuting behaviour. Growth and Change 48, 682–699.

Balbontin, C., Hensher, D.A., 2021. Understanding business location decision making for transport planning: An investigation of the role of process rules in identifying influences on firm location. Journal of Transport Geography 91, 102955.

Balbontin, C., Hensher, D.A., 2019. Firm-specific and location-specific drivers of business location and relocation decisions. Transport Reviews 39, 569–588.

Baptista, R., Mendonça, J., 2010. Proximity to knowledge sources and the location of knowledge-based start-ups. The Annals of Regional Science 45, 5–29.

Ben-Akiva, M.E., Lerman, S.R., Lerman, S.R., 1985. Discrete choice analysis: theory and application to travel demand. MIT press.

Bodenmann, B.R., 2004. Location choice of firms with special emphasis on spatial accessibility. ETH Zurich.

Bodenmann, B.R., Axhausen, K.W., 2012. Destination choice for relocating firms: A discrete choice model for the St. Gallen region, Switzerland. Studys in Regional Science 91, 319–341.

Breidert, C., Hahsler, M., Reutterer, T., 2006. A review of methods for measuring willingness-to-pay. Innovative marketing 2.

Breiman, L., 2001. Random forests. Machine learning 45, 5–32.

Brouwer, A.E., Mariotti, I., Van Ommeren, J.N., 2004. The firm relocation decision: An empirical investigation. The Annals of Regional Science 38, 335–347.

Buczkowska, S., 2017. Quantitative models of establishments location choices: spatial effects and strategic interactions. Université Paris-Est.

Cao, J., Ermagun, A., 2017. Influences of LRT on travel behaviour: A retrospective study on movers in Minneapolis. Urban Studies 54, 2504–2520.

Chen, X., Wang, E., Guo, J., Miao, C., 2021. Location choice and spatial distribution of the electronic information manufacturing industry in China. Growth and Change 52, 1410–1439.

Cheng, L., Chen, X., De Vos, J., Lai, X., Witlox, F., 2019. Applying a random forest method approach to model travel mode choice behavior. Travel behaviour and society 14, 1–10.

Chin, J.T., 2020. Location choice of new business establishments: Understanding the local context and neighborhood conditions in the united states. Sustainability 12, 501.

De Bok, M., Van Oort, F., 2011. Agglomeration economies, accessibility, and the spatial choice behavior of relocating firms. Journal of Transport and Land Use 4, 5–24.
Ding, C., Mishra, S., Lu, G., Yang, J., Liu, C., 2017. Influences of built environment characteristics and individual factors on commuting distance: A multilevel mixture hazard modeling approach. Transportation Research Part D: Transport and Environment 51, 314–325.
Elgar, I., Farooq, B., Miller, E.J., 2015. Simulations of firm location decisions: Replicating office location choices in the Greater Toronto Area. Journal of choice modelling 17, 39–51.
Gabe, T.M., Bell, K.P., 2004. Tradeoffs between Local Taxes and Government Spending as Determinants of Business Location*. Journal of Regional Science 44, 21–41. https://doi.org/10.1111/j.1085-9489.2004.00326.x
Greenwell, B.M., 2017. pdp: An R package for constructing partial dependence plots. R J. 9, 421.
Guimaraes, P., Figueirdo, O., Woodward, D., 2003. A tractable approach to the firm location decision problem. Review of Economics and Statistics 85, 201–204.
Haque, K., Mishra, S., Golias, M.M., 2021. Multi-period transportation network investment decision making and policy implications using econometric framework. Research in Transportation Economics 89, 101109.
Hawkins, J., Nurul Habib, K.M., 2022. A Latent Auction Model of Firm Establishment Location Choice in the Greater Toronto Area. A Latent Auction Model of Firm Establishment Location Choice in the Greater Toronto Area.
Hensher, D.A., Ho, C.Q., Ellison, R.B., 2019. Simultaneous location of firms and jobs in a transport and land use model. Journal of Transport Geography 75, 110–121.
Hensher, D.A., Teye, C., Ellison, R.B., Ho, C.Q., 2017. Integrating an aggregate model of intrametropolitan business location choices into a location-based employment model. Institute of Transport and Logistics Studies (ITLS), University of Sydney (Working Study).
Holl, A., Mariotti, I., 2018. The geography of logistics firm location: the role of accessibility. Networks and Spatial Economics 18, 337–361.
Jiang, H., Xu, W., Zhang, W., 2018. Transportation accessibility and location choice of Japanese-Funded electronic information manufacturing firms in Shanghai. Sustainability 10, 390.
Jiang, Y., Timmermans, H.J., Yu, B., 2018. Relocation of manufacturing industry from the perspective of transport accessibility–An application of percolation theory. Transport Policy 63, 10–29.
Kang, S., 2020. Warehouse location choice: A case study in Los Angeles, CA. Journal of Transport Geography 88, 102297.
Kimelberg, S.M., Williams, E., 2013a. Evaluating the Importance of Business Location Factors: The Influence of Facility Type. Growth and Change 44, 92–117. https://doi.org/10.1111/grow.12003
Kimelberg, S.M., Williams, E., 2013b. Evaluating the importance of business location factors: The influence of facility type. Growth and Change 44, 92–117.
Lampón, J.F., González-Benito, J., García-Vázquez, J.M., 2015. International relocation of production plants in MNEs: Is the enemy in our camp? Studys in Regional Science 94, 127–139.
Lee, K.-D., Hwang, S.-J., 2016. Regional characteristics, industry agglomeration and location choice: Evidence from Japanese manufacturing investments in Korea. Asian Economic Journal 30, 123–145.

Limtanakool, N., Dijst, M., Schwanen, T., 2006. The influence of socioeconomic characteristics, land use and travel time considerations on mode choice for medium-and longer-distance trips. Journal of transport geography 14, 327–341.

Malecki, E.J., 2009. Geographical environments for entrepreneurship. International Journal of Entrepreneurship and Small Business 7, 175–190.

Manski, C.F., 1977. The structure of random utility models. Theory and decision 8, 229.

McFadden, D., 1977. Modelling the choice of residential location.

Nilsen, Ø.L., Díez-Gutiérrez, M., Andersen, S.N., Tørset, T., Cherchi, E., 2020. Where and why do firms choose to move? Empirical evidence from Norway. Journal of Transport and Land Use 13, 207–225.

Paleti, R., Mishra, S., Haque, K., Golias, M.M., 2021. Latent class analysis of residential and work location choices. Transportation Letters 13, 696–706.

Paleti, R., Mishra, S., Haque, K., Golias, M.M., 2020. Latent class analysis of residential and work location choices. Transportation Letters 1–11. https://doi.org/10.1080/19427867.2020.1783610

Pellenbarg, P.H., Van Wissen, L.J., Van Dijk, J., 2002. Firm relocation: state of the art and research prospects. Citeseer.

Pineda-Jaramillo, J.D., 2019. A review of Machine Learning (ML) algorithms used for modeling travel mode choice. Dyna 86, 32–41.

Rashidi, T.H., Auld, J., Mohammadian, A.K., 2012. A behavioral housing search model: Two-stage hazard-based and multinomial logit approach to choice-set formation and location selection. Transportation Research Part A: Policy and Practice 46, 1097–1107.

Rau, H., Popp, M., Namberger, P., Mögele, M., 2019. Short distance, big impact: The effects of intra-city workplace relocation on staff mobility practices. Journal of Transport Geography 79, 102483.

Risselada, A., Schutjens, V., Van Oort, F., 2013. Real estate determinants of firm relocation in urban residential neighbourhoods. Tijdschrift voor economische en sociale geografie 104, 136–158.

Rossi, F., 2019. Identifying Factors Relevant for Firms' Location and Relocation. The Case of Ticino, in: Relocation of Economic Activity. Springer, pp. 109–123.

Sakai, T., Beziat, A., Heitz, A., 2020. Location factors for logistics facilities: Location choice modeling considering activity categories. Journal of Transport Geography 85, 102710.

Samani, Ali Riahi, Mishra, S., Dey, K., 2022. Assessing the effect of long-automated driving operation, repeated take-over requests, and driver's characteristics on commercial motor vehicle drivers' driving behavior and reaction time in highly automated vehicles. Transportation research part F: traffic psychology and behaviour 84, 239–261.

Samani, Ali R., Mishra, S., Lee, D.J., Golias, M.M., Everett, J., 2022. A new approach to develop large-scale land-use models using publicly available data. Environment and Planning B: Urban Analytics and City Science 2399808321999399.

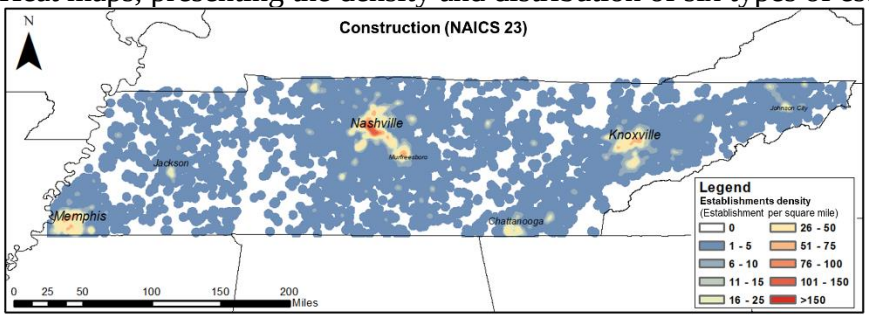
Scheuer, S., Haase, D., Haase, A., Wolff, M., Wellmann, T., 2021. A glimpse into the future of exposure and vulnerabilities in cities? Modelling of residential location choice of urban population with random forest. Natural Hazards and Earth System Sciences 21, 203–217.

Sharma, I., Mishra, S., 2022. Quantifying the consumer's dependence on different information sources on acceptance of autonomous vehicles. Transportation research part A: policy and practice 160, 179–203.

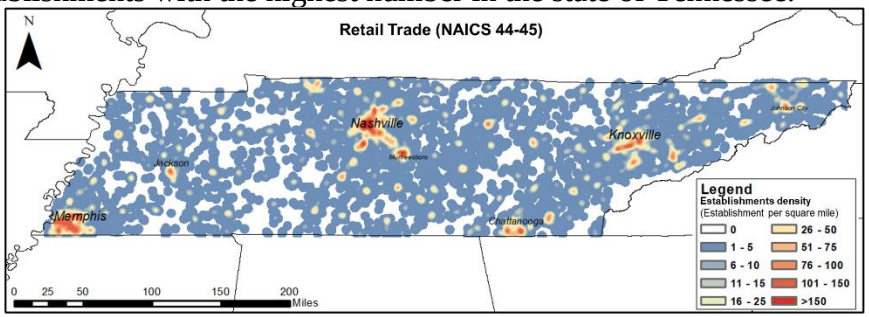
Targa, F., Clifton, K.J., Mahmassani, H.S., 2006. Influence of transportation access on individual firm location decisions. Transportation research record 1977, 179–189.
van der List, C., 2022. How do Establishments Choose their Location? Taxes, Monopsony, and Productivity.
Washington, S., Karlaftis, M.G., Mannering, F., Anastasopoulos, P., 2020. Statistical and Econometric Methods for Transportation Data Analysis. CRC Press.
Weterings, A., Knoben, J., 2013. Footloose: An analysis of the drivers of firm relocations over different distances. Studys in Regional Science 92, 791–809.
Willigers, J., Van Wee, B., 2011. High-speed rail and office location choices. A stated choice experiment for the Netherlands. Journal of Transport Geography 19, 745–754.
Wu, K., Wang, Y., Ye, Y., Zhang, H., Huang, G., 2019. Relationship between the built environment and the location choice of high-tech firms: Evidence from the Pearl River Delta. Sustainability 11, 3689.
Xue, F., Yao, E., 2022. Adopting a random forest approach to model household residential relocation behavior. Cities 125, 103625.
Ye, Y., Wu, K., Xie, Y., Huang, G., Wang, C., Chen, J., 2019. How firm heterogeneity affects foreign direct investment location choice: Micro-evidence from new foreign manufacturing firms in the Pearl River Delta. Applied geography 106, 11–21.
Yi, Y., 2018. Firm relocation and age-dependent reliance on agglomeration externalities. The Annals of Regional Science 61, 439–456.
Yuan, Q., 2021. Location of warehouses and environmental justice. Journal of Planning Education and Research 41, 282–293.
Zhu, S., Pickles, J., He, C., 2017. Global, Regional, and Local: New Firm Formation and Spatial Restructuring in China's Apparel Industry, in: Geographical Dynamics and Firm Spatial Strategy in China. Springer, pp. 73–96.

5.0 APPENDIX A

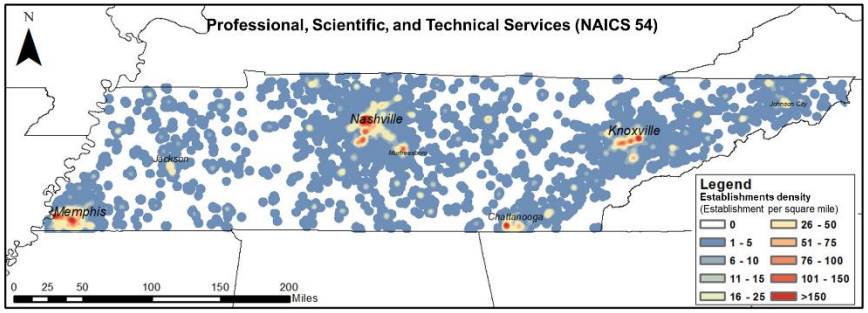
Heat maps, presenting the density and distribution of six types of establishments with the highest number in the state of Tennessee.



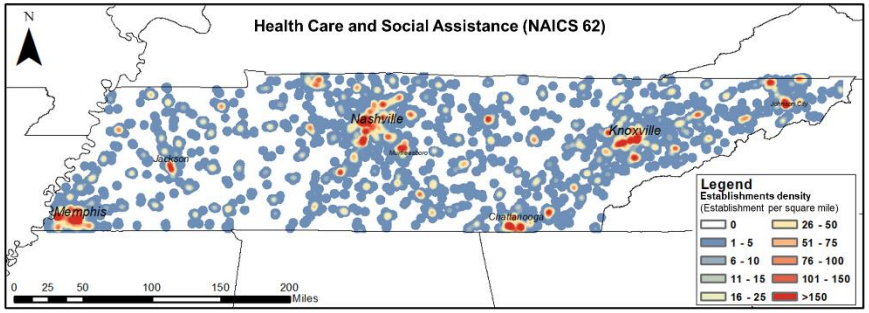
a. Distribution and the density of establishments related to construction (NAICS 23)



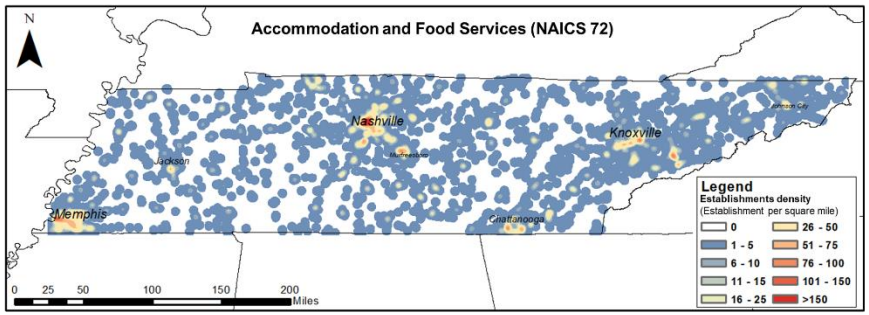
b. Distribution and the density of establishments related to retail trade (NAICS 44-45)



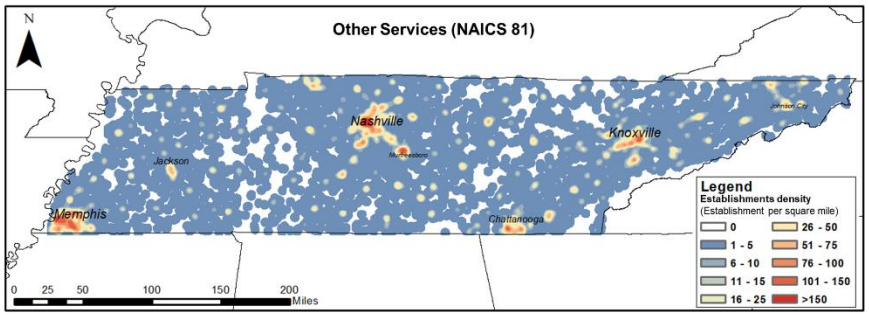
c. Distribution and the density of establishments related to professional, scientific, and technical Services (NAICS 54)



d. Distribution and the density of establishments related to health care and social assistance (NAICS 62)



e. Distribution and the density of establishments related to accommodation and food service (NAICS 72)



f. Distribution and the density of establishments related to other services (except public administration) (NAICS 81)

6.0 APPENDIX B

Results of developing MNL on location choices of different types of establishments, coefficient (*t-value*)

Variables	Agri, Forest, Fishing	Mining	Utilities	Construction	Manufacturing
Intercept 2					
Intercept 3					0.29 (2.27)*
Intercept 4			-0.44 (-1.74).	0.25 (0.85).	
Intercept 5	2.99 (2.78)**				
Office Profile					
Land Value	-1.24 (-1.79).			-1.57 (-13.11)***	-0.65 (-5.62)***
Square Feet	0.9 (6.75)***		1.09 (2.50)*	-0.37 (-11.89)***	0.16 (2.56)*
Office Size			-0.99 (-3.52)***	-1.65 (-7.73)***	-0.74 (-13.64)***
Accessibility					
Interstate		-0.25 (-1.80).		-0.06 (-2.52)*	
Urban Highway		-0.32 (-2.24)*			
Highways			-0.38 (-2.01)*	-0.05 (-2.03)*	
Major Arterials				-0.06 (-2.49)*	-0.06 (-2.10)*
Neighborhood attributes					
Population					-0.08 (-4.86)***
Unemployment		0.84 (1.86).			
Large Size HH					
Highly Educated Pop.					
High Income HH	-0.29 (-2.66)**		1.15 (3.22)**		
Poverty rate < 1		1.27 (2.84)**		0.03 (3.65)***	
Pop. < 18 years old		-2.05 (-2.60)**			
Pop. > 65 years old		-1.65 (-3.72)***			
Commercial Area	-1.31 (-2.07)*			-0.68 (-11.66)***	-0.65 (-7.11)***
Industrial Area			-2.3 (-2.22)*	0.27 (3.70)***	1.06 (10.78)***
Agricultural Area	1.18 (3.50)***				0.28 (4.83)***
Metropolitan				0.26 (12.05)***	
CSA		-1.41 (-3.45)***			-0.11 (-3.15)**
Interaction between establishments					
Agri, Forest, Fishing	6.47 (23.70)***				-0.11 (-2.57)*
Mining		10.57 (8.98)***			-0.12 (-2.02)*
Utilities		-1.1 (-1.82).	6.46 (17.96)***		-0.2 (-4.44)***
Construction	-0.85 (-5.10)***			1.88 (81.93)***	-0.07 (-2.25)*
Manufacturing	-0.33 (-1.85).			0.1 (5.06)***	1.07 (40.02)***
Wholesale Trade			-0.51 (-2.43)*	-0.06 (-3.35)***	
Retail Trade		-2.13 (-3.43)***		-0.12 (-6.06)***	-0.12 (-3.58)***
Transport & Ware.				0.18 (10.15)***	-0.12 (-3.83)***
Information				-0.1 (-5.13)***	-0.22 (-6.98)***
Finance & Insurance				-0.36 (-18.15)***	-0.33 (-9.92)***
Real Estate				-0.1 (-4.97)***	-0.11 (-3.30)***
Technical Services	-0.73 (-3.70)***	-2.64 (-3.82)***		-0.22 (-10.47)***	-0.11 (-3.09)**
Management				-0.16 (-5.53)***	-0.32 (-7.52)***
Admin & Support			-0.51 (-2.37)*	0.12 (6.50)***	-0.11 (-3.59)***
Educational Services	-0.5 (-2.84)**	-1.4 (-2.96)**		-0.18 (-9.93)***	-0.27 (-9.44)***
Health & Social Assis	-0.94 (-4.90)***		-0.92 (-3.52)***	-0.45 (-21.62)***	-0.32 (-9.53)***
Arts & Recreation					-0.16 (-5.20)***
Accom. & Food	-0.72 (-3.60)***		-1.16 (-4.25)***	-0.2 (-9.79)***	-0.28 (-8.20)***
Other Services	-1.11 (-5.90)***	-1.67 (-2.74)**	-1.11 (-3.96)***	-0.49 (-21.97)***	-0.31 (-8.66)***
Public Admin	-0.56 (-3.30)***			-0.18 (-10.15)***	-0.19 (-6.81)***
Model specifics					
Log-Likelihood:	-682	-993	-294.3	-4067	-15907
McFadden R^2:	0.784	0.74	0.823	0.2368	0.306
AIC	1,190.3	267.9	658.2	81,527.9	31,969.3
Num. of Observation	9,850	2,395	5,155	165,585	70,675

$. p < .1, * p < .05,\ ** p < 0.01, \text{and } *** p < .001$

Variables	Wholesale Trade	Retail Trade	Transport & Ware.	Information	Finance & Insurance
Intercept 2					
Intercept 3		0.04 (1.75).		0.16 (2.02)*	
Intercept 4		0.05 (2.13)*	0.17 (3.38)***		
Intercept 5	-0.09 (-2.51)*				
Office Profile					
Land Value	-0.22 (-5.88)***	-0.17 (-15.73)***	-0.17 (-3.28)**	-0.33 (-7.30)***	-0.53 (-11.84)***
Square Feet	2.04 (8.28)***		1.75 (6.65)***	2.05 (6.23)***	-0.11 (-1.66).
Office Size	-0.89 (-13.69)***	-2.01 (-14.84)***	-1.59 (-3.34)***	-1.14 (-18.63)***	-0.7 (-18.14)***
Accessibility					
Interstate		-0.03 (-3.34)***			0.08 (3.78)***
Urban Highway	-0.04 (-2.69)**		-0.01 (-1.65).		
Highways	-0.04 (-2.13)*				
Major Arterials	-1.96 (-2.55)*	-0.03 (-1.81).	-0.05 (-1.82).		-0.03 (-1.81).
Neighborhood attributes					
Population	-0.29 (-5.04)***	-0.13 (-8.62)***	0.04 (2.27)*	0.07 (4.88)***	-0.07 (-4.39)***
Unemployment					-0.05 (-4.09)***
Large HH					
Highly Educated Pop.		-0.04 (-6.13)***	-0.14 (-6.85)***		
High Income HH		0.04 (4.89)***			0.02 (2.33)*
Poverty rate < 1		0.06 (8.96)***			
Pop. < 18 years old					
Pop. > 65 years old				-0.11 (-6.29)***	0.07 (6.85)***
Commercial Area		-0.18 (-5.20)***			-0.16 (-2.72)**
Industrial Area	1.47 (14.23)***	-0.13 (-2.73)**	1.3 (10.49)***		-1.05 (-11.29)***
Agricultural Area		0.22 (9.17)***	0.55 (8.073)***		-0.08 (-1.87).
Metropolitan	-0.11 (-2.71)**	0.05 (3.29)***	-0.23 (-4.85)***		
CSA		-0.03 (-2.24)*			
Interaction between establishments					
Agri, Forest, Fishing		0.04 (1.75).		0.2 (4.39)***	
Mining		-0.13 (-7.61)***			0.16 (4.80)***
Utilities	-0.15 (-3.34)***	-0.1 (-8.29)***	-0.24 (-4.22)***		0.09 (3.69)***
Construction	-0.09 (-2.47)*	0.12 (10.58)***	-0.1 (-2.49)*	-0.11 (-2.63)**	-0.13 (-6.59)***
Manufacturing	0.2 (7.04)***	-0.02 (-1.79).	-0.22 (-5.89)***		
Wholesale Trade	3.01 (68.56)***	1.25 (71.29)***	-0.22 (-5.80)***	-0.2 (-5.24)***	
Retail Trade		0.75 (55.93)***	-0.18 (-4.30)***	-0.13 (-2.68)**	
Transport & Ware.			2.72 (65.60)***	0.08 (2.17)*	
Information	-0.07 (-2.15)*	-0.11 (-8.13)***	-0.22 (-5.83)***	4.06 (58.78)***	0.12 (6.25)***
Finance & Insurance	-0.31 (-8.79)***	-0.03 (-2.58)**	-0.32 (-7.81)***	-0.3 (-5.86)***	3.19 (72.99)***
Real Estate	-0.07 (-2.19)*	-0.21 (-15.07)***	-0.21 (-5.05)***	-0.19 (-4.18)***	
Technical Services	-0.24 (-6.28)***	-0.08 (-4.68)***	-0.25 (-5.72)***	-0.25 (-4.60)***	-0.14 (-5.20)***
Management	-0.11 (-2.54)*	-0.07 (-5.74)***	-0.29 (-5.54)***	0.15 (3.53)***	0.21 (9.24)***
Admin & Support		-0.13 (-11.89)***	-0.13 (-3.14)**		
Educational Services	-0.22 (-7.38)***	-0.32 (-22.90)***	-0.24 (-6.68)***	-0.15 (-4.12)***	
Health & Social Assis	-0.44 (-12.06)***		-0.42 (-10.07)***	-0.29 (-4.94)***	-0.28 (-9.59)***
Arts & Recreation			-0.13 (-3.46)***		-0.04 (-2.44)*
Accom. & Food	-0.39 (-11.23)***	-0.25 (-15.24)***	-0.14 (-3.24)**	-0.22 (-4.05)***	
Other Services	-0.53 (-12.49)***	-0.22 (-20.89)***	-0.22 (-4.84)***	-0.52 (-7.70)***	-0.33 (-9.83)***
Public Admin	-0.16 (-5.59)***				
Model specifics					
Log-Likelihood:	-1,388	-10,333	-10,090	-8,886	-3,3816
McFadden R^2:	0.371	0.173	0.35851	0.353	0.221
AIC	27,918.3	206,776.3	20,273.37	17,889.41	67,761.06
Num. of Observation	68,520	346,625	48,866	42,705	134,885

$. p < .1, * p < .05$, $** p < 0.01$, and $*** p < .001$

Variables	Real Estate	Technical Services	Management	Admin & Support	Educational Services
Intercept 2					
Intercept 3	-0.19 (-1.76).		8.87 (2.49)*		0.65 (2.51)*
Intercept 4	-0.24 (-2.21)*		5.1 (1.8).	0.27 (2.04)*	0.65 (2.48)*
Intercept 5	-0.19 (-1.68).				0.51 (1.94).
Office Profile					
Land Value	-0.33 (-10.44)***	-0.39 (-15.22)***		-0.33 (-8.73)***	-0.14 (-7.95)***
Square Feet	-0.42 (-16.01)***	-0.16 (-5.30)***	-1.24 (-2.81)**		
Office Size	-0.54 (-14.28)***	0.51 (28.78)***		0.11 (4.44)***	
Accessibility					
Interstate			-0.3 (-2.04)*		
Urban Highway			-0.25 (-1.72).		
Highways				-0.4 (-3.77)***	
Major Arterials		-0.94 (-2.74)**		-0.3 (-2.38)*	
Neighborhood attributes					
Population		0.03 (3.80)***		0.06 (3.08)**	0.11 (2.75)**
Unemployment	0.03 (3.44)***				0.04 (2.20)*
Large HH				0.06 (6.05)***	-0.04 (-2.11)*
Highly Educated Pop.		0.04 (6.09)***			0.06 (3.45)***
High Income HH	0.06 (5.10)***	0.08 (10.65)***	0.42 (3.60)***	0.08 (5.59)***	
Poverty rate < 1		-0.03 (-4.05)***		-0.06 (-3.77)***	-0.07 (-3.22)**
Pop. < 18 years old					0.19 (6.55)***
Pop. > 65 years old			-0.63 (-3.56)***		
Commercial Area	-0.3 (-5.10)***		-2.81 (-3.74)***	-0.35 (-4.89)***	-0.76 (-7.28)***
Industrial Area	-0.39 (-4.27)***			0.17 (1.78).	
Agricultural Area		-0.2 (-5.58)***			
Metropolitan	0.07 (2.59)**	-0.06 (-2.66)**			
CSA		0.08 (3.70)***		0.2 (5.90)***	
Interaction between establishments					
Agri, Forest, Fishing	0.09 (3.23)**	0.05 (2.54)*			0.12 (2.49)*
Mining		0.17 (6.63)***			
Utilities		0.12 (5.57)***		-0.12 (-2.92)**	
Construction	-0.1 (-4.42)***	-0.09 (-5.02)***		-0.12 (-4.25)***	-0.23 (-6.16)***
Manufacturing	-0.04 (-1.93).	0.18 (12.12)***		0.09 (3.57)***	-0.09 (-2.56)*
Wholesale Trade	0.07 (3.19)**	-0.06 (-3.65)***	-1.58 (-3.93)***	-0.14 (-5.84)***	-0.19 (-5.65)***
Retail Trade		-0.12 (-6.44)***		-0.11 (-3.86)***	-0.12 (-2.90)**
Transport & Ware.	-0.07 (-3.23)**	0.08 (4.99)***		0.17 (7.50)***	-0.1 (-2.67)**
Information	-0.05 (-2.25)*	0.13 (7.57)***		-0.09 (-3.72)***	-0.21 (-5.86)***
Finance & Insurance	-0.13 (-5.00)***	-0.12 (-6.00)***		-0.31 (-10.96)***	-0.37 (-8.98)***
Real Estate	3.01 (80.99)***	-0.07 (-3.94)***		-0.14 (-5.09)***	-0.21 (-5.51)***
Technical Services	-0.29 (-10.51)***	2.28 (81.55)***		-0.22 (-6.80)***	
Management	0.19 (6.92)***	0.3 (16.20)***	10.15 (10.32)***	0.12 (3.47)***	
Admin & Support	-0.06 (-2. 70)**		0.95 (1.91).	2.81 (77.79)***	-0.12 (-3.20)**
Educational Services	-0.06 (-3.08)**		-1.63 (-3.97)***	-0.13 (-5.46)***	3.61 (69.95)***
Health & Social Assis	-0.33 (-11.52)***	-0.41 (-18.69)***	-1.57 (-2.43)*	-0.42 (-13.74)***	-0.34 (-7.47)***
Arts & Recreation		0.05 (3.22)**			
Accom. & Food	-0.17 (-6.51)***	-0.2 (-10.02)***		-0.28 (-9.49)***	-0.37 (-8.75)***
Other Services	-0.4 (-12.14)***	-0.48 (-19.88)***		-0.6 (-17.23)***	-0.37 (-7.38)***
Public Admin	-0.1 (-5.18)***	0.08 (5.50)***		-0.12 (-5.18)***	
Model specifics					
Log-Likelihood:	-277,784	-69,821	-121.6	-2,129	-10,438
McFadden R^2:	0.2025	0.4956	0.8597	0.2523	0.36046
AIC	55,723.65	106,100.8	395.954	42,768.17	21,025.74
Num. of Observation	108,235	191,440	2,770	88,480	50,705

$.p < .1, * p < .05,\ ** p < 0.01,$ and $*** p < .001$

Variables	Health & Social Assist	Arts & Recreation	Accom. & Food	Other Services	Public Admin
Intercept 2			0.05 (2.13)*	0.05 (1.94).	
Intercept 3		0.18 (2.26)*		0.06 (2.56)*	
Intercept 4	0.15 (2.35)*	0.14 (1.77).		0.05 (2.1)*	
Intercept 5			0.06 (2.55)*	0.07 (3.18)**	
Office Profile					
Land Value	-0.12 (-30.40)***	-0.4 (-8.14)***	-0.29 (-15.94)***	-0.72 (-14.68)***	-0.92 (-6.75)***
Square Feet	-0.65 (-36.00)***				4.22 (5.88)***
Office Size	0.51 (41.80)***	-0.19 (-5.16)***	-2.1 (-10.7)***	-2.21 (-13.86)***	1.62 (32.00)***
Accessibility					
Interstate	-0.09 (-7.39)***			-0.6 (-4.48)***	
Urban Highway	-0.03 (-1.93).	-0.03 (-2.06)*		-0.3 (-3.16)**	
Highways			-0.6 (-2.93)**		-0.99 (-3.89)***
Major Arterials	-0.03 (-1.88).		-0.03 (-1.69).		-0.9 (-3.05)**
Neighborhood attributes					
Population	-0.05 (-5.25)***		0.04 (2.10)*		0.06 (2.03)*
Unemployment	-0.04 (-6.78)***	0.07 (4.65)***		0.02 (3.86)***	
Large HH	-0.07 (-13.73)***				
Highly Educated Pop.	0.02 (3.58)***				
High Income HH	-0.08 (-10.72)***		0.03 (2.02)*		
Poverty rate < 1	0.13 (19.21)***		0.03 (2.71)**	0.02 (3.73)***	0.08 (3.75)***
Pop. < 18 years old			-0.02 (-2.03)*		
Pop. > 65 years old	0.06 (13.22)***	-0.08 (-4.43)***	-0.08 (-7.66)***	-0.04 (-5.93)***	-0.15 (-7.01)***
Commercial Area	0.72 (26.39)***	-0.33 (-3.25)**	0.36 (6.83)***	-0.39 (-10.70)***	-1.03 (-9.61)***
Industrial Area	-1.19 (-23.59)***		-0.46 (-5.76)***	-0.6 (-11.43)***	
Agricultural Area	-0.26 (-10.29)***		0.31 (7.58)***		
Metropolitan	0.12 (5.25)***	0.12 (2.58)**	0.15 (6.49)***	-0.05 (-3.08)**	
CSA	-0.06 (-3.32)***		0.1 (4.55)***		-0.25 (-7.06)***
Interaction between establishments					
Agri, Forest, Fishing	-0.14 (-9.57)***	0.19 (4.10)***		0.05 (2.84)**	
Mining	-0.15 (-7.74)***		0.09 (2.83)**		
Utilities	-0.16 (-10.49)***		-0.05 (-2.16)*		0.32 (8.35)***
Construction	-0.11 (-9.75)***	-0.3 (-6.89)***	-0.24 (-13.14)***	-0.11 (-8.99)***	-0.16 (-4.06)***
Manufacturing	-0.19 (-19.49)***	-0.12 (-3.32)***	0.12 (7.63)***	0.04 (3.045)**	0.17 (4.87)***
Wholesale Trade	0.08 (8.17)***	-0.15 (-3.84)***	-0.12 (-7.32)***	-0.06 (-5.45)***	-0.15 (-4.24)***
Retail Trade	-0.2 (-16.88)***	-0.14 (-2.92)**	0.09 (4.54)***	-0.07 (-5.319)***	-0.14 (-3.22)**
Transport & Ware.	-0.03 (-3.06)**			0.03 (2.98)**	0.31 (8.90)***
Information	-0.06 (-6.11)***	-0.08 (-2.07)*	0.16 (9.56)***		0.11 (2.89)**
Finance & Insurance	0.21 (16.66)***	-0.26 (-5.55)***		-0.08 (-5.45)***	
Real Estate	-0.07 (-5.67)***	-0.15 (-3.28)**	-0.04 (-1.90).	-0.08 (-5.82)***	-0.18 (-4.47)***
Technical Services		-0.34 (-6.62)***	-0.22 (-10.08)***	-0.21 (-14.74)***	
Management	-0.04 (-3.17)**				0.42 (9.61)***
Admin & Support	-0.08 (-7.53)***		-0.09 (-5.18)***		-0.14 (-3.72)***
Educational Services	0.21 (21.98)***	-0.11 (-3.12)**	-0.12 (-7.33)***	0.04 (3.07)**	0.11 (3.12)**
Health & Social Assis	3.07 (103.94)***	-0.5 (-9.54)***	-0.28 (-11.98)***	-0.29 (-19.37)***	-0.29 (-5.66)***
Arts & Recreation	-0.06 (-6.15)***	3.84 (63.16)***	0.09 (5.54)***	0.06 (5.33)***	
Accom. & Food	-0.05 (-3.99)***	-0.17 (-3.35)***	2.94 (86.17)***	-0.11 (-7.54)***	-0.27 (-5.84)***
Other Services	-0.27 (-16.39)***	-0.63 (-10.60)***	-0.42 (-15.20)***	1.39 (76.31)***	-0.6 (-10.31)***
Public Admin	0.04 (4.26)***		-0.26 (-16.16)***	-0.07 (-6.58)***	4.13 (68.72)***
Model specifics					
Log-Likelihood:	-127,440	-9,125.9	-51,596	-88,479	-9258.6
McFadden R^2:	0.18	0.339	0.215	0.175	0.574
AIC	253,311.5	18,391.82	86,500.58	177,114.2	18,645.19
Num. of Observation	486,750	42,925	161,710	297,350	67,520

$.p < .1, * p < .05, \quad ** p < 0.01, \text{and } *** p < .001$