



# Accessibility, Network Structure, and Consumers' Destination Choice: A GIS Analysis of GPS Travel Data and the CLUSTER Simulation Module for Retail Location Choice

**Final Report**

*Prepared by:*

Arthur Huang  
David Levinson

**Department of Civil Engineering  
University of Minnesota**

CTS 12-14

## Technical Report Documentation Page

1. Report No. CTS 12-14	2.	3. Recipients Accession No.	
4. Title and Subtitle Accessibility, Network Structure, and Consumers' Destination Choice: A GIS Analysis of GPS Travel Data and the CLUSTER Simulation Module for Retail Location Choice		5. Report Date October 2012	
		6.	
7. Author(s) Arthur Huang and David Levinson		8. Performing Organization Report No.	
9. Performing Organization Name and Address Department of Civil Engineering University of Minnesota 500 Pillsbury Drive SE Minneapolis, MN 55455		10. Project/Task/Work Unit No. CTS Project #2011064	
		11. Contract (C) or Grant (G) No.	
12. Sponsoring Organization Name and Address Intelligent Transportation Systems Institute Center for Transportation Studies University of Minnesota 200 Transportation and Safety Building 511 Washington Ave. SE Minneapolis, MN 55455		13. Type of Report and Period Covered Final Report	
		14. Sponsoring Agency Code	
15. Supplementary Notes <a href="http://www.its.umn.edu/Publications/ResearchReports/">http://www.its.umn.edu/Publications/ResearchReports/</a>			
16. Abstract (Limit: 250 words) Anecdotal and empirical evidence has shown strong associations between the built environment and individuals' travel decision. To date, data about individuals' travel behavior and the nature of the retail environment have not been linked at the fine-grained level for verifying such relationships. GPS and GIS have revolutionized how we measure and monitor land use and individual travel behavior. Compared with traditional travel survey methods, GPS technologies provide more accurate and detailed information about individuals' trips. Based the GPS travel data in the Twin Cities we analyze the impact of individuals' interactions with road network structure and the destinations' accessibility on individuals' destination choice for home-based non-work retail trips. The results reveal that higher accessibility and diversity of services make the destination more attractive. Further, accessibility and diversity of establishments in a walking zone are often highly correlated. A destination reached via a more circuitous or discontinuous route dampens its appeal. In addition, we build an agent-based simulation tool to study retail location choice on a supply chain network consisting of suppliers, retailers, and consumers. The simulation software illustrates that the clustering of retailers can emerge from the balance of distance to suppliers and the distance to consumers. We further applied this tool in the Transportation Geography and Networks course (CE 5180) at the University of Minnesota. Student feedback reveals that it is a useful active learning tool for transportation and urban planning education. The software also has the potential of being extended for an integrated regional transportation-land use forecasting model.			
17. Document Analysis/Descriptors Shopping trips, Travel behavior, Origin and destination, Destination choice, Global Positioning System, Networks, Industrial location, Retail location choice, Simulation		18. Availability Statement No restrictions. Document available from: National Technical Information Services, Alexandria, Virginia 22312	
19. Security Class (this report) Unclassified	20. Security Class (this page) Unclassified	21. No. of Pages 46	22. Price

# **Accessibility, Network Structure, and Consumers' Destination Choice: A GIS Analysis of GPS Travel Data and the CLUSTER Simulation Module for Retail Location Choice**

## **Final Report**

*Prepared by:*

Arthur Huang  
David Levinson

Department of Civil Engineering  
University of Minnesota

**October 2012**

*Published by:*

Intelligent Transportation Systems Institute  
Center for Transportation Studies  
University of Minnesota  
200 Transportation and Safety Building  
511 Washington Avenue SE  
Minneapolis, Minnesota 55455

The contents of this report reflect the views of the authors, who are responsible for the facts and the accuracy of the information presented herein. This document is disseminated under the sponsorship of the Department of Transportation University Transportation Centers Program, in the interest of information exchange. The U.S. Government assumes no liability for the contents or use thereof. This report does not necessarily reflect the official views or policies of the University of Minnesota.

The authors, the University of Minnesota, and the U.S. Government do not endorse products or manufacturers. Any trade or manufacturers' names that may appear herein do so solely because they are considered essential to this report.

# Acknowledgments

The study was funded by the Intelligent Transportation Systems (ITS) Institute, a program of the University of Minnesota's Center for Transportation Studies (CTS). Financial support was provided by the United States Department of Transportation's Research and Innovative Technologies Administration (RITA).

The authors would also like to thank their reviewer, Dr. Susan Handy, for her valuable comments which helped them improve the report.

# Contents

<b>1</b>	<b>Introduction</b>	<b>1</b>
<b>2</b>	<b>Accessibility, Network Structure, and Consumers' Destination Choice: A GIS Analysis of GPS Travel Data</b>	<b>2</b>
2.1	Literature Review .....	2
2.1.1	Travel Behavior and Land Use .....	2
2.1.2	Travel Behavior and Transportation Networks .....	4
2.1.3	GPS in Collecting Travel Data .....	5
2.1.4	Modeling Shopping Destination Choice .....	6
2.1.5	Summary .....	6
2.2	A Conceptual Framework .....	6
2.3	Empirical Data .....	8
2.3.1	Study Area .....	8
2.3.2	GPS Travel Data in the Twin Cities .....	9
2.3.3	Dun & Bradstreet Data Set .....	9
2.3.4	Street Network Data .....	11
2.3.5	US Census 2000 .....	11
2.4	Consumers' Shopping Destination Choice .....	11
2.4.1	Research Question and Hypothesis .....	11
2.4.2	Define Home-based Non-work Retail Trips .....	11
2.4.3	Variables in the Model .....	12
2.4.4	Choice Alternatives and Choice Set Construction .....	13
2.4.5	Discrete Choice Modeling .....	14
2.4.6	Results and Analysis .....	15
2.4.7	Discussion .....	15
<b>3</b>	<b>CLUSTER Simulation Module for Retail Location Choice</b>	<b>17</b>
3.1	Literature on Modeling Retail Location Choice .....	17
3.2	CLUSTER: An Agent Location Choice Model .....	18
3.2.1	The Model .....	18
3.2.2	The User Interface .....	20
3.2.3	The Variables Panel .....	22
3.2.4	Control Panel .....	22
3.2.5	Visualization Panel .....	23
3.3	Use the Module in Teaching .....	23
3.3.1	Evaluation .....	26
<b>4</b>	<b>Discussion and Conclusions</b>	<b>28</b>
	<b>References</b>	<b>30</b>

# List of Tables

2.1	Summary of selected studies from literature .....	7
2.2	Two types of GPS devices used in this study .....	9
2.3	Definition of home-based non-work trip trips in this research .....	12
2.4	Modeling individuals' destination choice for home-based non-work shopping trips after the opening of the new I-35 W Bridge (number of subjects: 141; number of trips: 2643) .....	16
3.1	Values of parameters (Model 1: homogenous goods; Model 2: heterogeneous goods) .....	22

# List of Figures

2.1	The land use-transport feedback cycle (source: Levinson and Krizek (2008)).	3
2.2	A diamond of action for travel decision-making (source: Levinson and Krizek (2008)).	4
2.3	A conceptual framework of the relationships among transportation networks, land use patterns, consumers' destination choice, and retailers' location choice.	8
2.4	An individual's shopping trip trajectory captured by the in-vehicle GPS device on October 5th, 2008.	10
2.5	An illustration of location choice for home-based non-work retail trips.	15
3.1	The user interface for the CLUSTER applet.	21
3.2	A window showing a summary of variables and results.	23
3.3	Visualization of the retail spatial pattern in equilibrium for Model 1.	24
3.4	Visualization of the retail spatial pattern in equilibrium for Model 2.	25

# Executive Summary

Anecdotal and empirical evidence has shown strong associations between the built environment and individuals' travel decisions. Nevertheless, data about individuals' travel behavior and the nature of the retail environment have yet not been linked at the fine-grained level for verifying such relationships. GPS and GIS have revolutionized how we measure and monitor land use and individual travel behavior, and thus have provided opportunities for filling these research gaps. Compared with traditional travel survey methods, GPS technologies provide more accurate and detailed information about individuals' trips. Based on the GPS travel data of 141 subjects in the Twin Cities, we analyze the impact of individuals' interactions with road network structure and the destinations' accessibility on individuals' destination choice for home-based non-work shopping trips. The results reveal that higher accessibility and diversity of services make a destination more attractive. Further, accessibility and diversity of establishments in a walking zone are often highly correlated. In terms of network structure, a destination reached via a more circuitous or discontinuous route dampens its appeal. In addition, we build an agent-based simulation tool to study retail location choice on a supply chain network consisting of suppliers, retailers, and consumers. The simulation software illustrates that the clustering of retailers can emerge from the balance of distance to suppliers and the distance to consumers. This tool is further applied in the Transportation Geography and Networks course (CE 5180) at the University of Minnesota. Students' feedback reveals that it is a useful active learning tool for transportation and urban planning education. In addition to planning education, the software also has the potential of being extended and calibrated as an integrated regional transportation-land use forecasting model.



# Chapter 1

## Introduction

The American urban life has a strong connection with automobiles. According to the 2009 National Household Travel Survey, 91.4% of work trips nationwide are conducted by vehicles, with 87.8% for family/personal errands (shopping, personal businesses, etc.) and 76.9% for social and recreational activities (Santos et al., 2011). In comparison, public transit only accounts for 3.7% of work trips, 1.4% of family/personal errands, and 1.3% for social and recreational trips. The average vehicle trip length is 12.2 miles for work trips, 6.7 miles for shopping trips, and 11.9 miles for social and recreational activities. In 2009 the average annual vehicle shopping trips per household reached 459, about 58% higher than 1983 and 8.6% higher than 1990. The above statistics posit that (1) automobile is the dominant mode choice, and that (2) people's destinations are distant from origins.

Coupled with auto dominance is the development of low-density suburbs where the road network structure features curved roads and cul-de-sacs, with low network connectivity. In terms of land use, the number of shopping centers witnessed 264% growth during the period 1960-1975 and 113 percent growth during the period 1975-1991 (Feinberg and Meoli, 1991).

This research aims to understand how individuals' non-work destination choice behavior depends on the built environment characterized by land use patterns and network structure. To contribute to planning research along these lines, this research adopts two components: empirical analysis and agent-based simulation. The first component of this research examines the impact of land use and transportation networks on vehicle travelers' shopping destination choice based on GPS travel data of 141 subjects in the Twin Cities. In addition, we propose a conceptual framework to disclose the connections between individuals' shopping location choice and retail location choice. The second component introduces a simulation tool CLUSTER (Clustered Locations of Urban Services, Transport, and Economic Resources) which is created to illustrate the principles of retail location choice for planning education and transportation policy analysis.

The report is divided into three parts. Chapter 2 addresses the questions of accessibility, network structure, and destination choice. Chapter 3 shows the agent-based retail location choice module and several examples. Chapter 4 concludes the report and identifies future work.

# Chapter 2

## Accessibility, Network Structure, and Consumers' Destination Choice: A GIS Analysis of GPS Travel Data

This chapter is organized as follows. First, we summarize existing research on the relationship between land use and travel behavior, the connection between transportation network and travel behavior, the application of GPS technologies in studying travel behavior, and discrete choice models for location choice. Second, a conceptual model on the nexus among travel behavior, land use, and transportation networks is proposed. Third, we further introduce the empirical data used in this research and model individual drivers' retail destination choice.

### 2.1 Literature Review

Boelter and Branch (1960) contend that a city can be conceptualized in two parts: (1) the human factors (such as historical, legal, traditional, psychological factors) that influence the evolution of the city and (2) the physical form of the city, consisting of three types of nodes (the domicile, place of work, and service). There is an extensive body of literature examining the nexus between transportation and land use. This report reviews four research topics most relevant to our research.

#### 2.1.1 Travel Behavior and Land Use

Travel behavior and land use have long been found to be closely related (Handy, 1996; Schwanen et al., 2004; Scott and Horner, 2008). Accessibility is one key concept linking the two together. Accessibility refers to the ease of accessing activities given a place, considering “the desire of people or firm to overcome spatial segregation” (Hansen, 1959). Accessibility implies two aspects of the transport-land use connection. First, it evaluates the spatial distribution of amenities (e.g., stadiums and retail stores) and opportunities (e.g., jobs and services). Second, it concerns the ease of access, decided by transportation networks, transportation mode, and distance between places. Accessibility measures can be categorized into three types: cumulative opportunities, gravity-based measures, and random utility theory (Handy and Clifton, 2001). These measures serve to characterize the relationships among opportunities to access, the cost, and transportation and land use systems. Con-

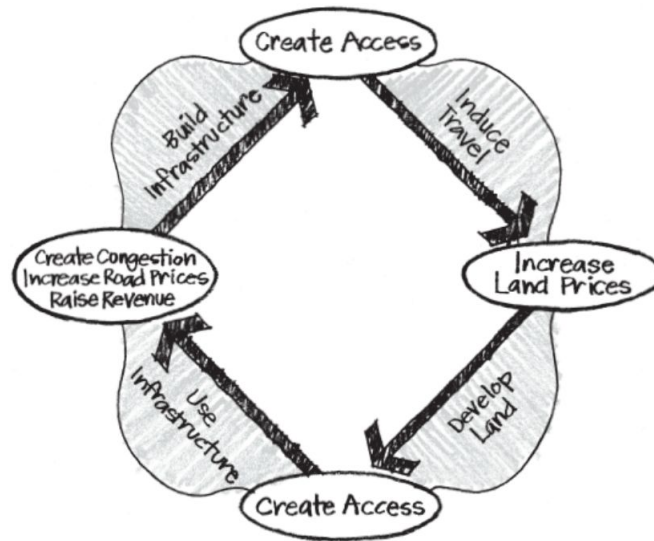


Figure 2.1: The land use-transport feedback cycle (source: Levinson and Krizek (2008)).

ceptually, such relationships can be illustrated in Figure 2.1, where creating congestion, creating access, increasing land prices, building infrastructure, and individuals' travel decisions constitute a closed loop (Levinson and Krizek, 2008). Figure 2.2 further shows that an individual makes travel decisions based on existing opportunities and constraints (e.g., quality of service, travel cost) from competitors and complements of the services they desire to patronize (Levinson and Krizek, 2008). A series of models such as ITLUP (Putman, 1983), IRPUD (Wegener, 1982), and TRANUS (de la Barra, 1989) have incorporated such relationships in modeling. Wegener (2004) and Iacono et al. (2008) perform a comprehensive review on the state-of-art transportation-land models. Most of the models consider the connection between land use dynamics and travel behavior at the aggregate or disaggregate level.

A plethora of efforts have been made to identify such connections empirically. The general approach to study the impact of land use on travel behavior is to regress travel behavior measures (e.g., VMT or mode choice) on the built environment measures (e.g., population density, mixed use, connectivity, and accessibility) and travelers' demographic features. At the aggregate level, some studies have found that neighborhoods with greater land use mix, greater employment and population densities, or job-house balance, all else equal, are associated with lower auto use, higher trip auto/transit mode share, or reduced trip length (Cervero, 1989; Cervero and Kockelman, 1997; Ewing et al., 1995; Handy, 1993; Schimek, 1996). However, Boarnet and Sarmiento (1998) argue that (1) such aggregated studies do not sufficiently control for the demographic differences of these neighborhoods, and that (2) the sample size of the studied neighborhoods is too small. Thus, analyzing travel data at the disaggregate level may provide new insights. Such examples include Boarnet and Crane (2001); Cervero (1996); Crane and Crepeau (1998), which control for socio-demographics of the trip maker and household characteristics. Cao et al. (2006); Handy et al. (2006) further inspect individuals' attitudes toward travel and find it statistically significant in explaining the choice of walking to a destination. With a comprehensive review and meta-analysis of studies on the built environment and travel behavior, Ewing and Cervero (2010) conclude that VMT is

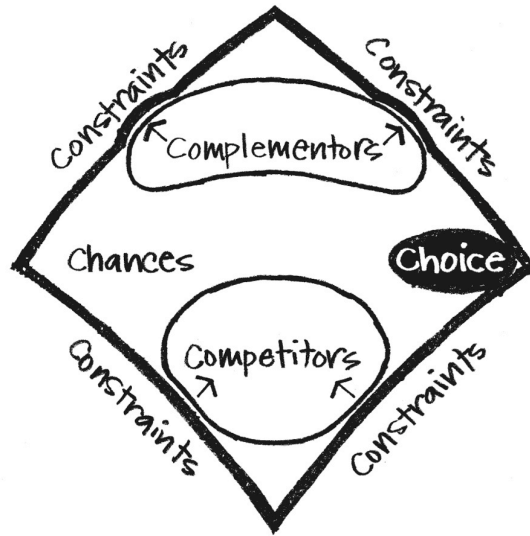


Figure 2.2: A diamond of action for travel decision-making (source: Levinson and Krizek (2008)).

most strongly associated with accessibility to destination and secondly with street network design.

The individual/household travel data in the majority of the above studies are based on recall activity surveys and self-reported trip diaries. The disadvantage of this method is that trips may be underreported or misreported (Bachu et al., 2001). Houston et al. (2011) in a study to use GPS devices to track 47 residents' diurnal patterns and traffic exposure in Long Beach, CA, find that respondents underreported 49% of the locations and trips recorded in GPS devices. Therefore there is a call for new technologies to collect individuals' travel data which can provide more complete activity profiles.

### 2.1.2 Travel Behavior and Transportation Networks

Transportation network structure is also related with travel behavior. Traditional interests in transportation networks are more in the fields of geography (seeing networks as an input to regional development) (Taaffe et al., 1996) and physics (focusing on the topology and spatial evolution of the networks) (Gastner and Newman, 2006). Yet the connection between transportation network structure and travel behavior has been insufficiently investigated, particularly at the microscopic level. For example, the 3D principles (density, diversity, and design) by Cervero and Kockelman (1997) lack a clear description of transportation networks.

With the availability of fine-grained network data, the past decade has witnessed a growing interest in studying the connection between travel behavior and transportation networks. Xie and Levinson (2007) propose a set of new network measures (such as ringness, webness, beltiness, circuitness, and treeness) from travelers' perspective. Hess (1997) uses block size, length and completeness of sidewalk networks to indicate street network connectivity to explain the pedestrian volumes between two neighborhoods. Jiang et al. (2009) study human mobility patterns in the context of a large street network. Levinson and El-Geneidy (2009) use network circuitry (the ratio of

road network distance to Euclidean distance) to understand the choice of home-work pairs. Based on the data in the Twin Cities and Portland, they find that road network pattern transitions from grid-like to tree-like as moving away from the city center. In terms of travel, workers are more likely to commute with lower circuitry for saving travel time (Levinson and El-Geneidy, 2009). Derrible and Kennedy (2009, 2010) quantify the network structure of 33 metro systems, and find that the metro network structure and transit ridership are closely associated. Additionally, the hierarchy, topology, morphology, and scale of road networks are found to be associated with household spatial activities, road congestion levels, trip distance, and daily vehicle kilometers traveled per capita (Parthasarathi et al., 2011; Parthasarathi and Levinson, 2011a,b). Based on a 2009 survey of students at Davis High School in Davis, CA, Emond and Handy (2012) conclude that students' perception of distance from home to school (which implies the effects of biking network) strongly affects their choice of biking even after controlling for actual travel distance.

### **2.1.3 GPS in Collecting Travel Data**

The first few studies using GPS to collect travel data date to the 1990s (Casas and Arce, 1999; Draijer et al., 2000; Wagner, 1997). The GPS technologies have gained popularity in collecting travel data as its precision improves over time. Schönfelder et al. (2005) contend that instrumented vehicle GPS data can provide unique insights into the “structure, size, and stability of human activity spaces”. Its advantages over traditional paper-and-pencil diary methods include (Draijer et al., 2000): (1) real-time spatial and temporal information about a trip is available, such as distance, travel times, travel speed, and route information; (2) no misreporting or underreporting of trips; (3) the data are stored in digital formats; (4) the burden of the correspondent is reduced (some researchers further make phone interviews with the subjects or require them to fill online surveys). In addition, the procedure to draw trip trajectories based on GPS points in GIS can be automated. The readers may refer to Li (2004), Quddus et al. (2005), Quddus et al. (2007), and Zhu (2010) for details about processing GPS travel data in the GIS environment.

Some previous research using GPS to study travel behavior include: Li et al. (2004) (an inspection of the travel time variability in commute trips, and its effects on departure time and route choice, including cases with trip-chaining), Li et al. (2005) (an analysis of attributes determining whether to choose one or more routes in the morning commute), Wolf et al. (2003) (capturing under-reporting trips in household travel surveys), and Zhang and Levinson (2008) (an estimation of the value of information for travelers, and a comparison of the impact of information with other variables such as travel time, distance, and aesthetics).

GPS has also been used to particularly investigate individuals' shopping behavior. Yue et al. (2011) use GPS-based taxi trajectories to calibrate Huff's model. Based on the taxi trips to major shopping malls (choice-based samples) in Wuhan, China, they calculate the proportion of visits for each center, and regress it on stores' attractiveness measures (such as size and distance). Nevertheless, their model does not incorporate individuals' socio-demographics and does not model each trip's destination choice. Kawasaki and Axhausen (2009) collect person-based GPS data for 3,521 individuals in Zurich and Winterthur, Switzerland for one week. A shopping trip is extracted from the GPS data if the destination is within 30 meters from a grocery store, and the nearest store is seen as the destination. They use the  $k$ -means clustering algorithm to partition stores in the two cities into different clusters and randomly sample clusters as alternative destinations for each trip. However, this research does not consider the fact that an individual may visit multiple stores in one

trip, and the clustering algorithm does not consider the road network.

### **2.1.4 Modeling Shopping Destination Choice**

In research on location choice, shopping destination choice has been a topic of keen interest. Shopping destination choice is often modeled with discrete choice models. Such models are based on the random utility maximization hypotheses, meaning that in a choice set of all available alternatives an agent selects one with the highest utility. A basic structure is the multinomial logit model (MNL). McFadden (1978) shows that the MNL model can consistently estimate parameters from a sample of alternatives through maximizing the conditional likelihood function, a feature that makes MNL widely used in modeling discrete choices. One important hypothesis about MNL is the independence of irrelevant alternatives property (IIA): the random components of the utilities are independently and identically distributed (i.i.d.) (Ben-Akiva and Lerman, 1985). Since this assumption does not hold in many cases (such as the red bus/blue bus paradox), a variety of extensions have been developed; two classes of models receive the most attention: generalized extreme value (GEV) class and mixed multinomial logit (MMNL) class of models (Bhat and Guo, 2004). In addition, because retail services are often clustered, there are two schools of thoughts: (1) a consumer considers all possible alternatives, (2) individuals initially evaluate clusters of alternatives and then evaluate alternatives in a cluster (Fotheringham, 1988). Table 2.1 summarizes literature on discrete choice models applied in studying shopping destination choice. These studies tend to use traffic zones or specific stores (such as big supermarkets or malls) as destinations.

### **2.1.5 Summary**

With the above literature review, we identify the following research gaps:

- There is a lack of empirical study in properly defining a traveler's destination choice set using GPS travel data.
- There is a lack of theory on defining a place/destination from a behavioral perspective.
- There is a lack of study on examining the impact of road network structure on individuals' retail shopping behavior using GPS travel data.

## **2.2 A Conceptual Framework**

To shed light on the micro-foundation of the nexus between the built environment and individuals' destination choice, we propose a conceptual framework (Figure 2.3) which features a feedback loop.

The first pathway is the impact of transportation networks and land use on individuals' destination choice. Transportation networks can be measured by topology and the hierarchy of roads. Land use patterns impact the number of potential activities and types of activities one can engage in. Transportation networks and land use patterns together affect distances between trip origins and destinations. Further, network topology and network hierarchy influence travel time reliability and perceived reachability of a destination. Distance between origins and destination and network

Table 2.1: Summary of selected studies from literature

Study	Data	topic	Model	Results
Timmermans (1996)	Self-completion questionnaires in Eindhoven, Netherlands	Sequential mode and destination choice of shopping trips	MNL	The choice of transport mode does not influence the choice of shopping centers. Shopping centers with greater size are more attractive than those of smaller size.
Pellegrini et al. (1997)	Phone survey data on shopping trips in Gainesville FL	Parameter sensitive to choice set specification for shopping destination choice	MNL	The stability of parameter estimates can be sensitive to choice set size and composition.
Bhat (1998)	1990 San Francisco Bay Area Household Travel Survey	Travel mode and departure time choice of shopping trips	MNL for mode choice and MNL-OGEV for departure time choice	In estimating travel time choice, nested logit model outperforms the MNL model and MNL-OGEV model outperforms the nested logit model in terms of data fit.
Pozsgay and Bhat (2001)	1996 Dallas-Fort Worth household activity survey	Destination choice for home-based recreational trips	nonlinear-parameter MNL	Agglomeration effects are prominent in affecting recreational attraction-end choice.
Bernardin et al. (2009)	Destination choice of home based maintenance trips and home-based other trips	Household survey in Knoxville, Tennessee	MNL, agglomerating and competing destination choice models (ACDC)	The ACDC model reflects the effects of trip chaining and spatial agglomeration whereas MNL cannot.
Newman and Bernardin (2010)	Mode choice and destination choice for work tours	2000 Knoxville Urban Area Household Travel Behavior data	(Hierarchical ordering) nested logit	Hierarchical ordering of decision nesting trees is important for modeling location and mode choice; employing a reverse ordering can be a good choice.
Leszczyc et al. (2000)	Grocery shopping data in Springfield, MO	Consumers' store choice and trip time choice	Hazard model and MNL	Store choice and shopping time choice are interdependent. Spatial competition between stores affect consumers' store choice and switching behavior.
Wang and Lo (2007)	Stated shopping preference data in Chinese immigrants in Toronto	The preferences of Chinese immigrants between the fast-growing Chinese supermarkets and competing mainstream supermarket chains.	MNL	Immigrant shoppers' socio-demographic attributes more affect their shop destination choice than distance and accessibility.

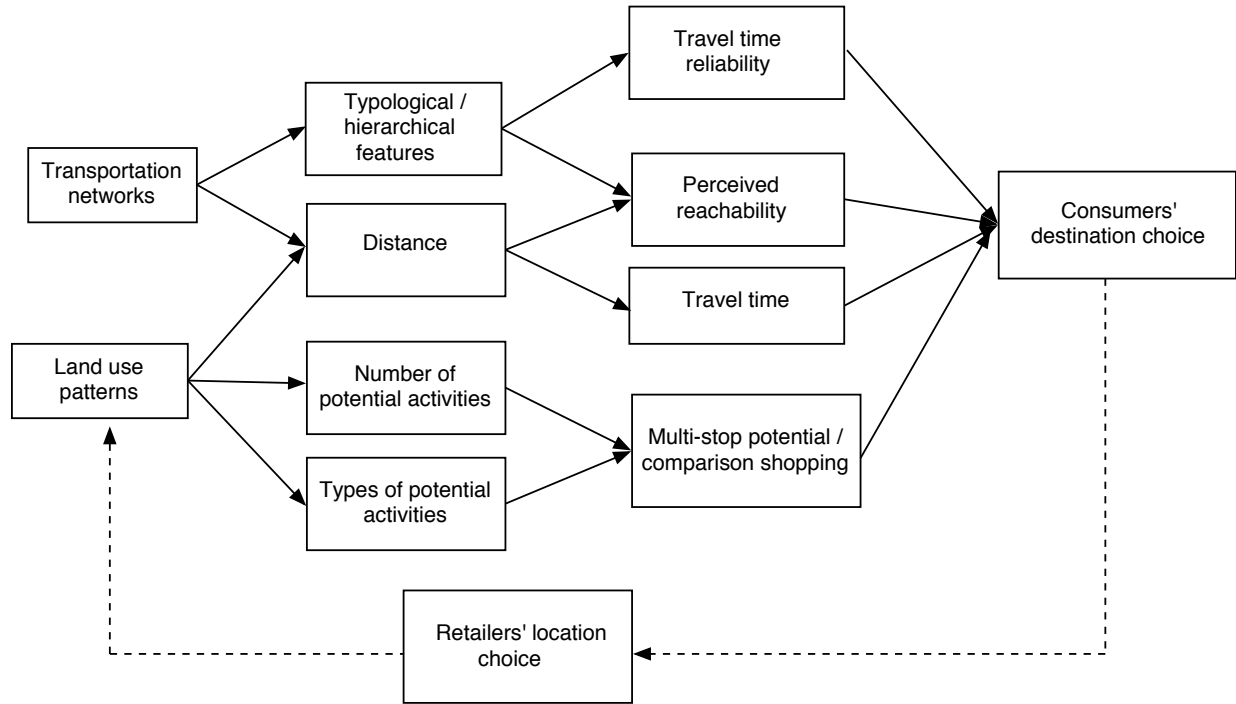


Figure 2.3: A conceptual framework of the relationships among transportation networks, land use patterns, consumers' destination choice, and retailers' location choice.

topology/hierarchy affect travel time and individuals' perceived reachability of destinations. On the land use side, the number and types of potential activities provide consumers incentives for multi-stop shopping and comparison shopping. At the end of the first pathway, consumers' perceptions of networks and land use, such as travel time reliability, perceived reachability, perceived travel time, and opportunities for comparison shopping, serve as inputs to one's travel decisions.

The second pathway is the influence of individuals' destination choice on retail destination choice, which ultimately shapes land use. We use dashed lines to indicate that (1) retail destination choice is induced by cumulative decision making on all consumers' part, and that (2) there exist spatio-temporal differences between these two pathways.

To examine the hypotheses in the framework, we employ the GPS travel data to study the first pathway and use simulation to illustrate the second pathway.

## 2.3 Empirical Data

### 2.3.1 Study Area

The Minneapolis-St. Paul Metropolitan Area is selected as the study area because of the availability of the unique GPS travel data. This area typifies many regions, including high-density CBDs and low-density suburbs, allowing comparison of the difference between suburban non-work trips and urban non-work trips.



Table 2.2: Two types of GPS devices used in this study

Product	Feature	Frequency	No. of subjects	Diaries
Ostarz VT-1000	Logging	One point per 25 meters	97	No
VMT, Inc	Real time	One point per second	44	Yes for partial trips

### 2.3.2 GPS Travel Data in the Twin Cities

GPS travel data in the Twin Cities were collected before and after the collapse and reopening of the I-35 W Bridge in Minneapolis from October 2008 to December 2008 for 13 weeks. The original objective was to understand commuters' travel behavior change after the reopening of the replacement of the I-35 W bridge (Zhu, 2010). The project was funded by the MnDOT, NSF, and Oregon Transportation Research and Education Consortium.

The targeted subjects in this study have the following characteristics: (1) age between 25 and 65, (2) legal drivers, (3) have a full-time job and follow a "common" work schedule, (4) drive alone to work, and (5) are likely to be affected by the opening of the replacement I-35W Mississippi River Bridge. The data collection process lasted from September to December of 2008, during which about 100 surveyed subjects made over 20,000 trips. The process included three stages (Zhu, 2010): (1) Recruiting the subjects. People interested in the study were told to finish an online survey about their demographics, home and work addresses, and work schedule, and only those with the required characteristics were selected, of which 147 participated and finished the whole data collection period. (2) Installing GPS devices in selected subjects' vehicles. (3) Collecting the subjects' vehicle travel data for 13 weeks. The subjects were told to follow their regular travel routes and to complete a series of surveys for several weeks about more trip information. There were two types of GPS devices used in his study which are compared in Table 2.2. The first type was the real-time tracking GPS devices provided by the subcontractor Vehicle Monitoring Technologies (VMTInc). The GPS frequency was one point per second and was transmitted to the server in real time. Professional installers installed the real-time GPS devices in 43 subjects' vehicles (28 females and 15 males). The subjects were required to complete an online travel diary once a week. The second type of devices was the logging GPS (QSTARZ BT-Q1000p GPS Travel Recorders). In contrast with the previous type, the data can only be exported manually at the end of the study. The GPS frequency was one point per 25 meters. In all, 97 subjects' vehicles were equipped with this type of GPS devices. Throughout the study period, the subjects were not required to report travel diaries. (3) Creating GPS trip trajectories. The trip trajectories are drawn based on the GPS points in the underlying the Metropolitan Planning Road Network of the area. The details on creating the trajectories can be found in Zhu (2010). Figure 2.4 shows an exemplary shopping trip trajectory from an in-vehicle GPS device.

### 2.3.3 Dun & Bradstreet Data Set

Purchased from Dun & Bradstreet, Inc, the Dun & Bradstreet data set (2005) documents establishment-level data containing attribute information on location, sales, employees. The establishment-level

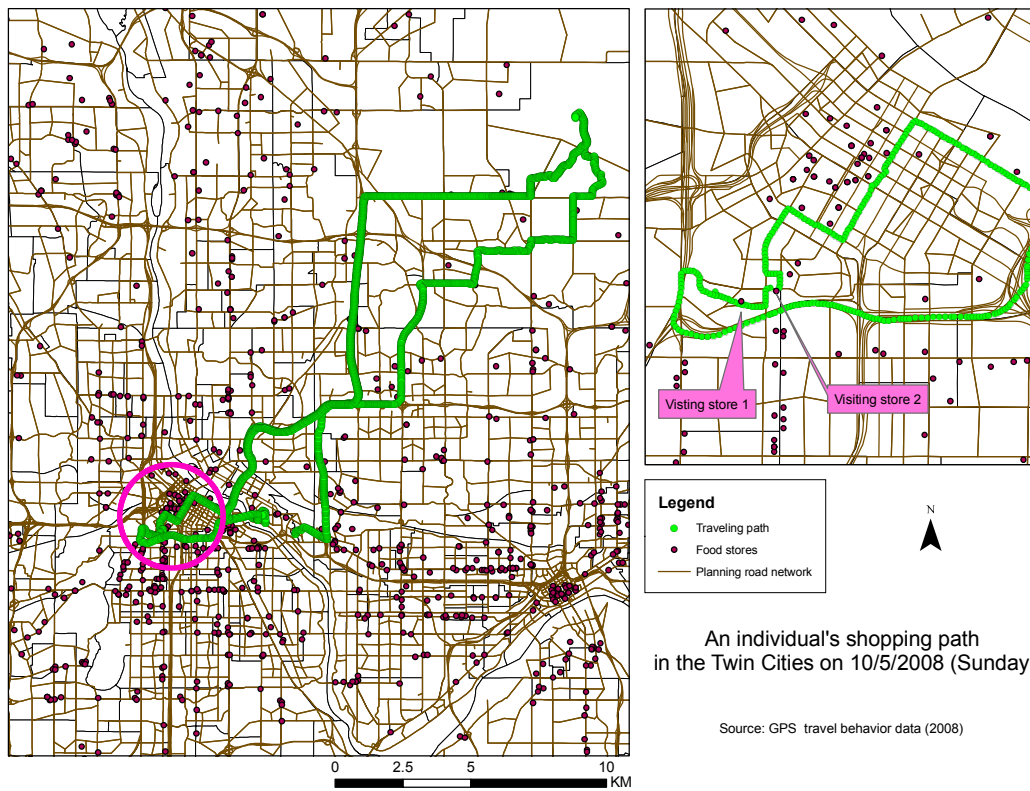


Figure 2.4: An individual's shopping trip trajectory captured by the in-vehicle GPS device on October 5th, 2008.

data were grouped into 17 categories based on the 6-digit classifications of the North American Industry Classification System (NAICS). The data are cleaned and merged with parcel-level land use data from the Metropolitan Council. Details about cleaning and merging the two data sets can be found in Horning et al. (2008).

### **2.3.4 Street Network Data**

The street data used in this research include the metropolitan planning network in the Twin Cities and the TLG road network data. The metropolitan planning network is simpler than the TLG network. The GPS trip trajectories are created based on the metropolitan network to reduce computational burden. Aggregating the speeds of travelers' vehicles across multiple observations per link allows us to estimate weekday average travel speed in three time periods (morning peak hours, afternoon peak hours, and midday) (Zhu et al., 2010).

### **2.3.5 US Census 2000**

This data set incorporates income and population information at the census block/block group level.

## **2.4 Consumers' Shopping Destination Choice**

### **2.4.1 Research Question and Hypotheses**

The research question, stated simply, is why do consumers drive to certain places for shopping? Considering the effects of land use and transportation networks on individuals' shopping trips, we hypothesize that:

1. Destinations with greater accessibility and diversity of services, all else equal, are more attractive.
2. Destinations farther away, all else equal, are less attractive.
3. Destinations reached via a more circuitous or discontinuous route, all else equal, are less attractive.

### **2.4.2 Define Home-based Non-work Retail Trips**

This research focuses on home-based non-work retail shopping trips which are defined as trips starting from home and ending on a retail parcel which is not the workplace. The reason for focusing on such trips is that destination choice in this case is less likely to be affected by work trips. The first task is to find such trips from the whole GPS data set. To this end we need first to understand how far people typically park from home/work when driving home or to work. We do this by using the trips that have both GPS trajectories and travel diaries from our data.

The processing procedure is described as follows:

1. Add the trip purpose attribute from travel diaries to the GPS data by matching the time stamp of the GPS trips and that of the trips noted in the individual's diaries.

Table 2.3: Definition of home-based non-work trip trips in this research

Point	Selection criteria
GPS origin	$\leq 800$ meters from home
GPS destination	<ul style="list-style-type: none"> <li>• <math>\geq 1000</math> meters from work</li> <li>• On a mixed-use or commercial land use parcel</li> </ul>

2. Select trips whose trip purpose is going to work or going home.
3. If a trip is a home trip, calculate the Euclidean distance between the trip destination and the home address. If a trip is a work trip, calculate the Euclidean distance between the trip destination and the work address.

We further examine the distribution of the distances between parking locations and home/work addresses. Some distances are too large to be defined as reasonable. It could be due to wrong trip purpose information from the travel diaries. We make an educated guess and use the 80 % threshold (after the 80 % threshold there are big gaps among the distances). Therefore, based on the threshold, we posit that if a trip is to be defined as a home trip, the parking location should be within 800 meters from the individual's home address. If a trip is to be defined as a work trip, the parking destination should be within 1000 meters from the individual's work address.

Defining whether a trip is a shopping trip depends on the land use of the trip destination. Wolf et al. (2001) have shown that it is feasible and reliable to derive trip purpose from GPS data using spatially-accurate land use parcel data. Therefore, in our study if a trip ends on a commercial or mixed-use parcel, it is considered as a shopping trip.

Based on the above analysis, a home-based non-work retail shopping trip is defined as a trip whose origin is within 800 meters from the individual's home address and whose origin is at least 1000 meters away from the individual's work address and end on commercial or mixed-use land use parcels (Table 2.3). Overall, we identify 141 subjects' 2643 home-based non-work retail trips in our data.

### 2.4.3 Variables in the Model

The dependent variable is destination choice. In a choice set, if a destination was actually visited by a traveler, the destination is marked by 1; otherwise 0. Independent variables have three categories: land use, network structure, and socio-demographics.

The land use category includes measuring accessibility and diversity of services around a parking destination. A destination's accessibility in this research is measured by cumulative opportunities within 10-min walking from a parking destination. The 10-min walking area is created with the Network Analyst tool in ArcGIS based on the TLG road network, with the assumption that people's average walking speed is 5.47 km per hour (3.4 miles/hr) (Krizek et al., 2009). A destination's accessibility a destination in this research is measured by cumulative opportunities.

The diversity of services is measured by the entropy of services around the parking destination. The entropy of services around destination  $k$  equals:

$$H_k = 1 - \sum_{s=1}^S p_{ks} \ln(p_{ks}) \quad (2.1)$$

Where  $p_{ks}$  is the proportion of service type  $s$ . The service type of a store is decided by the six-digit North American Industry Classification System (NAICS) code.  $S$  is the total number of services at destination  $k$ . The greater  $H_k$  is, the more diverse a zone is.

The network structure variables focus on the interaction of vehicle trips with road networks. The variables include travel time, discontinuity, circuitry, and roadness. Travel time for an actual trip can be calculated based on GPS points. Travel time for an alternative destination is the shortest network travel time from origin to destination. Depicting the degree of road hierarchy in a trip, discontinuity of a trip is calculated as the sum of the changes of road hierarchies by every 5 mph on the route. The greater the value, the more discontinuous a trip is. Circuitry aims to quantify how circuitous a trip is. It is computed as the ratio of the network distance of a trip to the Euclidean distance of a trip. The greater the value, the more circuitous a trip is. Roadness is a measure of the number of turns a driver needs to make in a trip. In this report it is calculated as the sum of the number of roads along a route. A road is defined based on consecutive intersections. If the acute angle of every three intersections is between 170 degrees and 180 degrees, they constitute one road; otherwise two roads. It is hypothesized that greater roadness of a route increases the complexity of a trip and thus the mental transaction costs required to travel to that route, and therefore reduces the attractiveness of a destination.

#### 2.4.4 Choice Alternatives and Choice Set Construction

The alternative destinations of a trip are generated from the Census blocks that the trip does not end in. The Census blocks are used as alternative destinations as they are the finest spatial units where socio-demographic and economic data are available. Note that there are 42,243 blocks in the Seven-County Twin Cities region, and to incorporate all of them in an individual's choice set is computationally infeasible. Traditional methods to create choice set include one actually visited destination and some number of randomly selected destinations from the region. Yet this approach biases toward choosing more destinations that are far away than nearer, and does not consider individuals' perception of the area they would consider traveling to. In this study, we propose a new method to construct individuals' destination choice set.

This method is a combination of Monte Carlo simulation and random drawing, which considers both the distribution of trips by travel time and the differences among individuals. Consistent with Kawasaki and Axhausen (2009), we set the choice set for each individual to be 20. Other sizes of the choice are also tested in the analysis. The steps to construct the choice set are as follows:

1. All trips are divided into 12 categories based on a 5 minutes' interval ranging from 0 to 60 minutes. The proportion of trips within each interval is calculated and ranked from 1 to 12. We can estimate the function of the total number of trips (represented by  $y$ ) in each interval based on the rank order (represented by  $x$ ). Based on this function, the estimated proportion of trips in each time interval can be obtained.

2. Based on Monte Carlo simulation, given the choice set size, we generate the number of destinations in each geographical area from one's home by the travel time intervals.
3. The centroids of Census blocks in the Twin Cities are generated from the Census blocks, where each centroid is seen as a possible destination.
4. Given the existing road network in the Twin Cities with estimated travel speed on each link, the travel region from each individual's home by the 5 minutes' time interval (from 5 minutes to 60 minutes) can be generated using the ArcGIS Network Analyst tool. The centroids of blocks in different travel regions from one's home can be found.
5. The alternative 19 destinations for each trip are randomly selected from each time interval based on the outputs of Monte Carlo simulation.

The advantage of this method is that it considers the trip distribution for all individuals and narrows the search area down to all equivalently distant areas as the distribution of actual trips, while incorporating random choice within the search area.

## 2.4.5 Discrete Choice Modeling

In modeling destination choice, the multinomial logit model and its variations are most widely used. An illustration of a traveler's destination choice for trips ending on non-work non-home alternative destinations is shown in Figure 2.5. Since the GPS data are panel data with repeated choices for individuals, this research applies the mixed-logit model to model shopping destination choice. Let  $Y_{ik}$  be the destination chosen by individual  $i$  on choice  $k$  and  $b_i$  be the random term associated with individual  $i$ . For each choice presentation, the mixed-effects logit model is:

$$(Y_{ik}|b_{ik}) \text{ binomial}(1, p) \quad (2.2)$$

$$\text{Logit}(p_{ik}) = U_{ik} + b_i \quad (2.3)$$

$$U_{ik} = V_{ik} + \epsilon_{ik} \quad (2.4)$$

$$b_i = N(0, \delta^2) \quad (2.5)$$

Consistent with random utility theory, the utility for consumer  $i$  to visit destination  $k$ ,  $U_{ik}$ , is defined as:

$$U_{ik} = f(\mathbf{L}_k, \mathbf{R}_{ik}, \mathbf{S}_{ik}) \quad (2.6)$$

where  $\mathbf{L}_k$ : a vector of a destination  $k$ 's land use measures.

$\mathbf{R}_{ik}$ : a vector of trip-related road network measures from  $i$ 's home to destination  $k$ .

$\mathbf{S}_{ik}$ : a vector of individual  $i$ 's socio-demographic characteristics interacting with land use variables of destination  $k$ .

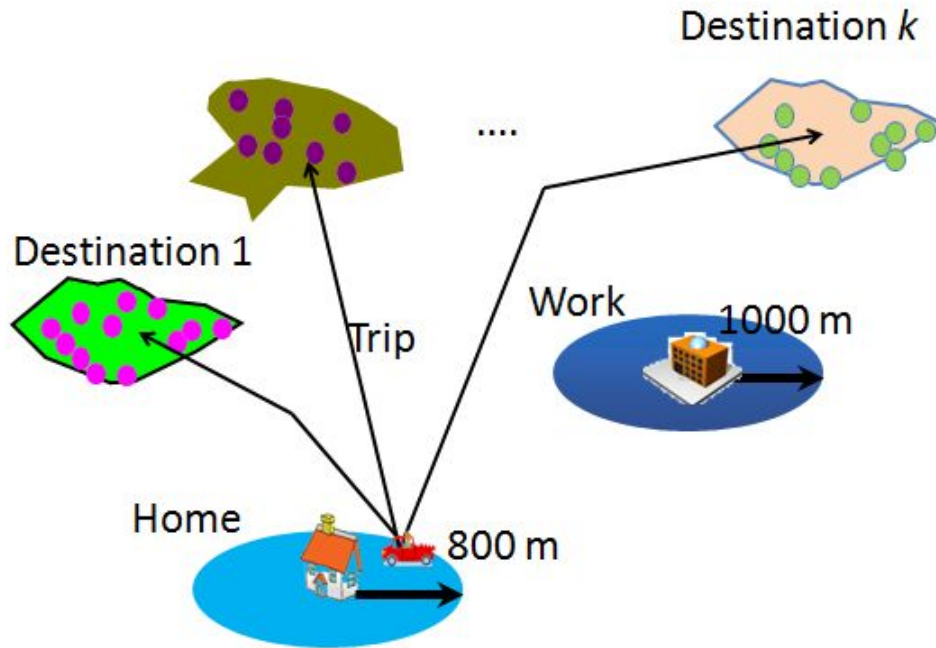


Figure 2.5: An illustration of location choice for home-based non-work retail trips.

## 2.4.6 Results and Analysis

First we test the correlation between accessibility and diversity. The Pearson correlation of the two variables equals 0.97, meaning that they are highly correlated. Considering the high correlation of these two variables, we run the model separately for accessibility and diversity. The results are shown in Table 2.4. Model 1 only uses accessibility measures to represent the built environment. Model 2 only used diversity to measure the built environment. The AIC value in Model 1 is lower than Model 2, indicating that Model 2 has a higher goodness of fit than Model 1.

The positive coefficients of diversity and accessibility around the destinations show that, as hypothesized, more services or more types of services around a location, the more favorable it is for travelers to choose as a destination.

Network structure affects a destination's attractiveness. In addition to travel time, circuitry, and roadness of paths have negative impacts on destination choice, signaling that a destination accessed via roads with circuitry and more turns will lower travelers' odds to visit it. This finding corroborates that road network topologies not only can impact travelers' perceived travel time (Emond and Handy, 2012; Parthasarathi and Levinson, 2011a), but also can influence travelers' perception on the ease of reaching destinations.

## 2.4.7 Discussion

The study analyzes GPS travel data in the Twin Cities to examine the impact of destination accessibility and road network structure on travelers' destination choice for home-based non-work trips. We hypothesize that (1) destinations with greater accessibility and diversity of services, all

Table 2.4: Modeling individuals' destination choice for home-based non-work shopping trips after the opening of the new I-35 W Bridge (number of subjects: 141; number of trips: 2643)

Model type		Mixed-effects logit model	
Number of subjects		141	
Number of trips		2643	
Dependent variable		Trip Destination choice	
Independent variables		Model 1	Model 2
Built environment	Accessibility	0.746 **	
	Male * accessibility	-0.026 ***	
	Diversity of services		0.405 ***
	Male * diversity		-0.024 **
Travel/network features	travel time (ln)	-0.471 ***	-0.466 ***
	discontinuity (ln)	0.018	0.020 *
	circuitry (ln)	-0.090 ***	-0.090 ***
	roadness (ln)	-0.125 ***	-0.106 ***
	$\delta$	0.323 ***	0.336 ***
	AIC	15044	14789

\*\*\* statistically significant at the 1% level.

\*\* statistically significant at the 5% level.

\* statistically significant at the 10% level.

else equal, are more attractive to drivers, and that (2) destinations reached via a more circuitous or discontinuous route, all else equal, are less attractive to drivers. The results have corroborated our hypotheses. Cumulative accessibility and diversity of services around destinations, all else equal, promote a destination's appeal. In addition, a destination reached via a route of greater circuitry, discontinuity, and roadness lessens its attractiveness to drivers. The results indicate that both land use patterns and road network structure can affect individuals' shopping destination choice. This research also posits that ITS technologies such as GPS can provide unique insights into travel behavior at the microscopic level.



# Chapter 3

## CLUSTER Simulation Module for Retail Location Choice

### 3.1 Literature on Modeling Retail Location Choice

Cumulatively, travel decisions can bear upon the built environment. The connection between consumers' location choice and retail location choice has been established in theory. Traditional economic geography theories explain stores' spatial distribution patterns mainly through differences in geographical, demand, supply, and cost conditions (Christaller, 1933; Lösch, 1940; Marshall, 1890; Weber, 1909).

Central place theory (CPT) argues for a casual relationship between population and central functions serving the population (Christaller, 1933; Forbes, 1972). In CPT, goods are categorized into higher-order and lower-order, depending on how frequently they are purchased. Examples of lower-order goods include daily groceries and food, which are replenished quickly; such retail functions exist widely in communities, towns, and cities because there is enough population base to support them. Examples of higher-order goods include automobiles, jewelry, and household appliances; these goods are replaced less often and therefore are farther away from small towns in order to serve larger populations in the region. This theory further claims that functions in a hexagonal market size work most efficiently.

Empirically, it has been demonstrated that consumers' multi-purpose shopping behavior and retailers' profit-maximizing locational choice can lead retailers selling different products to cluster together (Eaton and Lipsey, 1982; Huang and Levinson, 2011). Yet we rarely find such strict orderly-arranged urban hierarchy as posited by CPT for several reasons. Individuals seldom purchase goods from the nearest available location; multipurpose shopping decreases average transportation costs per product. Furthermore, the variety of retail in one location, enabling comparison shopping, is an important factor attracting consumers. CPT requires that transportation cost is equal in all directions and proportional to distance. Geographical constraints and other factors shaping network structure deny this symmetry, while transportation cost varies by mode. The assumption of an evenly distributed population is equally improbable. In addition, CPT serves only to explain the static spatial patterns and cannot reveal stores' dynamic location choice process.

Over the past decade, agent-based modeling has gained popularity in revealing individual agents' dynamic behavior in location choice. In modeling land use dynamics, this approach has been used to model evolution of land-cover systems human settlement patterns (Brown et al., 2005;

Evans and Kelley, 2004; Sanders et al., 1997; Webster, 2003; Wu and Webster, 1998, 2000). Given three sets of agents (suppliers, retailers, and consumers), Huang and Levinson (2009, 2011) model the location choice of retailers on a circle of locations by giving each retailer a set of rules for location choice. Such tools have the potential of being extended for planning education to educate transportation engineering and planners on the micro-foundations of individual stores' location choice.

Our review suggests that the relationship between consumers' destination choice and retailers' destination choice has not been sufficiently modeled at the microscopic level and visualized for policy analysis and planning education.

## **3.2 CLUSTER: An Agent Location Choice Model**

As a web-based simulation module, the CLUSTER program (Clustered Locations of Urban Services, Transport, and Economic Resources) is an agent-based prototype that models retail location choice on supply chains consisting of consumers, retailers, and suppliers. Based on Huang and Levinson (2011), this module aims to investigate the effects of transportation cost, consumers' travel behavior, and suppliers' locations on retail location choice. It can be used as pedagogical tool to help students visualize the relationship between business spatial distributions and the firm-related or consumer-related factors at the micro level. The goals of this module are to:

- Provide an interactive webpage where different urban policy parameters can be tested and results can be visualized.
- Familiarize students and planners with basic urban patterns through visualization.
- Help transportation students and planners to understand the micro-foundations of urban structures in terms firm spatial distribution vis-à-vis supplier and consumer locations.

As part of the STREET tools (Simulating Transportation for Realistic Engineering Education and Training), the CLUSTER module is freely available for non-commercial use at its home page: <http://www.street.umn.edu/CLUSTER.html>.

### **3.2.1 The Model**

The structure and algorithm of the model are delineated in Huang and Levinson (2011).

#### **Assumptions and Definition of a Retail Cluster**

In the simplified three-layer supply chains, products flow from suppliers, via retailers, to consumers; money flows in the opposite direction. All agents are presumed to own perfect information; they locate at a circular area of discrete locations. The idea of a circle has the following advantages: (1) one-dimension; (2) providing an enclosed area (which is similar as a *de facto* geographical region and limits location choices for retailers). A location on a circle is intended to provide conceptual understanding. It can be extrapolated to a regional level or local level to represent its relative distance to the market or to suppliers.

Two kinds of markets are tested based on this framework: first, a market of homogenous goods; second, a market of two complementary goods with consumers' trip chaining behavior in shopping. The computational models are programmed in java, where each agent is modeled as an object. In the beginning of each round, consumers patronize retailers based on their rules to meet their needs on the product; after consumers finish shopping, retailers calculate their profits (revenue - cost) and assess the profitability of other locales. At the end of each round, assuming other retailer locations are fixed, each retailer moves to the locale that can provide the highest profit. The locales and profits of retailers are updated for each round.

A cluster is defined as an agglomeration of retailers which are geographically adjacent or co-located. The density of a cluster is calculated as the number of retailers in a cluster divided by the number of locations in the cluster. The average cluster density of  $n$  retailers,  $\varphi_n$ , is formulated as:

$$\varphi_n = \frac{1}{M} \sum_{i=1}^M \frac{\varepsilon_i}{\tau_i} \quad (3.1)$$

where  $\varepsilon_i$  is the number of numbers in cluster  $i$ ;  $\tau_i$  is the number of locales covered by cluster  $i$ ;  $M$  is total number of clusters.

## Consumers

In a market of homogenous goods (named  $x$ ) with  $W_x$  total number of retailers, a consumer selects a retailer to patronize based on its attractiveness, which depends on the observable shortest distance between the consumer and the retailer and other unobservable factors. For example, for consumer  $p$ , the attractiveness index  $A_{pi}$  of Retailer  $R_{xi}$  (the  $i$ th number of retailers of product  $x$ ) is represented as:

$$A_{pi} = k_1 \cdot d_{pi}^{-\beta} + \epsilon_p \quad (3.2)$$

Where  $d_{pi}$  is the shortest distance between consumer  $p$  and retailer  $i$ ;  $k_1$  and the scaling parameter  $\beta$  are positive constants. The function indicates that longer travel distance would generally diminish consumers' willingness to patronize. White noise  $\epsilon_p$  shows a certain degree of randomness.

In a market of two complementary goods sold by two kinds of retailers, let  $R_{xi}$  indicate retailer  $i$  of product  $x$ , and  $R_{yj}$  indicate retailer  $j$  of product  $y$ . A trip is defined as a round-trip for a consumer from home to visit  $R_{xi}$  and  $R_{yj}$ . Given  $W_x$  number of  $R_{xi}$  and  $W_y$  number of  $R_{yj}$ , there are in total  $W_x \cdot W_y$  trip candidates.

The utility for consumer  $p$  to patronize retailer  $R_{xi}$  and  $R_{yj}$  (indicated by Pair  $t$ ) equals:

$$A_{pt} = \sum_{t=1}^{W_x \cdot W_y} k_1 \cdot d_t^{-\beta} + \epsilon_p \quad (3.3)$$

After calculating all retailers' attractiveness indexes, a consumer probabilistically selects a retailer to patronize. The probability for consumer  $p$  to patronize retailer  $R_{xi}$ ,  $\rho_{pi}$ , is calculated based on a simplified version of Huff's model (Huff, 1964):

$$\rho_p = \frac{e^{A_{pi}}}{\sum_{i \in W_x} e^{A_{pi}}} \quad (3.4)$$

In the market of two goods, the probability for consumer  $p$  to visit  $R_{yj}$  can be similarly calculated.

The Roulette Wheel Selection method is adopted for a consumer to select a retailer in each round. This approach indicates that retailer  $i$  with higher  $\rho_{pi}$  for consumer  $p$  has a greater chance to be selected by this consumer. A consumer's probabilities of patronizing all retailers comprise a wheel of selection, which is updated for every round. A spin of the wheel selects a retailer; once a retailer is selected, a consumer buys all needed products from this retailer. The sequence for consumers to patronize retailers is randomly decided for each round.

## Retailers

Retailers connect suppliers and consumers on supply chains. In each round, a retailer evaluates expected profits of all locales and moves to the locale of the highest profit. For example, retailer  $R_{xi}$ 's expected profit in locale  $m$ ,  $\Pi_{xm}$ , is calculated as:

$$\Pi_{xm} = \left( \sum_{p=1}^N \lambda_x \cdot \rho_{pm} \right) \cdot [\theta_x - \sum_{k=1}^K (\delta_x + u \cdot \sigma_{mk}) d_{mk}] \quad (3.5)$$

Where  $\lambda_x$  indicates individual customer's demand on product  $x$  (with total  $N$  customers);  $\rho_{pm}$  stands for the probability for consumer  $p$  to patronize the retailer in locale  $m$ ;  $\theta_x$  means retail unit sales price of product  $x$  (a constant in the model);  $\delta_x$  means supplier unit sales price of  $x$  (a constant);  $u$  is the transport cost per unit distance per product;  $\sigma_{mk}$  indicates the shortest distance between supplier  $k$  of product  $x$  and locale  $m$ ;  $d_{mk}$  is a binary variable, which equals 1 if a retailer in locale  $m$  patronizes supplier  $k$ .  $\sum_{p=1}^N \lambda \cdot \rho_{pm}$  represents total expected sales of products in locale  $m$ . The part in brackets refers to expected profit per product, equaling sales price minus cost. A retailer's cost includes the purchasing cost of products from a supplier and the shipping cost which is proportional to shipping distance and quantity of products. Here we assume a retailer patronizes its closest supplier. After evaluating profits of all the  $C$  locales on the circle, retailer  $R_{xi}$  moves to the locale that provides the highest expected profit  $\Pi_{xi}$ , given others are geographically fixed at that time. Each retailer can only move once per round; the sequence of moving is randomly decided.

## Suppliers

We assume that all suppliers keep the same unit sales price. Moreover, they are evenly distributed on the circle and are fixed in all rounds. Further, in the market of two complementary goods, suppliers of the two products co-locate. It is presumed that suppliers can always produce enough goods to meet market demand.

### 3.2.2 The User Interface

The CLUSTER java applet has three panels: the variables panel, the control panel, and the visualization panel, which are shown in Figure 3.1. The following section introduces the function of each panel.

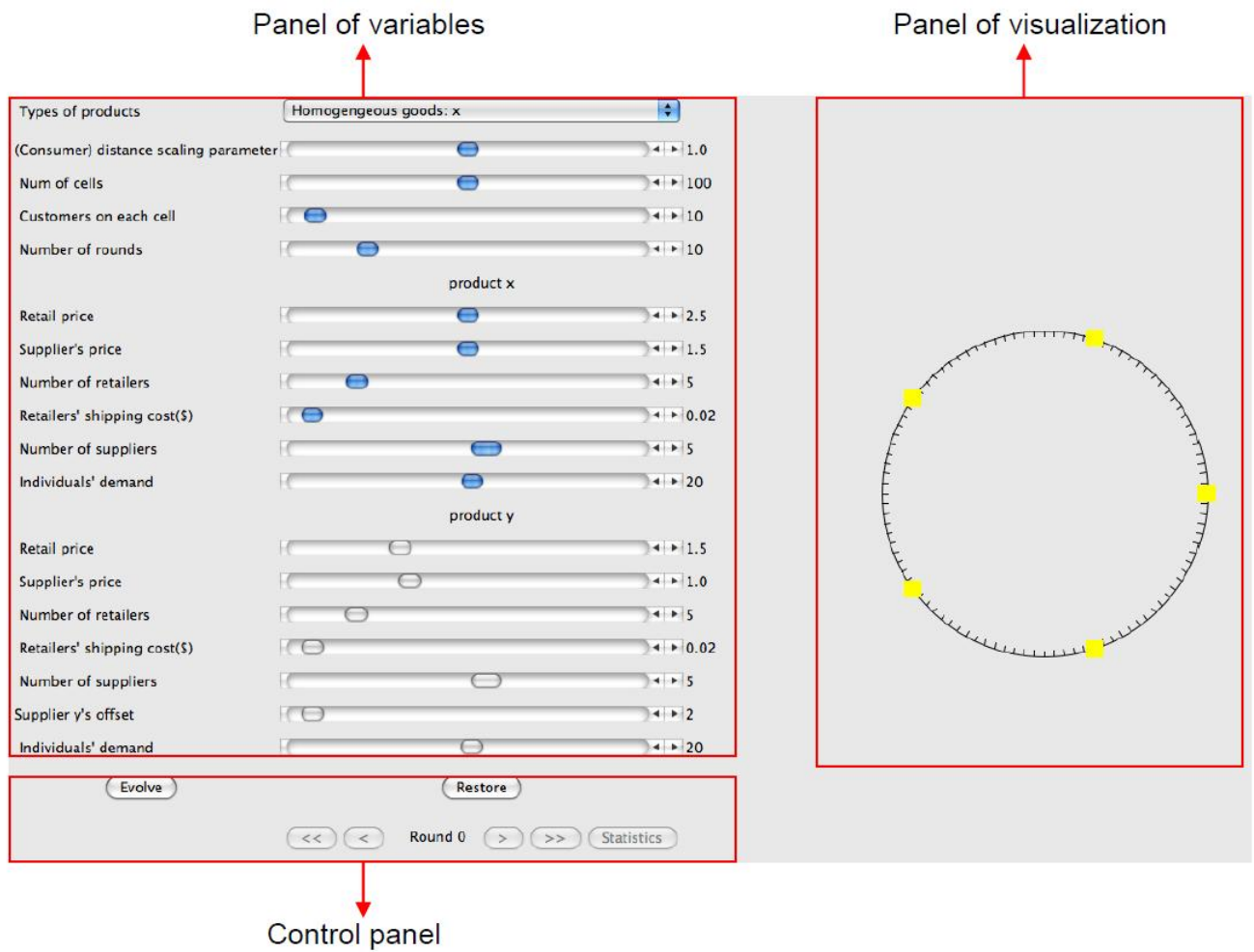


Figure 3.1: The user interface for the CLUSTER applet.

Table 3.1: Values of parameters (Model 1: homogenous goods; Model 2: heterogeneous goods)

Variables	Descriptions	Model 1	Model 2
$\beta$	exponent of distance decay	1.0	1.0
$k_1$	constant	1	1
C	# of locales on the circle	100	100
N	total number # of consumers	10000	10000
K	# of suppliers of $x$	5	5
L	# of suppliers of $y$		5
$u_x$	unit shipping cost per locale distance of product $x$ (\$ )	0.02	0.02
$u_y$	unit shipping cost per locale distance of product $y$ (\$ )		0.02
$\theta_x$	retail unit sales price of $x$ (\$ )	2.5	2.5
$\theta_y$	retail unit sales price of $y$ (\$ )		1.5
$\delta_x$	supplier unit sales price (\$ )	1.5	1.5
$\delta_y$	supplier unit sales price (\$ )		1.0
$\lambda_x$	individual consumer demand on $x$	20	20
$\lambda_y$	individual consumer demand on $y$		20

### 3.2.3 The Variables Panel

In the first row the user can choose the type(s) of products: “homogeneous goods:  $x$ ” or “heterogeneous goods:  $x$  and  $y$ ”. If the scenario of homogeneous goods is chosen, the variables for product  $y$  are disabled, and are enabled otherwise.

The consumer distance scaling parameter ( $\beta$ ) ranges from 0 to 2.0. The greater the parameter is, the faster the consumer’s utility decreases with distance. The variable of the number of cells shows the number of discrete cells on the circle. On each cell there are a number of customers who can be placed by changing the number of customers on each cell (ranging from 1 to 200). The marks of each cell on the visualization panel will be painted darker as the number of customers on each cell increases. The number of rounds (iterations), in the next row, ranges from 1 to 50, with the default value 10.

For retailers of product  $y$ , there are several variables: retail unit sale price, supplier’s unit sales price, number of retailers, retailers’ shipping cost (shipping products from suppliers to retailers), number of suppliers, and individuals’ demand on product  $x$ . The explanation of the parameters and the base case scenarios are shown in Table 3.1.

Suppliers’ locations are evenly distributed. Given different numbers of suppliers, suppliers’ space between each other on the circle varies. In the scenario of heterogeneous products  $x$  and  $y$ , users can also set suppliers of product  $y$ ’s spatial offset from suppliers of product  $x$ ; suppliers’ locations will immediately be plotted as the number of suppliers and the value of offset are given.

### 3.2.4 Control Panel

There are two buttons on the control panel: “Evolve” and “Restore”. The Evolve button, once clicked, runs the program. As the program is running, there will be a message box at the end of

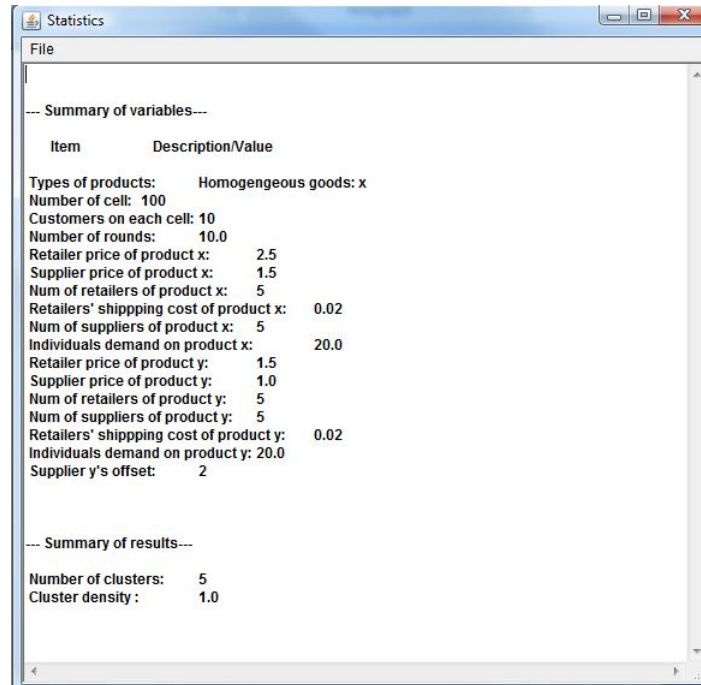


Figure 3.2: A window showing a summary of variables and results.

the panel showing the progress. Retailers' locations at the end of each round will be plotted on the panel of visualization. Once the program ends, users can check the locations of retailers in each round by clicking <, <<, >, >>. The restore button can restore the values of parameters to the base case. In addition, when the simulation ends, the user can click "statistics" button to see a summary of variables and basic analysis of the retail spatial distribution pattern. An example is shown in Figure 3.2.

### 3.2.5 Visualization Panel

The visualization panel visualizes the locations of retailers and suppliers. Suppliers are marked by rectangles; retailers are represented by circles. Objects of yellow color indicate objects of product  $x$ . Objects of red color indicate those of product  $y$ . Figure 3.3 and Figure 3.4 display the spatial distribution of retailers for Model 1 and Model 2. In model 1, we can see that every two retailers sit on a supplier's location. In model 2, the co-location of retailers of complementary goods has emerged.

## 3.3 Use the Module in Teaching

The CLUSTER module was applied in CE 5180 (Transportation Geography) at the Department of Civil Engineering in the 2011 Spring Semester. Most of the students were graduate students from the Department of Civil Engineering and Urban and Regional Planning Program at the University of Minnesota. A few others were undergraduate students in civil engineering and graduate students in geography and industrial engineering.

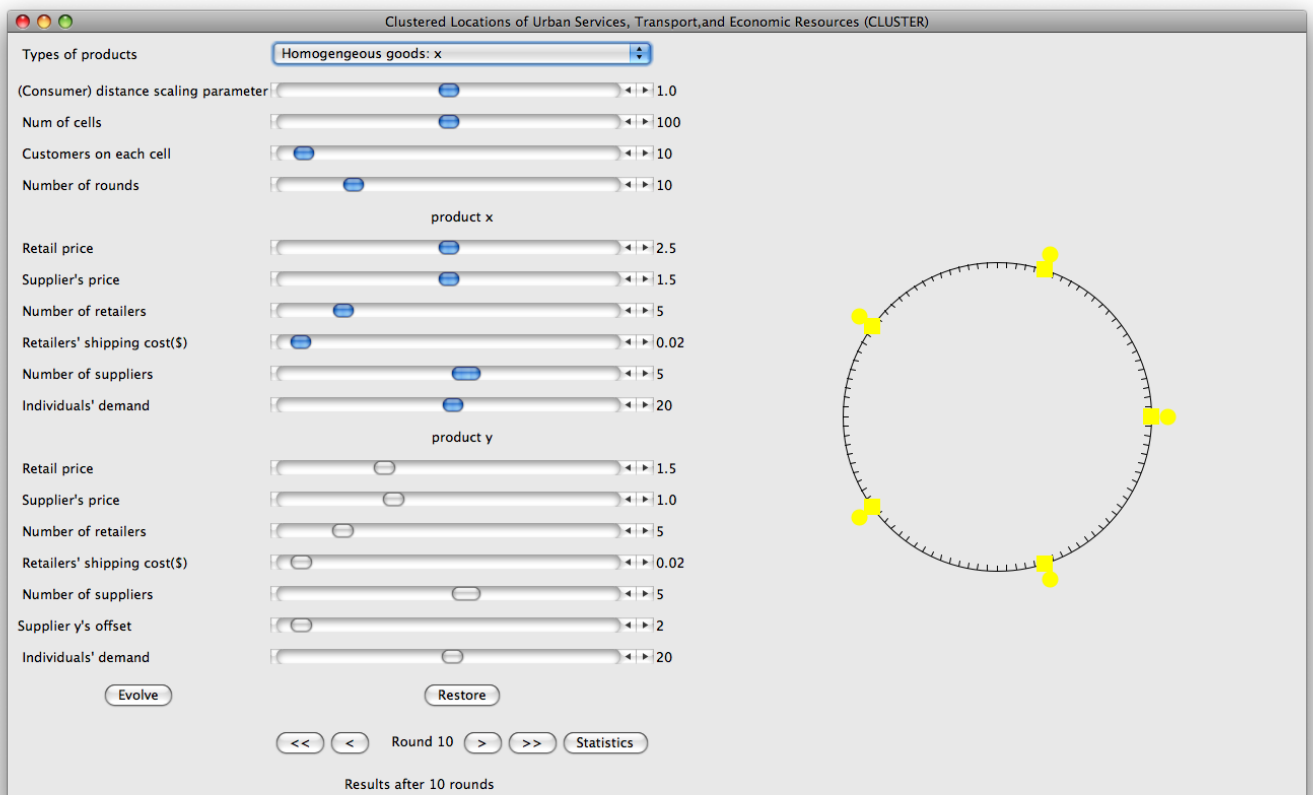


Figure 3.3: Visualization of the retail spatial pattern in equilibrium for Model 1.



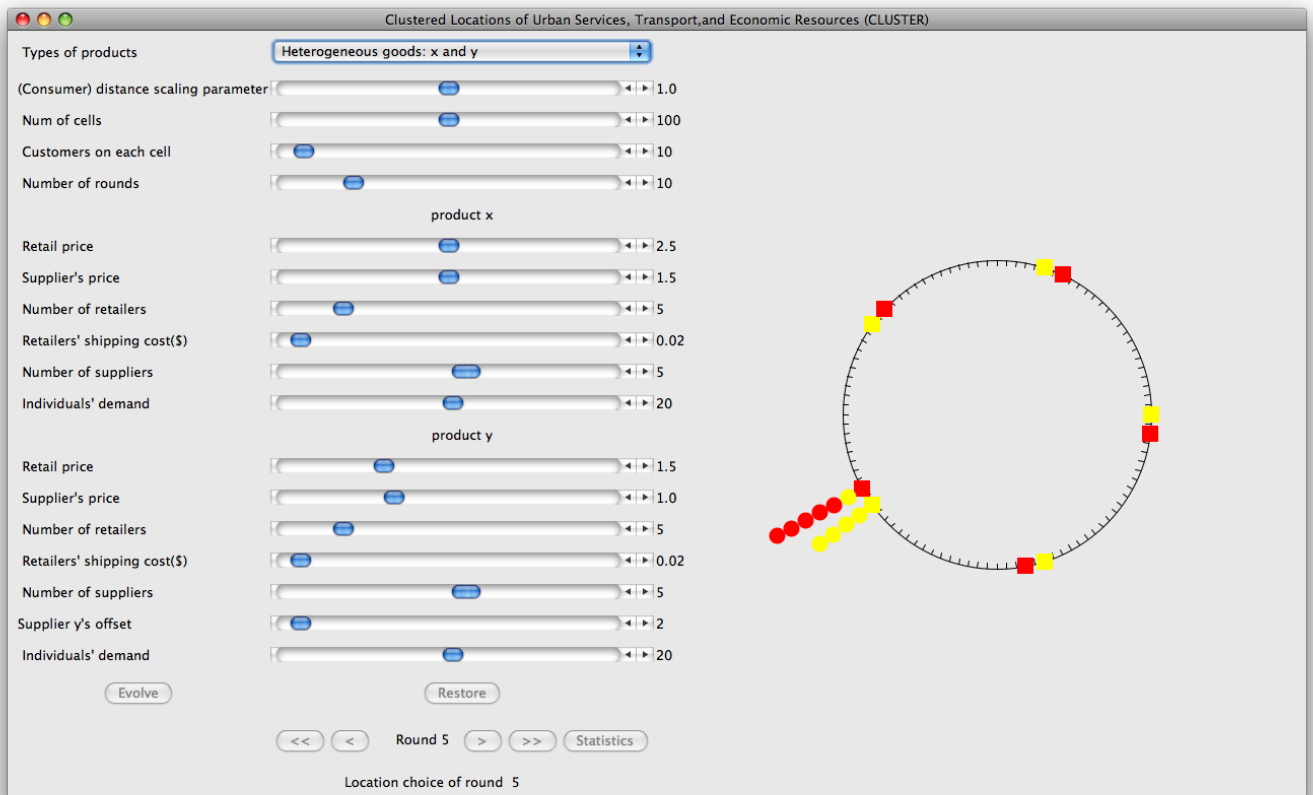


Figure 3.4: Visualization of the retail spatial pattern in equilibrium for Model 2.

In the assignment, students were asked to finish a few tasks by using the module and finish a report to summarize their findings. The assignment we created is shown in the Appendix.

### 3.3.1 Evaluation

Our assessment on students' reports shows that students have achieved the pedagogical goal of applying this module. The average score of the assignment equals 95/100. The students became not only familiar with the module, but also obtained a deeper understanding about the effects of different agents' relationships on the distribution of retailers. Some examples of students' summaries and discussion are shown as follows:

#### Example 1:

The analysis of the number of retailers relative to other variables proved educational. A relatively small number of retailers relative to suppliers results in the both entities located in close proximity. As the number of retailers grows, a few retailers slowly reposition themselves further away from the supplier to gain a geographic advantage that exceeds the marginal shipping costs. As the market becomes saturated with retailers, their location becomes more seemingly random as other variables gain greater influence. But, no matter what the number, retailers form clusters, some tight and some loose, and all centered around suppliers.

#### Example 2:

The CLUSTER simulation tool could successfully capture the relationships between different parameters of agent based cluster model such as number of suppliers/retailers/consumers, shipping cost, and scale parameter and simulate the distribution of the retailers. The purpose of this study is mainly focused on the relationship between the shipping cost and the retailer distribution. In conclusion, when the shipping cost per product is higher, the average cluster density will be higher too which mean retailers located more close to the supplier to reduce shipping cost.

High profit margins, low shipment costs, high  $\beta$  coefficients, and low demand encourage retailers to de-cluster in order to be more convenient and attract customers. Retailers are very sensitive to the cost of shipment.

#### Example 3:

The cluster model mapped how clustering occurs for retailers, suppliers and consumers. In a scenario with few retailers, retailers were observed to cluster around the existing suppliers. In the cluster model it was assumed that supplier locales were fixed. For computation purposes this is more feasible as compared to real world situations where due to market pressure some suppliers have been forced to relocate. The cause of relocation can be attributed to suppliers facing stiff competition and end up being forced to combine operation with its neighboring location/branch. Such relocations have been attributed to suppliers effort to cut spending on fixed costs of operation. Shipping cost and consumers willingness to travel emerged as one of the main driving forces behind clustering of retailers around. Also the gap between the number of

complementary goods is also attributed to cause clustering; the higher the gap more clustering was observed.

Some students further provided suggestions on the improvement of the program. For example, one student pointed out that “some clustering effects likely rely on the all-or-nothing strategy that agents use to select the very maximum expected profit no matter what the outcome, rather than weighing other factors into a decision between what may be virtually indistinguishable locales with respect to expected profit.” In addition, students indicated that there are other incentives that attract customers to go to certain retail clusters, such as brand loyalty and environmental externalities. Students’ feedback posits that they have gained deeper understanding of the economic principles of retail location choice in the active and fun learning process.

Overall, using the CLUSTER module in CE 5180 stimulated students’ thinking on the spatial and economic relationships among consumers, retailers, and suppliers. This simulation tool has proven to be a useful active learning tool to help students to understand the micro-foundations of retail location choice in the context of supply chains.

# Chapter 4

## Discussion and Conclusions

This project takes empirical and theoretical approaches to examine the bi-directional relationship between the built environment and individuals' shopping destination choice. Empirically, based on GPS travel data in the Twin Cities, this study employs GIS analysis to examine the impact of destination accessibility and road network structure on travelers' destination choice for non-work non-home trips. We hypothesize that a destination's accessibility, diversity of services, and road network topology influence travelers' destination choice. The results have confirmed our hypotheses. Cumulative accessibility and diversity of services in the adjacent walking zones, all else equal, promote a destination's appeal. In addition, a destination reached via a route of greater circuitry and more turns lessens its attractiveness to visitors. The results indicate that both land use patterns and road network structure affect individuals' shopping destination choice. This research also posits that ITS technologies such as GPS can provide unique insights into microscopic travel behavior.

We further build an agent-based CLUSTER simulation tool to visualize how the demand (indicated by distance from consumers) and cost (indicated by distance from suppliers) affect retail stores' location choice. The simulation module is used for teaching in transportation and land use planning as an interactive learning tool, and has the potential of being extended for policy analysis and scenario testing. The application of this tool in CE 5180 Transportation Geography and Network Science has proved to be successful in engaging students and encouraging them to think in-depth about the micro-foundations of retail location choice through active learning.

The future work of this research will take three directions. The first direction is to use person-based GPS devices to collect travel data combined with surveys. Our empirical study is based on in-vehicle GPS devices to collect travel data, and the in-vehicle GPS devices have the following limitations: (1) non-motorized trips are not captured; (2) which establishment a correspondent visited is unknown; (3) trip purpose is not known for all GPS trips. Several strategies can be adopted to overcome such limitations: (1) use person-based GPS devices or smart phone with GPS functions to track individuals' all trips and more precise locations of trip origins and destinations (for example, we may know what stores are visited in one trip.); (2) require individuals to fill surveys for more trip information (such as trip purpose, travel mode, perception of travel/destination) to complement the GPS data. Second, we would examine the effects of trip tours on individuals' destination choice. It is important to further understand how a trip tour with multiple purposes affects travelers' destination choice. This is thus a more complex decision-making process. Third, when applying the CLUSTER simulation tool in the future courses, we would conduct pre- and

post-surveys about students' understanding of the subject to investigate the role this tool plays in enhancing students' learning outcomes. It is also of interest to incorporate scenario-based environment (such as the actual transportation network and land use data) into the CLUSTER module for building a policy analysis tool.

# References

- Bachu, P., Dudala, T., and Kothuri, S. (2001). Prompted recall in global positioning system survey: Proof-of-concept study. *Transportation Research Record*, 1768:106–113.
- Ben-Akiva, M. and Lerman, S. R. (1985). *Discrete Choice Analysis*. Cambridge, MA: MIT Press.
- Bernardin, V., Koppelman, F., and Boyce, D. (2009). Enhanced destination choice models incorporating agglomeration related to trip chaining while controlling for spatial competition. *Transportation Research Record*, 2132:143–151.
- Bhat, C. (1998). Analysis of travel mode and departure time choice for urban shopping trips. *Transportation Research Part B: Methodological*, 32(6):361–371.
- Bhat, C. and Guo, J. (2004). A mixed spatially correlated logit model: Formulation and application to residential choice modeling. *Transportation Research Part B: Methodological*, 38(2):147–168.
- Boarnet, M. and Crane, R. (2001). The influence of land use on travel behavior: specification and estimation strategies. *Transportation Research Part A: Policy and Practice*, 35(9):823–845.
- Boarnet, M. and Sarmiento, S. (1998). Can land-use policy really affect travel behaviour? a study of the link between non-work travel and land-use characteristics. *Urban Studies*, 35(7):1155.
- Boelter, L. M. K. and Branch, M. C. (1960). Urban planning, transportation, and systems analysis. *Proceedings of the National Academy of Sciences*, 46(6):824.
- Brown, D., Page, S., Riolo, R., Zellner, M., and Rand, W. (2005). Path dependence and the validation of agent-based spatial models of land use. *International Journal of Geographical Information Science*, 19(2):153–174.
- Cao, X., Handy, S. L., and Mokhtarian, P. (2006). The influences of the built environment and residential self-selection on pedestrian behavior: Evidence from Austin, TX. *Transportation*, 33(1):1–20.
- Casas, J. and Arce, C. H. (1999). Trip reporting in household travel diaries: A comparison to gps-collected data. Presented at the 78th Transportation Research Board Annual Meetings, Washington, D.C.
- Cervero, R. (1989). *America's Suburban Centers: The Land Use-Transportation Link*. Winchester, MA: Unwin Hyman.

- Cervero, R. (1996). Mixed land-uses and commuting: evidence from the American Housing Survey. *Transportation Research Part A: Policy and Practice*, 30(5):361–377.
- Cervero, R. and Kockelman, K. (1997). Travel demand and the 3Ds: Density, diversity, and design. *Transportation Research Part D: Transport and Environment*, 2(3):199–219.
- Christaller, W. (1933). *Die zentralen Orte in Süddeutschland*. Gustav Fischer Verlag, Jena. English translation: *The Central Places of Southern Germany*. Englewood Cliffs, NJ: Prentice-Hall.
- Crane, R. and Crepeau, R. (1998). Does neighborhood design influence travel?: A behavioral analysis of travel diary and gis data. *Transportation Research Part D: Transport and Environment*, 3(4):225–238.
- de la Barra, T. (1989). *Integrated land use and transport modelling. Decision chains and hierarchies*. Cambridge, UK: Cambridge University Press.
- Derrible, S. and Kennedy, C. (2009). Network analysis of world subway systems using updated graph theory. *Transportation Research Record*, 2112:17–25.
- Derrible, S. and Kennedy, C. (2010). Characterizing metro networks: State, form, and structure. *Transportation*, 37(2):275–297.
- Draijer, G., Kalfs, N., and Perdok, J. (2000). Global positioning system as data collection method for travel research. *Transportation Research Record*, 1719:147–153.
- Eaton, B. C. and Lipsey, R. G. (1982). An economic theory of central places. *Economic Journal*, 92(365):56–72.
- Emond, C. R. and Handy, S. L. (2012). Factors associated with bicycling to high school: Insights from Davis, CA. *Journal of Transport Geography*, 20:71–79.
- Evans, T. and Kelley, H. (2004). Multi-scale analysis of a household level agent-based model of landcover change. *Journal of Environmental Management*, 72(1/2):57–72.
- Ewing, R. and Cervero, R. (2010). Travel and the built environment. *Journal of the American Planning Association*, 76(3):265–294.
- Ewing, R., Haliyur, P., and Page, G. (1995). Getting around a traditional city, a suburban planned unit development, and everything in between. *Transportation Research Record*, 1466:53–62.
- Feinberg, R. A. and Meoli, J. (1991). A brief history of the mall. *Advances in Consumer Research*, 18(1):426–427.
- Forbes, J. D. (1972). Central Place Theory: An analytical framework for retail structure. *Land Economics*, 48(1):15–22.
- Fotheringham, A. (1988). Consumer store choice and choice set definition. *Marketing Science*, 7(3):299–310.

- Gastner, M. and Newman, M. (2006). The spatial structure of networks. *The European Physical Journal B*, 49(2):247–252.
- Handy, S. (1993). Regional versus local accessibility: implications for nonwork travel. *Transportation Research Record*, 1400:58–66.
- Handy, S. (1996). Understanding the link between urban form and nonwork travel behavior. *Journal of Planning Education and Research*, 15(3):183–198.
- Handy, S., Cao, X., and Mokhtarian, P. (2006). Self-selection in the relationship between the built environment and walking: evidence from northern california. *Journal of the American Planning Association*, 72(1):55–74.
- Handy, S. and Clifton, K. (2001). Evaluating neighborhood accessibility: Possibilities and practicalities. *Journal of Transportation and Statistics*, 4(2):67–78.
- Hansen, W. (1959). How accessibility shapes land use. *Journal of the American Planning Association*, 25(2):73–76.
- Hess, P. (1997). Measures of Connectivity - Comparison of pedestrian environments in Seattle and Bellevue Suburbs, Washington. *Places*, 11:58–65.
- Horning, J., El-Geneidy, A., and Krizek, K. J. (2008). Access to destinations: Parcel level land use data acquisition and analysis for measuring non-auto accessibility. Access to Destinations Study Series Report: Mn/DOT 2008-19. St. Paul, MN: Minnesota Department of Transportation.
- Houston, D., Jaimes, G., Ong, P., and Winer, A. (2011). Daily activity patterns, travel behavior, and traffic exposure in communities near the Los Angeles-Long Beach Port Complex. *Journal of Transport Geography*, 19:1399–1409.
- Huang, A. and Levinson, D. (2009). Retail Location Choice with Complementary Goods: An Agent-Based Model. In Zou, J., editor, *Complex Sciences, Lecture Notes of the Institute for Computer Sciences, Social Informatics and Telecommunications Engineering*.
- Huang, A. and Levinson, D. (2011). Why retailers cluster: an agent model of location choice on supply chains. *Environment and Planning B*, 38(1):82–94.
- Huff, D. (1964). Defining and estimating a trading area. *The Journal of Marketing*, 39:41–57.
- Iacono, M., Levinson, D., and El-Geneidy, A. (2008). Models of transportation and land use change: a guide to the territory. *Journal of Planning Literature*, 22(4):323–340.
- Jiang, B., Yin, J., and Zhao, S. (2009). Characterizing the human mobility pattern in a large street network. *Physical Review E*, 80(2):021136.
- Kawasaki, T. and Axhausen, K. (2009). Choice set generation from GPS data set for grocery shopping location choice modelling in canton Zu-rich: Comparison with the Swiss Microcensus 2005. *Arbeitsberichte Verkehrs - und Raumplanung*, 595.



- Krizek, K., Iacono, M., El-Geneidy, A., Liao, C., and Johns, R. (2009). Access to destinations: Application of accessibility measures for non-auto travel modes. Access to Destinations Study Series Report: Mn/DOT 2009-24. St. Paul, MN: Minnesota Department of Transportation.
- Leszczyc, P., Sinha, A., and Timmermans, H. (2000). Consumer store choice dynamics: an analysis of the competitive market structure for grocery stores. *Journal of Retailing*, 76(3):323–345.
- Levinson, D. and El-Geneidy, A. (2009). The minimum circuitry frontier and the journey to work. *Regional Science and Urban Economics*, 39(6):732–738.
- Levinson, D. and Krizek, K. (2008). *Planning for Place and Plexus: Land Use and Transport*. New York: Routledge.
- Li, H. (2004). *Investigating Morning Commute Route Choice Behavior Using Global Positioning Systems and Multi-day Travel Data*. PhD thesis, Georgia Institute of Technology, Atlanta, GA.
- Li, H., Guensler, R., and Ogle, J. (2005). Analysis of morning commute route choice patterns using global positioning system-based vehicle activity data. *Transportation Research Record*, 1926:162–170.
- Li, H., Guensler, R., Ogle, J., and Wang, J. (2004). Using global positioning system data to understand day-to-day dynamics of morning commute behavior. *Transportation Research Record: Journal of the Transportation Research Board*, 1895:78–84.
- Lösch, A. (1940). *Die räumliche Ordnung der Wirtschaft*. Fischer, Jena. English translation: *The Economics of Location* (1954). New Haven, CT: Yale University Press.
- Marshall, A. (1890). *Principles of Economics*. London, UK: Macmillan & Co. (8th edition published in 1920).
- McFadden, D. (1978). Modeling the choice of residential choice. In Karlqvist, A., Lundqvist, L., Snickars, F., and Weibull, J., editors, *Spatial Interaction Theory and Planning Models*, pages 75–96.
- Newman, J. and Bernardin, V. (2010). Hierarchical ordering of nests in a joint mode and destination choice model. *Transportation*, 37(4):677–688.
- Parthasarathi, P., Hochmair, H., and Levinson, D. (2011). Network structure and spatial separation. *Environment and Planning B*, 39(1):137–154.
- Parthasarathi, P. and Levinson, D. (2011a). Metropolitan mobility. Presented at the 90th Annual Transportation Research Board Conference, #11-3511, Washington, D.C.
- Parthasarathi, P. and Levinson, D. (2011b). Network structure and activity spaces. Presented at the 90th Annual Transportation Research Board Conference, #11-3905, Washington, D.C.
- Pellegrini, P., Fotheringham, A., and Lin, G. (1997). An empirical evaluation of parameter sensitivity to choice set definition in shopping destination choice models. *Papers in Regional Science*, 76(2):257–284.

- Pozsgay, M. and Bhat, C. R. (2001). Destination choice modeling for home-based recreational trips: Analysis and implications for land use, transportation, and air quality planning. *Transportation Research Record*, 1777:47–54.
- Putman, S. (1983). *Integrated Urban Models: Policy Analysis of Transportation and Land Use*. London: Pion.
- Quddus, M., Noland, R., and Ochieng, W. (2005). Validation of map matching algorithms using high precision positioning with gps. *Journal of Navigation*, 58(2):257–271.
- Quddus, M., Ochieng, W., and Noland, R. (2007). Current map-matching algorithms for transport applications: State-of-the art and future research directions. *Transportation Research Part C: Emerging Technologies*, 15(5):312–328.
- Sanders, L., Pumain, D., Mathian, H., Guerin-Pace, F., and Bura, S. (1997). SIMPOP: a multiagent system for the study of urbanism. *Environment and Planning B*, 24:287–306.
- Santos, A., McGuckin, N., Nakamoto, H., Gray, D., and Liss, S. (2011). Summary of travel trends of 2009 national household travel survey. U.S. Department of Transportation/Federal Highway Administration, Report No. FHWA-PL-II-022. Washington, D.C.:
- Schimek, P. (1996). Household motor vehicle ownership and use: How much does residential density matter? *Transportation Research Record*, 1552:120–125.
- Schönfelder, S., Li, H., Guensler, R., Ogle, J., and Axhausen, K. (2005). Analysis of commute atlanta instrumented vehicle gps data: Destination choice behavior and activity spaces. *Arbeitsberichte Verkehrs-und Raumplanung*, 303.
- Schwanen, T., Dijst, M., and Dieleman, F. (2004). Policies for urban form and their impact on travel: the Netherlands experience. *Urban Studies*, 41(3):579.
- Scott, D. and Horner, M. (2008). Examining the role of urban form in shaping peoples accessibility to opportunities: An exploratory spatial data analysis. *Journal of Transport and Land Use*, 1(2):90–119.
- Taaffe, E., Gauthier, H., and Morton, E. (1996). *Geography of Transportation*. Upper Saddle River, NJ: Prentice-Hall.
- Timmermans, H. (1996). A stated choice model of sequential mode and destination choice behaviour for shopping trips. *Environment and Planning A*, 28:173–184.
- Wagner, D. P. (1997). Lexington area travel data collection test: Gps for personal travel surveys. Final Report. Federal Highway Administration, Columbus, OH.
- Wang, L. and Lo, L. (2007). Immigrant grocery-shopping behavior: Ethnic identity versus accessibility. *Environment and Planning A*, 39(3):684.
- Weber, A. (1909). *über den Standort der Industrien*. Mohr, Tübingen. English translation: Theory of the Location of Industries (1929). Chicago, IL: University of Chicago Press.

- Webster, C. (2003). The nature of the neighborhood. *Urban Studies*, 40(13):2591–2612.
- Wegener, M. (1982). Modeling urban decline: A multilevel economic-demographic model for the dortmund region. *International Regional Science Review*, 7(2):217–241.
- Wegener, M. (2004). Overview of land use transport models. In Hensher, D. A., Button, K. J., Haynes, K. E., and Stopher, P., editors, *Handbook of Transport Geography and Spatial Systems*.
- Wolf, J., Guensler, R., and Bachman, W. (2001). Elimination of the travel diary: Experiment to derive trip purpose from global positioning system travel data. *Transportation Research Record*, 1768:125–134.
- Wolf, J., Oliveira, M., and Thompson, M. (2003). Impact of underreporting on mileage and travel time estimates: Results from global positioning system-enhanced household travel survey. *Transportation Research Record*, 1854:189–198.
- Wu, F. and Webster, C. (1998). Simulation of natural land use zoning under free-market and incremental development control regimes. *Computers, Environment and Urban Systems*, 22(3):241–256.
- Wu, F. and Webster, C. (2000). Simulating artificial cities in a GIS environment: urban growth under alternative regulation regimes. *International Journal of Geographical Information Science*, 14(7):625–648.
- Xie, F. and Levinson, D. (2007). Measuring the structure of road networks. *Geographical Analysis*, 39(3):336–356.
- Yue, Y., Wang, H., Hu, B., Li, Q., Li, Y., and Yeh, A. (2011). Exploratory calibration of a spatial interaction model using taxi gps trajectories. *Computers, Environment and Urban Systems*.
- Zhang, L. and Levinson, D. (2008). Determinants of route choice and the value of traveler information: A field experiment. *Transportation Research Record: Journal of the Transportation Research Board*, 2086:81–92.
- Zhu, S. (2010). *The Roads Taken: Theory and Evidence on Route Choice in the Wake of the I-35 W Mississippi River Bridge*. Ph.D. Dissertation, Department of Civil Engineering, University of Minnesota, Minneapolis, MN.
- Zhu, S., Levinson, D., and Liu, H. (2010). Measuring winners and losers from the new I-35 W Mississippi river bridge. Presented at the 89th Transportation Research Board Annual Meetings, #10-2298, Washington, D.C..

# **Appendix**

Assignment for using the CLUSTER Module in teaching CE 5180 at the University of Minnesota

Assignment 1: The agent-based CLUSTER (Clustered Location of Urban Services, Transport, and Economic Resources) model <sup>1</sup>.

The main objectives of this assignment are to help students:

- understand retail location choice in the context of supply chains
- understand the mechanism of the evolution of retail clusters
- understand the impact of economic and geographical factors on retail clusters

Empirical studies have found that hierarchical distributions of economic activities and resources exist in almost every city, region, and nation (such as the US carpet production industry concentrated in Dalton, Georgia and the Italian textile industry in Prato, Italy. The mechanism of the clustering of industries and service as well its impact on is not yet sufficiently examined. The economic division of the Metropolitan Council started a pilot project to understand how individual business owners choose locations and the policy factors that impact the distribution of retailers. You are hired as an economic analyst to study the effects of possible policy initiatives or alternative decision-making assumptions in retail distribution. Parts of your responsibility are to provide theoretical groundings on the mechanism of clustering and to explore the implications of the following changes in economic policy and individual retailers' decision-making assumptions.

The changes of the assumptions include: (1) products' shipping cost, (2) number of suppliers, (3) number of customers, (4) geographical factors (such as distance scaling parameter).

The theory underlying the CLUSTER model is described in the following paper: Huang A, Levinson D, 2011, "Why retailers cluster: an agent model of location choice on supply chains" *Environment and Planning B: Planning and Design* 38(1) 82-94.

This agent-based model is employed to study retail location choice in a market of homogeneous goods and a market of complementary goods. On a circle comprised of discrete locales, retailers play a non-cooperative game by choosing locales to maximize profits which are impacted by their distance to consumers and to suppliers. A brief in-class demonstration will be given to familiarize you with the underlying model of CLUSTER.

Your Tasks In completing this project, you must fulfill the following tasks:

Task 1: Understand the simulator Run simulations under default values as well as one alternate set of scenarios (by sliding the scroll bars of some variables):

- distance scaling parameter: 1.0
- number of suppliers: 5
- number of retailers: 10
- retail sales price of product x (\$): 2.5
- suppliers' price of product x (\$): 1.5
- individual consumer demand on product: 20

---

<sup>1</sup><http://www.street.umn.edu/CLUSTER.html>

Task 2: Run the simulation under different economic scenarios of interest.

You can adjust values of parameters to reflect different assumptions. Copy the graphic output for your report. (You can use copy screen function of you computer). Compare the retail distribution patterns with the result of the base case.

Task 3: Submit a memo reporting your findings.

An recommended outline for your report is as follows:

1. Problem statement

2. Methodology

- Simulation (briefly describe CLUSTER and report your results from Task 1)
- Analysis methodology (stating what and why you choose a particular method)

3) Evaluation and Analysis

4) Results and Findings

5) Discussion of limitations

6) Conclusion

The report must be no more than 2500 words. Electronic submission required in PDF.