What impacts travel mode choice of the South Side of Chicago?

UP431 Grad-term project
Qiushi Huang
May 12th

Introduction

Southside Chicago refers to the south of Chicago's loop, and it is the largest of three Sides of the city. This area's unemployment rate and income level is relatively low compared to other sides of the city. My internship in 2021 summer was to survey STEM employers in nine neighborhoods of the South Side. The survey received 270 responses. With the knowledge of travel model, I got interested into how the socio-economic factors would affect the travel choice in this area? To be more specific, I am interested in how do: (1) Education, (2) Employment Status, (3) Work hours per week, (4) number of jobs a person has, and (5) the industry a person works in affect their travel choice?

I choose the study area as the nine neighborhoods I surveyed, which are highlighted with blue in Figure 1. Figure 2 is the zoomed-in map showing the names of neighborhoods.

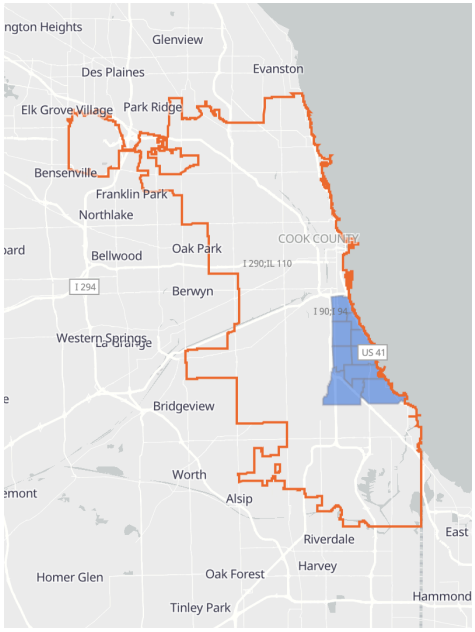

Figure 1 Study Area in Chicago

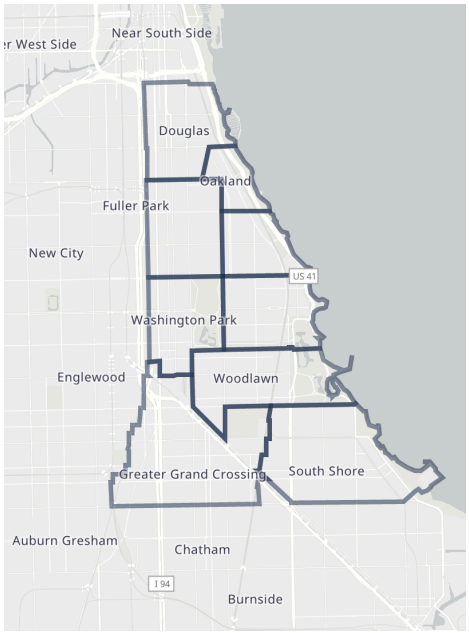

Figure 2 Nine Neighborhoods in the Study Area

Data source and methodology

The data sources I used in this study are:

1. My Daily Travel Survey 2018-2019 from CMAP Data Hub
2. Transit Availability Index from CMAP Data Hub
3. City of Chicago data portal hub
4. STEM Employer Survey data from my internship. It has 270 responses.

The main data I worked with is the My Daily Travel 2018-2019. I used the Census Tracts ID to first filter out the respondents from the study area, then extract the independent and dependent variables from the survey. The study area has 10338 observations. The independent variables are: education attainment, the employment status, the number of jobs a person has, and work hours per week. The dependent variable is the travel mode choice at person level. Then I put the data into RStudio to run the multinominal model and interpret the coefficients. The table below shows the meaning of each variable and data type.

Variable name	*Meaning*	*Type*	*Values*
Mode_imputed	The travel mode the respondent chose	Categorical	Auto, walk, train, bike, other (contains School bus, Taxi / limo (including Uber / Lyft), Rental car (Including Zipcar / Car2Go), Airplane, Boat / ferry / water taxi, and Something Else)
Educ	What is the highest grade or degree that you have earned?	Categorical	1,2,3,4,5,6 six levels from under grade 12 to graduate degree
Emply_ask	Employed or not	Dichotomous	1=yes, 2=no
Wrkhrs	How many hours do you work in a typical week at your primary workplace?	Numeric	Integers range from 0 to 80
Jobs	How many jobs do you work?	Numeric	Integers from 1 to 7
Indus	Industry that the respondent works for by the first 2 digits of the NAICS code	Categorical	"51", "54", "92", "61", "62","31-33", "56", "-1", "81", "48-49", "42","44-45", "72", "52", "71","53", “97"

Findings and analysis

1. Mode choice by education and employment status

In this model, I used the travel mode choice as dependent variable, while education attainment and employment status as independent variable. Auto is the reference level since it is the most common mode choice. I used the chi-square test to test the relationship between the two independent variables, which is shown in Figure 3. Generally, the higher level of education a person received, the lower their probability to be unemployed. Figure 4**Error! Reference source not found.** shows the summary statistics of the linear model between employment status and education level. The coefficient is negative and significant, the F-test Is significant too. All these shows that the relationship between the two variables is significant.

Table 1 shows the mode choice by education attainment (educ) and employment status (employ_ask). When not considering the two factors, the utility of walk is significantly higher than auto, while the utility of train and other are significantly lower. For the education attainment, the coefficient of train is 0.213, which means that one higher education level a person has, the 0.213 greater utility for train compared to auto, and this is significant at the 1% level. Both walk and other have positive coefficients, but the bike has a negative one. That

means the higher education a person has, the 0.055 less utility for bike compared to auto. Only the coefficient of train is at significant level.
For the employment status, the coefficients of walk, train, and bike are positive, but only the other is negative. That means when a person is unemployed, the utility of walk, train, and bike will be less, but the utility of other modes will be decreased, compared to an employed person. The employed people have more travel needs (for instance HBW trips), so they tend to combine various modes to fulfill their needs.
The positive utility of the level other might because of the school bus, whose passengers are unemployed school age students.
Overall, since the two independent variables are correlated, it should be more careful to interpret the coefficients of the multinominal model.

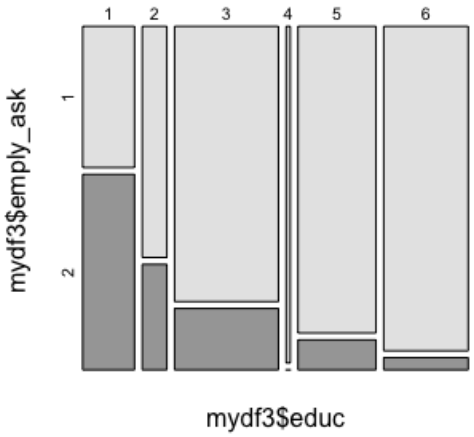

Figure 3 Chi-sq test for education and employment status

Employment Status by Education Level

	Dependent variable:
	emply_ask
educ	-0.093***
	(0.002)
Constant	1.547***
	(0.008)
Observations	10,338
R^2	0.172
Adjusted R^2	0.172
Residual Std. Error	0.359 (df = 10336)
F Statistic	2,145.178*** (df = 1; 10336)
Note:	*p<0.1; **p<0.05; ***p<0.01

Figure 4 Employment status by education level

Table1: Mode Choice by Education and Employment Status

	Dependent variable:			
	2	3	4	5
	Walk	Train	Bike	Other
	(1)	(2)	(3)	(4)
educ	0.006	0.213***	-0.055	0.147
	(0.013)	(0.025)	(0.053)	(0.145)
emply_ask	-0.573***	-0.760***	-3.132***	1.329**
	(0.061)	(0.137)	(0.719)	(0.531)
Constant	0.215**	-1.996***	-0.136	-8.041***
	(0.105)	(0.219)	(0.791)	(1.163)
Akaike Inf. Crit.	19,509.050	19,509.050	19,509.050	19,509.050
Note:				*p<0.1; **p<0.05; ***p<0.01

Table 1 Mode Choice by Education and Employment Status

2. Mode choice by work hours and number of jobs

I tested the relationship between these two variables, and the summary statistics are shown in **Error! Reference source not found.**. The coefficient of jobs is not significant, and the R2 is only 0.002 which means the model only explains 0.2% variations in the working hours. Therefore, these two variables do not have strong correlation.

Error! Reference source not found. is the output of the multinominal model that using work hours and number of jobs to predict the mode choice. When ignoring the two dependent variables, the utilities of all other four modes are significantly lower than auto.

For the working hours, the coefficients for train and other are positive, and walk and bike are negative. This means that one more working hour a person has in a typical week at their primary job, the 0.021 greater utility of train, and 0.061 utility of other modes. On the opposite, one more working hour a person has at a typical week at their primary job, 0.013 less utility for walk, and 0.017 less utility for bike. For the number of jobs, the coefficients of walk and bike are positive, while the utilities of train and other are negative. Each one more job a person has, the 0.153 greater utility of walk, and 0.173 greater utility of bike, compared to auto. On the opposite, each more job a person has, the 0.046 less utility for bike, and -1.70 less utility for other modes, compared to auto.

What's interesting is that the +/- sign of coefficients of the two variables are exactly opposite. Six out of eight coefficients are statistically significant at 1% level, one is significant at 5% level, only one coefficient is not significant. This might be the result of the weak relationship between the two variables as shown in Figure 5.

Working hours by number of jobs

	Dependent variable:
	wrkhrs
jobs	-1.142
	(0.858)
Constant	35.113***
	(1.094)
Observations	1,050
R^2	0.002
Adjusted R^2	0.001
Residual Std. Error	14.841 (df = 1048)
F Statistic	1.771 (df = 1; 1048)
Note:	*p<0.1; **p<0.05; ***p<0.01

Figure 5 Working hours by number of jobs

Table2-1: Mode Choice by Work Hours and Number of Jobs

	Dependent variable:			
	2	3	4	5
	Walk	Train	Bike	Other
	(1)	(2)	(3)	(4)
wrkhrs	-0.013***	0.021***	-0.017***	0.061**
	(0.002)	(0.003)	(0.006)	(0.024)
jobs	0.153***	-0.046	0.172**	-1.700***
	(0.019)	(0.034)	(0.079)	(0.546)
Constant	-0.281***	-2.531***	-3.527***	-6.454***
	(0.035)	(0.077)	(0.140)	(0.534)
Akaike Inf. Crit.	19,599.800	19,599.800	19,599.800	19,599.800
Note:				*p<0.1; **p<0.05; ***p<0.01

Table 2 Mode Choice by Work Hours and Number of Jobs

3. Mode choice by top 3 industries

The Table 3 shows the top 3 industries where most respondents work in study area in CAMP's my Daily Travel Survey 2017-2018, and the top 3 industries in my internship's STEM employer survey. The former one reflects the employees, while the last one reflects the employers.

Since the University of Chicago locates in Hyde Park, which is one of the nine neighborhoods in the study area, that might be the reason why Education services is the top 1 across all the industries. For the STEM employer survey, the top 1 is Ambulatory Health Care Services, which might be the impacts of the Medical School of University of Chicago and its affiliated hospitals or clinics. Also, since the professionals in health care industry tend to be more educated, they are more likely to fully understand then answer our survey, hence being overrepresented in the findings.

TOP 3 in CMAP survey		**Top 3 in STEM employer survey**	
First 2 digits of 2017 NAICS code	**Definition**	**First 2 digits of 2017 NACIS code**	**Definition**
61	Education services	62	Ambulatory Health Care Services
54	Professional, Scientific, and Technical Services	54	Professional, Scientific, and Technical Services
56	Administrative and Support and Waste Management and Remediation Services	31-33	Repair and Maintenance

Table 3 Top 3 industries in the My Travel Data Survey 2017-2018

Table 4 shows the mode choice by top 3 industries across all fields. The reference level is indus54: Professional, Scientific, and Technical Services. When ignoring all other variables, the utilities of train, bike, and other, are significantly lower than the utility of auto. The coefficients of other is even as high as -17. The only positive coefficient is for walk, which is 0.22, but not significant.

For the industry 56, which is Administrative and Support and Waste Management and Remediation Services, people who work in this industry have significant lower utility of walk, train, and other, compared to the people who work in industry 54. Their utility of bike is lower than auto too but not significant. For the people who work in industry 61, which is Education services, their utilities of walk, bike, and other modes are significantly higher than auto. However, the utility of train is significantly lower than auto. That might be the results of University of Chicago, whose faculty, staff, and students live near to the campus.

The model shows the disparity of the mode choice behaviors across the industries. People who work in industry 56: Administrative and Support and Waste Management and Remediation Services generally like to use auto more, but people who work in industry 61: Education services have more diverse mode choice. I also found that for all the three industries in this model, the utility of other modes is extremely different from auto, and the utilities of train are all significantly lower than auto.

Table3-1: Mode Choice by Top 3 Industries

	Dependent variable:			
	2 Walk	3 Train	4 Bike	5 Other
	(1)	(2)	(3)	(4)
indus56	-0.921***	-1.888***	-10.428	-9.097***
	(0.094)	(0.226)	(107.947)	(0.000)
indus61	0.161**	-0.503***	2.676***	10.820**
	(0.067)	(0.114)	(0.588)	(0.354)
Constant	0.022	-1.343***	-5.344***	-17.338**
	(0.056)	(0.088)	(0.579)	(0.354)
Akaike Inf. Crit.	10,661.500	10,661.500	10,661.500	10,661.50
Note:				*p<0.1; **p<0.05; ***p<0.0

Table 4 Mode Choice by Top 3 industries among the CMAP survey

Table3-2: Mode Choice by Top 3 Industries in STEM Survey

	Dependent variable:			
	2 Walk	3 Train	4 Bike	5 Other
	(1)	(2)	(3)	(4)
indus54	11.806	11.092	8.459***	-117.293
	(65.028)	(90.036)	(0.414)	
indus62	10.790	11.025	9.428***	-1.450
	(65.029)	(90.036)	(0.357)	(1.170)
Constant	-11.784	-12.434	-13.803***	-3.436***
	(65.028)	(90.036)	(0.244)	(1.017)
Akaike Inf. Crit.	4,132.569	4,132.569	4,132.569	4,132.569
Note:				*p<0.1; **p<0.05; ***p<0.01

Table 5 Mode Choice by Top 3 industries in the STEM survey

Table 5 shows the multinominal model between the mode choice and the top 3 industries in the STEM survey. The reference level is industry 31-33: Repair and Maintenance. When ignoring the other variables, the utility of all alternative modes is lower than auto, and the utilities of bike and other are significantly lower. For industry 54: Professional, Scientific, and Technical Services, the utility of bike is significantly higher than auto, so do the utilities of walk and train though they are not significant. The coefficient of other is extreme, -117, but not significant. For industry 62: Ambulatory Health Care Services, the utility of bike is significantly higher than auto. The utility of walk and train are also higher than auto, but not significant. The utility of other modes is lower than auto.

From this model, I found that the utility of bike in both industry 54 and industry 62 are all significantly higher than auto. Similarly, their coefficients of walk and train are positive and big, but their coefficients for other modes are negative. Generally, the mode choices of these two industries are more diverse, while the mode choice of the industry 31-33: Repair and Maintenance is solely auto.

Furthermore, I tested the mode choice by all industries included in the CMAP My Daily Travel Survey. The summary statistics and the industry definition are shown in the appendix. The reference level is industry 31-33: manufacturing. Similarly, the mode choice behaviors vary across industries.

Conclusions and policy implementations

I built three multinominal models to test the relationship between mode choice and the variables I am interested in. For the education attainment and employment status, the employed people with higher education tend to use more alternatives compared to unemployed people with lower education. For the working hours and number of jobs, their impacts on mode choice are disparate: people work more hours will walk and bike less compared to auto, but people with more jobs will use these two modes more. When I breakdown the mode choice by industry, I see the disparity across industries too. Health industry professionals and Education service professionals use alternative modes more, Repair

& Maintenance professionals, as well as the people work in Administrative and Support and Waste Management and Remediation Services, rely more on cars. Industries could reflect the practitioners' income and education level to some extent, which are two important socioeconomic factors when analyzing travel behaviors.

The variation in mode choice behaviors across industries provides us a new approach to make policies. The industry might not be a weak indicator for research but is a very straightforward approach to draft policies through industry organizations. It is crucial to understand that each industry has its unique characteristics, commuting patterns, and transportation needs. By digging further into the industry differences, policy makers can tailor their policies to address specific difficulties. For instance, policymakers can work with union to subsidize transit pass for the industries that workers rely more on auto to encourage more use on transit. Planners can plan for a Park-And-Ride transit center at where the workers commute.

Such an approach also needs the understanding in land use. The concentration of Health care and Education service industries in Hyde Park motivates practitioners to use alternative modes more. Being a huge employer in the South Side, University of Chicago make its students, faculties, and staff to live and work around, ultimately encourage more active transportation. Overall, by understanding the industrial and land use context, policy makers can develop efficient policies to encourage people make diverse transit mode options.

Appendix

Table4: Mode Choice by all industries

	Dependent variable:			
	2 Walk (1)	3 Train (2)	4 Bike (3)	5 Other (4)
indus42	51.840***	50.729***	-1.825	-78.278***
	(1.085)	(0.981)		(0.000)
indus44-45	47.953***	44.945***	31.506***	-85.630***
	(0.192)	(0.238)	(12.033)	(0.000)
indus48-49	50.132***	47.035***	6.419***	-63.806***
	(0.164)	(0.187)	(0.00000)	(0.000)
indus51	50.212***	47.992***	33.895***	-82.522***
	(0.286)	(0.273)	(12.038)	(0.000)
indus52	49.802***	46.348***	-16.289	-0.032
	(0.247)	(0.363)		(1.244)
indus53	50.049***	-21.040	-15.410	-68.426***
	(0.584)			(0.000)
indus54	50.764***	47.085***	30.979**	-55.302***
	(0.123)	(0.128)	(12.038)	(0.000)
indus56	49.843***	45.197***	-29.017***	-51.693***
	(0.132)	(0.217)	(0.000)	(0.000)
indus61	50.934***	46.591***	33.663***	-3.076**
	(0.117)	(0.120)	(12.027)	(1.238)
indus62	49.749***	47.018***	31.948***	-1.450
	(0.144)	(0.144)	(12.033)	(1.170)
indus71	50.653***	48.074***	33.166***	-99.810***
	(0.227)	(0.233)	(12.044)	(0.000)
indus72	50.311***	47.106***	34.425***	0.033
	(0.222)	(0.280)	(12.031)	(1.245)
indus81	49.979***	46.457***	21.218	-81.788***
	(0.153)	(0.195)	(108.231)	(0.000)
indus92	50.577***	46.950***	34.286***	-81.534***
	(0.182)	(0.246)	(12.030)	(0.000)
indus97	48.951***	47.734***	-13.537***	-67.352***
	(1.017)	(0.666)	(0.000)	(0.000)
Constant	-50.742***	-48.427***	-36.323***	-3.434***
	(0.112)	(0.099)	(12.026)	(1.016)
Akaike Inf. Crit.	15,326.090	15,326.090	15,326.090	15,326.090
Note:				*p<0.1; **p<0.05; ***p<0.01

Appendix 1 Mode choice by all industries included in CMAP survey

First 2 digits in NAICS code	Definition
31-33	Manufacturing
42	Wholesale Trade
44-45	Retail Trade
48-49	Transportation and Warehousing
51	Information
52	Finance and insurance
53	Real estate and rental and leasing
54	Professional, scientific, and technical services
56	Administrative and Support and Waste Management and Remediation Services
61	Educational Services
62	Health Care and Social Assistance
71	Arts, Entertainment, and Recreation
72	Accommodation and Food Services
81	Other Services (except Public Administration)
92	Public Administration
97	Something else

Appendix 2 All industries included in the CMAP survey