UNIVERSITY OF SKÖVDE

School of Technology and Society

BACHELOR DEGREE

Gender differences in commuting

Study with Swedish data

Bachelor Degree Project in Economics
Level 30p
Spring term 2008

Annika Isotamm 761113-8624

Supervisor: Michael Olsson
Examiner: Ph.D. Yinghong Chen

Abstract

Commuting is becoming more and more important in every day life in the world of complex labour markets. If for hundred years ago most workers lived less than one hour walking distance from their workplace, then today they commute daily outside the cities and villages they live. Often is it a trade off between unemployment and employment. The subject of this thesis is to investigate if there are differences in female and male commuting behaviour. Especially it studies to what extent their spatial interaction is affected by variables such as labour in origin municipality, jobs in destination municipality and the distance between these two. Gravity models are used, and six different versions are estimated. The commuting data used in thesis comes from SCB (Statistics Sweden) and data about travel time collected by the Swedish Road Administration. Commuting data includes 1 003 771 people in Sweden who commute to a job located in another municipality than he or she was settled. The results imply that men are less sensitive to distance between two locations. Moreover, the amount of labour in origin and the amount of jobs in destination have higher effect on male commuting.

***Keywords**: Commuting, gender differences, gravity model*

Contents

1 Introduction 4
1.1 Previous research 5
2 Model and method 7
3 Data 13
4 Results 15
4.1 Outcome of Model 1 15
4.2 Results of Models 2-5 17
4.3 Outcome of Model 6 18
4.4 Discussion of the results 18
5 Conclusions and suggestions for future studies 20
References 22
Appendix 1 Model summary 23
Appendix 2 ANOVA (Analysis of variances) 23
Appendix 3 Normal distribution 24
Appendix 3 Collinearity statistics 26

1 Introduction

A commuting is the travel from one's residence to the working place. It is a phenomenon of industrialized society. During the past decades there have been tendency of increased commuting. It can be described both by the number of persons who commute daily and the average commute.

According to SCB (Statistics Sweden, 2006) there was 1 003 771 people in Sweden who commute to a job located in another municipality than he or she was settled. More than half of those commuters were male. It isn't only that there are more men who commute, but they also commute a longer distance than women do. Previous research implies that women are more sensitive for commuting distance from different reasons. Commuting differences between men and women are largest in households with children.

According to economic theory all individuals are utility maximizers. It means that they want to maximize their utility in every decision. How is this related to the commuting? Men commute longer distances and because of this they have to give up more leisure time. Women choose more likely job close to home even if the wages are lower than in the case of choosing to commute longer and they loose thereby some of their earnings. From earlier studies in the field, it is clear that male and female commuting differs from each other. The gender differences in Swedish commuting data will be examined in this thesis. Often in studies of commuting behaviour the female commuting is viewed from perspective of responsibilities at home and the labour market imperfections. Also distance and time factor have always been interesting. Socioeconomic factors give an idea how commuting is determined by education, salaries and family situation etc. This paper will investigate the commuting differences with starting point of accessibility to jobs in one municipality, accessibility to labour in another municipality and the distance between these municipalities.

The layout for the rest of the thesis is as follows. Chapter 1.1 presents the previous research. In chapter 2, the method of research is presented and the choice of model for estimation will be introduced. In chapter 4, the data set and limitations are presented. The results are summarised in chapter 5. Finally, chapters 6 will include conclusions and suggestion for future research.

1.1 Previous research

Commuting distance has been a popular subject for studies. Tkocz and Kristensen (1994) explain the distance by urban spatial structure and socioeconomic variables, i.e. how among other variables such as commuter's age, education, salaries, and family relationships affect the commuting decisions.

For households with two incomes residential and work place location have to be balanced. Usually one of them or sometimes both have to commute. Angel and Hyman (1972) imply that the best location for a job is close to urban centre and the best location for a residence is some distance outside. For labour, the location of firms is given but the location of the residence can be affected by utility maximization choices. Certainly firm's location is influenced by profit maximizing opportunities and therefore can not be seen as given from firm's standpoint. Thus individual can't affect firm's location but place of resident and therefore is work journey in many cases unavoidable. People choose their residences to maximize their utility. How utility is maximized differs from person to person. We all live somewhere and choices for residential location are made for different reasons. Utility can be maximized by a lake close to home or by nice neighbourhood, alternatively nearness to the work. Firm's location is dependent on profit maximizing prospects.

Women are relatively more sensitive for commuting distance, and there are several reasons for this. Camstra (1996) suggest that in urban regions females commute relatively more often and in suburban regions female commuting is relatively less frequent. The reason may be that in thinly populated areas, longer distances to possible work places make women choose not to travel. In urban environments the distance to work is shorter and therefore have women incentive to choose to work. According to Camstra (1996) there are some main reasons why female choose employment close to the home. First, women have on average weaker position in the labour market that finds expression in labour market imperfections. These imperfections imply for example that females work less hours and thus the travel time relative to the worked hours is too high. Furthermore women earn on average less than men do, and it makes commuting more expensive for females. Moreover, women relatively more often have lower status jobs, which are more common and it is easier to find a job close to home. Second, women do have responsibilities for the most of the housekeeping and because of this they prefer shorter commute. There is trade off between the household work and time to work and commute.

Commuting differences between men and women are largest in households with children. Accordingly women have generally the main responsibility for housekeeping; and children are one more reason to commute shorter distances. Turner and Nieminer (1997) present the “household responsibility hypothesis”, according to which housekeeping and children are motivation why females want to work close to residence.

It is said in the thesis of MacDonald (1999) that “male workplace changes lead to household residential location changes and female workplace changes and that this pattern leads to shorter female and longer male commute trips”. Consistent with this study it means that actually it is the occupation of the male worker that determines how long female in household will commute. The issue here can seem paradoxical. If males can choose work journey, why would they choose long commuting distance? It may have something to do with best choice of residential location. Attractive work places and attractive housing areas do not necessarily have the same location. Accordingly, residential choice is dependent on male work place and female work place is dependent on residential location. And as said earlier, females often choose work close to home for different reasons, hence shorter work journeys for females.

2 Model and method

The gravity model is often used for estimation of different patterns of spatial interaction. The model received its name from Newton's gravity model, that says following: attraction between two objects is proportional to their mass and inversely proportional to their respective distance. In investigation of spatial interaction, different variables are used dependent on which area of interaction is studied. Gravity models of trade forecast trade flows often based on GDP and distances between destinations. Models of migration are used to predict degree of interaction between two places such as traffic flows and migration between two areas. In this thesis, the gravity model is used to analyse commuting. Commuting pattern, C_{ij}, is a function of the amount of labour in given origin municipality, O_i, the amount of jobs in destination municipality, D_j, and distance between municipalities T_{ij}:

$$C_{ij} = f\left(O_i, D_j, T_{ij}\right) \qquad (1)$$

C_{ij} is amount of commuters from location i to location j, these streams of people can be related to the amount of jobs, D_j, in destination j and to amount of labour, O_i, in origin i. In addition, it has to be taken into consideration the distance, i.e. travel time between two given destinations, T_{ij}. The simplest formula of spatial interaction is given by:

$$C_{ij} = \alpha \frac{O_i D_j}{T_{ij}}, \qquad (2)$$

where α is a level constant. The formula says that the spatial interaction of two locations is proportional to their importance and inversely proportional to their respective distance. Interaction between locations declines when the distance, commuting time or commuting costs increase. In other words commuting between municipalities depends on the importance of the given origin municipality, expressed by the labour force settled in the municipality, on the importance of the given destination municipality, expressed by accessibility to the work places and the distance between these municipalities.

To make the model more applicable in reality, several parameters can be extended to the model:

$$C_{ij} = \alpha \frac{O_i^{\beta} D_j^{\gamma}}{T_{ij}^{\lambda}} \qquad (3)$$

β is the so called push factor. It means that the more labour there is in origin the more people will commute to all other destinations. γ is the attractiveness of destination related to workplaces, the so called pull factor. Look at Figure 1: if *j* is a given municipality with jobs, we can assume that the more jobs there are in municipality *j*, the more people will commute there, because the probability to find job is higher.

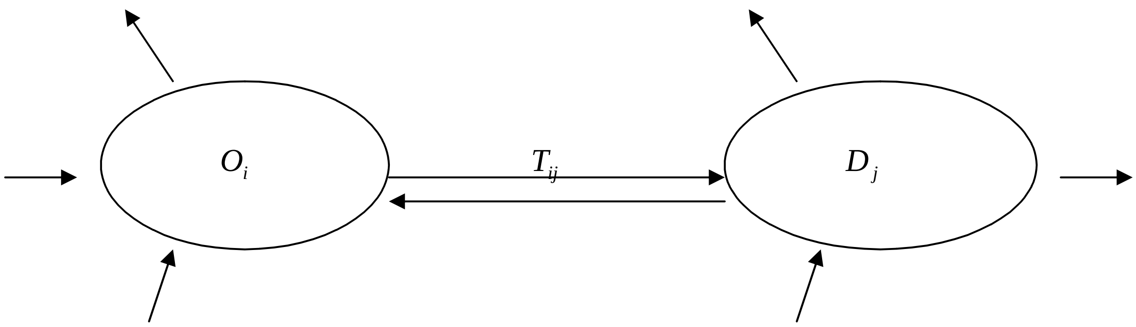


Figure 1: Spatial interaction between two destinations were:

- O_i – the amount of labour that is settled in given municipality *i*
- D_j – the amount of jobs in given municipality *j*
- T_{ij} – the distance between municipality *i* and municipality *j*

Note: O_i can be any given municipality with labour in Sweden, D_j can be any given municipality with workplaces. There is spatial interaction not only between two municipalities at the same time. If O_i is the origin municipality there can be more destination municipalities and this origin municipality can be destination municipality for commuters from any other given municipality. Arrows directed from O_i represent commuting to other communities, arrows directed to O_i represent commuting to municipality *i* from other municipalities. The same is intended for D_j: arrows directed away from circle represent commuting to other municipalities and arrows directed to circle represent commuting flows from other municipalities. Thus municipalities *i* and *j* can simultaneously be both origin and destination.

λ is called distance friction, i.e. it shows how the length of the work trip influences commuting. It is reasonable to assume that the longer trip to the work the less people will commute. Figure 2 shows that commuting and trip distance have a negative relationship. And more important, at a certain point the amount of commuters decreases remarkable. It is because average daily commuting trip is between 77 and 59 kilometres depending on the actual conditions of infrastructure and commuter's gender (SIKA, 2000). Accordingly, if the average length of trip is crossed, the amount of commuter decreases rapidly. The decrease is relatively drastic in relation to rise in travel time. In this thesis commuting over border of municipalities but not commuting within municipalities will be treated. If we would add the last variable, the fall of amount of commuters by increase of distance would be even larger because usually commuting trips within municipalities are shorter and there are more commuters. But as mentioned, commuting within municipalities is not included in this thesis.

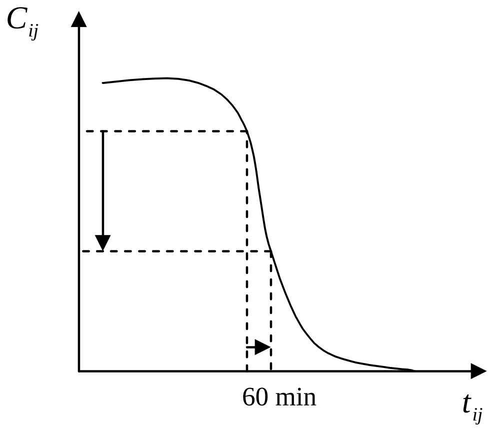


Figure 2: Relationship between commuting and trip distance.

By continuing the evolution of the model T_{ij}^{λ} from equation (3) will be replaced with $e^{\lambda T_{ij}}$, because according to studies is this expression better suitable to the data. By transformation of equation (3) we get following spatial interaction equation:

$$\hat{C}_{ij} = \alpha O_i^{\beta} D_j^{\gamma} \exp\{-\lambda T_{ij}\}, \text{ where } \exp\{-\lambda T_{ij}\} \text{ is the same as } \frac{1}{\exp\{\lambda T_{ij}\}}. \quad (4)$$

Shortly, the model says that the amount of commuters is the dependent variable and it depends on the amount of jobs and employees in all destinations, and commuting distance. Variables D_j, O_j and T_{ij} are given, parameters α, β, γ and λ have to be estimated. The values of $\alpha, \beta,$ and γ will influence the prediction of spatial interaction.

Equation (4) is not in linear form and it can not be used for estimation by linear regression. It has to be converted to the linear form by taking logarithm of equation (4) and then equation (5) can be used to estimate parameters in a regression model:

M1: $$\ln \hat{C}_{ij} = \alpha' + \beta \ln O_i + \gamma \ln D_j + \lambda T_{ij} \qquad (5)$$

where α' is the logarithm of α. M1 is given by equation (5) and it will be the basis to develop hypothesis about differences in male and female commuting. The model will be used on total data, data for respective men and women. The parameters will be estimated for each group. According to earlier research men commute longer both in kilometres and in time, following hypothesis can be considered:

H1: The distance friction λ is assumed be smaller for male than for female commuters, $\lambda_f < \lambda_m$, i.e., the distance influences work trips of men relatively less. If distance friction for male is larger than for female, a further hypothesis can be considered $\lambda_1 > 0$. λ_1 shows that extra effect to regression outcome if the commuter is male, and it will be presented in a later model, M3.

H2: Since relatively more men commute, it can be assumed that the pull factor γ men will be higher for men than for women, $\gamma_f < \gamma_m$. In consideration that women tend to work in the residence municipality, it can be assumed that γ_1 from M4, the extra amount of commuters in regression outcome that depends on male commuting, is larger than zero, $\gamma_1 > 0$.

H3: The amount of male labour in origin should have higher effect on commuting and thus the push factor β for men is assumed to be higher than for women, $\beta_f < \beta_m$. If the push factor is higher for men, then β_1 should accordingly be larger than zero because women tend to work where they live and commute relatively less. Thus hypothesis $\beta_1 > 0$ can be considered.

To summarize it, if the commuter is male the push and pull factors will be higher and the distance friction factor will be lower than for female. Moreover, all the extra effects to commuting that male commuting provides, are larger than zero. The hypothesis can be explained as follows. Women are more sensitive to distance because of their responsibilities in housekeeping and the weaker position in the labour market. H2 suggest that since there is higher number of men that commute everyday, it is reasonable to assume that if there are more work places in a given destination, even more males are attracted to commute. H3 relies on the same assumption that higher number of males commute. If there is more male labour force in a origin, thus there are more men that potentially can commute and hence higher push factor for men.

If we now extend the hypothesis, even if the assumptions are true we do not know exact how much higher are the effects from male commuting. One possibility to show differences more illustratively in female and male commuting is to add dummy variables to M1. Dummy variables are used to account for qualitative factors in econometric models. In general dummy variables increases model fit and its coefficient of determination. It indicates the non-existence or existence of some effect that may change the results of outcome. Example is dummy variable information about subgroups that may possibly have different intercepts. Here can gender be the basis for appearance such subgroups. First dummy variable will be treated as supplemental to intercept. Adding a dummy variable G (1 if the person is male; 0 otherwise) gives:

M2: $$\ln \hat{C}_{ij} = \alpha' + \beta \ln O_i + \gamma \ln D_j + \lambda T_{ij} + \alpha'_1 \mathrm{G} \quad (6)$$

M2 can be interpreted as follows: given the amount of labour in one municipality, jobs in another municipality and the distance between municipalities, the amount of commuters will be on average α_1' units different if the person is male.

Instead to assume that gender affects the value of intercept as in M2, there can also be changes in the slope of the relationship. Changes in slope can be allowed by including in the model an additional explanatory variable that is equal to the product of dummy variable and one continuous variable. In next model the variable added is the product of length of commuting trip and dummy variable gender, $[GT_{ij}]$:

M3: $$\ln \hat{C}_{ij} = \alpha' + \beta \ln O_i + \gamma \ln D_j + \lambda T_{ij} + \lambda_1 [GT_{ij}] \quad (7)$$

$[GT_{ij}]$ in this model is called interaction effect because it is the effect of interaction of two variables. The new variable is the extra distance friction that male commuting provides. The interaction variable takes value equal to commuting distance if $G = 1$, and zero otherwise. The parameter λ_1 in M3 shows that amount of extra effect on commuting pattern that male commuting gives. Related to H1 λ_1 is expected to be positive; $\lambda_1 > 0$.

Further interaction variable that will be added in the model is the product of amount of jobs in destination municipality and dummy variable gender, $[G \ln D_j]$:

M4: $$\ln \hat{C}_{ij} = \alpha' + \beta \ln O_i + \gamma \ln D_j + \lambda T_{ij} + \gamma_1 [G \ln D_j] \qquad (8)$$

γ_1 gives the amount of the so called extra pull factor that depends on male commuting, i.e. if the commuter is male the amount of jobs in destination municipality *j* makes the commuting more attractive. Consistent with H2: $\gamma_1 > 0$, γ_1 is expected to be positive.

Variable $[G \ln O_i]$ will be added in M5. It is the product of amount of labour in origin municipality and gender:

M5: $$\ln \hat{C}_{ij} = \alpha' + \beta \ln O_i + \gamma \ln D_j + \lambda T_{ij} + \beta_1 [G \ln O_i] \qquad (9)$$

This new variable says something about in which proportion the commuting pattern will be affected if the person who is settled in origin municipality, is male. There is no extra friction if the person is female. As expressed in H3: $\beta_1 > 0$, β_1 is expected to be positive.

M6 will include all dummy variables in one equation:

M6: $$\ln \hat{C}_{ij} = \alpha' + \beta \ln O_i + \gamma \ln D_j + \lambda T_{ij} + \alpha_1' G + \beta_1 [G \ln O_i] + \gamma_1 [G \ln D_j] + \lambda_1 [GT_{ij}] \qquad (10)$$

3 Data

The commuting data used in thesis comes from SCB (Statistics Sweden) and data about travel time collected by the Swedish Road Administration. The data used in the paper is commuting data over borders of municipalities i. e. commuting within municipalities is not included. Commuting time is restricted to 90 minutes single trip, which is reasonable commuting time for everyday commuting. Sure there are people that have longer work trips, but probably they do not commute every day and in this thesis will be examined only everyday journeys to the work. Thus these persons whose commuting time exceeds 90 minutes single trip have been outsorted and they will be not considered in this thesis as everyday commuters.

Data set includes 289 municipalities of Sweden. It contains total 1 003 771 individuals that have work journey from one municipality to another, 574 661 of them male and 429 110 female. According to given data set the distribution between male and female commuters is 57.2 and 42.7 per cent respectively for men and women if commuting time from municipality *i* to municipality *j* is less than 30 minutes, see Table1.

Table 1: Female, male and total commuting versus travel time

Travel time	Female			Male			Total	
	Quantity	%*	%**	Quantity	%*	%**	Quantity	%*
<30 minutes	350 621	81.71	44.82	431 746	75.13	55.18	782 367	77.94
31-40 minutes	36 482	8.50	37.16	61 965	10.74	62.84	98 177	9.78
41-50 minutes	18 887	4.40	35.87	33 764	5.88	64.13	52 651	5.25
51-60 minutes	11 389	2.65	34.17	21 946	3.82	65.83	33 335	3.32
61-70 minutes	5303	1.24	32.17	11 181	1.95	67.83	16 484	1.64
71-80 minutes	3798	0.89	31.15	8393	1.46	68.85	12 191	1.21
81-90 minutes	2630	0.61	30.70	5936	1.03	69.30	8566	0.85
Total	429 110	100	42.75	574 661	100	57.25	1 003 771	100

* The relative amount of respectively females, males and total commuters classified after commuting time.

** The relative amount of females and males in each group of travel time

The longer the trip in time is the smaller the share of female commuters. When the commuting time from origin *i* to destination *j* is between 80 and 90 minutes the share of women have decreased to 30.7 per cent. It is a decrease by 17.8 per cent. It clearly shows that more male commuters have longer work journey in time. Average length of commuting trips, as expected, differs for male and female commuters. Women's work journey is in average 4 minutes shorter that it is for men, respective 49 and 53 minutes. In Table 1 can be screened commuting times and number of commuters, female and male separately and all together. About 2/3 of all commuters, precisely 77.9 per cent, have daily work journey less than 30 minutes. If we look at men and women separately, than an amount of women who have commuting trip less than 30 minutes, is higher, about 81.7 per cent compared to men's 75 per cent. According to previous studies of commuting the fact that more women choose to work close to home will be confirmed here.

Data set used for estimation is tested for correlation and collinearity. These results are shown in Appendix. In actual case there is no significant correlation between variables. Also the assumption of normal distribution is fulfilled. Thus regression models can be used for estimation of commuting patterns.

4 Results

The presentation and analysis of results focuses on two important components, first differences in male and female commuting and second the proportion of the differences. The outcomes of 6 different regression models will be presented. Table 2 contains estimated parameters for models M1-M6 and respective t-values. Chapter 5.1 includes outcome for M1 using first the total data that includes both male and female data, and then results with either female or male data separated. Chapter 5.2 will describe the outcomes of models with dummy variables added one by one, respective models M2-M5. Finally chapter 5.3 consist the outcome of regression model M6 were all dummy variables are included to equation (5). For more detailed results see Appendix 1.

Table 2: Estimated parameters based on Models 1-6., respective t-values in parentheses.

	α'	β	γ	λ	α'_1	λ_1	γ_1	β_1
M1 T*	-6.091	0.488	0.831	-0.063				
	(-34.569)	(34.522)	(64.713)	(-105.038)				
M1 F*	-4.958	0.380	0.740	-0.061				
	(-25.812)	(23.807)	(50.023)	(-82.766)				
M1 M*	-5.845	0.472	0.786	-0.058				
	(-34.569)	(33.295)	(62.142)	(-97.340)				
M2	-5.646	0.427	0.764	-0.059	0.446			
	(-44.303)	(40.249)	(79.258)	(-127.491)	(22.206)			
M3	-5.439	0.430	0.765	-0.064		0.008		
	(-42.739)	(40.423)	(79.242)	(-121.446)		(21.457)		
M4	-5.408	0.427	0.738	-0.059			0.049	
	(-42.556)	(40.261)	(76.449)	(-127.505)			(22.369)	
M5	-5.417	0.400	0.764	-0.059				0.052
	(-42.644)	(37.301)	(79.315)	(-127.577)				(22.566)
M6	-4.958	0.380	0.740	-0.061	-0.887	0.002	0.045	0.093
	(-26.972)	(24.876)	(52.271)	(-86.484)	(-3.490)	(2.578)	(2.347)	(4.359)

* M1 Total, M1 Female, M1 Male

Significance level on the parameters is 5 %.

4.1 Outcome of Model 1

With help of M1 commuting patterns for total, female and male over border commuting between municipalities in Sweden can be estimated. The model says that given the amount of labour in one municipality, the amount of jobs in another municipality and the distance between these two municipalities, the outcome for commuting patterns can be calculated approximately. In Table 2 all parameters and t-values are presented, according which all parameters are significant.

By interpreting the models, it is important to note that if in M1 Total O_i stands for all labour in given municipality *i*, then in M1 Female and M1 Male, O_i stands for respective female and male labour settled in given municipality *i*. Comparing models M1-M3 gives us idea how commuting patterns in general are affected by given variables. There are differences in parameters that were expected. In general can be said that labour settled in municipality *i* and the work places in municipality *j* affect the commuting patterns positively and the distance between the municipalities have a negative effect on commuting. It can be seen that there are differences in parameters for male and female commuting models. Hypothesis H1 states that distance friction for male commuting model is smaller than for female. According to regression results is this hypothesis truth: $\lambda_f = -0.061$ and $\lambda_m = -0.058$. In other words λ in male commuting model have a less negative effect on commuting outcome. Looking at female commuting model, if time distance between two municipalities increases by one per cent, the amount of commuters will decrease, on average, by 0.061 per cent, if all other variables are held constant. Respective decrease in male commuting model is only 0.058 per cent. According to regression outcome, and as assumed in H1, women are relatively more sensitive to commuting distance. The distance friction influences the regression outcome negatively but the effect is smaller for male commuting in contrast to female. H2 states that the pull factor is higher for male commuting model. It is clear that estimated parameters satisfy the condition for H2 to be verified: $\gamma_f = 0.740$ and $\gamma_m = 0.786$, thus $\gamma_f < \gamma_m$. If amount of jobs in destination municipality will increase by one per cent the amount of commuters from location *i* to location *j* will increase, on average, by 0.786 per cent in the case of male commuting. Higher number of work places in destination municipality is more attractive for men than for women. H3 states that so called push factor is higher for men than for women. From Table 2 can be seen that $\beta_f = 0.380$ and $\beta_m = 0.472$. If amount of female labour in origin increases by one per cent the amount of women that commute from location *i* to location *j* increases, on average, by 0.380 per cent if all other variables are held constant. Respective increase in male commuting model is 0.472 per cent. Condition $\beta_f < \beta_m$ also holds.

4.2 Results of Models 2-5

Added dummy variables in M2-M5 indicate better model fit and higher coefficient of determination. Dummy variable in M2 illustrates the particular effect of gender on regression outcome and affects the position of intercept.

$$\text{M2:} \quad \ln \hat{C}_{ij} = -5.646 + 0.427 \ln O_i + 0.764 \ln D_j - 0.059 T_{ij} + 0.446 G \qquad (11)$$

α'_1 is the extra effect on commuting if the commuter is male. If the amount of labour in origin, jobs in destination and the distance are given, the amount of commuters from municipality *i* to municipality j will be on average 0.446 per cent higher if the commuter is male. M3 includes variable [GT_{ij}], the interaction between gender and distance:

$$\text{M3:} \quad \ln \hat{C}_{ij} = -5.439 + 0.430 \ln O_i + 0.765 \ln D_j - 0.064 T_{ij} + 0.008 [GT_{ij}] \qquad (12)$$

If the commuter is male, the extra friction from distance is on average 0.008 per cent, all other variables held constant. It states that the distance has relatively smaller negative effect on male commuting. And thereby hypothesis H1, $\lambda_1 > 0$, is verified. Variable [GD_j] is added to M4. It is the effect of relation between gender and given amount of work places in the destination municipality.

$$\text{M4:} \quad \ln \hat{C}_{ij} = -5.408 + 0.427 \ln O_i + 0.738 \ln D_j - 0.059 T_{ij} + 0.049 G [GD_j] \qquad (13)$$

As follows, if the commuter is male, the extra effect on commuting outcome is 0.049 per cent. i.e. hypothesis H2 is truth, $\gamma_1 > 0$. If all other variables are given, there are, on average, 0.049 per cent more commuters from *i* to *j* if the commuter is male. In M5 is shown the extra friction that interaction of gender and the amount of labour in given municipality, *i* gives:

$$\text{M5:} \quad \ln \hat{C}_{ij} = -5.417 + 0.400 \ln O_i + 0.764 \ln D_j - 0.059 T_{ij} + 0.052 [GOi] \qquad (14)$$

If the person is male, the regression outcome is on average 0.052 per cent higher if all other variables are held constant, and accordingly the hypothesis H3 is verified, $\beta_1 > 0$.

4.3 Outcome of Model 6

In Model 6 are included all variables discussed earlier in different models. It consist the parameters estimated from total commuting data.

$$\ln \hat{C}_{ij} = -4.958 + 0.380 \ln O_i + 0.740 \ln D_j - 0.061 T_{ij} - 0.887 G + 0.002[GO_i] + 0.045[GD_j] + 0.093[GT_{ij}] \qquad (15)$$

If we compare models M4-M5 where only one dummy variable is added in respective model and the last model M6 where the all dummy variables are included, can we see that the parameters are different. The regression outcome of the M6 takes in account the effects of all variables. If λ_1 in M3 is 0.008, then respective value in M6 is 0.002. All effects taken in consideration, the effect of λ_1, the extra distance friction if the commuter is male, is smaller. There is similar result by looking at parameter γ_1. In M6 is the extra effect from male commuter smaller than in M4. The opposite can be said about β_1, the extra effect from male commuter is higher in M6. By further examining of models can be seen that parameters β, γ, λ in M1 Female and M6 are equal. And by addition of parameters β_1, γ_1 and λ_1 from M6 and β , γ and λ from M1 Female gives the result equal to parameters β , γ and λ in M1 Total. It clearly shows that by adding dummy variables β_1, γ_1 and λ_1 in M6, we have separated the effect that male commuting provides. By representing the information in such form allows discerning the effects of different subgroups, like in this case gender dissimilarities.

4.4 Discussion of the results

To summarize, the results verify the hypothesis. By taking in account variables such as labour in origin municipality, jobs in destination municipality and the distance between these two, male commuting is in relatively higher degree affected by the two first variables and less affected by distance. The fact that the amount of male labour in origin affects the commuting relatively more may indicate that men are more ambitious and are relatively more willing to commute to get better job status. At the same time can be said that they relatively often choose not to take lower status jobs that are available close to home. Female labour in origin affects relatively less the commuting outside residence municipality. It may indicate that women have easier to find jobs in home municipality and by that they have fewer incentives to commute. Or it can depend on household's responsibilities that they choose not to commute. Results imply also that males are more attracted by amount of work places in other municipalities. The reasons can be the same as mentioned, men are more ambitious and have often higher wages and it gives motivation to commute. Whereas females commute relatively more close to home and are less attracted by work places in other municipalities. For females the household's responsibilities counter balance these higher earnings, which they would get by choosing to commute. Like Camstra (1996) said that women are relatively more affected by labour market imperfections. Finally, males commute relatively longer distances. We can see similar results in comparison with earlier researches. The women are relatively more sensitive to distance. Probably, even if the work places in other municipalities may effort some advantages such as higher salaries or better career opportunities, is the distance clear disadvantage. And in combination with household's responsibilities, women relatively often choose not to commute as long distances as men do. It can be understood as follows, men maximize their utility by higher earnings and job status, while women by having job and housekeeping in balance. Future trend can be expected that female and male commuting differences get smaller due the females in Sweden strive for same conditions on the labour market as males have, assuming that the women and men have the same incentives to commute.

5 Conclusions and suggestions for future studies

Main purpose of the paper was to examine the differences in male and female commuting in Sweden by using gravity model adjusted for spatial interaction studies. Estimated regression models show clearly that there are differences in female and male commuting. First, the amount of male labour in origin municipality influences in higher degree the commuting to all other municipalities. Second, the higher amount of jobs in destination municipality is more attractive for male commuters. And third, men are more indifferent for commuting distance. The dummy variables included to regression model allow easier interpretation of regression results.

In this thesis is the commuting in Sweden examined with help of gravity model, i.e. commuting that is affected by properties of two locations and distances between these locations. But it can be appropriate to study same issue with alternative spatial interaction models. One of them is competing origin/destination model, which says that spatial decision makers use hierarchical strategy, that is, at first choose a region for interaction and then the specific destination within preferred region (Akwawua and Pooler 2001). The destinations compete with each other, good qualities of one destination can compensate for bad qualities of another. It would be interesting to investigate how selections of broad regions in Sweden differ for men and women and what influences it. One other model for spatial interaction is intervening opportunity model which examines issues like why is one destination more attractive than another. It can be asked here, are the locations which are attractive for men as attractive also for women?

Furthermore commuting over time has to be taken into closer examination. In chapter 5 calculated parameters can differ from each other in different periods of time, this due the technological progress and better infrastructure conditions, access to workplaces and amount of labour in different locations. In different period of time the accessibility of workplaces in different destinations can vary and it influence parameters that determine attractiveness to commute and parameters which determine movements. Also the labour can fluctuate over time, decreases and increases in labour force affect the push factor – higher amount of labour means higher push factor and vice versa. Technological progress improves transport efficiency and thereby more people are willing to commute to the work and journey to the work will be longer because they can move longer distances with same time.

This paper consists of studies that include only commuting over borders of municipalities. One other interesting subject for future studies could be to include intra municipality commuting or introduce some other variables into the model, example gasoline price.

From macroeconomic point of view can be interesting to investigate how the new a-kassa rules in Sweden affect the commuting. If today an unemployed person in Sweden must be willing to take a job practically anywhere in Sweden, disregarding how long he or she has to commute then how will this new rule change the commuting behaviour by pushing people to choose to move closer to urban areas with better opportunities to find a job?

References

Andersson, G., Jorner, U. & Ågren, A. (1994) *Regressions- och tidsserieanalys*. Lund: Studentlitteratur.

Angel, S. & Hyman, G.M. (1972) Urban spatial interaction. *Environment and Planning*, *4*, 99–118

Camstra, R. (1996) Commuting and Gender in a Lifestyle Perspective. *Urban Studies*, *33*, 283-300

Hill, R. C., Griffits, W. E. & Judge, G.G. (2001) *Undergraduate econometrics* (2nd edition). Hoboken: John Wiley & Sons Inc.

Körner, S. & Wahlgren, L. (2006) *Statistisk dataanalys* (4th edition). Lund: Studentlitteratur.

MacDonald H.I. (1999) Women's Employment and Commuting: Explaining the Links. *Journal of Planning Literature*, *13*, 267-283

Olsson, M. (2002) *Studies of Commuting and labour market integration*. Jönköping International Business School

Sohn, J. (2005) Are commuting patterns a good indicator of urban spatial structure? *Journal of Transport Geography*, *13*, 306-317

Sermons, M.W. & Koppelman, F.S. (2001) Representing the differences between female and male commute behavior in residential location choice models. *Journal of Transport Geography*, *9*, 101-110

Turner, T. & Niemeier, D. (1997) Travel to work and household responsibility: new evidence. *Transportation*, *24*

Hu, P. & Pooler, J. (2002) An empirical test of the competing destinations model. *Journal of Geographical Systems*, *4*, 301-323

Siaw; A. & Pooler, J.A. (2001) The development of an intervening opportunities model with spatial dominance effects. *Journal of Geographical Systems*, *3*, 69-86

Rodrigue, J-P., Comtois, C. & Slack, B. (2006) *The Geography of Transport Systems.* New York: Routledge.
Accessible in the Internet: http://people.hofstra.edu/geotrans/index.html (from 2008-02-10)

Appendix 1 Model summary

Table A1: Summary of estimated regression models

Model	R	R^2	Adjusted R^2	Std.error of the estimate
M1 Total	0.837	0.701	0.701	1.0896345
M1 Female	0.794	0.631	0.631	1.1441368
M1 Male	0.824	0.679	0.679	1.0536352
M2	0.811	0.658	0.658	1.0962559
M3	0.810	0.657	0.657	1.0976697
M4	0.806	0.650	0.650	1.1087015
M5	0.811	0.658	0.658	1.0955600
M6	0.811	0.658	0.658	1.0949425

In Model Summary is seen R^2 that measures the relationship between the dependent and independent variables, i.e. how large share of variation of dependent variable can be explained by the linear regression equation, expressed as percentage of total variation. The higher the R^2, the higher percent of variation in dependent variable is explained by regression. In current case it cannot be interpret because interpretation of R^2 calculated in logarithm form is not correct. This because of sum of error terms calculated by the least squares estimator has to be zero. But sum of error terms calculated on values on logarithm form, like here, is not necessarily equal to zero. The variables have a relationship that reminds linear regression if they are in logarithm form but not if they are in original non-logarithm form. R^2 is only a descriptive measure, it does not say anything about the quality of the model. Problem with R^2 in multiple regression model can be that by adding more variables it can be made larger, even if the variables added have no economic importance. An alternative measure here can be adjusted R^2 that do not explain the percent of variation but says something about the proportion of the variation in the dependent variable explained by the explanatory variables.

Appendix 2 ANOVA (Analysis of variances)

The total sum of squares (SST) is a total variation of dependent variable. Sum of squares of error (SSE) measures a variation around the adjusted regression line that is not explained by regression. The sum of squares of regression (SSR) is the variation that is explained by regression. To every sum of squares belongs an amount of degrees of freedom and mean sum of squares (mean square) that given by the equivalent sum of squares divided by the amount of degrees of freedom it is build on. The mean square error is estimate of the error variance.

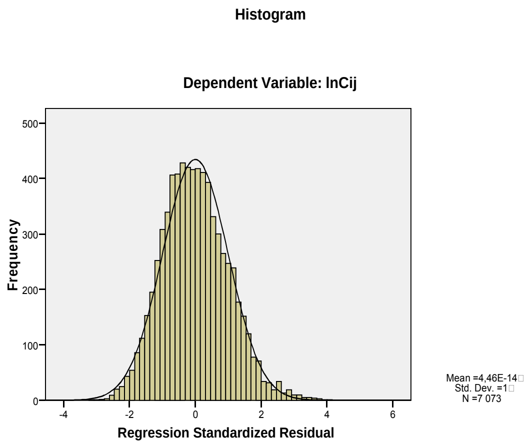
The square root of the last one is spread of residuals that can be described as a standard deviation around the regression line.

Table A2: ANOVA

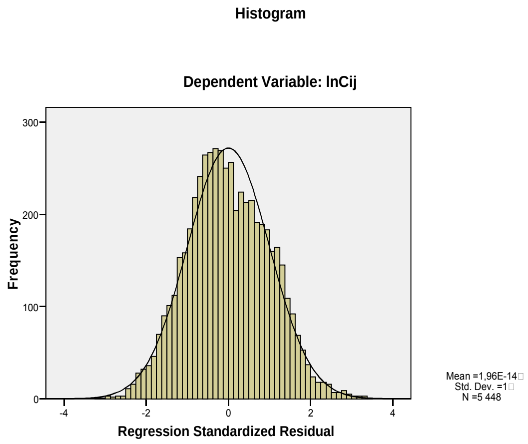
Model		Sum of squares	df	Mean square	F	Sig.
M1 Total	Regression	19664.499	3	6554.833	5520.774	0.000
	Residual	8393.047	7069	1.187		
	Total	28057.546	7072			
M1 Female	Regression	12181.651	3	4060.550	3101.909	0.000
	Residual	7126.462	5444	1.309		
	Total	19308.114	5447			
M1 Male	Regression	15899.367	3	5299.789	4773.952	0.000
	Residual	7500.155	6756	1.110		
	Total	23399.522	6759			
M2	Regression	28154,896	4	7038,724	5856,930	0,000
	Residual	14665,285	12203	1,202		
	Total	42820,181	12207			
M3	Regression	28117,045	4	7029,261	5833,999	0,000
	Residual	14703,135	12203	1,205		
	Total	42820,181	12207			
M4	Regression	27820,020	4	6955,005	5658,068	0,000
	Residual	15000,161	12203	1,229		
	Total	42820,181	12207			
M5	Regression	28173.510	4	7043.377	5868.251	0.000
	Residual	14646.671	12203	1.200		
	Total	42820.181	12207			
M6	Regression	28193.612	7	4027.659	3359.464	0.000
	Residual	14626.569	12200	1.199		
	Total	42820.181	12207			

Appendix 3 Normal distribution

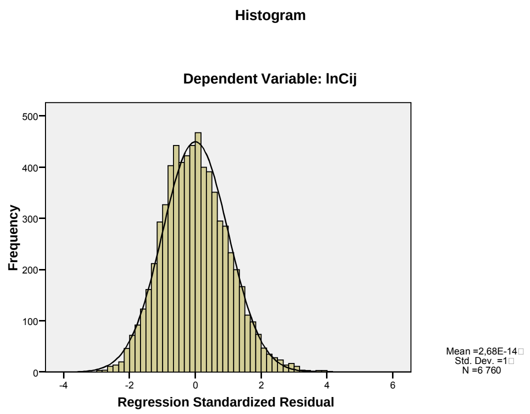
By choosing log-log model the assumption of normal distribution has to be fulfilled. As it can be seen in following figure the distribution of residuals of all estimated regression models remind normal distribution, i.e the assumption of normal distribution is fulfilled.



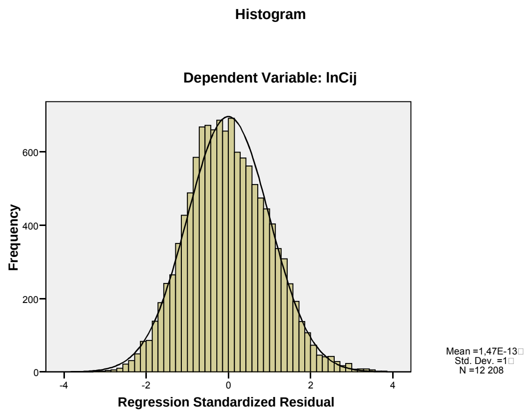
(a), Total model



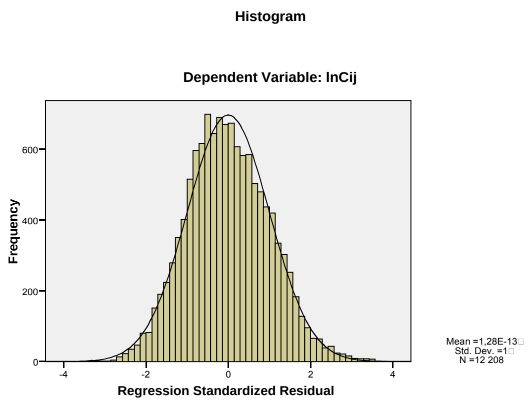
(b) Female



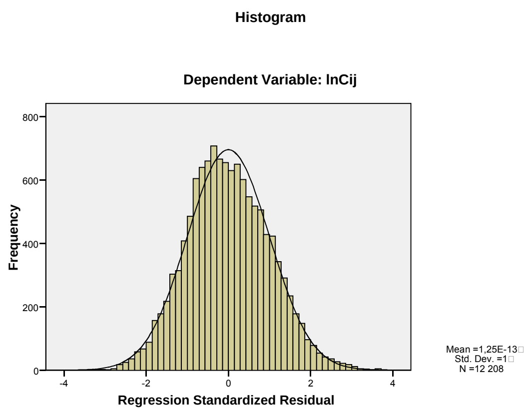
(c) Male



(d) Gender added



(e) GT_{ij} added



(f) GD_j added

Figure A2: Distribution of residuals

Appendix 3 Collinearity statistics

One problem with regression models can be that some of economic variables move together in systematic way and because of this there can be difficulties to sort out the effect of one particular explanatory variable on the dependent variable. It is called collinearity. The easiest way to detect collinearity is to calculate correlation between variables two by two. In the case of no correlation between two variables, the value of correlation coefficient is 0. The rule of thumb is that a correlations coefficient between two explanatory variables is greater than 0.8 or 0.9 indicates a strong linear association and it means a harmful collinear relationship. Problem arises if there are more then two variables that are involved in multicollinear relationship. It can not be discovered by examining the variables pair wise. To detect multicollinearity VIF-test can be used. VIF index can be easily calculated in SPSS. If collinearity exist, the information data gives about independent variable's is not enough to see individual effect of each particular explanatory variable and it most likely leads to incorrect outcome of regression. Next table includes correlation statistics that can be indicator of collinearity.

Table A3: Table of correlations

		$\ln C_{ij}$	$\ln O_i$	$\ln D_j$	T_{ij}	G	GT_{ij}	GD_j	GO_i
M1 Total	$\ln C_{ij}$	1.000	0.268	0.411	-0.683				
	$\ln O_i$	0.268	1.000	0.033	-0.043				
	$\ln D_j$	0.411	0.033	1.000	0.026				
	T_{ij}	-0.683	-0.043	0.026	1.000				
M1 Female	$\ln C_{ij}$	1.000	0.208	0.357	-0.647				
	$\ln O_i$	0.208	1.000	0.027	-0.001				
	$\ln D_j$	0.357	0.027	1.000	0.091				
	T_{ij}	-0.647	-0.001	0.091	1.000				
M1 Male	$\ln C_{ij}$	1.000	0.263	0.409	-0.662				
	$\ln O_i$	0.263	1.000	0.028	-0.031				
	$\ln D_j$	0.409	0.028	1.000	0.040				
	T_{ij}	-0.662	-0.031	0.040	1.000				
M2	$\ln C_{ij}$	1,000	0,238	0,382	-0,647	0,051			
	$\ln O_i$	0,238	1,000	0,026	-0,014	0,038			
	$\ln D_j$	0,382	0,026	1,000	0,059	-0,040			
	T_{ij}	-0,647	-0,014	0,059	1,000	0,086			
	G	0,051	0,038	-0,040	0,086	1,000			
M3	$\ln C_{ij}$	1,000	0,238	0,382	-0,647		-0,211		
	$\ln O_i$	0,238	1,000	0,026	-0,014		0,021		
	$\ln D_j$	0,382	0,026	1,000	0,059		-0,019		
	T_{ij}	-0,647	-0,014	0,059	1,000		0,461		
	GT_{ij}	-0,211	0,021	-0,019	0,461		1,000		
M4	$\ln C_{ij}$	1,000	0,238	0,382	-0,647			0,259	
	$\ln O_i$	0,238	1,000	0,026	-0,014			0,025	
	$\ln D_j$	0,382	0,026	1,000	0,059			0,449	
	T_{ij}	-0,647	-0,014	0,059	1,000			0,004	

		lnC_{ij}	lnO_i	lnD_j	T_{ij}	G	GT_{ij}	GD_j	GO_i
	GD_j	0,100	0,041	0,080	0,089			1,000	
M5	lnC_{ij}	1.000	0.238	0.382	-0.647				0.081
	lnO_i	0.238	1.000	0.026	-0.014				0.149
	lnD_j	0.382	0.026	1.000	0.059				-0.037
	T_{ij}	-0.647	-0.014	0.059	1.000				0.081
	GO_i	0.081	0.149	-0.037	-0.037				1.000
M6	lnC_{ij}	1.000	0.238	0.382	-0.647	0.051	-0.211	0.100	0.081
	lnO_i	0.238	1.000	0.026	-0.014	0.038	0.021	0.041	0.149
	lnD_j	0.382	0.026	1.000	0.059	-0.040	-0.019	0.080	-0.037
	T_{ij}	-0.647	-0.014	0.059	1.000	0.086	0.461	0.089	0.081
	G	0.051	0.038	-0.040	0.086	1.000	0.853	0.986	0.988
	GT_{ij}	-0.211	0.021	-0.019	0.461	0.853	1.000	0.845	0.841
	GD_j	0.100	0.041	0.080	0.089	0.986	0.845	1.000	0.975
	GO_i	0.081	0.149	-0.037	0.081	0.988	0.841	0.975	1.000

The table of correlations shows that all values, except some values in M6, are close to zero, thus there is no significant correlation between explanatory variables and accordingly no collineraty between variables. The high correlation between variables in M6 can be explained by that variables are products of gender and one continuous variable.

VIF (Variance inflation factor)-test is used to detect multicollinearity. It measures how much is the variance of one particular coefficient increased because of the collienarity. The square root of VIF index implies how much larger the standard error is compared with what it would be in the case of no correlation. If VIF > 5, then there is high multicollinearity. According to table of coefficients there is no significant multicollinearity , all VIF indexes are close to 1, except some variables in M6. It can be explained that these variables are products of two variables and because of that they move together in systematic way.

Table A4: Collinearity statistics, VIF index

	lnO_i	lnD_j	T_{ij}	G	GT_{ij}	GD_j	GO_i
M1 Total	1.003	1.002	1.003				
M1 Female	1.001	1.009	1.008				
M1 Male	1.002	1.002	1.003				
M2	1.003	1.006	1.012	1.011			
M3	1.002	1.007	1.279		1.275		
M4	1.003	1.010	1.011			1.016	
M5	1.025	1.006	1.011				1.032
M6	2.078	2.178	2.301	162.709	8.385	80.608	86.404