VRIJE UNIVERSITEIT AMSTERDAM

RESEARCH PAPER

SCHOOL OF BUSINESS & ECONOMICS

Suburbanites:

Analysing the Evolution of Commuting Dynamics in The Netherlands

Supervisor:

Prof. Henri L.F. de Groot

Author:

Andres Guzman Cordero

Editor:

Igo Dobbe

June 24, 2022

Abstract

Commuting is generally perceived as a necessary evil. Lying at the intersection of two different areas of economics, labour and urban, we find a complex system that interweaves housing and work locations. This paper studies the evolution of the transport behaviour of commuting agents in the last decade to observe the distribution of employees and employers. We also ascertain the presence of the COVID-19 effect on commuting behaviour, estimate its impact and provide an analysis of its implication. Stemming out of these objectives, we make use of spatial panel data from The Netherlands over the last decade. The gravity model of migration was chosen to perform this analysis. It was found that the previously observed trends have not been exacerbated nor slowed by the COVID-19 pandemic. Overall, a dispersion dynamic has been observed where agents have either relocated or changed jobs. Furthermore, the distance decay has been increasing, corroborating the dispersion phenomenon. Therefore, the pape provides a brief introduction to the analysis of distance decay regarding commuting behaviour in The Netherlands. Thereby, leaving the use of advanced panel data models to account for the time and spatial lags yet to come.

Contents

1 Introduction

The current generations have experienced many 'once in a lifetime' events and while we are still living through the first war of this decade, the last crisis was unprecedented in its consequences. In less than 3 months, the streets were deserted and our lifestyle transformed. The COVID-19 virus keeps morphing many aspects of our lives after more than 2 years since its discovery. One of the most important changes that the pandemic brought upon us is in our relationship with transportation and life in agglomerations. During the first year, we saw an exodus from urban areas to less densely populated ones, with this phenomenon becoming commonplace in developed nations. One might raise the question of the implication of this change in our work-life relationship. With people moving out of the city, our travel behaviour might be different in significant ways. Furthermore, the need for social distancing drove the companies to allow working from home, another important reason to see a change in the travel behaviour of workers.

At the heart of these changes lies one part of our travel behaviour, commuting patterns. Commuting is generally perceived as a necessary evil, that short moment of the day where we go from point A (Home) to point B (Workplace) trying to minimise time, stress and other troublesome experiences. Some do not find commuting as irritating as others do, appreciating the time to be productive in the meantime. The study of these patterns is no simple feat. Lying at the intersection of two different areas of economics, labour and urban, we find a complex system that interweaves housing and work locations. The distribution of these and many other variables can influence the pattern that unfolds around an urban agglomeration. White (1986) tried to disentangle this system and look at the influence that gender has in commuting. The results exemplified the evolving dynamics of this part of our travel behaviour. Two dimensions of this problem are of special relevance, the time and distance of each trip. Many studies have put special emphasis on the travel behaviour of commuters to analyse these dimensions, with conclusions such as the 'Commuting Time Paradox' having important implications (Van Ommeren & Rietfeld, 2005). This paradox describes how the travel time of workers has been relatively stable disregarding the increasing distances covered by workers. The reason behind the Paradox has been the increased use of faster methods of transportation over time.

No explanation can fit all the variables that commuting patterns have. To shine some light on this complex behaviour, many studies have looked into other of the agent's characteristics. Groot, De Groot & Veneri (2013) have found that a higher education level predicts longer travel times over longer distances. Many reasons have been proposed for this phenomenon, house ownership being one of them. Other studies such as Lee & McDonald

(2002) have tried to provide a broader view on this topic, finding that family structure has an impact on the commuting behaviour of people. We can see that the many variables could be introduced and be of interest in this topic. By definition, this specific travel behaviour of agents is dependent on the work conditions such as location and other studies turn their focus on the work aspect of this travel phenomenon, such as firm decentralisation (Dubin, 1991). Lastly, another focus lies in the spatial aspect of commuting and its relation to the agents. Transportation infrastructure and urban structure have been identified by the analysis of commuting data (Sohn, 2004) showing discrepancies between the spatial structure identified in the commuting pattern while the distribution of the employed residents is not. Furthermore, regional differences have been found to influence the intentions to migrate to urban agglomerations (Van Ham & Hooimijer, 2008).

Closely related to human migration and taking inspiration from its similitude, commuting patterns have been estimated by making use of the gravity and radiation models. The use of the physical model to approximate other phenomena is not new. The use of the gravity model in social dynamics goes back to Ravenstein (1885, 1889) who confirmed that the population movements resembled Newton's law of gravity. The insight is simple, the population of urban agglomerations behave in the same way that mass behaves when describing gravity. Many decades later, the use of models from Physics in the social sciences was popularised by John. Q Stewart (Stewart, 1950). The gravity model still stands as the model of choice of many researchers (Poot et al, 2016), and has been proven to be more accurate in densely populated areas, such as London (Masucci et al, 2013). It is worth mentioning that the model has been modified from its original, Newtonian form. The modification lies in the behaviour of distance decay. The strength of the 'gravitational' pull was observed to have an inverse relation with distance, and not the distance squared as in the physical model (Zipf, 1946). Other methods, such as the radiation model of migration were born out of the many limitations of the gravity model. The gravity model does not have a rigorous derivation of its results and lacks theoretical guidance. Moreover, it requires traffic data to fit the parameters used. On the other hand, the radiation model provides a modelling framework without free parameters by taking inspiration from the travel process of energetic particles or waves through a vacuum. The results of this model have rigorous derivation and incorporate the interaction of two populations of a region. Under certain circumstances, such as estimation of country-to-country migration in the US, the results are more accurate than those provided by the gravity model (Simini et al, 2012). This is in contrast to the systematic inconsistencies of the gravity model under the same circumstances.

The Netherlands provides a perfect scenario to examine the commuting behaviour of agents. It is a developed nation with a well-developed transportation system and abundant traffic data. Being a densely populated country with a contrasting distribution of urban agglomerations, the country lends itself to a unique opportunity to compare the different models and examine the long-term travel behaviour of agents. From these patterns, many authors have tried to identify important variables such as household attributes and residential context. Moreover, these variables have a strong impact on modal choice, distance travelled, and other important consequences (Dieleman, Dijst, & Burghouwt, 2001). We expect that the pandemic made significant changes in the transport behaviour of commuting agents. We want to ascertain the presence of the COVID-19 effect on commuting behaviour, estimate its impact and provide an analysis of its implication. Stemming out of these objectives, the purpose of this paper is to analyse the change in commuting patterns in The Netherlands over the last decade. The relevant literature is presented in section 2, with special attention put into previous research on the Dutch commuting patterns and the predictive models of human migration. The methodology used in this paper is presented in section 3, differentiating between the data and the econometric analysis applied to the data to find the results as shown in section 4. We provide an analysis of the implication of our results in section 5, and finally, we provide some concluding remarks in section 6.

2 Literature Review

The intention to identify the different influential variables in an agent's commuting pattern is indicative of its importance. We can divide the studies into two groups based on their focus; the first focuses on the agent's characteristics while the other focuses on the spatial characteristics. Both focuses have been at the centre of many studies that analysed the commuting patterns in The Netherlands.

An agent's characteristics have been studied in different ways. Ory et al (2004) have shown that the intrinsic variability of each agent's perception of commuting is high. Many workers enjoy commuting, with different factors influencing the enjoyment of the trip. The study made use of four models: Objective Mobility, Subjective Mobility, Travel Liking and Relative Desired Mobility. Each model gave insight into a different area of the agent's preferences. While the authors established the necessity of further study, the evidence for heterogeneity in commuting behaviour demonstrates the complexity of the issue at hand. Their result has special consequences in its policy implications, due to the possible lack of uniform acceptance. Other studies have analysed specific discrepancies in agent preference for mode of transportation. Burbidge, Goulias and Kim (2005) have shown that the agent's lifestyle has a strong influence on the choice of transportation. Classifying modes of transport into active and inactive modes, the study shows that the trips made by each group are significantly different. These results also have important policy implications due to the risk of inactive people rejecting policies aimed at active lifestyles because of not meeting their needs. It is worth mentioning that the study might have different results if repeated in a new urban agglomeration due to the diverse city design. Active lifestyle modes of transportation have different adoption over countries (Burbidge et al, 2005). In The Netherlands, active modes of transportation such as bicycles, are significantly higher than in other nations (European Commission, 2014).

Numerous other putative causal variables of commuting time and distance may be found in the substantial research on this travel behaviour. For example, Giuliano (1998) showed that longer commutes are connected to the male gender, greater income, house ownership, age, usage of public transportation, and work status in a study of commuting time in Los Angeles. Several additional empirical studies has focused on gender disparities in work trip time and discovered that women commute to work faster than males. White (1986) had previously shown significant commuting disparities between males and females in his study of urban commuting patterns. Variables, such as the number of children and secondary workers at home, have different responses. Male's trip length would decrease if there are children present and a secondary worker, while female's trip length was unresponsive to variables other than young children. Song Lee & McDonald (2002) corroborated the sex disparities in their study on the determinants of commuting time and distance for Seoul Residents. In comparison to other workers, married women had shorter commutes. Self-employed, unpaid family workers and part-time workers have lower commutes in terms of both time and distance when compared to full-time workers. Longer commutes are associated with higher wages (homeowners, employees with higher education, and older male workers in this study). Commuting time differences between male and female married workers are much greater among individuals with lower education levels. Access to manufacturing occupations and white-collar positions in non-manufacturing industries appears to be a problem for women. Women are more likely to take manufacturing jobs when they become available near their homes. Commuting highlights the fact that Korean women are still extensively employed in female-dominated occupations, such as service and sales employees.

One important characteristic that has been studied is educational background. While the link between the level of education and commuting behaviour has been treated as statistical control before, Groot, De Groot & Veneri (2012) made use of micro data from The Netherlands to study the specific role of education on commuting behaviour. Higher educated employees commute farther, both in terms of distance and time, according to their findings. Furthermore, highly educated workers display specific patterns. They utilise public transportation and bicycles more frequently, are more likely to commute to agglomerated regions and places with high incomes and are more likely to dwell in and commute from areas with higher land rents. A proposed explanation for these findings is the relatively higher search frictions. Because highly educated people tend to do more specialised jobs, the chances of finding a good match near to where they reside are relatively low. While the difference in the mode of transportation could link back to the analysis of commuter preferences. The results also corroborate the 'commuting time paradox', which is the observation that the average travel time has not varied over time (Van Ommeren & Rietfeld, 2005). The study proposes a model with one fundamental conclusion, in the long term, the ratio of average commuting expenses to average salaries remains constant if labour market tightness, defined as the ratio of unemployment to vacancies, is maintained. The implication of this finding is that greater earnings encourage workers to tolerate higher commuting expenses.

On the other hand, many authors have focused on the urban structure side of commuting. Contrary to the educational findings in The Netherlands, it has been shown that there are significant and systematic differences in commuting time in the vast majority of the 20 largest U.S. metropolitan regions. Although commuting times usually decrease as employees' home locations approach the urban centre, inhabitants in low-income census block groups in the central city have lengthier commuting times, according to thematic maps and summary statistics (Shen, 1993). And going back to The Netherlands, Van Ham & Hooimijer (2008) discovered that employees in the Randstad have longer commutes and are less ready to relocate for employment than workers in the rest of the country when it comes to the effects of the labour market's geographical structure on commuting and migration. Regional disparities in employment access can explain a portion of this effect. Living in the Randstad has no influence on motivation to relocate in search of employment when access to employment possibilities is taken into account, showing that access to employment prospects reduces the need to relocate. After adjusting for employment access, the effect of residing in the Randstad on commuting remained. The polynucleated urban nature of the Randstad might explain some of this. A long commute, as a result of this structure, is particularly successful in boosting the number of jobs that can be accessible and, as a result, the number of openings that can be applied for. Polycentric labour markets provide significant flexibility to individuals prepared to endure a long commute, but monocentric labour markets provide restricted possibilities, increasing the likelihood of job-related migration. Congestion is another element that may contribute to the Randstad residents' lengthy journeys. Furthermore, it has been shown that commuting distance boosts daily and weekly labour supply by a little amount while having no effect on the number of workdays (Gutiérrez-i-Puigarnau & Van Ommeren, 2009).

Public policy is widely used to influence some of these factors. However, even if successful in improving the degree of jobs-housing balance, Giuliano & Small (1993) found that attempts to change the metropolitan-wide structure of urban land use through policy intervention are likely to have disappointing effects on commuting patterns. Such strategies do not address the primary causes of location pattern dispersion. Furthermore, the usual economic approach to analysing urban location, which uses the trade-off between land and commuting costs as the key driver of residential placement, fails to explain actual location patterns adequately. Nonetheless, the simplicity of the modelling approach does not deprive the results of all validity. Sohn (2004) analysed commuting data to identify the urban structure. The result has shown discrepancies between the spatial structure identified in the commuting pattern while the distribution of the employed residents is not. These findings followed an agglomerating trend from 1987 to 1995, showing that the distribution of employment is less dispersed while the opposite is true for employed agents.

3 Methodology

3.1 Data

The empirical part of this paper builds upon linked micro data from Centraal Bureau voor de Statistiek (CBS). The source for data on commuter behaviour is the 2011 to 2020 cross-sections of the Dutch Onderzoek Verplaatsingen in Nederland (OViN) and Onderweg in Nederland (ODiN) . As worker and job characteristics are not available through this research, we have used data from the Dutch Labour Force Survey (Enquête beroepsbevolking, EBB) from 2018 to 2020, which is available through the CBS StatsLine. Following the methodology used by Groot, De Groot & Veneti (2012), we have modified the data in the following manner. We deleted several observations to ensure that only workers with a sufficiently strong link to the job market are included. Workers must be between the ages of 18 and 65 and work a minimum of 12 hours each week. The degree of education (there are three levels), age, municipality of residence, gender, and whether a worker is working part-time or full-time are all characteristics of workers. The municipality in which the person works is provided for each employee. On the commuting of each worker, we have data on the (self-reported) mode of transport, average travel distance and travel time.

3.1.1 Commuting Distance and Time

Table 1 presents descriptive statistics regarding the variables that are related to commuting behaviour. We find that the average distance for a one-way trip in 2018 and 2019 was around 19,2 kilometers with a steep decrease to 16,6 kilometers in 2020. The same pattern is observed in the average travel time where we observe an average time of around 29,8 minutes in 2018-2019 and a decrease to 26,1 minutes in 2020. Given that we are looking at the average distance and travel times per trip for commuting agents, the data could suggest that the workers have either relocated in residence or have changed to closer jobs.

Transport Mode	**Distance** (*km*)			**Time** (*min*)		
Year	2018	2019	2020	2018	2019	2020
Auto (Driver)	25.05	25.22	22.81	30.76	30.88	27.83
Auto (Passenger)	22.74	25.41	18.59	30.47	31.99	26.82
Bicycle	4.74	4.85	4.43	18.99	19.34	18.65
Walking	2.47	2.91	1.79	11.89	13.97	12.21
Other	24.76	21.58	16.97	32.83	30.16	24.59
Private Transport	15.95	15.99	12.92	24.99	25.27	22.02
Train	40.75	41.16	38.40	66.75	67.07	65.66
Bus/Tram/Metro	15.32	15.21	14.29	42.96	42.42	43.99
Public Transport	28.04	28.19	26.35	54.86	54.75	54.83
Total	19.23	19.26	16.66	29.60	29.92	26.13

Table 1: Descriptive Statistics 2018-2020

We can also observe that, on average, public transport is used for longer commutes while private transport is used for shorter commutes. There exists high variability between modes of transportation in each category. Within public transport, trains are used for commutes that are almost three times the distance that agents cover with other modes of public transportation. However, the time spent on these trips is on average 50% longer when using the train than other modes of public transportation. This observation is in line with the aforementioned 'commuting time paradox', and the phenomenon is present across all other modes of transportation in different proportions.

In Figures 1 and 2, we can observe the average distance between the centre of the region of work and the centre of the work region. The municipalities inside the metropolitan areas have on average shorter distances to work than the municipalities outside of them. We see a decrease in the distances between 2014 and 2020 which is in line with the data shown in Table 1. We can observe a dispersion of the commuting distances over time, with an increase in the bigger agglomerations and a decrease in the previous red agglomerations. This data coincides with the negative net migration rate of inter-municipality movements reported by the CBS (2020).

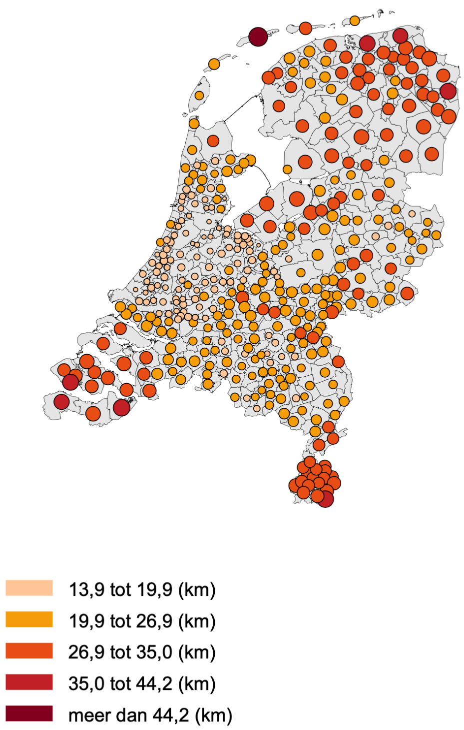

Figure 1: Distance to Job - 2014

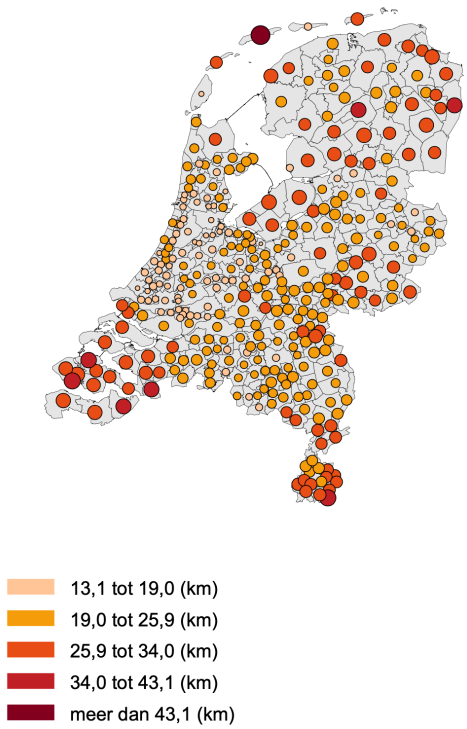

Figure 2: Distance to Job - 2020

3.1.2 Commuting Network

In Figure 3, we can better observe the commuting patterns of all municipalities in The Netherlands at the end of 2019. The doughnuts shown in the figure represent the work municipalities, and we can observe that it shows concordance with the distribution of travel distances shown in Figures 1 and 2. The relative size of the doughnuts represents the population of the municipality. The commuting network ignores international commuting, which could be present in municipalities close to the borders. The data shows that the network redistributed some of the commuting patterns from the urban agglomerations to the less densely populated municipalities surrounding them. The network shows the polycentric labour markets described in the literature (Van Ham & Hooimijer, 2008).

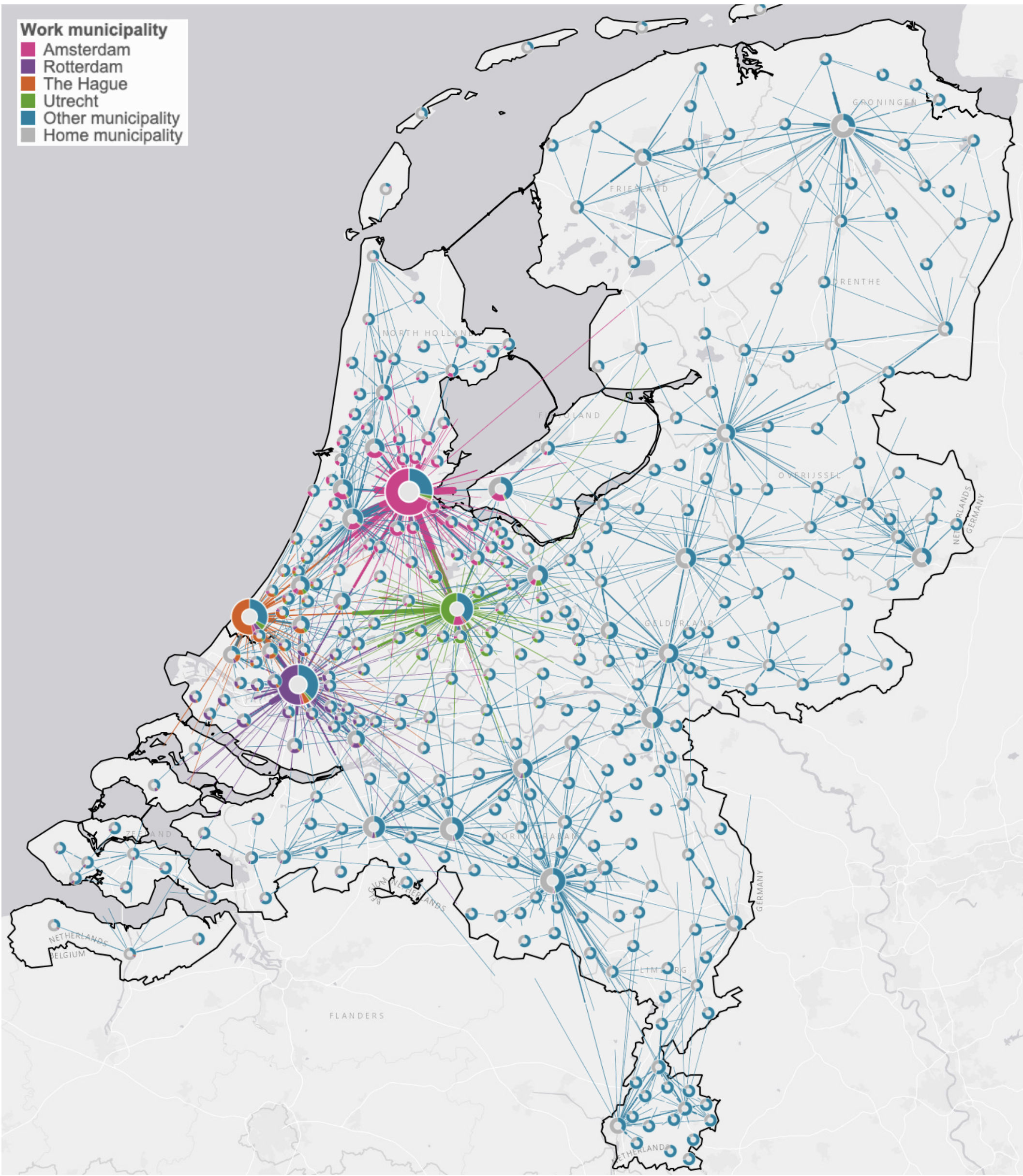

Figure 3: Commuting trends in Netherlands - 2019

In Figure 4, we can observe the trend of the different transportation methods over the last decade. We can see a clear downward trend in the use of public transportation in commuting, with an acceleration of the trend in 2020. Contrary to this, there used to be an upward trend in the use of private transport. The trend diverged between private methods of transportation. While bicycles (fiets) and walking trips continued to increase in distance for both sexes equally, the distance covered in cars as a driver (auto bestuurder) and as a passenger (auto passagier) had divergent upward trends, with women increasing faster. These observations are in line with the information presented in Table 1, where we observe an overall decrease in distance and time across the whole spectrum of transportation modes. In Figures 5 and 6, we can observe the discrepancies in the trends between males and females. The data coincides with the literature showing that men commute on average over longer distances. Due to a lack of data, there are some inconsistencies with the changes in 2020.

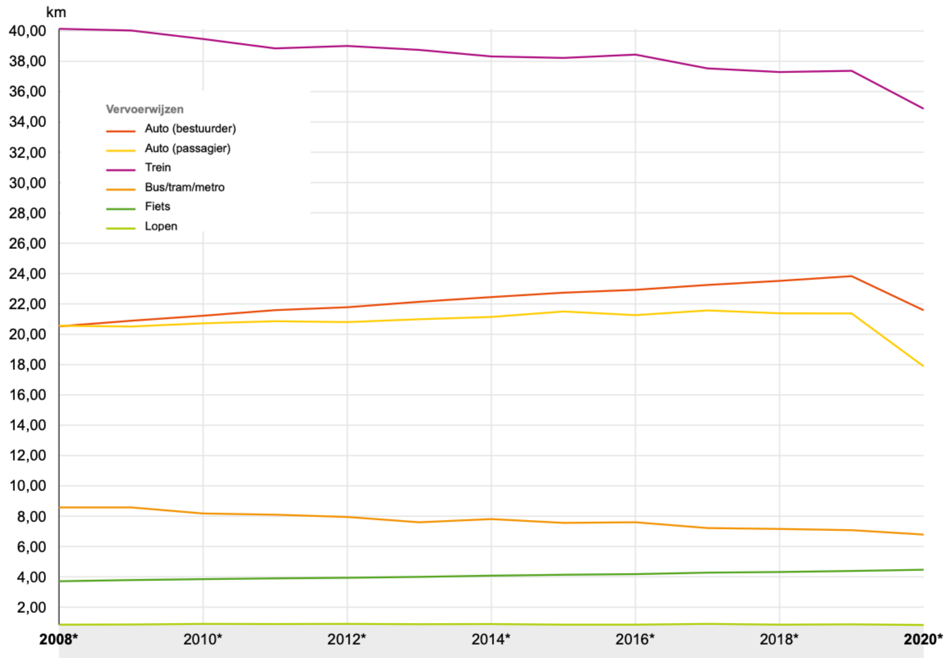

Figure 4: Mobility Trends

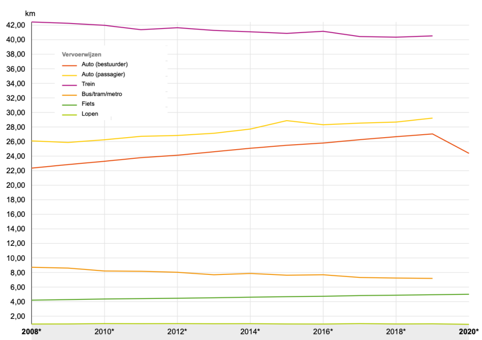

Figure 5: Mobility Trends - Women

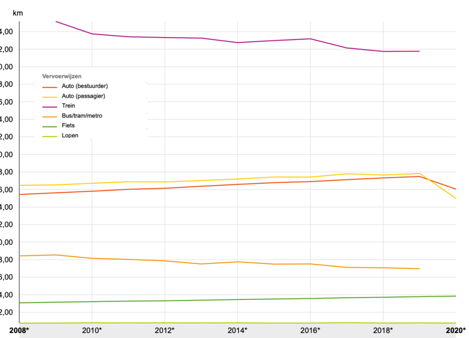

Figure 6: Mobility Trends - Men

3.2 Econometric Analysis

The analysis consists of the following stages. Firstly, we will estimate the gravity model for three instances: a general (1) model for all the data, a model for each year, and (2) a model for all data except the 2020 observations. Thereby analysing the evolution of distance decay between 2014 and 2020. Lastly, we will compare the predictions made by models to ascertain their efficiency and effectiveness in predicting inter-municipality migration in The Netherlands. Moreover, we examine if the commuting patterns and distance decay have significantly changed in 2020.

3.2.1 Gravity Model of Migration

The gravity model of migration is based on the empirical evidence that the commuting between two places i and j, with origin population m_i and destination population n_j, is proportional to the product of these populations and inversely proportional to a power law of the distance between them, r_{ij}. The intuition is simple: when you analyse the movement of agents between two places, the population size of each place acts as the mass of two objects in a vacuum. The object (city) produces a gravitational pull proportional to its mass (population) and the interaction between the two bodies' gravity will be inversely proportional to the second power of the distance between them. The farther away the bodies (cities) are, the weaker the interaction (commuters) is. We can immediately see the connection to migration such that migration is dependent on the distance between two places and the attractiveness of a destination is connected to its population size. The model is as follows:

$$G_{ij} = A\frac{m_i^{\alpha} n_j^{\beta}}{r_{ij}^{\gamma}} \tag{1}$$

where A is a normalisation factor and α, β and γ are the parameters of the model, which can be estimated by the following multiple regression analysis:

$$\ln G_{ij} = A + \alpha \ln m_i + \beta \ln n_j - \gamma \ln r_{ij} + \varepsilon_{ij} \tag{2}$$

As explained before, the model has certain limitations. First, it lacks a rigorous derivation for Equation (1). Second, although it is modelled after Newton's law of gravity, the model lacks theoretical guidance to define the parameters. Third, it predicts that the number of commuters increases without limit as the population of the destination increases. However, the number of commuters cannot exceed the origin's destination. Fourth, its deterministic nature does not leave room for fluctuations, as observed in the empirical data.

Even with these limitations, the gravity model is a particularly efficient tool for explaining patterns in observed gross migration flows. It connects subnational area demographics at the start of a period with later inward and outward migration. Therefore, it is not unexpected that while the gravity model is still being used to analyse internal migration movements, it has been increasingly employed to analyse international migration flows. While it is clear from the literature that econometric challenges persist once researchers use more complex versions of the model that incorporate systemic and dynamic effects along with numerous push and pull factors, spatial spillovers, and other factors, it is surprising how little is known about the potential application of the model to a multi-regional population projections methodology. The gravity model, despite being idealised and simplistic, fits internal migration data extremely well, frequently producing modified R-squared values between 0.8 and 0.9.

3.2.2 Distance Decay

Geographically speaking, the word "distance decay" refers to the impact of distance on social or physical relationships. According to the distance decay effect, interaction between two locations decreases as their distance grows. The interaction between the two locations starts to lessen once the distance is outside of the activity space. Thus, it is said that many geographic events may be explained mathematically by the inverse square rule of physics, which is one of the ways that physical concepts like gravity are frequently figuratively linked to geographical circumstances. In Figure 7, we can observe the distance decay effect in The Netherlands. The functional relation between the two variables is not linear but rather displays a power or exponential form.

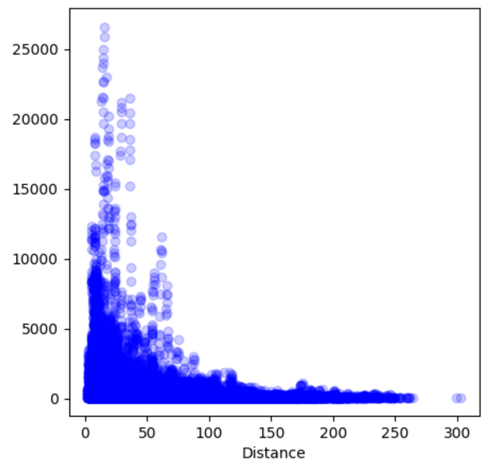

Figure 7: Distance Decay - Empirical Results

The distance-decay function of the geographical gravity model was initially an inverse power law, which suggests a scaling process in spatial interaction. However, the distance exponent of the model cannot be explained in terms of Euclidean geometry. The result is the dimension problem, a scenario. Particularly, it was unable to build the gravity model based on the power law using ordinary techniques and general concepts. In order to represent the distance-decay function for the gravity model, the inverse power function was altered to a negative exponential function. The exponential-based gravity model ignores the first law of geography, which states that “everything is related to everything else, but near things are more related than distant things” (Waldo, 1970).

As mentioned before, it is noticeable that the relation in our empirical data is not directly linear. To solve this problem, and already presented above, the natural logarithms of each value are used in the regression. The resulting relation presents itself with a more simple interpretation than the logarithmic relation that the original values suggest. Chen (2015) has established that the gravity model’s distance exponent resembles a fractal dimension. Thus, utilising ideas from fractal geometry, the dimensional conundrum of the power-based gravity model may be overcome. The exponential function also suggests location and localisation, which goes against the fundamental idea of spatial interaction. The power function suggests distant activity, which is a prerequisite for gravitational pull on earth. Therefore, a power-function is a better fit than an exponential-function for the gravity model. Finally, by adopting a suitable postulate, the gravity model based on power law decay may be derived from the entropy maximising principle.

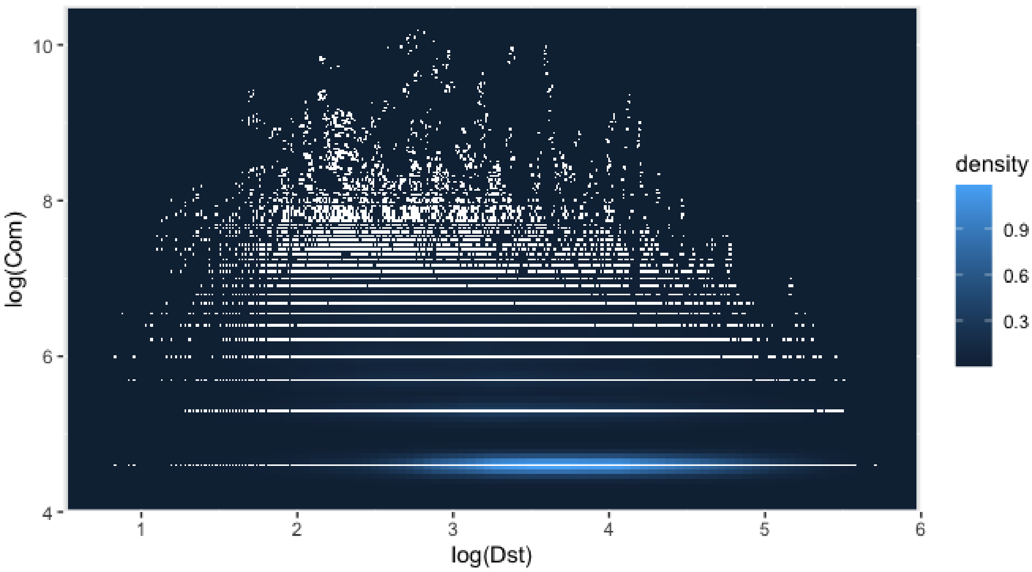

Figure 8: Distance Decay - Natural Logarithm

4 Results

The estimation results are presented in Table 2. As expected, distance has a strong, negative effect on the number of commuters between Dutch municipalities with a downward trend over the years and a deceleration of the trend during 2020. Also as expected are the effects of the populations, the origin's population exhibits a stronger effect than the destination. These results are in line with the literature, we expected the destination's population to have a stronger effect due to the attractive consequences of agglomerations. The two main factors influencing commuting patterns are the desirability of locations as places to live (due to amenities) and as places to work (due to increased productivity and earnings). Land rentals will rise and the net inflow of workers will decline as a location becomes, ceteris paribus, more desirable as a place to live. Land rents will increase as well if a place becomes more desirable to work in, but because of the trade-off between commuting and migration, the net inflow of employees will also increase.

While all of the variables in the models are statistically significant at $\alpha = 0.01$, we find that the R-squared is moderate at best, $R^2 = [0.497, 0.519]$. Each model predicts close to half of the commuting agents. In Figure 9, we can observe that there exists a relatively low number of commuters for many different trajectories and a high number of commuters in specific trajectories. These are the statistical results of the information displayed in Figures 1, 2 and 3. The regression results show the negative-power function of distance decay, $\delta = [-0.930, -1.022]$. This follows the expected result of a negative distance decay as delineated by the literature.

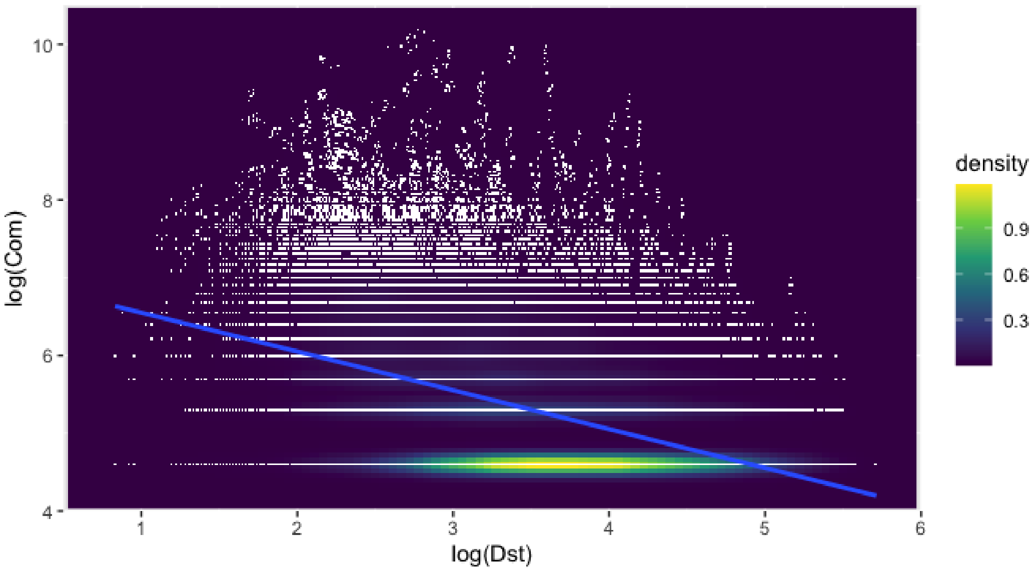

Figure 9: OLS Results - General Model

	Dependent variable:									
	(1)	2014	2015	2016	2017	2018	2019	2020	(2)	RLM
Destination Pop	0.559***	0.541***	0.543***	0.548***	0.563***	0.568***	0.577***	0.571***	0.556***	0.538***
	(0.003)	(0.007)	(0.007)	(0.007)	(0.007)	(0.007)	(0.007)	(0.007)	(0.003)	(004)
Distance	-0.981***	-0.930***	-0.939***	-0.962***	-1.007***	-0.993***	-1.019***	-1.022***	-0.974***	-0.964***
	(0.004)	(0.009)	(0.009)	(0.010)	(0.010)	(0.010)	(0.010)	(0.010)	(0.004)	(0.003)
Origin Pop	0.392***	0.390***	0.381***	0.382***	0.392***	0.398***	0.404***	0.396***	0.391***	0.373***
	(0.003)	(0.007)	(0.007)	(0.007)	(0.007)	(0.007)	(0.007)	(0.007)	(0.003)	(0.003)
const	-1.671***	-1.640***	-1.524***	-1.515***	-1.630***	-1.791***	-1.858***	-1.713***	-1.662***	-1.400***
	(0.040)	(0.103)	(0.104)	(0.104)	(0.103)	(0.106)	(0.112)	(0.112)	(0.043)	(0.041)
Observations	79,427	11,373	11,316	11,363	11,456	11,543	11,222	11,154	68,273	79427
R^2	0.509	0.498	0.502	0.504	0.515	0.511	0.519	0.519	0.508	-
Adjusted R^2	0.509	0.497	0.502	0.503	0.515	0.510	0.519	0.519	0.508	-
Residual Std. Error	0.664	0.656	0.656	0.662	0.661	0.668	0.672	0.670	0.663	
F Statistic	27495.553***	3752.906***	3801.757***	3840.687***	4061.099***	4012.314***	4038.032***	4012.458***	23483.397***	-
Notes:	*p<0.1; **p<0.05; ***p<0.01									

Table 2: OLS Estimation Results - Gravity Models

We cannot simply analyse individual bivariate plots to identify associations since we are performing a multivariate regression. Instead, we wish to examine how the dependent variable and independent variables relate to one another when reliant on another independent variable. For this purpose, we will use four different plots. First we display the data and the estimates evaluated on them. Second, we plot the residuals of the regressor that we want to observe. Third, we make use of the partial regression plot, sometimes referred to as additional variable plot, where we show the errors from the estimation results for one regressor and the regressand. At last, the CCPR plot offers a mechanism to assess the impact of one regressor on the response variable by considering the impacts of the other independent variables. $Residuals + \beta_i X_i$ against XI is the formula for the partial residuals graphic. To demonstrate where the fitted line would be, the component adds $\beta_i X_i$ against X_i. If X_i has a strong correlation with any of the other independent variables, caution should be used. If this is the case, the variation that is seen in the plot is likely to be lower than the actual variance.

The results presented in Figure 10 make use of the previously outlined plots and are regarding the general model (1) that pools all the data over the years. Interestingly, the results of all models follow the same trend and distribution, therefore they can be interpreted in one model and be extrapolated to the others. This similarity between the results was expected due to the nature of the commuting process. Nonetheless, the data does not present significant deviations year over year even when accounting for the COVID-19 pandemic. In the top left figure we observe the plot of the original data against the predicted values with the black line delineating the 95% confidence interval. The outcome is as the R-squared already suggested, there is a high level of accuracy but still lacks a complete overlap with the data. It is worth keeping in mind that the results are given as natural logarithm, not as the initial units, hence why we see a linear relation that is in reality a negative-power relation as shown in 7. The top right picture shows that there is a linear relation on the error terms and that they are homoskedastic which suggests that the Gauss-Markov assumptions hold. Next, we observe the partial regression plot where there exists a zone of densely packed observations around zero and more dispersion as the values increase.

In summary, results tell the same story as Table 2 but deconstructed by variable. We can observe the same relations shown in the table with the added value of observing the distribution of the results. We can observe the same segmentation that was present in Figure 8, however we also observe a second segmentation over the population of the origin and destination locations.

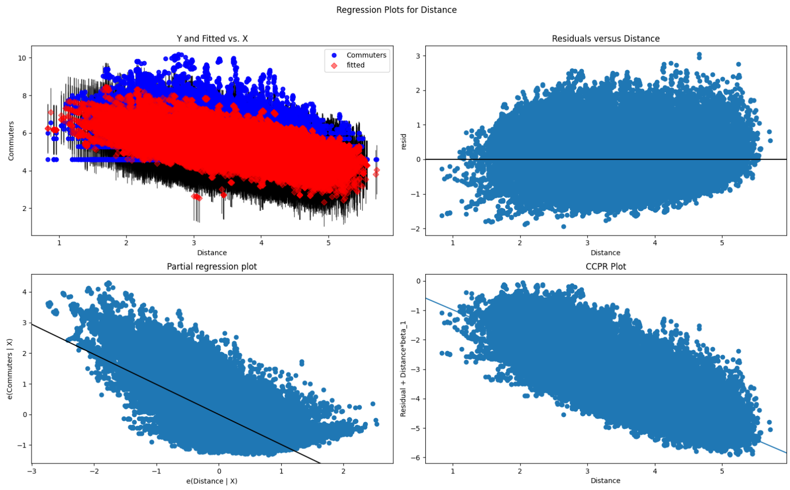

Figure 10: Error Grid - Gravity Model (1) - Distance

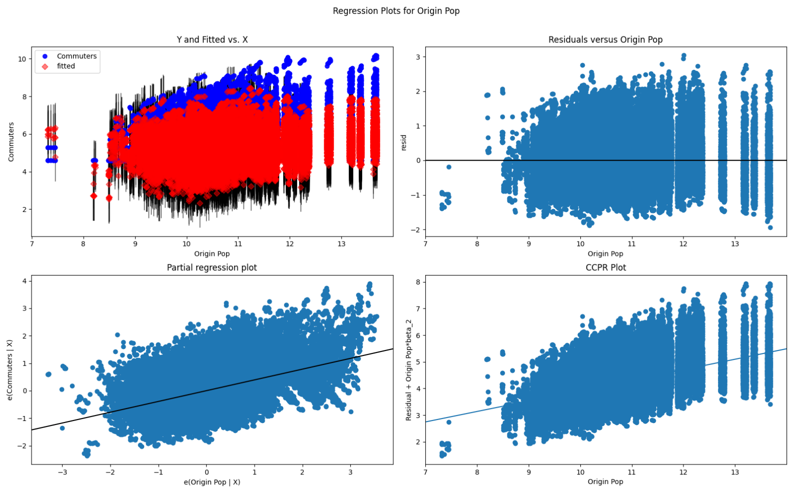

Figure 11: Error Grid - Gravity Model (1) - Origin Population

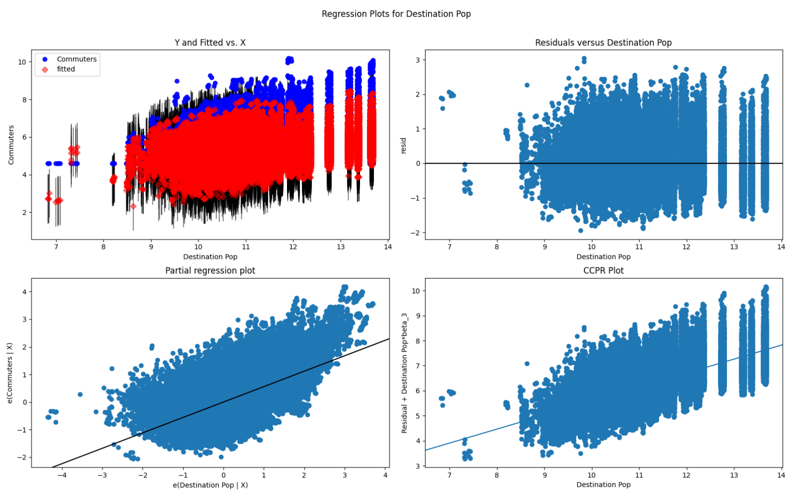

Figure 12: Error Grid - Gravity Model (1) - Destination Population

Following the results from the model estimated with heteroskedasticity robust (HC1) standard errors, we plotted the residuals in Figure 13. A method for spotting non-linearity is the Residuals vs Fitted graph. As evidenced by the graph, there seems to be a non-linear relation in the error terms that is segmented by groups. Next, to visually determine if residuals are regularly distributed, utilise the QQ-plot, and it shows that the distribution is relatively normal. Moreover, the residuals' homoskedasticity is examined using the scale-location plot; the same pattern as in the first graph is seen which would suggest a high level of heteroskedasticity with segmentation. However, further testing and Robust Linear Regression confirm that the errors are not heteroskedastic. At last, the residuals vs leverage tests for observations outside the Cook's distance curves because these points are thought to be important observations that have the potential to influence the fit. As observed in this last graph, we do not observe any relevant outlier that might be disturbing the estimation results.

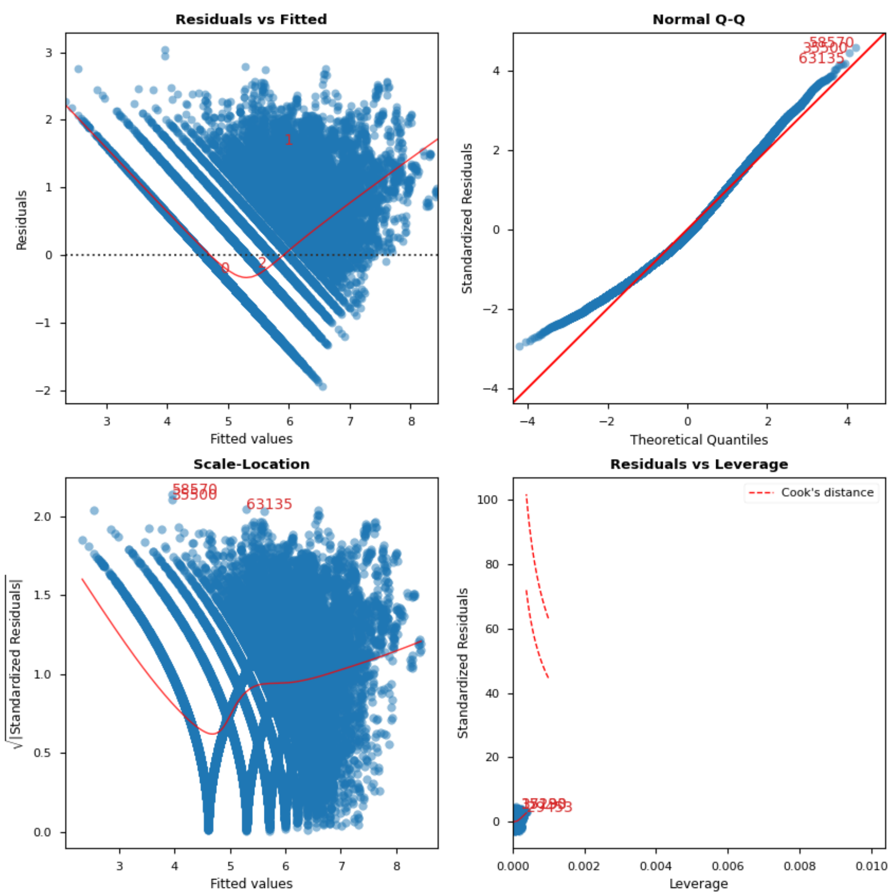

Figure 13: General Diagnostics - Gravity Model (1)

5 Discussion

Let us now turn to a discussion of the reported results, with special emphasis on the methodological caveats and the implications of the results. First, we have observed a decrease in the distance decay over time which suggests that in the time period from 2014 to 2020, the agents have chosen to work in closer locations to their place of residence or that they have relocated to be closer to work. Second, we showed that there was a slight disturbance in the constant trend on the time and distance covered per trip in the year 2020, which suggest the effect of the COVID-19 pandemic on commuting behaviour. These two pieces of information help form a clear story. In the aggregate, the commuting time paradox has held but there appears to be a slight decrease in the distance between an agent's job location and residence location. This is also corroborated by the dispersion phenomena observed in Figures 1 and 2. However, more research is needed to ascertain a permanent effect due to the lack of post-pandemic data. The evolution of the commuting dynamics appears to have been accelerated by the pandemic but there is not enough evidence yet for a conclusion to be drawn.

The trend in the choice of transport has been a constant decline in the use of public transport and an increase of private methods of transport. The information is in line with the aforementioned story, the use of long-distance (train) transport has declined over time which suggests that agents transport themselves over shorter distances. However, there has been a clear increase in the use of the car as a medium-distance transport, which would be connected with the increase of distance decay, thus more negative values in the γ coefficient from Equation (2). The perturbation observed during 2020 in the use of public transport might not have historical significance since the density of the choice was unattractive during the pandemic but might have reversed to previous numbers when the pandemic subsided. More research on these areas would shed light on the long term effects of the choice of public transport. This topic could be of special interest to policy makers due to the desire to have a higher portion of the population making use of public transport in order to optimise the flow of people and decrease the marginal external costs of time lost in transportation. Another line that could be pursued is the connection between road pricing and the downward trend on the use of public transportation combined with the upward trend in the use of cars. Therefore, a lower-than-optimal road pricing could be a reason for the trends observed in the choice of transportation, thus the evolution of road pricing could be integrated in the estimation to observe if it has an impact on the distance decay while also observing for possible endogeneity in the distance and road-pricing variable which could be solved through an Instrumental Variable (IV) approach.

On the analysis of the evolution of distance decay, we can observe that the results are not as significant as other studies have suggested, with our R-squared indicating a moderate fit. The literature has shown that the fit of the model is highly dependent on the spatial structure on which it is applied with special efficacy on the analysis of the distance decay between densely populated areas. This would suggest that an analysis on the Dutch commuting patterns where we group by population density or size of the urban agglomeration could improve the fit of the model and provide more accurate estimates to the policy maker. Moreover, a careful handling of the data where spatial autocorrelation is accounted for might improve the results and resolve some of the issues encountered. This phenomenon of spatial autocorrelation considers not only the geographic distribution of the agents and jobs but also the network of flows which is of special interest in commuting behaviour. Thereby, a better analysis could be made to improve the fit of the model and the accuracy of the estimate. On another note, the use of an alternative model - such as the radiation model - could provide a good start for a comparative analysis of the different population densities. An extension of the study could also provide an analysis of the network by making use of other stochastic network models, such as stochastic and time-dependent transport networks. Moreover, the time frame could be expanded to observe a more accurate depiction of the commuting behaviour of The Netherlands.

One connection that has been so far overlooked in this study is the dependence on the labour supply and the housing market. As mentioned in the literature review, other studies have incorporated the estimation of land rents, which could be a useful addition when accounting for the exodus seen in the metropolitan areas of the Netherlands. With the increasing prices on the Dutch housing market, this approach could also be improved by segmenting the groups by agent's characteristics, with education and income taking special relevance. Furthermore, the choice on residential location is altered by the constant merge of municipalities that has been seen in the last two decades. These events might increase the discrepancies between the chronological data and could disturb the results of the distance decay. If municipalities keep disappearing and being incorporated in other, bigger municipalities then the trend on commuting could be effectively muted.

Moving on to the methodological caveats, there exist a few points that must be discussed. First, there appears to be a non-linear relation in the error terms and heteroskedasticity as observed in Figure 13, however the results from the Robust Linear Model in Table 2 show results that are not statistically different from the ones offered by the OLS model. It appears that the observed segmentation is due to the existence of many commuters regardless of the distance. Therefore, the line patterns observed in the estimation results of Figure 13 are the corresponding patterns observed in commuting patterns from Figure 2 and in the population patterns from Figures 11 and 12. These observations suggest that the spatial autocorrelation previously suggested could be an important avenue to improve the model and the analysis. Second, a time-series analysis could improve the estimation results due to the nature of the data.

6 Conclusion

This paper describes our research on the evolution of commuting patterns in The Netherlands. Making use of the spatial data from 2014 to 2020, the paper ought to shed light on the evolution of the distance decay and the aggregate commuting behaviour between Dutch municipalities and describe the job-residence distribution. First, the data clearly show a trend in the choice of transportation while maintaining time constant, with a slight increase in the distances. These results connect back to the 'commuting time paradox' previously observed. Furthermore, there appears to be a dispersion of the residence location when controlling for job opportunities, which tend to agglomerate in densely populated urban areas. Moreover, the results suggest a clear trend on the evolution of the commuting behaviour, with no significant deviation when accounting for the COVID-19 pandemic.

Second, the empirical analysis making use of the gravity model to estimate the distance decay over time provided moderate R-squared values, but the results are a starting point for a further, more in-depth analysis of the spatial distribution of the labour force and the labour in The Netherlands. Within the limitation, the models provide a clear outline of the changes that have been seen over the past decade. The phenomenon of dispersion goes hand in hand with the increasing distance decay from the model. The estimation results also presented a challenge due the strong segmentation in commuter flows. This segmentation initially led to a possible non-linear relation in the error terms and heteroskedasticity. However, a Robust Linear Regression confirmed that these were the result of the segmentation.

Lastly, many caveats need to be further researched, therefore the results must be carefully interpreted and not overstated. The main highlights are the corroboration of the patterns observed in other research and the description of the dispersion that has taken place when controlling for job availability. The possible spatial autocorrelation of the data and the expansion to model the commuting network could be the next logical steps. The manifold insights that have been extracted provide a clear answer, commuting lies at the intersection of a complex system that overlaps between the labour market and the agent's choice of residential location.

The use of proper statistical and econometric methods is imperative in the analysis of spatial data. This paper has provided a brief introduction to the analysis of spatial decay regarding commuting behaviour in The Netherlands. Thereby, leaving the use of advanced panel data models to account for the time and spatial lags yet to come. While simplistic at best, the results are the first step in incorporating the heterogeneity and interdependence endemic of spatial data.

7 References

- Burbidge, S. K, Goulias, K. G, & Kim, T. (2006). Travel Behavior Comparisons of Active Living and Inactive Living Lifestyles. UC Berkeley: University of California Transportation Center. Retrieved from https://escholarship.org/uc/item/4j9602x6

- Chen, Y. (2015). The distance-decay function of geographical gravity model: Power law or exponential law? https://doi.org/10.48550/ARXIV.1503.02915

- Dieleman, F. M., Dijst, M., & Burghouwt, G. (2002). Urban Form and Travel Behaviour: Micro-level Household Attributes and Residential Context. Urban Studies, 39(3), 507–527. https://doi.org/10.1080/00420980220112801

- Dubin, R. (1991). Commuting Patterns and Firm Decentralization. Land Economics, 67(1), 15–29. https://doi.org/10.2307/3146482

- Giuliano, G. (1998) Information technology, work patterns and intrametropolitan location: a case study, Urban Studies, 35, pp. 1077-1095.

- Giuliano, G., & Small, K. A. (1993). Is the Journey to Work Explained by Urban Structure? Urban Studies, 30(9), 1485–1500. https://doi.org/10.1080/00420989320081461

- Gutiérrez-i-Puigarnau E, Van Ommeren, J.N. (2010) Labour supply and commuting. Journal of Urban Economics. 2010;68(1):82-89.

- Groot, S. P. T., de Groot, H. L. F., & Veneri, P. (2012). The Educational Bias in Commuting Patterns: Micro-Evidence for the Netherlands. (TI Discussion Paper; No. 12-080/3). Tinbergen Institute. http://www.tinbergen.nl/discussionpapers/12080.pdf

- Lee, B. S., & McDonald, J. F. (2003). Determinants of Commuting Time and Distance for Seoul Residents: The Impact of Family Status on the Commuting of Women. Urban Studies, 40(7), 1283–1302. http://www.jstor.org/stable/43084316

- Masucci, A. P., Serras, J., Johansson, A., & Batty, M. (2013). Gravity versus radiation models: On the importance of scale and heterogeneity in commuting flows. Physical Review E, 88(2), 022812. https://doi.org/10.1103/PhysRevE.88.022812

- Ory, D. T., Mokhtarian, P. L., Redmond, L. S., Salomon, I., Collantes, G. O., & Choo, S. (2004). When is commuting desirable to the individual? Growth and Change, 35(3), 334–359. https://doi.org/10.1111/j.1468-2257.2004.00252.x

- Poot, J., Alimi, O., Cameron, M.P., & Maré, D.C. (2016) The Gravity Model of Migration: The Successful Comeback of an Ageing Superstar in Regional Science. Social Science Research Network; 2016.

- Quality of Transport report (2015) (Report). European Commission. December 2014. p. 11. Archived (PDF) from the original on 8 September 2015. Retrieved 29 December 2016.

- Ravenstein, E.G. (1885, 1889) The laws of migration, Part 1 and Part 2. Journal of the Royal Statistical Society, 48(2): 167-235 and 52(2): 241-305.

- Shen, Q. (2000). Spatial and social dimensions of commuting. Journal of the American Planning Association, 66(1), 68–82. https://doi.org/10.1080/01944360008976085

- Simini, F., González, M. C., Maritan, A., & Barabási, A.-L. (2012). A universal model for mobility and migration patterns. Nature, 484(7392), 96–100. https://doi.org/10.1038/nature10856

- Sohn J. (2005) Are commuting patterns a good indicator of urban spatial structure? Journal of Transport Geography. 2005;13(4):306-317.

- Stewart, Q.J. (1950) The development of social physics. American Journal of Physics 18: 239-253.

- Van Ham, M., & Hooimeijer, P. (2009). Regional differences in spatial flexibility: Long commutes and job related migration intentions in the netherlands. Applied Spatial Analysis and Policy, 2(2), 129–146. https://doi.org/10.1007/s12061-008-9016-2

- Van Ommeren, J., & Rietveld, P. (2005) The commuting time paradox. Journal of Urban Economics. 2005;58(3):437-454.

- Tobler W., (1970) "A computer movie simulating urban growth in the Detroit region". Economic Geography, 46(Supplement): 234–240.

- White, M. J. (1986). Sex Differences in Urban Commuting Patterns. The American Economic Review, 76(2), 368–372. http://www.jstor.org/stable/1818798
- Zipf, G.K. (1946) The P1 P2/D hypothesis: On the intercity movement of persons. American Sociological Review, vol. 11, Oct.