Regional income differences elucidated by facilities:

A case study for the Netherlands

Master Thesis Utrecht School of Economics

Name	Gerrald Frens
Master programme	Economic Policy
Student number	6926592
Supervisors	Dr. Bas Machielsen,
	Dr. Yolanda Grift
Date	30th of June 2023

Acknowledgement

I am very delighted that Bas Machielsen supervised me during the writing-process. I am grateful for the topical conversations we had and his guidance was really helpful to me. Also Yolanda Grift, my second supervisor, gave me useful feedback and helped me improving my work. Special thanks to my family, who supported me during the four years I attended U.S.E.

JEL classifications D31, E64, R12, R23, R53

Keywords Income determinants, regional inequality, facilities, (internal) migration

Abstract

In the Netherlands, there are differences in average income levels and in the provision of facilities between neighbourhoods. This paper examines its relationship by using CBS StatLine data for the period 2016-2021. It finds that the average distance to facilities negatively affects future average income levels on the neighbourhood level via the channel of migration; as the average distance to facilities declines, the net migration rate rises. This causes average incomes to increase, indicating that rich people who can afford it to move, move to those neighbourhoods. In addition, it finds that the provision of any facility to a neighbourhood is more important than the provision of many facilities. Also, a retail facility and mobility with a car is the most vital for a neighbourhood and attracts high income migrants.

Table of content

1. Introduction

The Netherlands is only a small country. Many might think that regional differences in the Netherlands are negligible, as facilities are relatively nearby for everyone and every region enjoys equal opportunities. However, these days, the gap in society between urban and rural areas is a hot topic. Regional differences seem to play a role in the Netherlands. Political parties of all kinds promise in election campaigns to fight the regional inequalities and emphasize that societal discontent is to a large extent explained by rural areas having a disadvantage on urban areas (CDA, 2023; PvdA, 2023). For example, there are many more bus stops in big cities than in rural villages (Bastiaanssen & Breedijk, 2022). Such differences look innocent and even a bit charming in first instance, but they might be exemplary for deeper structural problems that the countryside faces. This paper aims to elucidate regional income differences in the Netherlands by the accessibility of facilities. It finds a causal relationship between the accessibility of facilities in a neighbourhood and the regional average future income level per income recipient and estimates its magnitude. Furthermore, it reveals its cause and scrutinizes which facilities are the most important.

In 2021, Josse de Voogd and René Cuperus published their 'Atlas van Afgehaakt Nederland', a study on societal differences in the Netherlands. According to their publication, regional inequality is an important element why people drop out of society. This is reflected by their voting preferences (De Voogd & Cuperus, 2021). The past decades, the situation has only worsened for the inhabitants of the rural areas and the gap between urban and rural regions has widened. For efficiency-considerations, the number of facilities all over the country fell since the 1980s, but the decrease was more prominent and visible in the rural areas. For instance, the number of police offices in Zeeland (mainly a rural province) fell with no less than 74 percent, whereas the number of police offices in South-Holland (mainly an urban province) only decreased with 27 percent over the period 1980-2021. Also, numerous facilities like libraries and primary schools merged. As a result, the average distances increased. Even entire municipalities merged, along with their town halls, undermining the ideal of a visible and accessible government. In the regions where the public facilities deteriorated the most and average distances increased the most, the aversion to the government is the strongest and right-wing populist parties[1] flourish (Van de Ven, 2021).

Also the financial gap between urban and rural areas has increased over the past decades. According to Caspar van den Berg and Bram van Vulpen, inequality between the European Union (EU) countries has decreased since the EU has been established. However, regional inequality within countries has increased at the same time. As a result of the globalization caused by the unification of EU countries, the urban growth regions in the EU benefited, whereas the rural regions were harmed by this process. This entails other negative implications for the society as a whole, such as a decrease in social cohesion (Van den Berg & Van Vulpen, 2019).

Figure 1 shows the average standardised income per person per Dutch municipality in 2020. Clearly, we observe regional income differences in the Netherlands and the urban and rural municipalities can be defined easily. The average income levels of many western municipalities and some southern municipalities, making up the Randstad (the highly urbanized region in the provinces Utrecht, North-Holland and South-Holland) and BrainPort (the technological region in North-Brabant), is higher than the average income in the other regions. The most obvious regions regarding lower income levels are the three northern provinces: Groningen, Friesland and Drenthe. On the other hand, the provinces Utrecht, North-Holland and South-Holland are outperforming the other regions in average standardised income levels. Gelderland and North-Brabant are in the middle of the income spectrum.

In short, we observe two trends, namely (1) deteriorating facilities in the rural areas and (2) an increasing income gap between regions, with urban areas having an advantage over rural areas. This paper aims to investigate its relationship. Although a lot of attention has been attributed to regional differences in the Netherlands and many reports on this topic have been written (see for instance: Van den Berg & Kok, 2021; Weterings, Buitelaar, & Edzes, 2019) the relationship between facilities and income in a broad sense is not yet investigated for the Dutch case.

[1] The author mentions PVV, FvD and BBB.

This paper proceeds in the following way. Section 2 discusses the relevant literature in this field and summarizes the main takeaways from the existing literature. Section 3 outlines the research methodology and describes the data that is used to estimate the effect. The results can be found in section 4. Section 5 checks the robustness of the models, whereas section 6 elaborates on the models. Section 7 lists the policy implications that the results of the analysis have and section 8 points at the limitations of the research. Section 9 concludes. Tables, figures and references can be found in the appendices.

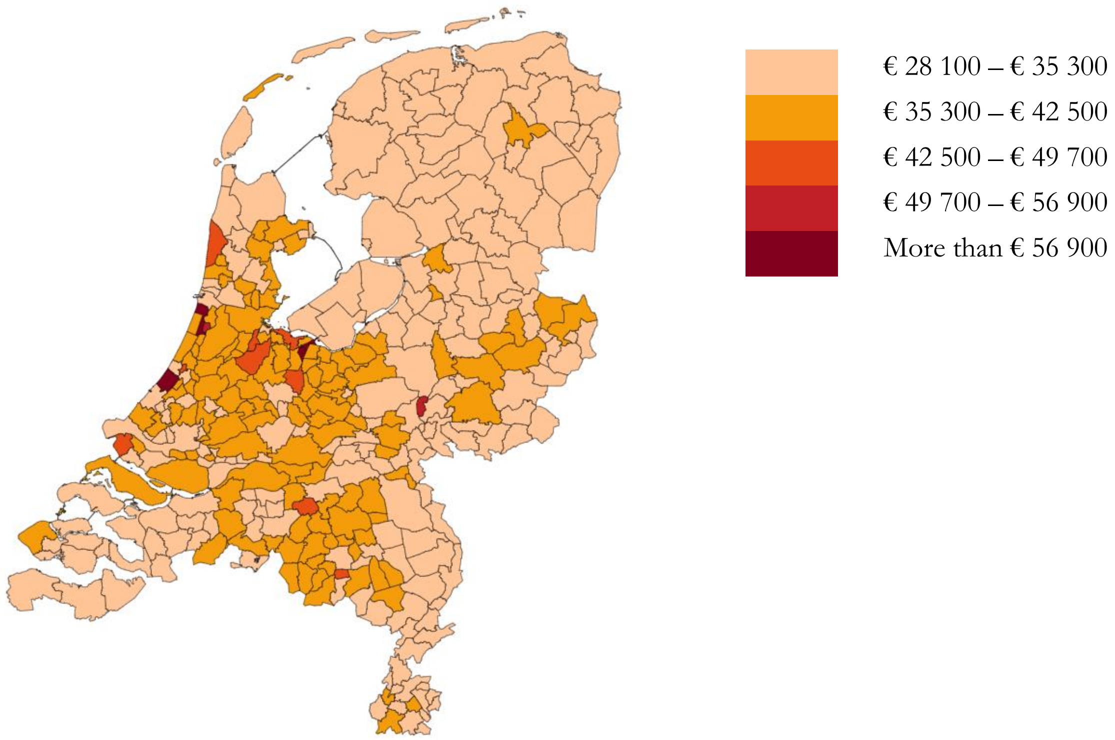


Figure 1: Standardised average income per municipality, 2020. Source: CBS StatLine

2. Literature review

The fact that neighbourhoods with a low average distance to facilities have on average higher income levels can be explained by two reasons. Either high income earners move to neighbourhoods with many facilities, or the provision of facilities in those neighbourhoods lead to an increase in the incomes of the initial inhabitants. Literature touching upon the effect of regional characteristics and the prevalence of facilities on incomes on the neighbourhood level can therefore be clustered into two groups. Firstly, a part of the literature indicates that high skilled people and high income earners concentrate around places with good facilities and migrate to those neighbourhoods. This causes an indirect effect of facilities on income, namely through migration. Secondly, a stream in the literature aims to prove the direct effect, namely, that the existence of decent facilities increases average income in a region.

2.1 High income earners move to facilities

Tolsma, et al. (2009) study the effects of different neighbourhood characteristics, such as the proximity to facilities, on social cohesion by making use of Dutch survey data from 2004. Theoretically, the main reason that individuals with a low income live in a neighbourhood where facilities are scarce is that they cannot afford it to move to places where facilities are available and nearby. For people with a high income who rather live in cities or villages with better facilities, it is a feasible option to move out. Also Arntz (2010) confirms this and adds that high-skilled people have the incentive to move to places where their human capital investments get rewarded. The existence of facilities in a broad definition – schools, infrastructure, safe environment, etc. - is a necessary condition for social cohesion to originate, so the poorer neighbourhoods are the ones where social cohesion lacks, deteriorating living conditions, as the absence of social cohesion has in turn a negative impact on the neighbourhood, urging high income earners to move to better neighbourhoods. This creates a vicious circle for those neighbourhoods. The main finding of their research is that average neighbourhood income levels appear to be the most important and decisive factor for social cohesion to arise; the effect of average neighbourhood income on trust levels and participation in volunteering is positive and significant. So, mean income is positively correlated with social cohesion. Following this line of reasoning, policymakers should target mean income when they aim to increase social cohesion. Since our paper finds that lowering the average distance to facilities is a way of targeting mean income, it is expected that social cohesion increases as a positive externality.

By using survey results amongst ICT workers in the Netherlands, Van Oort, et al. (2003) find that high income knowledge workers prefer proximity to facilities over proximity to work when making residential choices. Evidence from Germany shows that high skilled people and high income earners tend to move to the western and richer part of Germany. This is mainly due to economic conditions and better wages that are available in western Germany. However, also the availability of facilities may contribute to this internal migration (Arntz, 2010). Research from Belgium shows that proximity to facilities significantly affects internal migration. People who can afford it, tend to move to places with better facilities. This pattern is also visible in the United States, where researchers found evidence for 'reverse commuting'; high income earners prefer to live in cities, with facilities nearby, while they are working outside the city centres (Glaeser, Kolko, & Saiz, 2001). In certain cities, the facilities are better available in the suburbs, rather than in the city centres, for instance in Detroit. In these cities, the richer people are more likely to live in the suburbs (Brueckner, Thisse, & Zenou, 1999). This suggests that high income earners decide to live near facilities, regardless where they work.

Not only natives with a high income level converge around regions with decent facilities, also foreigners move there. Beckers and Boschman (2019) assess the inflow of foreign highly skilled workers with a high income to the Netherlands over the period 2000 to 2009 and in particularly the residential neighbourhood choices that these highly skilled workers make. They find that there is an over-representation of foreign highly skilled workers in city centres. Moreover, the incidence of these skilled migrants in a certain neighbourhood correlates with the number of catering sale points (e.g. restaurants) in that neighbourhood. Also, foreign highly skilled workers tend to live closer to international schools. Further, they state that households with a higher income or with a higher education level are more successful on the housing market. Therefore, they are better able to move to attractive neighbourhoods, i.e. safe and accessible neighbourhoods with decent facilities nearby. Foreign knowledge workers who migrate to the Netherlands are more likely to move to cities in the Randstad or BrainPort, mainly because of the facilities that are available there (Raspe, et al., 2014). Hunt and Mueller (2004) find a similar effect in the United States and Canada: high skilled migrants attach a higher value to facilities in their proximity than less skilled labourers.

2.2 Facilities generate high income earners

Sá, et al. (2006) scrutinize the effect of proximity to post-secondary education – being an important facility – on the likelihood of receiving post-secondary education in the Netherlands, using survey and spatial data from 1998-2000. They find that high school graduates living outside of the Randstad are more likely to attend a professional college (hbo) instead of a university, ceteris paribus, because people in the Randstad generally live closer to a university. More importantly, high school graduates' decision whether to continue their education or not is greatly influenced by the post-secondary institution's geographic accessibility. This implies that the proximity to a post-secondary education institution affects future income levels.

When it comes to primary education, children generally attend a primary school that is in their direct neighbourhood. Better performing schools can be found in wealthier neighbourhoods, so school segregation between neighbourhoods can be explained by income differences between neighbourhoods. This seems obvious, but once again allows the gap in society to persist, since wealthier families receive education of higher quality, leading to better labour market perspectives. This effect is found within Dutch cities, including smaller regional cities. This suggests that not only the proximity to a facility matters, but also the quality of the facility. Thus, neighbourhoods with a lower average income level do not only have a higher average proximity to a primary school, but also a lower quality of their primary education. Geographical factors thus have large explanatory power in the impact the school system has on social inequality (Boterman, 2019). This is also confirmed by Oberti and Savina (2019), who performed similar research in French cities. They find strong inequalities between schools and confirm that school segregation and school success are related.

2.3 Theoretical framework

Following the current literature, it is expected that the average distance to facilities has a negative effect on average income levels. This works through two channels. Firstly, neighbourhoods with a low average distance to facilities attract high income earners, either Dutch or foreign. This channel unfolds an indirect effect of the distance to facilities on income levels, that is, through migration. An increase in the distance to facilities has a negative effect on migration to a neighbourhood and an increase in migration to a neighbourhood positively influences average income levels. Secondly, the prevalence of facilities itself leads to a rise in the income level of the people living in the neighbourhoods, so an increase in the average distance reduces future income levels (see figure 2).

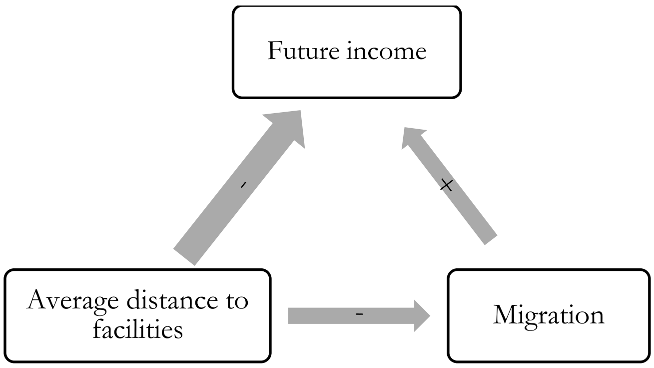


Figure 2: Box-arrow diagram

3. Methodology and data

As mentioned before, the objective of this research is to examine the effect that regional differences regarding facilities have on interregional income differences in the Netherlands. The data that is required to answer this question is retrieved from CBS StatLine (Statistics Netherlands) for the years 2016 until 2021 for all neighbourhoods in the Netherlands. There are approximately 14,000 neighbourhoods in the Netherlands and the geographical location of these neighbourhoods is defined by the municipalities. The CBS has created a guideline for the Dutch municipalities on how to divide their municipality into neighbourhoods (Centraal Bureau voor de Statistiek, 2020). Neighbourhoods that are reclassified throughout the years of observation and thus cannot be compared over time are excluded from the analysis. The variable that measures education is only available for three years: 2019, 2020 and 2021. Also, the year 2021 lacks many observations, as the variable that captures average income per income recipient per neighbourhood is not completely available for 2021 at the moment of writing.

All the following equations are tested and controlled for heteroskedasticity by making use of robust standard errors.

3.1 The effect of proximity to facilities on income

In order to test the effect of average distance to facilities on income levels, the dependent variable of our analysis is average gross income per income recipient (*income*) per year. This is measured in thousands of euros for the next period (t + 1), year in this case. The independent variable of the analysis represents the average distance to facilities (*X*). This is an average of different variables measuring the proximity to a specific facility. This variable contains average distance to nearest GP practice, average distance to nearest large supermarket (with a surface of more than 150m^2), average distance to nearest childcare centre, average distance to nearest primary school, average distance to nearest train station and average distance to nearest ramp to a main road. The first four variables are listed in the regional core statistics by the CBS and the latter two are variables concerning mobility and are therefore essential for people to commute to (better-paid) jobs. These six averages measure the sum of the distance to the nearest specific facility for every inhabitant of the neighbourhood, divided by the number of inhabitants. For privacy reasons, these variables are only available when a neighbourhood has 10 or more inhabitants.

The bivariate regression model we described above only contains one explanatory variable and chances are high that variables are omitted. Therefore, we should not completely rely on this analysis. However, the added value of this bivariate regression model is that this model has many observations, as for instance *education* is only available for a limited number of years. Also, it shows the importance of including control variables, as the interpretation changes substantially after adding the control variables. Income level thus does not solely depend on proximity to facilities. We know from the literature that education level is correlated with earnings (Card, 2018; Krueger & Mikael, 2001). Therefore, education level serves as one of the control variables (*education*). CBS lists the number of people per neighbourhood that are low, middle or highly educated. According to CBS, someone is middle educated when the highest attained education program is high school or vocational education (havo, vwo, mbo 2-4) and highly educated if someone attended professional college or university (hbo, wo). The remainder is regarded as low educated. In order to make an index, the share of people that are highly educated is used as the *education* variable. This variable thus takes on values between 0 and 1. Also urbanization level (*urbanization*) is one of the control variables, as this variable might be correlated with proximity to facilities, since facilities are better available in urban areas. This variable takes on values between 1 and 5, where 1 is very urban and 5 is not urban. The index is directly retrieved form CBS and is based on the housing density per neighbourhood. Another variable that might drive the results too is the population of the neighbourhood (*population*), as an investment in a facility has a higher return in a populous neighbourhood. This is measured by the number of inhabitants, rounded to 5. Another control variable that is somewhat related to the previous one is the surface of the neighbourhood (*size*), since certain neighbourhoods have for instance nature reserves, where no people live and no facilities are needed. This is measured by the number of hectares, excluding water surface. *Migration* is also a control variable that is included in the model, capturing the number of people that migrated from and to the specific neighbourhood. This is measured by the difference in population minus the natural population growth (*⊿Population – (birth – death)*), thus a negative number corresponds with a net migration from the neighbourhood to other neighbourhoods (in the Netherlands or abroad) and vice versa. The birth rates and death rates are also retrieved from CBS and are rounded to 5. The regression analysis also controls for age differences across neighbourhoods. Since the average income per income recipient is used, the model already controls for the number of children (that do not receive an income) in a certain neighbourhood. However, pensioners do receive an income, albeit in general lower than people of the working age population. Therefore, we include the share of people that are aged between 15 and 65, thus belong to the working age population, in the regression (*working*) as well as the share of the population that is 65 years and older (*elderly*). The last control variable that is added to the model is *year*, in order to avoid spurious regression due to a time trend that might be present in the data. In conclusion, the basic model is specified in the following way:

$$\text{Income}_{it+1} = \beta_0 + \beta_1 X_{it} + \beta_2 \text{Education}_{it} + \beta_3 \text{Urbanization}_{it} + \beta_4 \text{Population}_{it} + \beta_5 \text{Size}_{it} + \beta_6 \text{Migration}_{it} + \beta_7 \text{Working}_{it} + \beta_8 \text{Elderly}_{it} + \beta_9 \text{Year}_{it} + \alpha_i + \varepsilon_{it} \tag{3.1}$$

Where α_i is a neighbourhood specific effect that is consistent over time and ε_{it} is the error term. A descriptive table of the observations of the variables used in equation 3.1 is shown in table 1 (appendix 1).

The basic model will be estimated with and without control variables using a pooled Ordinary Least Squares (OLS) estimator. In both cases, β_1 measures whether average distance to facilities is correlated with income levels one year later. However, this paper aims to find a causal relationship and the basic model may not expose a causal relationship between proximity to facilities and income levels, as exceptions might drive the results to a large extent, exposing a positive correlation. Therefore, as a variation on the basic model, X is replaced by ΔX, suggesting that a change in average distance to facilities will cause regional income differences.

Existing literature expects that a negative β_1 will be found, indicating that a lower average distance to facilities increases income levels in the next period. The literature explains this relationship by the effect of migration and by the effect of rising incomes for the initial inhabitants.

3.2 The effect of proximity to facilities on migration

In order to measure to what extent the effect of proximity to facilities on income is caused by migration, a similar approach will be followed. However, we only consider ΔX here, as migration is also a flow variable. This leaves us with the following model:

(3.2) $$Migration_{it} = \beta_0 + \beta_1 \Delta Xit + \beta_2 Education_{it} + \beta_3 Urbanization_{it} + \beta_4 Population_{it} + \beta_5 Size_{it} + \beta_6 Working_{it} + \beta_7 Elderly_{it} + \beta_8 Year_{it} + \alpha_i + \varepsilon_{it}$$

This model is carried out both with control variables and without control variables, in order to test the importance of control variables and to have many observations. These regressions test whether a difference in proximity to facilities results in a difference in population due to migration. *Income* is excluded from the list of control variables, since the expected migrants are high income earners. Leaving *income* in the regression would bias the results, as the income differences is the underlying mechanism that hypothetically leads to migration. A significant coefficient for β_1 proves that a change in proximity to facilities leads to migration, either internal or foreign. The literature expects this coefficient to be negative, meaning that an increase in the average distance to facilities in a neighbourhood leads to emigration and vice versa.

3.3 The first difference model

So far, we only considered models that explain income levels by (differences in) distance to facilities and other variables. However, we are also interested in income differences. The main issue with the approach in sections 3.1 and 3.2 is that there is a risk of reverse causality. Technically, future income levels cannot influence the provision of facilities in hindsight. However, scatter plot 1 (appendix 2) shows that income levels of the future ($income_{it+1}$, denoted by *F.Income*) are to a large extent correlated with past income levels ($income_{it-1}$, denoted by *L.Income*). It could be the case that the distance to facilities decreased on average in the rich neighbourhoods, whereas the distance increased for the poorer neighbourhoods. For instance, a lack of demand for a child care centre for financial reasons can lead to an increased distance to this facility. So, not facilities makes a neighbourhood richer, but being richer creates facilities.

In order to control for reverse causality, the first differences of the basic models are taken. However, we do not take the first difference of *education*. Since this variable is only available for a limited number of years, we would lose too many observation. *Δpopulation* is replaced by *migration*, since these two are highly correlated and we are particularly interested in migration rather than in population change. *Size* is removed, since there is barely any variation in *Δsize*. Of course, the time trend is also not in first differences. Also, we control for municipality specific effects, by using dummies for the municipalities. This results in the following first difference model:

(3.3) $$\Delta\text{Income}_{it+1} = \beta_0 + \beta_1\Delta X_{it} + \beta_2\text{Education}_{it} + \beta_3\Delta\text{Urbanization}_{it} + \beta_4\text{Migration}_{it} + \beta_5\Delta\text{Migration}_{it} + \beta_6\Delta\text{Working}_{it} + \beta_7\Delta\text{Elderly}_{it} + \beta_8\text{Year}_{it} + \text{municipality dummies} + \varepsilon_{it}$$

In order to align the model that estimates the effect of migration on income with this approach, we also take the first difference of equation 3.2. For the same, above-mentioned reasons, we do not take the first difference of *education* and *year* and we remove *size*. Also, *Δpopulation* is not included, since this variable highly correlates with *migration*. Again, we also make use of municipality dummies. The first difference migration-model is specified in the following way:

(3.4) $$\text{Migration}_{it} = \beta_0 + \beta_1\Delta X\text{it} + \beta_2\text{Education}_{it} + \beta_3\Delta\text{Urbanization}_{it} + \beta_4\Delta\text{Working}_{it} + \beta_5\Delta\text{Elderly}_{it} + \beta_6\text{Year}_{it} + \text{municipality dummies} + \varepsilon_{it}$$

4. Results

4.1 The effect of proximity to facilities on income

The outcomes of the pooled OLS estimations of equation 3.1 are shown in column 2 of table 2 (appendix 1). Variations of the basic model can be found in column 1, 3 and 4. These models analyse the effect of the proximity to facilities on future income levels. Column 1 shows that there is no significant effect of average distance to facilities on average income per income recipient for the next year if control variables are not taken into account. When control variables are added (column 2), a positive relationship between average distance to facilities and future income is observed, meaning that neighbourhoods with a 1 kilometre higher average distance to the six facilities captured by X have a € 233 higher average income per income recipient the next year, holding all other conditions constant. Existing literature contradicts this finding. This coefficient might therefore reveal a correlation rather than a causal relationship, since certain exceptions might drive the results to a large extent. For instance, there may be certain rich neighbourhoods lacking facilities and vice versa that are affecting the direction of the coefficient and are concealing a negative causal relationship. Therefore, this paper also looks into the effect of the change of the proximity to facilities on future income levels. To put it another way, it examines the effect on income in the next year if the average proximity to the six facilities changes. Column 3 and 4 of table 1 show the results. When the control variables are ignored, a positive relationship is observed, meaning that if the average distance to facilities increase by 1 kilometre, the average income per income recipient increase by € 832, ceteris paribus. This coefficient is significant at a 1% level, however, the fitting of the model extremely poor, since the R-squared is only 0.000 (rounded). Therefore, we do not attach any weight to this coefficient, other than that it demonstrates the importance of the inclusion of control variables. If the control variables are added, the effect of an increase in the average distance to facilities by 1 kilometre results in a decrease of average income per recipient of € 2868 per year, ceteris paribus, with an R-squared being increased to 0.539. Also the opposite holds: if the distance to facilities decreases in a neighbourhood by 1 kilometre on average, it would yield the average income recipient an additional € 2868 for the next year. This finding is in line with the expectations based on current literature. In this model, the effect of net migration on future income per income recipient is not significant.

Regional income differences cannot solely be explained by regional differences regarding proximity to facilities. Also other effects play a role. In this analysis, we observe a positive and significant effect of the share of highly educated people in a neighbourhood (*education*) and the time variable (*year*) on future average income per income recipient. The positive effect of education on earnings is in line with the expectations based on previous research that has been scrutinizing the returns to schooling. On the other hand, a negative and significant effect of housing density (*urbanization*, due to scaling, the coefficient is positive), the number of inhabitants per neighbourhood (*population*) and the share of people aged 15 and older (*working* and *elderly*) on income is observed. The negative effect on housing density and the number of inhabitants might be explained by the fact that people with a higher income tend to have bigger houses, reducing housing density and the population rate in a neighbourhood. However, a caveat is that the *population* variable is not taken into account when its value is smaller than 10, since these observations are missing for X. This biases the estimation of the coefficient. A remarkable finding in this analysis is that the share of people older than 15 in a neighbourhood negatively affects future income levels. In other words, neighbourhoods with a larger share of infants tend to have higher income levels, other things equal. A possible explanation for this finding is that the working population in neighbourhoods with more children might work more hours per week in order to feed their families, gaining extra income. Since *income* is defined as the average income per income recipient, it does not capture the amount of children in a neighbourhood.

4.2 The effect of proximity to facilities on migration

Since we know that there is evidence that if facilities move away from certain neighbourhoods the income level will fall in the future (ceteris paribus) and vice versa, the question of how this can be explained and which mechanisms are at stake here arises. According to the literature, the explanation might be twofold. On the one hand, a living area with better and more facilities in the direct neighbourhood is more attractive to migrants, especially to those with a high income, i.e. those who can afford it. On the other hand, the presence of facilities in the direct neighbourhood will enable the initial inhabitants to reach higher income levels. For instance, the occurrence of a child care centre nearby may stimulate a parent to work an extra day per week, gaining additional earnings. A train station in the direct proximity enables someone to commute easily to better-paid positions, for example.

Hence, equation 3.2 tests whether a change in the proximity to facilities leads to migration. Table 3 of appendix 1 shows the results in column 2. Column 1 shows the results when control variables are excluded from the regression model. Here, we do not observe a significant effect of a change in distance facilities on migration. However, if the control variables are added, we observe a negative relationship between a change in distance to facilities and migration, indicating that if the average distance to facilities increases, inhabitants migrate to other neighbourhoods. If the average distance to facilities increases by 1 kilometre, 15.77 persons away from the neighbourhood, ceteris paribus. The magnitude of this coefficient is less relevant, since neighbourhoods differ in population rate and this coefficient has a relatively large standard error of 3.2. However, the direction of the coefficient is clear and significant, confirming the findings of the existing literature that high income earners move to facilities.

Also, the effect of *education*, *population*, *size*, *working*, *elderly* and *year* is significant and partially explain migration, so there are many reasons for people to migrate to certain neighbourhoods, apart from the availability of facilities. According to the model, people tend to migrate to the higher educated, more populous and larger neighbourhoods with a relatively larger share of children.

4.3 The first difference model

In order to control for reverse causality, equations 3.3 and 3.4 are carried out. The outcomes can be found in table 4 (appendix 1). From the first difference model, we can infer that the effect of a difference in average distance to facilities does not significantly cause income differences to change. This contradicts the finding in section 4.1. However, the effect of migration on income differences is positive and significant, implying that if migration increases with one (so one person moves from another neighbourhood to the specific neighbourhood), the average gross income per income recipient in the neighbourhood rises with € 1,48. We have found evidence for the claim that high income earners move to neighbourhoods were facilities are nearer on average.

By comparing the findings of the first difference model with the pooled OLS models analysed in section 4.2, we do not observe significant differences between the effect of proximity to facilities on migration. Only the magnitude differs slightly. Therefore, we can conclude from the first difference model that a change in proximity to facilities significantly affects migration and migration significantly causes income differences to rise. So, we revealed the indirect channel through which proximity to facilities affects income. In this model, we did not find evidence for the direct effect of proximity to facilities on income. The magnitude of the indirect effect is -0.02237.[2] So, if the change in average distance to facilities increases by 1 kilometre, average income per income recipient decreases with € 22,37 in a neighbourhood.

[2] The effect of *migration* on Δ*income* (appendix 1, table 3, column 1) multiplied by the effect of ΔX on *migration* (appendix 1, table 3, column 2), i.e. 0.00148 * -15.11. This calculation is computed by Stata and rounded to five decimal places.

5. Robustness check

In all the regression models so far, the variable that measures proximity to facilities is defined as the mean value of six different numerical variables. Four of these variables are regarded as core statistics by the CBS and the other two variables concern mobility. Although the choice of these variables is motivated, we still may argue that this can be seen as rather arbitrary. Also, all the variables have equal weight. Therefore, as a robustness check, we make use of the minimum value of the six variables that measure proximity to facilities, as well as the maximum value of the six separate variables. The rationale behind this approach is that not only average distance matters, but also the prevalence of any facility in a neighbourhood at all – as a proxy for a certain degree of liveliness in a neighbourhood – expressed by the minimum value of X, *minX*. This approach attaches the most weight to the facility that is in the closest proximity to the neighbourhood's inhabitants. Also, the prevalence of all of the six facilities in the direct proximity matters, represented by the maximum value of X, *maxX*. On the contrary, this method values the facility with the highest average distance as the most important. The descriptive statistics of *minX* and *maxX* can be found in table 5 (appendix 1).

In section 4, we attach the most explanatory value to the first difference models. Also the models with ΔX as the independent variable and were control variables are included are relevant, although these models might reveal a reverse causal relationship. Therefore, we only consider similar regression models as a robustness check. These regression models are tested for heteroskedasticity and for all of them, robust standard errors are applied.

The results of the basic models are shown in table 6. If we take either the minimum value or the maximum value of the average distance to one of the six facilities as the independent variable (*minX* or *maxX*), we notice that the magnitude of the effect decreases from -2.868 to -1.614 and -1.221, respectively (column 1 and 2), albeit that the direction and the significance of the effect does not change. This can easily be explained by the fact that in these equations, the extreme values of the proximity to facilities are considered. Although it is important to take these into account for the above mentioned reasons, it is only one out of six components and the magnitude is therefore smaller than if the six components are jointly taken into account. By comparing the magnitude of the coefficients of $\Delta minX$ and $\Delta maxX$, we observe that the prevalence of any facility in a neighbourhood is of greater importance than the prevalence of all the six facilities, since the absolute value of the coefficient of $\Delta minX$ is larger than the absolute value of the coefficient of $\Delta maxX$. The control variables do not change in direction and significance compared to the model in table 2 (column 4) and also the magnitude does not change substantially.

The outcomes of the migration-model can be found in column 3 and 4 (table 6, appendix 1). In general, the coefficients can be interpreted in a similar way as the interpretation of the robustness check of the income-model. Again, we observe that the effect of a change in the minimum value of the proximity to a facility has a smaller effect in magnitude on migration than the average value of the six variables that represent proximity to facilities have (column 2, table 3, appendix 1), meaning that the effect of one extreme value is smaller than the joint effect of the six components. Also, the prevalence of any facility in the direct neighbourhood is a more important determinant for migration than the prevalence of all of the six facilities, since the coefficient of $\Delta maxX$ is statistically not significantly different from 0.

Table 7 shows the results of the robustness check in a first difference model. As in the main first difference model, we do not observe a significant effect of a change in proximity to facilities on a change in future income. The effect of migration on income differences does not differ in significance and magnitude. Also, the effect of a change in proximity to facilities on migration has the same direction and significance. Only the magnitude differs slightly. In line with the previous findings of the robustness check, we also notice here that the absolute value of the coefficient of $\Delta minX$ is larger than the coefficient of $\Delta maxX$ and it is significant.

In conclusion, we can infer from the results of the robustness check that the outcomes of the models in section 3 are found to be robust and reliable. By changing the definition of *proximity to facilities*, the initial coefficients change in magnitude. However, this is logically explained by the fact that only one aspect of *proximity to facilities* is taken into account here instead of six different aspects. In addition, we find that it is more important to have any facility at all in a direct neighbourhood than having them all nearby. However, having them all nearby also increases future income levels, although the magnitude is smaller. This suggests that there are decreasing marginal returns on the amount of facilities.

6. Additional research

We know from this research that a change in proximity to facilities has a causal effect on future average earnings per income recipient on the neighbourhood level via the channel of migration. Now, we may wonder which of the six facilities taken into account are the most vital for income levels to improve or deteriorate in a neighbourhood. In order to investigate this question, we take the individual six components of the variable that measures proximity to facilities separately and consecutively in a model as the independent variable. These six variables are average distance to nearest GP practice (*GP*), average distance to nearest large supermarket (*supermarket*), average distance to nearest childcare centre (*childcare*), average distance to nearest primary school (*primaryschool*), average distance to nearest train station (*train*) and average distance to nearest ramp to a main road (*ramp*). Since the first four variables are regarded as the core statistics by the CBS, we also take the average distance of these four variables (*core*) and compare them with the average of the other two variables concerning mobility (*mobility*). The models are carried out by using a pooled OLS estimator as well as a first difference estimator. These regression models are tested for heteroskedasticity and for all of them, robust standard errors are applied.

6.1 The effect of proximity to facilities on income

Table 8 of appendix 1 shows the results of the pooled OLS estimator. Whether the components are separately or jointly taken into account, we observe in both cases that only the change in average distance to a large supermarket, the change in average distance to a train station and the change in average distance to a ramp to a main road has a significant effect on future income levels. This means that the availability of a large supermarket in the neighbourhood and mobility are the most important facilities affecting future income levels. In magnitude, the two components regarding mobility have the largest effect on future income, especially the proximity to a train station. There is no evidence that the other three components of *proximity to facilities* separately have a significant effect on future income levels, although the coefficients are negative. Both the averages of the core statistics as well as the mobility variables are significantly affecting income levels.

6.2 The effect of proximity to facilities on migration

We are also interested in which facilities have the potential to influence migration. The approach to examine this is similar to the previous one. The results of these pooled OLS regressions (table 9, appendix 1) confirm that a supermarket in the neighbourhood and a ramp to a main road are not only increasing income outlooks, but are also motives for people to migrate. Surprising is that a change in the average distance to a train station is the most important facility that affects future income levels, but does not significantly influence migration. We further observe that *train* and *ramp* are jointly not significantly affecting migration.

6.3 The first difference model

By estimating a first difference model on the effect of differences in proximity to facilities on income differences (table 10, appendix 1), we do not find any significant effects for the separate components. This is not surprising and in line with the previous findings in section 4.3. The effect of migration on future income differences is still significant and ranging from 0.00147 to 0.00149 among the models. In the initial first difference model, this coefficient is 0.00148.

The effect of a difference in proximity to one facility on migration is significant for a supermarket and a ramp to a main road (table 11, appendix 1). This finding is in accordance with the finding in section 6.2 and with Bastiaanssen and Breedijk (2022), who find that people in the Netherlands obtain most of their mobility by cars, rather than by public transport. Also, the effect of the joint core statistics is significant, whereas the mobility variables are not.

To summarize, the accessibility of a supermarket and a ramp to a main road positively influences future income levels through the migration channel. There seems to be no facility that targets future incomes directly if we rely on the first difference model.

7. Policy implications

According to the analysis in section 4, reducing the average distance to facilities by 1 kilometre in a neighbourhood leads to an increase in income through the channel of migration. For equality considerations, it might be desirable that the government would invest more in accessibility to facilities in the relatively poor neighbourhoods with high average distances to facilities in order to raise the income levels and reduce regional income differences across neighbourhoods. Improving the proximity to facilities can be regarded as a strategy to reduce regional inequality and to enhance the perspectives of the rural neighbourhoods, since social cohesion is also expected to rise (Tolsma, et al., 2009). However, some side notes need to be taken into account.

Firstly, a 1 kilometre decrease in the average distance to facilities is rather optimistic and hard to realize. If we look at table 5 (appendix 1), we notice that the average distance to five of the six facilities is smaller than 2 kilometres and the maximum value is for four of them smaller than 12 kilometres. A 1 kilometre decrease is then a substantial difference that might not be easily reached. Also, the decision on where to open a facility can be a private sector decision, for instance when it comes to supermarkets and (to a lesser extent) a child care centre. The government can barely influence such decisions, if it would want to do so. Therefore, it is more reasonable to expect a smaller effect on average income levels if the government would aim to tackle regional (in)accessibility.

Secondly, the income variable used in this analysis entails gross income per income recipient. So, the effect on disposable (net) income is smaller. Also, the ex-ante net income differences are smaller than the gross income differences among neighbourhoods, due to progressive taxation in the Netherlands. All in all, the direction of the effect is not questioned, but the magnitude and the feasibility aspect need to be taken into consideration by policymakers.

Section 5 teaches policymakers that aim to reduce regional inequality that the largest effect can be retrieved by starting at investing in the neighbourhoods that have no facility at all nearby. For those neighbourhoods, there is the most to gain. Also, it is relatively easy to substantially reduce the average distance to only one facility in those neighbourhoods.

The analysis in section 6 shows that the main instruments for reducing regional inequality by means of migration are reducing the average distance to a supermarket or reducing the average distance to a ramp to a main road. The maximum value of the average distance to ramp is 43.7 kilometres for a neighbourhood in the municipality of Terschelling, Friesland (table 11, appendix 1). Although Terschelling is a Waddenzee island, the mean value and the standard deviation of this variable are the second highest of the variables taken into account, so there is room for improvement here. The maximum value of the distance to a large supermarket is way smaller, so it might be hard to realize large declines in these average distances. Moreover, the private sector decides whether to open a retail facility or not.

The average distance to a supermarket or a ramp is measured by the distance by road (or water). Hence, decreasing the average distance to a facility does not necessarily imply that there need to be more supermarkets or ramps in the Netherlands in order to increase average future income levels by migration and reduce regional income differences. Making them better accessible by road might suffice. This is exactly what the government is able to do and where the government is responsible for (Van den Berg & Kok, 2021). Also, the core statistics that touch upon liveliness in a neighbourhood jointly seem to be of greater importance than the accessibility of other neighbourhood, expressed by *mobility*. Unfortunately, in this analysis we did not find an instrument that only causes incomes to rise, without stimulating people to migrate.

In conclusion, if the (local) government would aim to decrease regional income differences, it should do so by starting at the neighbourhoods with the least facilities. In order to give every neighbourhood equal opportunities and to reduce income inequality between neighbourhoods, the government should consider *proximity to facilities* as an important policy instrument.

8. Limitations

This research paper finds evidence for the claim that the average distance to facilities is negatively related with future incomes via the migration channel. However, this analysis would gain strength if the number of observations would be increased. For *education*, there are only three years of observation: 2019, 2020 and 2021. For 2021, there are many observations lacking for *income*. Since in the analysis the income level for the next year is used ($income_{it+1}$), only two years are taken into consideration here, with one year lacking observations. Also the neighbourhoods that cannot be compared over time, due to reclassifications (for instance a street that was first part of neighbourhood A and later part of neighbourhood B) are excluded from the dataset, albeit that this effect is small. Redoing the analysis after a few years to include more data might shed some light on how sustained the findings are. Nevertheless, there are still more than 10,000 observations for the main models.

For neighbourhoods with less than 10 inhabitants, the variables that measure the proximity to facilities are missing from the data for privacy considerations, as people could possibly be traced back to their home address. This might bias the results, since the provision of facilities is most likely not normally distributed over the neighbourhoods with less than 10 inhabitants. These neighbourhoods are mainly the rural neighbourhoods that are expected to have lower income levels per income recipient and have probably less facilities nearby.

Also some statistical problems might potentially cause problems. First of all, *population* contains a unit root. This means that the distribution of this variable follows a random walk and causal relationships cannot be inferred from this variable. Also, given the fact that small neighbourhoods are excluded from the dataset, the observations with less than 10 inhabitants are excluded from the model, biasing the estimate of *population*.

However, the fact that *population* contains a unit root does not cause statistical issues in the first difference models. Second, certain variables could be omitted that are correlated with one of the variables, causing a bias. We strived to include all the relevant control variables, yet some might be overlooked.

9. Concluding remarks

We have performed several pooled OLS and first difference regressions in order to analyse the effect of proximity to facilities on future income levels. The dataset contains Dutch neighbourhoods for the period 2016-2021. The empirical research shows that there is supportive evidence for the claim that a change in average distance to facilities is negatively correlated with income levels per income recipient on the neighbourhood level. The existing literature also finds this effect and attributes this relationship to an indirect effect through migration and to a direct effect. This paper confirms that the effect of proximity to facilities on income is caused through the migration channel and the results are found to be robust. We found little evidence for the claim that facilities directly affect income levels. Only a pooled OLS model indicates that this relationship is present, although the causal effect may be reversed. In addition, this research shows that it is more important to have any facility nearby than having many facilities, indicating that there are decreasing marginal return on facilities. If we break down the facilities in components, we find that having a supermarket and a ramp to a main road in a neighbourhood have a significant effect on future income levels through migration. A GP practice, a childcare centre, a primary school and a train station seem to be of less importance.

There are some key point that policymakers should take into account. First of all, there is the most to gain if policymakers start by reducing the minimum value of the average distance to a facility for the relatively poor neighbourhoods. Secondly, it is important to invest in mobility and in retail facilities. Finally, the proximity to facilities is not the only income determinant. Other factors, like education, are also important factors to take into account and are probably more straightforward policy objectives.

Further research can be performed on the municipality level. This paper shows that there is an effect observed on the neighbourhood level, but testing whether this effect would persist on the municipality level would be of added value. Especially for policymakers, it would be more feasible to make policy for municipalities rather than for neighbourhoods. However, it would be a waste of resources to make policy on the municipality level if the effect is not visible there.

Furthermore, additional research can be done by breaking down the effect in hourly wages and in working hours. We observed that decreasing the average distance to facilities in a neighbourhood will increase future income levels due to migration. This increase in income levels can be caused in two ways. Either migrants have higher hourly wages or migrants are working more hours per week. Of course, it could also be a combination of these two. It would add to the literature if this would be scrutinized and especially the implications this potentially has for policymakers. For instance, certain facilities may have an effect on hourly wages, whereas other may affect working hours.

In the end, it is upon society to decide upon the desirability of equal opportunities between neighbourhoods and at which costs they are willing to provide them. Some might argue that inhabitants who choose to live in neighbourhoods lacking facilities do this willingly and are free to move elsewhere. Others adhere to the equality principle and argue that it is not feasible for everyone to move to neighbourhoods with better facilities. This paper only tries to provide objective background information for both the viewpoints.

10. Appendices

Appendix A ~ Tables

Table 1: Descriptive Statistics

Variable	Obs	Mean	Std. Dev.	Min	Max
Income	10472	34.33	8.523	8.6	172.3
X	10522	2.01	1.414	.35	16.117
ΔX	10522	-.011	.097	-2.567	1.367
Education	10522	.298	.139	0	.895
Urbanization	10522	3.188	1.506	1	5
Population	10522	2081.657	2173.981	40	28870
Size	10522	222.954	505.35	1	12813
Migration	10522	9.042	64.428	-1095	1220
Working	10522	.647	.076	.043	1
Elderly	10522	.2	.088	0	.957
Year	10522	2019.125	.331	2019	2020

Table 2: Regression results

VARIABLES	(1) $Income_{t+1}$	(2) $Income_{t+1}$	(3) $Income_{t+1}$	(4) $Income_{t+1}$
X	-0.0366	0.233***		
	(-1.587)	(5.126)		
ΔX			0.832***	-2.868***
			(3.171)	(-5.021)
Education		45.26***		45.10***
		(59.13)		(58.99)
Urbanization		0.854***		0.960***
		(14.96)		(18.53)
Population		-0.000267***		-0.000279***
		(-9.259)		(-9.591)
Size		5.35e-05		0.000167*
		(0.569)		(1.781)
Migration		0.00141		0.00148
		(1.481)		(1.536)
Working		-40.96***		-40.69***
		(-20.17)		(-19.87)
Elderly		-31.18***		-31.05***
		(-21.20)		(-20.88)
Year		0.520***		0.563***
		(3.559)		(3.827)
Constant	34.22***	-997.1***	34.49***	-1,085***
	(439.2)	(-3.384)	(671.3)	(-3.654)
Observations	38,026	10,533	29,238	10,522
R-squared	0.000	0.540	0.000	0.539
R2_adj	1.23e-05	0.539	0.000191	0.539

Robust t-statistics in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 3: Migration

	(1)	(2)
VARIABLES	Migration	Migration
ΔX	-0.637	-15.77**
	(-0.236)	(-2.190)
Education		41.16***
		(14.04)
Urbanization		0.288
		(1.082)
Population		0.00153***
		(3.606)
Size		0.00150**
		(1.991)
Working		-114.1***
		(-9.310)
Elderly		-52.05***
		(-4.834)
Year		-1.258***
		(-3.206)
Constant	6.779***	2,617***
	(29.95)	(3.299)
Observations	59,353	32,571
R-squared	0.000	0.022
R2_adj	-1.41e-05	0.0217

Robust t-statistics in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 4: First difference model

VARIABLES	(1) $\Delta Income_{t+1}$	(2) Migration
ΔX	0.302	-15.11**
	(1.088)	(-2.092)
Education	0.802***	50.48***
	(4.237)	(14.78)
ΔUrbanization	-0.129	-41.47***
	(-1.255)	(-2.915)
Migration	0.00148***	
	(5.456)	
ΔMigration	-2.43e-06	
	(-0.0748)	
ΔWorking	-1.083	-214.5***
	(-0.476)	(-5.055)
ΔElderly	1.381	-534.6***
	(0.582)	(-11.38)
Year	0.268***	-1.419***
	(7.007)	(-3.528)
Municipality dummies	…	…
	…	…
Constant	-541.3***	2,879***
	(-6.999)	(3.547)
Observations	10,083	32,571
R-squared	0.082	0.053
R2_adj	0.0504	0.0424

Robust t-statistics in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 5: Descriptive statistics of the components of *X*

Variable	Obs	Mean	Std. Dev.	Min	Max
minX	10522	.634	.536	0	10.2
maxX	10522	6.041	6.153	.6	58.4
Gp	10522	1.354	1.233	0	11.8
Supermarket	10522	1.253	1.182	0	11.7
Childcare	10522	.895	.838	.1	10.2
Primaryschool	10522	.934	.788	.1	10.2
Train	10522	5.813	6.287	.2	58.4
Ramp	10522	1.81	1.398	.1	43.7

Table 6: Robustness check

VARIABLES	(1) $Income_{t+1}$	(2) $Income_{t+1}$	(3) Migration	(4) Migration
ΔMinX	-1.614***		-11.31**	
	(-3.126)		(-2.387)	
ΔMaxX		-1.221***		-0.318
		(-5.407)		(-0.212)
Education	45.07***	45.13***	41.08***	41.09***
	(58.91)	(58.96)	(14.01)	(14.01)
Urbanization	0.956***	0.951***	0.257	0.242
	(18.46)	(18.41)	(0.962)	(0.900)
Population	-0.000280***	-0.000282***	0.00153***	0.00153***
	(-9.607)	(-9.710)	(3.615)	(3.618)
Size	0.000177*	0.000172*	0.00153**	0.00159**
	(1.886)	(1.823)	(2.030)	(2.076)
Migration	0.00149	0.00155		
	(1.555)	(1.615)		
Working	-40.74***	-40.57***	-114.4***	-114.1***
	(-19.90)	(-19.85)	(-9.341)	(-9.298)
Elderly	-31.07***	-30.99***	-51.97***	-51.83***
	(-20.89)	(-20.85)	(-4.815)	(-4.797)
Year	0.529***	0.577***	-1.448***	-1.491***
	(3.598)	(3.903)	(-3.811)	(-3.856)
Constant	-1,016***	-1,112***	3,002***	3,087***
	(-3.425)	(-3.731)	(3.909)	(3.953)
Observations	10,522	10,522	32,571	32,571
R-squared	0.538	0.539	0.022	0.021
R2_adj	0.538	0.539	0.0214	0.0210

Robust t-statistics in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 7: Robustness check in first differences

	(1)	(2)	(3)	(4)
VARIABLES	$\Delta \text{Income}_{t+1}$	$\Delta \text{Income}_{t+1}$	Migration	Migration
ΔMinX	-0.0932		-10.000**	
	(-0.481)		(-2.146)	
ΔMaxX		0.103		-0.673
		(1.025)		(-0.405)
Education	0.805***	0.804***	50.63***	50.56***
	(4.251)	(4.246)	(14.84)	(14.81)
ΔUrbanization	-0.122	-0.124	-41.67***	-41.66***
	(-1.189)	(-1.206)	(-2.907)	(-2.902)
Migration	0.00148***	0.00148***		
	(5.427)	(5.429)		
ΔMigration	-1.46e-06	-1.49e-06		
	(-0.0447)	(-0.0458)		
ΔWorking	-1.122	-1.092	-215.1***	-214.3***
	(-0.493)	(-0.480)	(-5.098)	(-5.077)
ΔElderly	1.291	1.339	-534.5***	-534.5***
	(0.544)	(0.565)	(-11.39)	(-11.43)
Year	0.275***	0.270***	-1.613***	-1.631***
	(7.251)	(7.073)	(-4.171)	(-4.140)
Municipality dummies	…	…	…	…
	…	…	…	…
Constant	-553.9***	-543.8***	3,271***	3,308***
	(-7.243)	(-7.065)	(4.191)	(4.160)
Observations	10,083	10,083	32,571	32,571
R-squared	0.082	0.082	0.053	0.052
R2_adj	0.0502	0.0503	0.0421	0.0418

Robust t-statistics in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 8: The effect of the six components on income

VARIABLES	(1) $Income_{t+1}$	(2) $Income_{t+1}$	(3) $Income_{t+1}$	(4) $Income_{t+1}$	(5) $Income_{t+1}$	(6) $Income_{t+1}$	(7) $Income_{t+1}$	(8) $Income_{t+1}$	(9) $Income_{t+1}$
ΔGP	-0.433								-0.106
	(-1.297)								(-0.312)
ΔSupermarket		-0.635**							-0.521**
		(-2.538)							(-2.065)
ΔChildcare			-0.298						-0.165
			(-1.246)						(-0.622)
ΔPrimaryschool				-0.293					-0.0957
				(-0.973)					(-0.285)
ΔTrain					-1.047***				-0.966***
					(-5.418)				(-4.943)
ΔRamp						-0.800***			-0.595**
						(-3.005)			(-2.235)
ΔCore							-1.191**		
							(-2.399)		
ΔMobility								-1.758***	
								(-5.778)	
Education	45.09***	45.08***	45.09***	45.08***	45.12***	45.12***	45.08***	45.14***	45.12***
	(58.91)	(58.94)	(58.89)	(58.90)	(58.96)	(58.88)	(58.95)	(58.96)	(58.98)
Urbanization	0.956***	0.958***	0.955***	0.956***	0.948***	0.962***	0.957***	0.957***	0.957***
	(18.44)	(18.45)	(18.42)	(18.44)	(18.37)	(18.55)	(18.47)	(18.50)	(18.49)
Population	-0.000279***	-0.000279***	-0.000279***	-0.000279***	-0.000282***	-0.000278***	-0.000279***	-0.000281***	-0.000281***
	(-9.568)	(-9.555)	(-9.585)	(-9.590)	(-9.720)	(-9.550)	(-9.566)	(-9.656)	(-9.660)
Size	0.000185**	0.000188**	0.000185**	0.000184*	0.000174*	0.000183*	0.000180*	0.000171*	0.000169*
	(1.966)	(2.000)	(1.961)	(1.959)	(1.842)	(1.939)	(1.920)	(1.811)	(1.797)
Migration	0.00149	0.00141	0.00152	0.00152	0.00156	0.00146	0.00149	0.00150	0.00146
	(1.555)	(1.470)	(1.582)	(1.585)	(1.622)	(1.520)	(1.549)	(1.564)	(1.515)
Working	-40.70***	-40.66***	-40.68***	-40.67***	-40.57***	-40.66***	-40.71***	-40.58***	-40.59***
	(-19.89)	(-19.86)	(-19.88)	(-19.88)	(-19.84)	(-19.88)	(-19.88)	(-19.85)	(-19.82)
Elderly	-31.00***	-30.95***	-30.99***	-30.99***	-30.98***	-30.96***	-31.02***	-30.96***	-30.98***
	(-20.86)	(-20.82)	(-20.84)	(-20.86)	(-20.85)	(-20.84)	(-20.86)	(-20.85)	(-20.83)
Year	0.526***	0.518***	0.522***	0.521***	0.571***	0.535***	0.526***	0.579***	0.580***
	(3.577)	(3.520)	(3.550)	(3.543)	(3.865)	(3.641)	(3.582)	(3.929)	(3.924)
Constant	-1,010***	-993.4***	-1,002***	-999.6***	-1,101***	-1,028***	-1,011***	-1,118***	-1,119***
	(-3.404)	(-3.347)	(-3.376)	(-3.370)	(-3.692)	(-3.468)	(-3.409)	(-3.756)	(-3.752)
Observations	10,522	10,522	10,522	10,522	10,522	10,522	10,522	10,522	10,522
R-squared	0.538	0.538	0.538	0.538	0.539	0.538	0.538	0.539	0.540
R2_adj	0.538	0.538	0.538	0.538	0.539	0.538	0.538	0.539	0.539

Robust t-statistics in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 9: The effect of the six components on migration

VARIABLES	(1) Migration	(2) Migration	(3) Migration	(4) Migration	(5) Migration	(6) Migration	(7) Migration	(8) Migration	(9) Migration
ΔGP	-4.031*								-3.526*
	(-1.875)								(-1.726)
ΔSupermarket		-8.098***							-7.586***
		(-3.537)							(-3.389)
ΔChildcare			-0.760						-0.175
			(-0.508)						(-0.162)
ΔPrimaryschool				-1.042					0.334
				(-0.294)					(0.107)
ΔTrain					-0.879				-0.0589
					(-0.484)				(-0.0345)
ΔRamp						-7.832**			-7.408**
						(-2.557)			(-2.473)
ΔCore							-12.44**		
							(-2.232)		
ΔMobility								-4.507	
								(-1.421)	
Education	40.96***	41.06***	41.09***	41.06***	41.12***	41.17***	40.99***	41.22***	41.06***
	(13.97)	(14.02)	(14.01)	(13.99)	(14.03)	(14.06)	(14.00)	(14.08)	(14.03)
Urbanization	0.292	0.276	0.243	0.244	0.240	0.252	0.300	0.239	0.326
	(1.104)	(1.028)	(0.906)	(0.911)	(0.893)	(0.940)	(1.130)	(0.891)	(1.230)
Population	0.00153***	0.00153***	0.00153***	0.00153***	0.00153***	0.00153***	0.00153***	0.00153***	0.00153***
	(3.614)	(3.606)	(3.619)	(3.616)	(3.618)	(3.620)	(3.606)	(3.617)	(3.606)
Size	0.00155**	0.00158**	0.00158**	0.00158**	0.00158**	0.00156**	0.00152**	0.00156**	0.00152**
	(2.042)	(2.073)	(2.076)	(2.084)	(2.073)	(2.052)	(2.011)	(2.053)	(2.036)
Working	-114.0***	-114.4***	-114.2***	-114.2***	-114.1***	-114.2***	-114.3***	-114.0***	-114.3***
	(-9.290)	(-9.339)	(-9.305)	(-9.310)	(-9.286)	(-9.309)	(-9.330)	(-9.286)	(-9.328)
Elderly	-51.82***	-51.82***	-51.85***	-51.86***	-51.85***	-51.93***	-52.00***	-51.90***	-51.90***
	(-4.798)	(-4.801)	(-4.800)	(-4.811)	(-4.801)	(-4.807)	(-4.824)	(-4.809)	(-4.824)
Year	-1.462***	-1.416***	-1.492***	-1.496***	-1.471***	-1.383***	-1.375***	-1.386***	-1.270***
	(-3.841)	(-3.734)	(-3.909)	(-3.929)	(-3.804)	(-3.625)	(-3.595)	(-3.570)	(-3.242)
Constant	3,029***	2,936***	3,089***	3,099***	3,046***	2,869***	2,853***	2,875***	2,640***
	(3.938)	(3.833)	(4.007)	(4.027)	(3.899)	(3.722)	(3.691)	(3.665)	(3.336)
Observations	32,571	32,571	32,571	32,571	32,571	32,571	32,571	32,571	32,571
R-squared	0.022	0.022	0.021	0.021	0.021	0.022	0.022	0.021	0.023
R2_adj	0.0213	0.0218	0.0210	0.0210	0.0210	0.0213	0.0216	0.0211	0.0222

Robust t-statistics in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 10: The effect of the six components on income in first difference

VARIABLES	(1) $\Delta Income_{t+1}$	(2) $\Delta Income_{t+1}$	(3) $\Delta Income_{t+1}$	(4) $\Delta Income_{t+1}$	(5) $\Delta Income_{t+1}$	(6) $\Delta Income_{t+1}$	(7) $\Delta Income_{t+1}$	(8) $\Delta Income_{t+1}$	(9) $\Delta Income_{t+1}$
ΔGP	0.103								0.0724
	(0.872)								(0.636)
ΔSupermarket		0.0658							0.0518
		(0.688)							(0.541)
ΔChildcare			0.0879						0.0648
			(0.802)						(0.620)
ΔPrimaryschool				0.0721					0.0270
				(0.635)					(0.248)
ΔTrain					0.0857				0.0784
					(0.921)				(0.872)
ΔRamp						0.00347			-0.0152
						(0.0334)			(-0.151)
ΔCore							0.221		
							(1.047)		
ΔMobility								0.106	
								(0.738)	
Education	0.802***	0.805***	0.803***	0.805***	0.804***	0.804***	0.804***	0.802***	0.803***
	(4.232)	(4.255)	(4.243)	(4.252)	(4.248)	(4.252)	(4.246)	(4.242)	(4.240)
ΔUrbanization	-0.125	-0.125	-0.125	-0.124	-0.124	-0.123	-0.128	-0.125	-0.129
	(-1.220)	(-1.219)	(-1.216)	(-1.208)	(-1.204)	(-1.200)	(-1.245)	(-1.214)	(-1.247)
Migration	0.00148***	0.00149***	0.00147***	0.00147***	0.00148***	0.00148***	0.00148***	0.00148***	0.00148***
	(5.441)	(5.461)	(5.410)	(5.407)	(5.426)	(5.432)	(5.451)	(5.439)	(5.418)
ΔMigration	-1.50e-06	-1.38e-06	-2.38e-06	-2.35e-06	-1.87e-06	-1.57e-06	-2.46e-06	-1.63e-06	-2.54e-06
	(-0.0457)	(-0.0423)	(-0.0733)	(-0.0726)	(-0.0573)	(-0.0482)	(-0.0756)	(-0.0501)	(-0.0783)
ΔWorking	-1.091	-1.096	-1.095	-1.097	-1.099	-1.106	-1.077	-1.109	-1.069
	(-0.480)	(-0.482)	(-0.481)	(-0.482)	(-0.483)	(-0.486)	(-0.473)	(-0.488)	(-0.469)
ΔElderly	1.345	1.332	1.317	1.329	1.349	1.308	1.369	1.327	1.406
	(0.567)	(0.563)	(0.556)	(0.560)	(0.569)	(0.552)	(0.577)	(0.560)	(0.591)
Year	0.272***	0.274***	0.273***	0.273***	0.270***	0.274***	0.271***	0.270***	0.268***
	(7.170)	(7.237)	(7.187)	(7.203)	(7.071)	(7.261)	(7.140)	(7.100)	(6.987)
Municipality dummies	…	…	…	…	…	…	…	…	…
	…	…	…	…	…	…	…	…	…
Constant	-548.3***	-552.2***	-549.8***	-551.1***	-544.9***	-552.3***	-547.3***	-545.6***	-540.6***
	(-7.162)	(-7.229)	(-7.179)	(-7.195)	(-7.064)	(-7.253)	(-7.132)	(-7.092)	(-6.980)
Observations	10,083	10,083	10,083	10,083	10,083	10,083	10,083	10,083	10,083
R-squared	0.082	0.082	0.082	0.082	0.082	0.082	0.082	0.082	0.082
R2_adj	0.0502	0.0502	0.0502	0.0502	0.0503	0.0501	0.0503	0.0502	0.0500

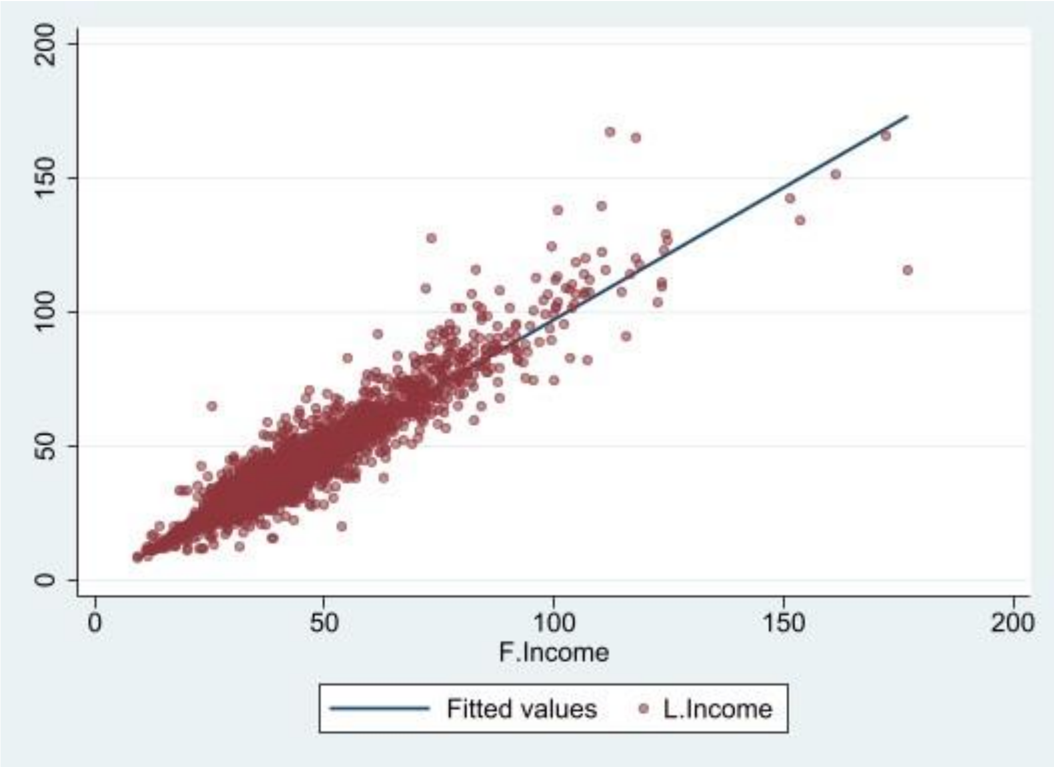
Robust t-statistics in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 11: The effect of the six components on migration in first difference

VARIABLES	(1) Migration	(2) Migration	(3) Migration	(4) Migration	(5) Migration	(6) Migration	(7) Migration	(8) Migration	(9) Migration
ΔGP	-4.047*								-3.562*
	(-1.823)								(-1.692)
ΔSupermarket		-7.271***							-6.782***
		(-3.214)							(-3.075)
ΔChildcare			-0.247						0.310
			(-0.168)						(0.285)
ΔPrimaryschool				-0.899					0.157
				(-0.268)					(0.0533)
ΔTrain					-1.440				-0.624
					(-0.734)				(-0.340)
ΔRamp						-6.830**			-6.373**
						(-2.196)			(-2.087)
ΔCore							-11.41**		
							(-2.063)		
ΔMobility								-5.054	
								(-1.504)	
Education	50.46***	50.48***	50.56***	50.55***	50.55***	50.64***	50.46***	50.57***	50.46***
	(14.76)	(14.80)	(14.81)	(14.79)	(14.80)	(14.85)	(14.77)	(14.82)	(14.77)
ΔUrbanization	-41.54***	-41.58***	-41.67***	-41.66***	-41.65***	-41.69***	-41.47***	-41.64***	-41.49***
	(-2.913)	(-2.900)	(-2.904)	(-2.907)	(-2.905)	(-2.904)	(-2.913)	(-2.905)	(-2.919)
ΔWorking	-213.9***	-214.7***	-214.3***	-214.4***	-214.3***	-214.5***	-214.5***	-214.4***	-214.4***
	(-5.048)	(-5.068)	(-5.080)	(-5.101)	(-5.080)	(-5.075)	(-5.055)	(-5.078)	(-5.049)
ΔElderly	-534.3***	-534.7***	-534.6***	-534.7***	-534.5***	-533.7***	-535.0***	-534.2***	-533.6***
	(-11.40)	(-11.41)	(-11.44)	(-11.51)	(-11.43)	(-11.38)	(-11.42)	(-11.40)	(-11.38)
Year	-1.613***	-1.579***	-1.654***	-1.652***	-1.602***	-1.555***	-1.540***	-1.522***	-1.431***
	(-4.163)	(-4.091)	(-4.263)	(-4.265)	(-4.056)	(-4.005)	(-3.945)	(-3.835)	(-3.566)
Municipality dummies	…	…	…	…	…	…	…	…	…
	…	…	…	…	…	…	…	…	…
Constant	3,272***	3,203***	3,354***	3,351***	3,249***	3,154***	3,123***	3,088***	2,904***
	(4.183)	(4.112)	(4.284)	(4.286)	(4.076)	(4.026)	(3.965)	(3.855)	(3.586)
Observations	32,571	32,571	32,571	32,571	32,571	32,571	32,571	32,571	32,571
R-squared	0.053	0.053	0.052	0.052	0.052	0.053	0.053	0.053	0.054
R2_adj	0.0422	0.0424	0.0418	0.0418	0.0418	0.0420	0.0423	0.0419	0.0428

Robust t-statistics in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Appendix B ~ Figures



Scatter plot 1: the correlation between income$_{it+1}$ and income$_{it-1}$

Appendix C ~ References

Arntz, M. (2010). What Attracts Human Capital? Understanding the Skill Composition of Interregional Job Matches in Germany. *Regional Studies, 44*(4), pp. 423-441.

Bastiaanssen, J., & Breedijk, M. (2022). *Toegang voor iedereen? Een analyse van de (on)bereikbaarheid van voorzieningen en banen in Nederland.* The Hague: PBL.

Beckers, P., & Boschman, S. (2019). Residential choices of foreign highly skilled workers in the Netherlands and the role of neighbourhood and urban regional characteristics. *Urban Studies, 56*(4), pp. 760-777.

Boterman, W. R. (2019). The role of geography in school segregation in the free parental choice context of Dutch cities. *Urban Studies, 56*(15), pp. 3074-3094.

Brueckner, J. K., Thisse, J.-F., & Zenou, Y. (1999, January 10). Why is central Paris rich and downtown Detroit poor?: An amenity-based theory. *European Economic Review, 43*(1), pp. 91-107.

Card, D. (2018, February 15). Returns to Schooling. *The New Palgrave Dictionary of Economics*.

CDA. (2023, February 2). *CDA wil aanpak van regionale ongelijkheid 'voor heel Nederland'*. Retrieved from cda.nl: https://www.cda.nl/actueel/nieuws/cda-wil-aanpak-van-regionale-ongelijkheid-voor-heel-nederland

Centraal Bureau voor de Statistiek. (2020). *Richtlijnen voor gemeenten bij het vaststellen van een indeling naar wijken en buurten – versie 2.4*. The Hague: CBS.

De Voogd, J., & Cuperus, R. (2021). *Atlas van Afgehaakt Nederland.*

Devogelaer, D. (2004). *Interne migraties in België: wie, waarom en naar welke gemeenten?* Brussels: Federaal Planbureau.

Glaeser, E. L., Kolko, J., & Saiz, A. (2001, January 1). Consumer City. *Journal of Economic Geography* , pp. 27-50.

Hunt, G. L., & Mueller, R. E. (2004, November). North American Migration: Returns to Skill, Border Effects, and Mobility Costs. *The Review of Economics and Statistics, 86*(4), pp. 988-1007.

Krueger, A. B., & Mikael, L. (2001). Education for Growth: Why and for Whom? *Journal of Economic Literature, 39*(4), pp. 1101-1136.

Oberti, M., & Savina, Y. (2019). Urban and school segregation in Paris: The complexity of contextual effects on school achievement: The case of middle schools in the Paris metropolitan area. *Urban Studies, 56*(15), pp. 3117-3142.

PvdA. (2023, February 12). *Samen met GroenLinks presenteren we onze plannen voor bereikbaar, betaalbaar en duurzaam openbaar vervoer.* Retrieved from pvda.nl: https://www.pvda.nl/nieuws/iedereen-stapt-in/

Raspe, O., Groot, S., Boschman, S., Beckers, P., Sleutjens, B., Boterman, W., & van Gessel, G. (2014). *Buitenlandse kenniswerkers in Nederland. Waar werken en wonen ze en waarom?* The Hague: PBL.

Sá, C., Florax, R. J., & Rietveld, P. (2006). Does Accessibility to Higher Education Matter? Choice Behaviour of High School Graduates in the Netherlands. *Spatial Economic Analysis, 1*(2), pp. 155-174.

Tolsma, J., Van der Meer, T., & Gesthuizen, M. (2009). The impact of neighbourhood and municipality characteristics on social cohesion in the Netherlands. *Acta Politica, 44*(3), pp. 286-313.

Van de Ven, C. (2021, October 13). *Hoe Den Haag uit Nederland verdween*. Retrieved from De Groene Amsterdammer: https://www.groene.nl/artikel/hoe-den-haag-uit-nederland-verdween?gclid=EAIaIQobChMIk8j3lpvt_QIVU_J3Ch1PCg-HEAAYASAAEgLOffD_BwE

Van den Berg, C., & Kok, A. (2021). *Regionaal maatschappelijk onbehagen.* Groningen: Rijksuniversiteit Groningen.

Van den Berg, C., & Van Vulpen, B. (2019, January 26). De kloof tussen bruisgebieden en krimpregio's blijft groeien. *Trouw.*

Van Oort, F., Weterings, A., & Verlinde, H. (2003, September). Residential amenities of knowledge workers and the location of ICT-firms in the Netherlands. *Tijdschrift voor Economische en Sociale Geografie, 94*(4), pp. 516-523.

Weterings, A., Buitelaar, E., & Edzes, A. (2019). *Regionale verschillen in werkzekerheid: de rol van beleid.* The Hague: PBL.