Transport Poverty in the Amsterdam Metropolitan Area: Relationships with Socioeconomics and the Built Environment at the Neighborhood Level

Delft University of Technology

Department of Transport and Planning

Faculty of Civil Engineering and Geosciences

CIE5050-09 Additional Graduation Work, Research Project

Thijs Bon - 4342623

Main supervisor: dr. Matthew Bruno

Chair: prof. Niels van Oort

December 2022

Summary

Limiting transport poverty is consequential in improving well-being and employment levels, which play meaningful roles in deciding public policy. We analyze how different socioeconomic and built environment factors are related to the transport poverty environment for car and public transport in terms of strength, significance, and direction for neighborhood zones within the Amsterdam Metropolitan Area—our study area. Additionally, we provide policy recommendations for the study area.

To this end, we use spatial distributions of environmental transport poverty indicators and perform weighted least-squares regression analyses, where we regress each transport poverty indicator on all built environment and socioeconomic variables. Our regression analyses are preceded by a combination of logarithmic variable transformation, insignificant variable elimination, and data normalization.

The environmental transport poverty indicators consist of average travel times, average single-trip travel costs, and the number of accessible jobs within thirty minutes. Regarding the built environment; we include inhabitant density and job density, whereas the socioeconomic characteristics in our analyses consist of household size, five age cohorts (0-18, 18-34, 35-54, 55-64, and 65^{+}), gender, car ownership, and income.

Our results indicate that levels of environmental car transport poverty are fairly low over the whole study area when compared to the public transport poverty environment—the highest car transport poverty levels among all zones correspond with the lowest levels of public transport poverty. Regression results demonstrate that differences in the transport poverty environment are substantially correlated with differences in the zonal built environment and socioeconomic characteristics only for public transport travel times and public transport job accessibility, which are also the only transport poverty indicators to exhibit considerable variation among the study area zones in general.

Furthermore, the strength and significance of job density almost invariably greatly exceed those of inhabitant density and the socioeconomic variables—in relation to the transport poverty environment indicators. Inhabitant density is, despite being overshadowed by job density, also deemed to exhibit a marked correlation to the public transport measures of travel

time and job accessibility. Both built environment variables are related to favorable transportation conditions. Remarkably, all socioeconomic variables display either insignificant or rather weak correlations to the transport poverty environment.

Consequently, we suggest focusing mainly on public transport measurements of travel times and job accessibility when analyzing differences in the transport poverty environment at a neighborhood level and the roles of job density and (to a lesser extent) inhabitant density herein. We visualize the combined spatial distributions of both job density and average public transport travel times, and inhabitant density and average public transport travel times for our study area. Areas potentially interesting for public transport improvement are identified based on the simultaneous occurrence of both relatively high levels of job or inhabitant density and relatively high travel times.

We note that our results are limited to transport poverty resources and opportunities in the context of neighborhood zonal averages within the Amsterdam Metropolitan Area. Further research might investigate the strength, significance, and direction of the relationships between socioeconomic and built environment characteristics, and transport poverty at the individual level—which could demonstrate the purportedly high influence of socioeconomic characteristics in relation to transport poverty differences among people. It would also be interesting to determine how widely applicable or location dependent our results are based on other studies which employ a different study area. Other suggestions are to include measures of subjective perception, transport poverty outcomes, and transport externalities.

Contents

1 Introduction

An extensive body of literature has addressed the ability of different groups of people to fulfill their mobility needs. This ability and related concepts are named and described in various manners. Moreover, generally agreed-upon definitions are rare throughout the literature among these concepts. We discuss several concept and definition pairs.

The first of these pairs is accessibility poverty, defined by Martens & Bastiaanssen (2019) as relatively high experienced travel times (compared to the entire population of the region in question) to key destinations such as health care services, leisure, jobs, and social activities, where low-income population groups are at a greater risk of experiencing accessibility poverty due to their dependence on the public transport system or active modes, which generally leads to higher travel times compared to the car. Fan & Huang (2011) define transportation affordability by the extent to which financial means are necessary to make use of transportation services. Transport-related social exclusion identifies ten dimensions, related to the capability of humans to access the transportation system, that are connected to societal participation for individuals (e.g., economic-, physical-, and informational exclusion) (Luz & Portugal, 2022).

A rather extensive notion is transport equity (Martens et al., 2019), which comprises available transportation modes, travel time to key destinations (e.g., jobs, health care, parks), observed travel patterns, well-being measures related to travel and activities, air and noise pollution, traffic safety, and attractiveness of active mode usage. Another concept definition pair is transportation disadvantage, described as "limits on or barriers to access to destinations such as employment, education, health care and nutritious foods (Shay et al., 2016)." The 15-minute city (Moreno et al., 2021) entails the transformation of contemporary car-dependent cities into cities where walking and cycling are viable options to satisfy most of the daily travel needs.

Finally, Awaworyi Churchill & Smyth (2019) take note of mobility poverty, accessibility poverty, and transport affordability as constituent parts of transport poverty. Here, mobility is defined by the available transportation options, whereas accessibility refers to the difficulty to reach key destinations such as work, shops, and social contacts. Affordability constitutes the extent to which financial means available to a person are sufficient to make use of the available transportation options.

There exist more concepts and definitions that could be added to this list, however, the ones mentioned here provide sufficient background information on the issue we want to address.

1.1 Research Relevance and Aim

Among all concept and definition pairs, we focus on transport poverty in this paper and split it into its three constituent parts (mobility poverty, accessibility poverty, and transport affordability) as given by Awaworyi Churchill & Smyth (2019). The importance of research into transport poverty is related to outcomes influencing public policy-making, such as well-being and employment levels. A higher level of transport poverty has been shown to negatively affect well-being (Awaworyi Churchill & Smyth, 2019; Awaworyi Churchill, 2020; Delbosc & Currie, 2011), and Bastiaanssen et al. (2020) find a negative association between transport poverty and employment outcomes by merging 93 studies with quantitative estimates on this association. Concluding, it is valuable to inform policy-making decisions by investigating the dynamics at play between transport poverty and concepts expected to affect it, or vice versa.

How socioeconomic and built environment characteristics are related to transport poverty is a much-researched topic (Martens et al., 2019; Fan & Huang, 2011; Shay et al., 2016; Jomehpour Chahar Aman & Smith-Colin, 2020; Voerknecht, 2020; Kampert et al., 2019), where careful consideration to the selection of the (combination of) indicators is often given based on logical reasoning. Nonetheless, there is next to no information on how different socioeconomic and built environment factors are related to each of the distinct dimensions of transport poverty in terms of strength, significance, and direction.

Lucas et al. (2018) do obtain such statistics on correlations between socioeconomic and built environment variables, and trip frequency (at the individual geographical level). However, they note that their analysis is somewhat hampered by their small sample size. Besides, they solely focus on trip frequency and there are no results reported on certain interesting socioeconomic and built environment variables such as job density and income.

Information on observed correlations between transport poverty, and socioeconomic and built environment variables can inform research into causal mechanisms potentially underlying the correlations, and facilitate the identification of people or environments at risk of exhibiting transport poverty. Subsequently, policy measures may be designed to reduce transport poverty by addressing its causes, or by targeting people or environments (e.g., compensation measures, urban planning, further research) more effectively.

The main aim of this study is to provide policy recommendations and inform further research into transport poverty by uncovering, at the neighborhood level, the strength, significance, and direction of the links between built environment and socioeconomic characteristics, and the three considered dimensions of transport poverty (mobility, accessibility, and affordability) for both car and public transport.

We limit the scope of the research by solely looking at resources and opportunities existing within the given transport environment in relation to transport poverty. It follows that any results do not pertain in any way to outcomes and subjective measures of transport poverty and leave heterogeneity between people in terms of experienced travel times and affordability measures unaccounted for.

1.2 Approach

Our research approach is visualized in a flowchart process scheme in Figure 1. To determine the strength, significance, and direction of the links between built environment and socioeconomic characteristics, and the transport poverty environment, we first conduct a literature review to deduce which socioeconomic and built environment characteristics are often linked to transport poverty and hence interesting to include in our analysis.

The strength, significance, and direction of relationships between built environment and socioeconomic characteristics (independent variables) and the transport poverty dimensions for car and public transport (dependent variables) are analyzed using linear regression analyses because this will provide such information. Next to a collection of potentially interesting independent variables to include resulting form the literature review, further desirable data properties originate from the choice for linear regression analyses. In linear regression analysis, each independent variable should exhibit low levels of correlation with all other independent variables to facilitate accurate estimation of the correlational links. Therefore,

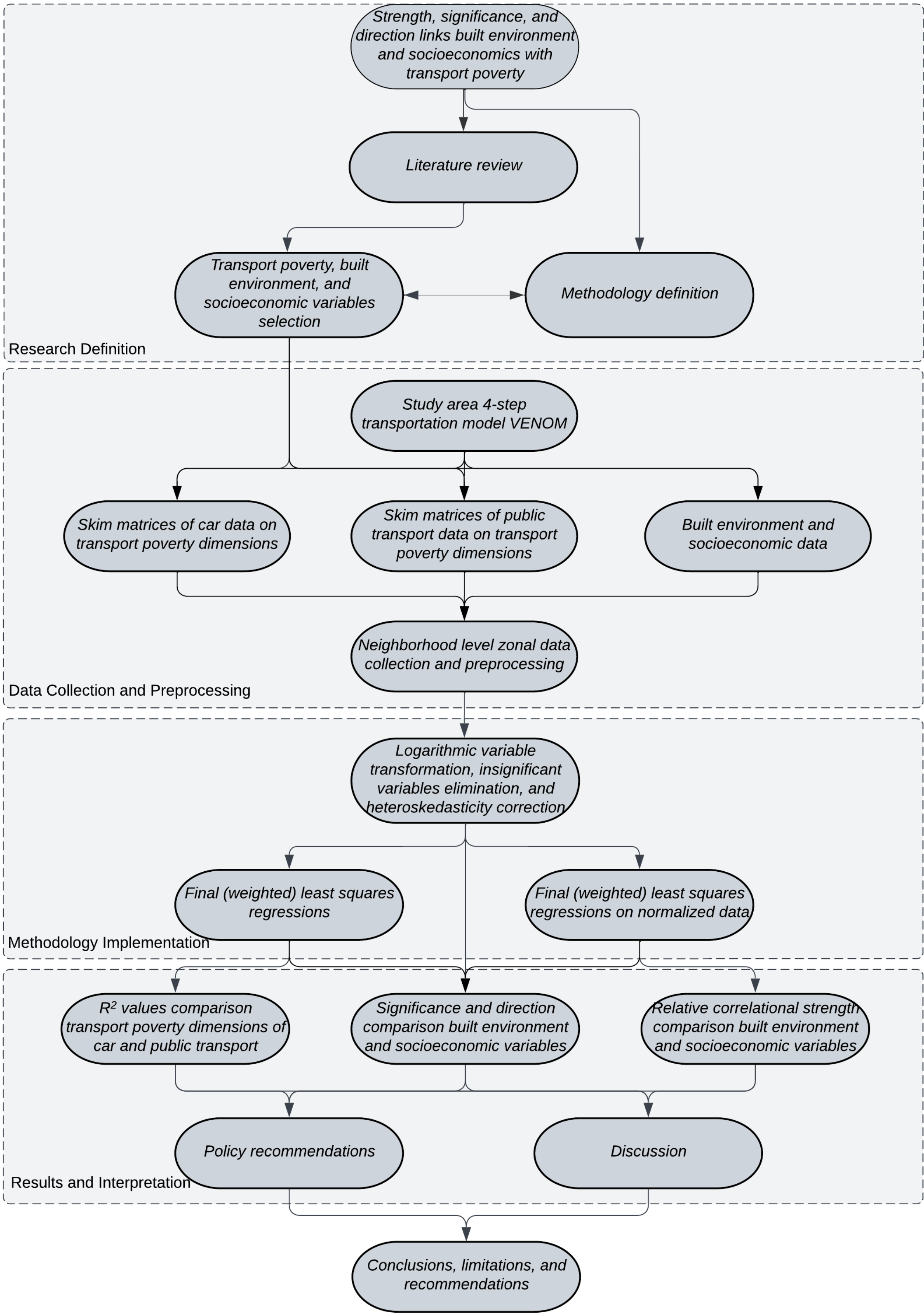

Figure 1: Research process scheme.

it should be averted to include closely related independent variables.

The resulting eleven included independent variables (built environment and socioeconomic characteristics) are inhabitant density, job density, household size, five age cohorts (0-18, 18-34, 35-54, 55-64, and 65^+), male, car ownership, and income. The three dimensions of transport poverty are mobility, accessibility, and affordability. These dimensions are respectively quantified through average inter-zonal travel times, the number of available jobs within half an hour of travel time, and average inter-zonal single trip costs. The three quantitative indicators are derived from the number of jobs for each zone as stated in the database of a study area specific 4-step transportation model, and from inter-zonal skim matrices on travel times and travel costs—which resulted from the application of this same transportation model. The three quantitative indicators are computed for car and public transport, we consequently end up with six dependent variables.

Our study area comprises the Amsterdam Metropolitan Area; the constituent neighborhood zones of the study area specific transportation model are visualized Figure 2. The socioeconomic, built environment, and transport poverty variables are computed on the level of these neighborhood zones. In Appendix A, Study Area and the Netherlands, the location of the study area within the Netherlands is depicted.

Before conducting the final linear regression analyses, we log-transform any variable when it is judged from histogram plots of its empirical and log-transformed empirical distributions that a logarithmic transformation substantially normalizes the data. Furthermore, the final linear regression analyses are preceded by eliminating insignificant independent variables through a systematic backward elimination method, and we correct for heteroskedasticity if statistical evidence for its presence is found. Afterward, a linear regression analysis is conducted for each dependent variable, regressing the dependent variable on all independent variables. Moreover, as additional analyses all variables are standardized before performing all regressions again, imposing the same scales for all variables.

Through the final regression analyses, the first intended main result is R^2 values for each of the six dependent variables, showing the amount of covariation between the estimated linear combination of the selected built environment and socioeconomic characteristics for each transport poverty indicator. In other words, the R^2 values will show the extent to which

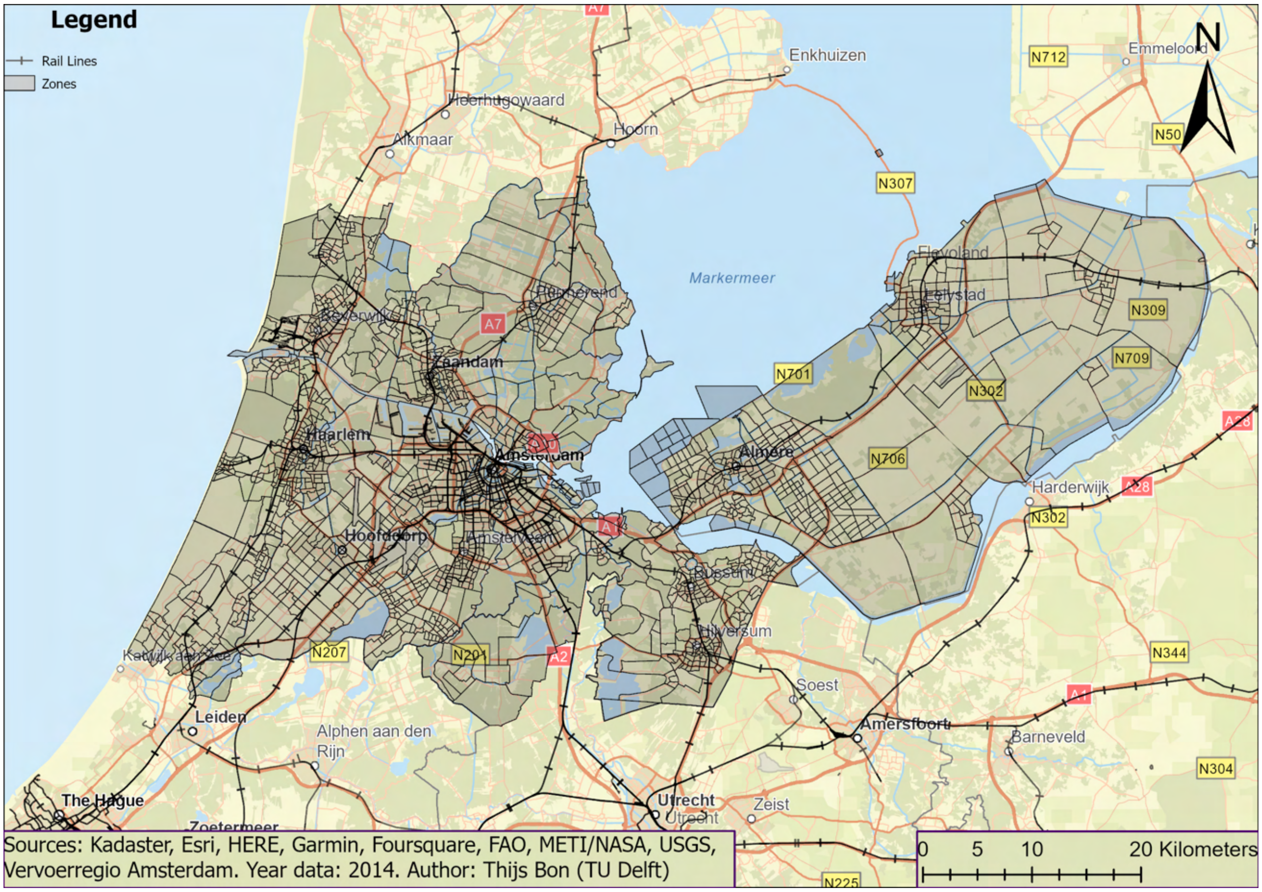

Figure 2: Amsterdam Metropolitan Area overview.

differences in the built environment and socioeconomic characteristics may be related to differences in the transport poverty indicator, for each indicator. Secondly, levels of significance and direction of the relationships between the built environment and socioeconomic characteristics, and the transport poverty indicators are identified from the regression results and from the backward elimination method. Thirdly, relative correlational strength measures follow from the coefficient estimates of the regressions on the normalized data. From the results, we provide policy recommendations for the study area, establish conclusions, and discuss implications and recommendations for further research as well as identify limitations of this research.

2 Literature Review

A literature review is conducted to decide which socioeconomic and built environment factors would be relevant to include in our analysis and to determine further how to shape the methodology so that the contribution to the existing body of literature is maximized. We use Google Scholar as our primary search engine and search for (combinations of) the keywords "transport poverty", "determinants", "factors", "socioeconomic(s)", "demographic(s)", "built environment", "land use", "mobility", "accessibility", and "transport(ation) affordability". Furthermore, sources referenced within the research found through the Google Scholar search are often also included in our literature review, as well as research that references the research found through the Google Scholar search.

From our obtained literature collection we find that concepts that have been shown to have a connection with transport poverty are the built (transportation) environment (Lucas et al., 2018; Godfrey et al., 2015; Martens, 2013), spatial distribution of the social network (Martens, 2013), fuel price (Godfrey et al., 2015), and predominantly the personal environment as expressed by socioeconomic indicators (Martens et al., 2019; Fan & Huang, 2011; Shay et al., 2016; Jomehpour Chahar Aman & Smith-Colin, 2020; Voerknecht, 2020; Kampert et al., 2019). According to Lucas et al. (2018), trip-making patterns of city inhabitants in Merseyside, England are influenced by socioeconomic characteristics to a lesser degree than by the built (transportation) environment. However, they note that their paper should be viewed as an exploration of their methodological approach, rather than as a robust causal analysis of the relationship between social disadvantage and transport at the micro level. Still, it illustrates the importance to remain aware of all the complex interactions at play between the built environment, the personal environment, and the transport environment.

2.1 Socioeconomics and Transport Poverty

Which socioeconomic indicators relate to transportation poverty (or overlapping notions) and in what way they interact has been researched extensively (Martens et al., 2019; Fan & Huang, 2011; Shay et al., 2016; Jomehpour Chahar Aman & Smith-Colin, 2020; Voerknecht, 2020; Kampert et al., 2019). Martens et al. (2019) establish relevant socioeconomic determinants of transport equity from an extensive literature review, namely: income, car ownership, age, gender, (dis)ability, and ethnicity. Fan & Huang (2011) argue for segregation based on household socioeconomic characteristics regarding transport affordability because this would define individuals' financial and time limitations. The household categories are defined by a multitude of combinations of the following socioeconomic factors: gender, marital status, occupational status, car ownership, and presence or absence of children. Furthermore, Shay et al. (2016) use low income, car ownership, mobility limitation, youth (non-driving age), seniors (over age 62), ethnic minority, and low English proficiency as socio-demographic measures influencing transportation disadvantage. This set was constructed from a comprehensive literature review, and by consulting the project sponsors—North Carolina Department of Transportation staff. Jomehpour Chahar Aman & Smith-Colin (2020) display an overview of socioeconomic factors influencing transit dependency from previous literature, they mostly coincide with the aforementioned measures.

For the Netherlands specifically, Hans Voerknecht extensively investigates the distribution of transport poverty among different socioeconomic groups at different locations and how this is impacted by policy decisions. For example, Voerknecht (2020) uses combinations of car ownership, household size, driving licence status, subsidized use of public transport, income, and preference for car or public transport to identify groups of people who might be analyzed separately regarding transport poverty. Furthermore, Kampert et al. (2019) establish threshold levels or categories from socioeconomic and built environment variables to identify the risk of transport poverty among households within neighborhood zones for two municipalities in the Netherlands. The used socioeconomic variables are car ownership, income, origin of income, migrational background, health status, and certain household characteristics including the age of the household members.

The link between socioeconomics and transport poverty at a neighborhood zone spatial level might also be viewed from the perspective of residential location choice as an intermediate factor. Zondag & Pieters (2005) find that, for the Netherlands, accessibility appears to have a significant but relatively low influence on residential location choice (locations with higher accessibility levels are preferred, ceteris paribus) compared to socioeconomic factors, neighborhood characteristics, and dwelling features, which is in line with their literature review. Accessibility is defined here as a household-type specific aggregate logsum value, founded in microeconomic utility maximization theory, that expresses the utility of diverse alternatives such as modes, destinations, and time-of-day options. They obtain these results from a quasi-dynamic land-use and transport interaction model (TIGRIS XL) for six predefined household types, partly at the four-digit postal code level but mostly at less detailed spatial levels. Thus, evidence is found for an interplay between certain transport poverty dimensions, socioeconomics, and the built environment through residential location choice, where accessibility is relatively unimportant.

2.2 Identified Gap

Much of the research thus far has employed socioeconomic and built environment data to segregate a population into various sub-groups and identify people or locations at risk of experiencing transport poverty. In the whole body of literature, used socioeconomic and built environment indicators generally appear well-founded. Even combinations have been made (Fan & Huang, 2011), accounting for the complex interplay of the indicators. Even though other scholars often give careful consideration to the selection of the (combination of) indicators based on logical reasoning (Martens et al., 2019; Fan & Huang, 2011; Shay et al., 2016; Jomehpour Chahar Aman & Smith-Colin, 2020; Voerknecht, 2020; Kampert et al., 2019), an analysis of how different socioeconomic and built environment factors are related to each of the distinct dimensions of transport poverty in terms of strength, significance, and direction, has yet to be conducted.

Lucas et al. (2018) do obtain such strength, significance, and directional information, at the individual geographical level, but they note that their estimates are hampered by their small sample size. Besides, they solely focus on trip frequency and there are no results reported on certain interesting socioeconomic and built environment variables such as job density and income. Therefore, we propose a quantitative analysis of the multidimensional relationships between built environment and socioeconomic characteristics, and transport poverty.

3 Methods

At the core, we use linear regression analyses to obtain the desired results on significance, strength, and direction among the built environment and socioeconomic characteristics in relation to transport poverty. Linear regression analysis is chosen because it delivers the desired results, outcomes are easily interpretable, and its steps are straightforward and transparent, making our analysis easily reproducible.

To perform linear regression analysis, we necessarily assume a linear mathematical model relating transport poverty to the built environment and socioeconomic characteristics. It is specified as

$$y_{ij} = \beta_j' x_i + \varepsilon_i, \quad i = 1, 2, \ldots, N, \quad j = 1, 2, \ldots, M, \tag{1}$$

where y_{ij} is the scalar observation on the j^{th} transport poverty measure in the i^{th} zone. The number of zones is given by $N = 1704$, while the number of transport poverty measures $M = 6$. The built environment and socioeconomic characteristics observations in zone i are represented by x_i, a $k \times 1$ vector, where $k-1$ gives the number of characteristics and k equals the number of independent variables (built environment and socioeconomic characteristics plus a constant). For transport poverty measure j, there is a corresponding $k \times 1$ vector of coefficients β_j which relates the built environment and socioeconomic characteristics to transport poverty measure j, and the scalar idiosyncratic error term (with a mean of zero) is ε_{it}, it has a variance of σ^2.

The sequential steps involved in our methodology consist respectively of logarithmic variable transformation, insignificant variables elimination, and heteroskedasticity correction. After performing these steps we obtain all the required information for the final regression models specifications. We conduct twelve final regression analyses, six on the (log-transformed) data and six on the normalized data. Normalized data is used to obtain the relative correlational strength statistics, and the regression analyses with original data provide coefficient estimates in terms of the (log-transformed) scales of measurement. Both sets of the six regression analyses are comprised of three car analyses and three public transport analyses, one for each of the three transport poverty indicators.

3.1 Logarithmic Variable Transformation

A logarithmic transformation allows for a multiplicative model structure to be estimated linearly. For instance, suppose that we have $log(y_{i1})$ and $log(x)$, then

$$log(y_{i1}) = \beta_j' log(x_i) + \varepsilon_i, \tag{2}$$

which is equal to

$$y_{i1} = x_{i1}^{\beta_{11}} x_{i2}^{\beta_{12}} \cdots x_{ik}^{\beta_{1k}} exp(\varepsilon_i), \tag{3}$$

where x_{ik} is the i^{th} observation on independent variable k and β_{1k} is the coefficient relating to the first transport poverty measure and the k^{th} independent variable. As formulas (2) and (3) illustrate, logarithmic variable transformation allows for substantially more flexibility in model specification. In the example above, all regressors are log-transformed, but this may even vary per regressor. In cases where a multiplicative specification is substantially more reflective of the observed relationship, this will lead to a greatly improved estimation of the respective assumed relationships.

From a visual inspection, it is decided for all regression variables (transport poverty, built environment, and socioeconomic) whether a logarithmic transformation substantially normalizes the data and thus would improve parameter estimates. If this is deemed to be the case, the variable will be log-transformed in the regression analyses. Visual inspection is performed by plotting a histogram for both the original and log-transformed data and visually comparing the shapes of the histograms to the typical shapes of an exponential and normal distribution.

When applying a logarithmic transformation to a variable that contains observations equal to zero, unrealistic and unworkable observations result because $\lim_{x\to 0^+} \ln(x) = -\infty$. Therefore, we change all zero values into one thousandth before applying the transformation. This way, the relatively more valuable information from these observations at the extreme of the observed range will be reasonably reflected.

3.2 Insignificant Variables Elimination

The elimination of insignificant variables in each regression analysis is carried out through backward elimination. This a well-known method to determine a set of significant independent variables to include in the final regression model. It starts by incorporating all variables in the model and then consecutively drops the least significant variable from the regression model estimated with the remaining variables until all remaining variables conform to a predetermined threshold level of statistical significance. We use the generally applied significance level of 5% for this threshold and display the results of the final regression models. The resulting regression models inform about the strength, significance, and direction of the correlation between the remaining socioeconomic and built environment measures and the transport poverty indicators in terms of the units of the variables.

3.3 Heteroskedasticity Correction

Heteroskedasticity, i.e., inconstant variance among observations, might be caused by outliers or omitted confounding variables. For our model, it would mean that $Var(\varepsilon) = \sigma_i^2 \neq \sigma^2$ for some $i \in \{1, \ldots, N\}$. Heteroskedastic observations with a relative high level of variance would obscure the relationship between the built environment and socioeconomic characteristics, and the transport poverty measures. Therefore, we limit the influence of such observations in proportion to their relative excess level of variance. To account for the possibility that an outlier might actually reveal interesting information contradicting the relationship established on the basis of the other observations, we plot studentized residuals for each regression against all the respective significant independent variables and visually determine from these plots whether any potential relationships might exist that we did not account for.

After the visual check for model misspecification through plotting of the studentized residuals against all of the respective significant independent variables, we carry out a White test (White, 1980) to statistically check for heteroskedasticity in the error terms. When the null hypothesis of homoskedasticity is rejected with at least 95% confidence, we estimate the linear regression model through weighted least squares (White, 1980) for improved efficiency of parameter estimates and to obtain correct standard errors. It works by making the influence on the parameter estimates of each observation inversely proportional to its estimated variance. When performing weighted least squares with the previously selected variables, it may happen that evidence for a coefficient estimate possibly being equal to zero is stronger than 5%. Hence, we conduct backward elimination again to ensure the significance of the coefficient estimates.

3.4 Relative Correlational Strength Estimation

We compare the independent variables in terms of the degree of correlational strength with each of the six transport poverty measures by normalizing the data and rerunning the (weighted) linear regression models with the explanatory variables as given by the backward elimination process on the nonnormalized data. Normalization is carried out by first subtracting the sample mean and subsequently dividing the outcome by the sample standard deviation, for all regressants and regressors. Measures of statistical significance are unaffected by data normalization, so it is unnecessary to redo the backward elimination with the normalized data. Normalization results in the same scales for all variables, as such the original units of measurement do not apply anymore. For model interpretability, we consequently carry out regressions with the normalized data as additional analyses, next to the regressions on the nonnormalized data. Since all normalized variables exhibit the same scale, the coefficient estimates represent how comparable changes in the independent variables relate to changes in the dependent variables. Comparing the coefficient estimates resulting from performing weighted least squares on the normalized data thus reveals the relative correlational strength levels.

4 Data

In this section, the used data is discussed. Starting with the socioeconomic and built environment data and followed by the transport poverty data. The spatial resolution of both is at a neighborhood level and the data stems from the year 2014, the most recent available year. Therefore, our results will pertain to this rather distant moment in time and any changes in the built environment, socioeconomic makeup of the neighborhood zones, and the transport poverty environment in the meantime will not be reflected. For example, the North-South metro line opening and subsequent extensive Amsterdam metro schedule changes occurred in 2018. However, no other transportation infrastructure change happened at a scale similar to the North-South metro line in the meantime, and the opening of the North-South metro line still presents only a relatively minor change when compared to the extensiveness of the public transport network of the whole study area.

Regarding changes in the built environment and socioeconomic makeup of neighborhoods from 2014 till now, it seems unlikely that these changes meaningfully impact the results because the vast majority of the zones likely did not see any property development projects large enough to substantially change the average levels pertaining to the built environment and socioeconomic characteristics. We thus expect that only a small number of zones that were largely impacted by property or transport infrastructure development saw considerable changes in the transport poverty measures or the built environment and socioeconomic characteristics, such as zones in the North of Amsterdam with easy access to a new metro station or zones that experienced big property development projects. Since we utilize average levels for transport poverty, built environment, and socioeconomic characteristics in each zone and use 1704 neighborhood zones for the regression, we expect that the results still mostly hold now.

There are 1941 neighborhood zones situated in the Amsterdam Metropolitan Area. After filtering out observations with missing data, 1704 neighborhood observations remain. Furthermore, there are 1959 zones situated outside the study area, of which 159 are located outside the Netherlands. These zones are important to include due to their influence on transport poverty within the study area. Table 1 describes the study area zonal data used (directly or indirectly) as input for our regression analyses, whereas Figure 3 visualizes the study area zonal geography. A geographical overview of all 3900 model zones is given in Figure 4. You can clearly see in Figure 4 how the model zones become increasingly fine-meshed when located closer to the Netherlands. In the Netherlands, the decrease of model zone area occurs too with decreasing distance to the study area zone, which can already be deduced from the information that the number of zones inside the study area is almost equal to the number of zones outside of the study area (1941 and 1959, respectively).

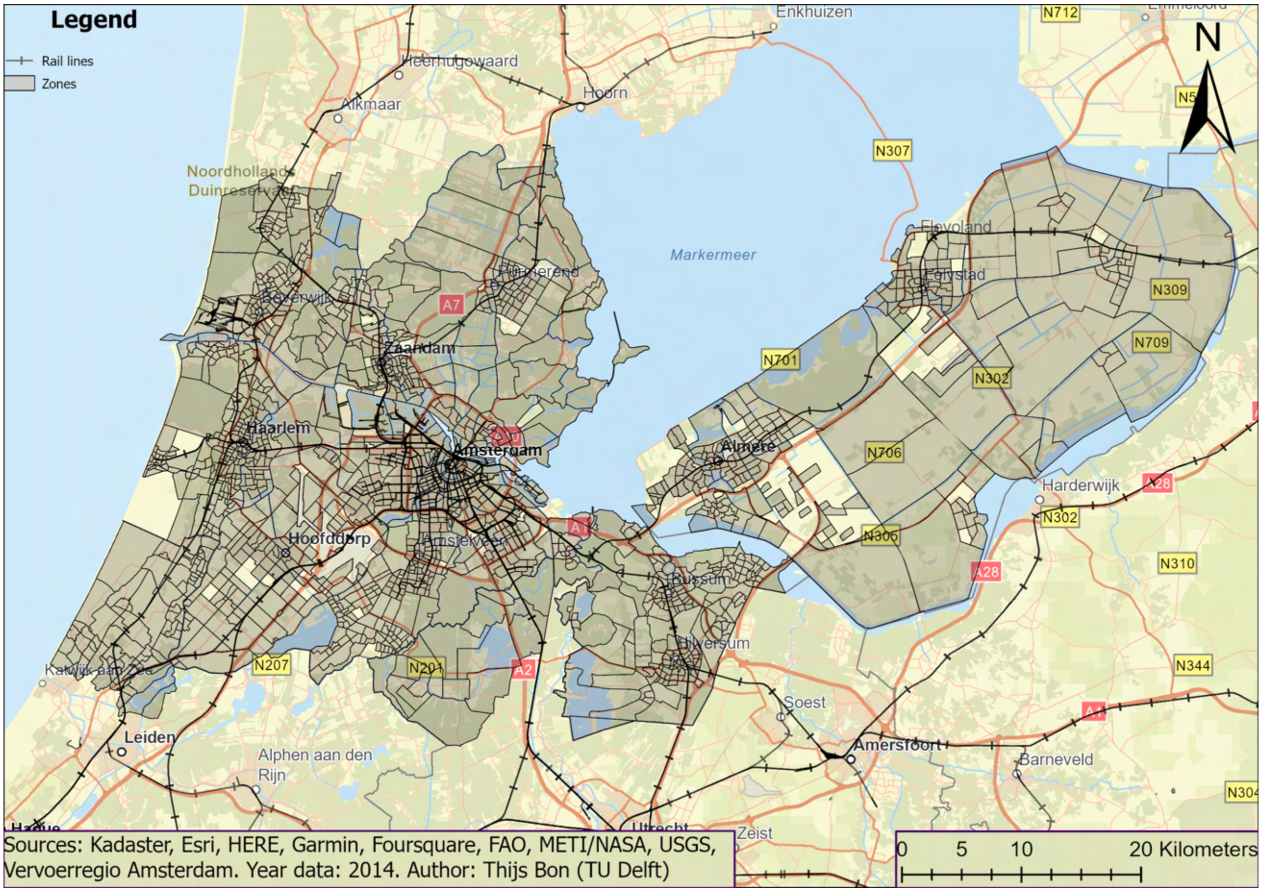

Figure 3: Geography of the 1704 study area zones.

The raw socioeconomic and built environment data consists (for each zone) of inhabitant age cohort totals, inhabitant gender (female and male) totals, total inhabitants, total number of households, area in hectares, number of jobs in seven sectors (agriculture, industry, retail, services, government, self-employed, and other), number of cars registered, and average income level. Most of the socioeconomic data originate from Statistics Netherlands, a Dutch governmental agency. Additionally, there are some other sources and some modification is performed by a research agency (ABF Research) tasked with providing the socioeconomic measures for the national and regional Dutch transportation models. For a complete description of this dataset see Groenemeijer et al. (2022).

To obtain efficient correlational statistics from the linear regression analyses, each independent variable should display only a limited amount of correlation with all other independent variables. Otherwise, the variance of the coefficient estimate will inflate, inhibiting accurate estimation of the correlational links. Hence, we want to limit the number of indepen-

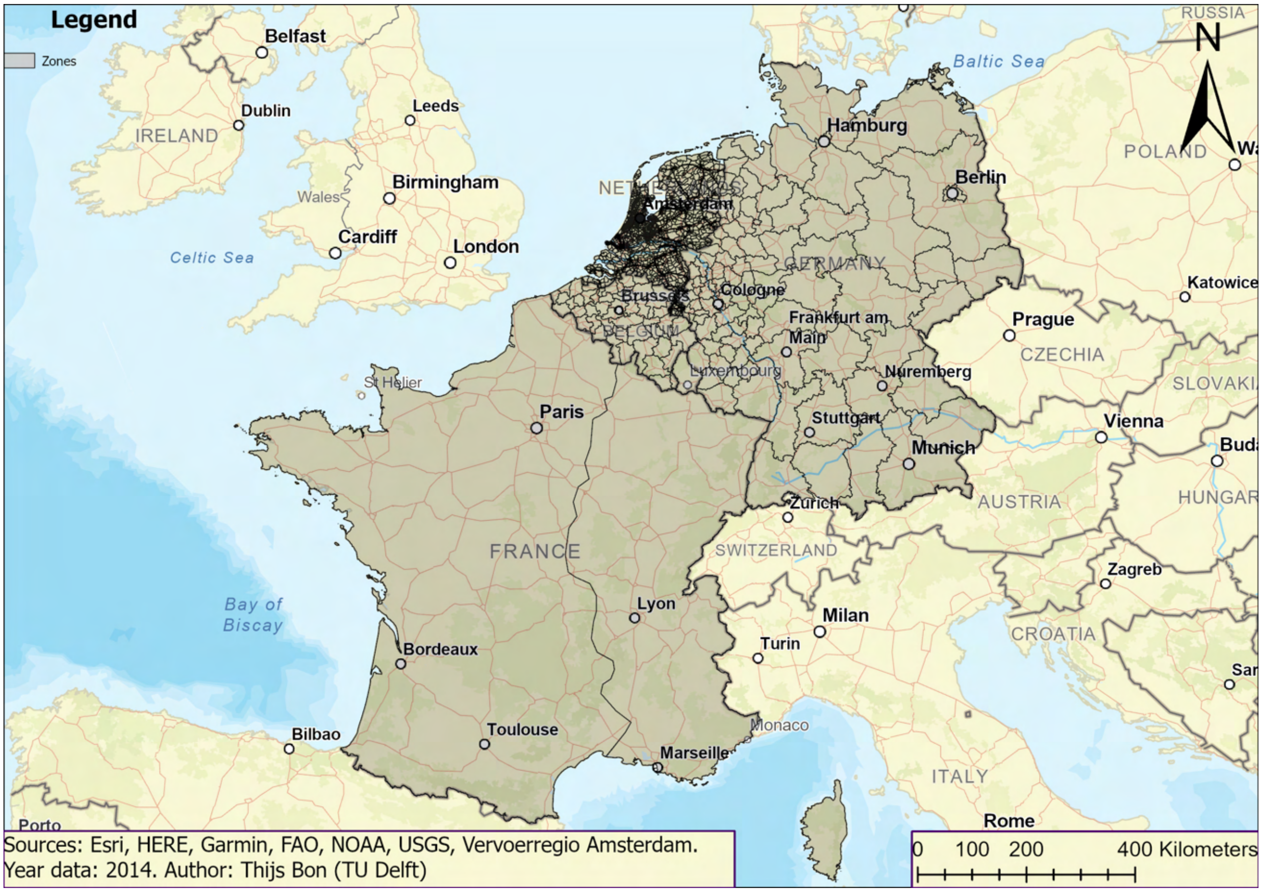

Figure 4: Geography of the 3900 model zones.

dent variables which are closely related to each other, which influences the built environment and socioeconomic variables selection process.

4.1 Socioeconomic Data

Age, gender, income, car ownership, and household size are included as socioeconomic explanatory variables in the analyses. Income is reflected by the average yearly income in euros among the inhabitants of a zone, whereas car ownership is represented by the average number of registered passenger cars per inhabitant. For the potential relation between household size and transport poverty, we include the average household size in the analysis. A household is defined as one or more persons who either share or do not share a dwelling with one or more other households and who provide in their livelihood for themselves as a single unit (CBS, 2003).

Table 1: Descriptive statistics of the 1704 study area zones.

	Mean	SD	Median	Min	Max	Description
Area	1.47	4.44	0.44	0.03	65.6	in square kilometers
Inhabitants	1.55	1.84	1.00	0.00	13.9	number of people ($\times 10^3$)
Jobs	0.77	1.42	0.30	0.00	14.5	number of jobs ($\times 10^3$)
Mobility level PT	2.32	0.29	2.28	1.75	4.44	average travel time in hours*
Mobility level car	1.20	0.08	1.19	1.06	1.46	average travel time in hours*
Job accessibility PT	35.2	57.9	13.7	0.00	387	jobs within 30 minutes ($\times 10^3$)
Job accessibility car	746	280	844	102	1223	jobs within 30 minutes ($\times 10^3$)
Costs PT	9.00	1.36	8.95	3.13	14.5	average travel costs in €*
Costs car	8.90	0.78	8.82	7.76	12.0	average travel costs in €*
Inhabitant density	3.97	4.23	3.25	0.00	27.8	people ($\times 10^3$) per square kilometer
Job density	1.97	3.74	0.75	0.00	44.26	jobs ($\times 10^3$) per square kilometer
Average household size	2.30	1.23	2.24	1.00	36.6	people per household
Age 0-18	0.19	0.08	0.19	0.00	0.85	fraction relative to all inhabitants
Age 18-34	0.20	0.13	0.18	0.00	1.00	fraction relative to all inhabitants
Age 35-54	0.29	0.09	0.29	0.00	1.00	fraction relative to all inhabitants
Age 55-64	0.13	0.07	0.13	0.00	1.00	fraction relative to all inhabitants
Age 65^+	0.18	0.12	0.17	0.00	1.00	fraction relative to all inhabitants
Male	0.51	0.07	0.50	0.00	1.00	fraction relative to all inhabitants
Average car ownership	0.48	0.17	0.47	0.00	2.00	cars per inhabitant
Average income	38.4	9.53	37.2	11.4	101	yearly income (€$\times 10^3$) per capita

Note: ** arithmetic average for a particular zone is computed over the travel times (or costs) to all other model zones (comprising the whole of the Netherlands and a small fraction of zones outside of the Netherlands).*

To limit the amount of correlation among the socioeconomic variables and consequently improve the final parameter estimates, we aggregate seven age cohorts into two age cohorts; all cohorts below 18 years, and all cohorts above 64 years. These ranges are chosen based on previous research, Martens et al. (2019) and Shay et al. (2016) both identity young people (under driving age) and seniors (62 years and older) as having a higher risk of experiencing transportation poverty. Instead of 62 years and older, we choose 65 years and older since this is the closest available separation in the data. The age cohorts aggregation means that we implicitly assume there are no differences in the experienced transport poverty environment between people of different ages who are above 64 years of age and similarly for all people under 18 years of age. We think that these are reasonable assumptions as household composition changes leading to a change of residence (e.g., marriage, divorce, parenthood, moving out from parental home) arguably do not occur as much for the aggregated age cohorts. Moreover, as aforementioned, previous research has also aggregated these age intervals (Martens et al., 2019; Shay et al., 2016). We include the other three age cohorts in the raw data (18 to 34 years, 35 to 54 years, and 55 to 64 years) as separate cohorts in our analysis. For all cohorts, we use fractions relative to the total number of inhabitants of the respective neighborhood.

Regarding gender, we use the fraction of males among all inhabitants, which means that we don't account for gender and age combinations but instead focus on the correlations between transport poverty, and age and gender in general. This aggregation of all the male age cohorts is motivated again by improved parameter estimates from limiting the amount of correlation among the socioeconomic variables.

It can be argued that, in our context, household structure is to a certain extent reflected through household size. The number of people comprising a household has a very direct relationship to household structure because 93% of households in the Netherlands consist of either a single person, a partner duo living together, or a partner duo with one or more children (CBS, 2022). However, the average household size might not give a good indication of household structure if a neighborhood is relatively heterogeneous in terms of household sizes and structures because, for example, single and three-person households in similar proportions could give the false impression that the two-person partner duo household is an appropriate characterization for the households within a zone. Concluding, the relationships between household size and transport poverty indicators might also include correlations relating to household structure, to a somewhat limited extent.

4.2 Built Environment Data

We use job density and inhabitant density to reflect the relationship between the built environment and the transport poverty environment. To obtain job density, we first aggregate all seven job sectors by computing the total number of jobs for each zone to exclude any correlation among the job sector variables and hence improve the final parameter estimate. This way we can accurately determine the relationship between general job density and transport poverty, without obscurement when multiple job sectors display similar spatial distributions.

A limitation is that we consequently neglect the potential heterogeneous location preferences among job categories. For example, industrial jobs might want to locate further away from residential areas to limit air and noise pollution, whereas service industry jobs would want to be located close to residences to increase the number of potential customers. However, job density would still give a good indication of a substantial amount of the travel demand associated with a zone in the sense that workers will move to and from work. This, in turn, naturally correlates with mobility, accessibility, and affordability measures. We divide the total number of jobs over the respective neighborhood area to acquire job density levels.

Another measure used in the Netherlands attempting to reflect the amount of human activity within a certain region is the address density (Den Dulk et al., 1992). Address density, defined as the number of addresses per square kilometer, is shown to significantly inversely affect the likelihood that an individual will work outside of their home municipality, for the Netherlands (Susilo & Maat, 2007). In line with this, Susilo & Maat (2007) find evidence that higher address density significantly reduces commuting distance. Furthermore, they observe that female gender, a higher income level, and a high education level are all correlated with higher commuting travel time. Because address density is unavailable at the considered neighborhood zone level, we include job density and inhabitant density in the regression analyses.

4.3 Transport Poverty Data

The transport poverty data consists of a car measure and a public transport measure for each of the three transport poverty dimensions (mobility, accessibility, and affordability), thus six in total. To construct these indicators, we use skim matrices of travel times and travel costs between zones, and the total number of jobs attributed to each zone. The considered period for the skim matrices is a morning rush hour (7:00-9:00) on an average workday.

The skim matrices result from a local traffic model used by transportation planners in the Amsterdam Metropolitan Area. The model is called VENOM (Vervoerregio Amsterdam, 2022) and it relies on traffic modeling software named OmniTRANS developed by Dat.mobility (Goudappel, 2022). OmniTRANS is a traditional 4-step transportation model and includes freight traffic, passenger car, bicycle, train, bus, tram, and metro as modes, walking is the access and egress mode when applicable. Multi-modality is modeled and its results are expressed as general public transport skim matrices as if it were a single mode. Additionally, results for car and bicycle users are generated. The skim matrices for financial costs and travel times consist of the respective values between every zonal pair, where costs concern single trip costs and thus do not take ownership of transport means into account. Since ownership costs are not included, the costs of using a car would be higher in reality whereas the used public transport costs do reflect the experienced costs. For more details see Smits (2011).

The indicators are computed for all zones within the study area, zones outside of the study area have a certain influence depending on which indicator is concerned. Firstly, for expressing mobility, we follow Martens & Bastiaanssen (2019). This entails a simple arithmetic average of the zone-to-zone travel times to express potential mobility for a particular zone. All model zones (both inside and outside of the study area) are involved in the computation of this indicator to evaluate mobility for the study area zones as all zones contain potential destinations, but not every unit of area affects the average travel times equally.

The zonal area is generally smaller for zones located in urban areas and these areas exist mostly in the middle western part of the study area. Zonal density in these urban areas is thus relatively high. Since every zone is weighted equally in the average travel time computation, travel times to an urban area will generally outweigh travel time to a rural area of the same surface area. Similarly, it holds in general that the further away a zone is from the study area, the larger its size and the lower the importance of locations within the zone. Hence, we do account for the superior relative importance of urban areas as opposed to rural areas because urban zones generally contain more potentially interesting destinations (e.g., shops, leisure, social contacts) per unit of land area than rural zones and we account for the lower importance of destinations located further away.

The geographical distribution of the mobility indicator for public transport (i.e, average travel times), as depicted in Figure 5, shows that lower travel times mostly exist along the rail lines and in the vicinity of the larger cities, whose names are shown. Moreover, the eastern part of the study area exhibits substantially higher public transport travel times. This eastern disadvantage also shows for the average car travel times, whose spatial distribution among the study zones is visualized in Figure 6. In contrast to public transport, lower car travel times are less concentrated along the main infrastructure (highways for cars) and the travel times in most of the southern part of the study area are at the lower end of the spectrum. By comparing the legends of Figure 5 and Figure 6, the substantial relative advantage of the car in terms of travel times becomes apparent, which is generally only half of the public transport travel time.

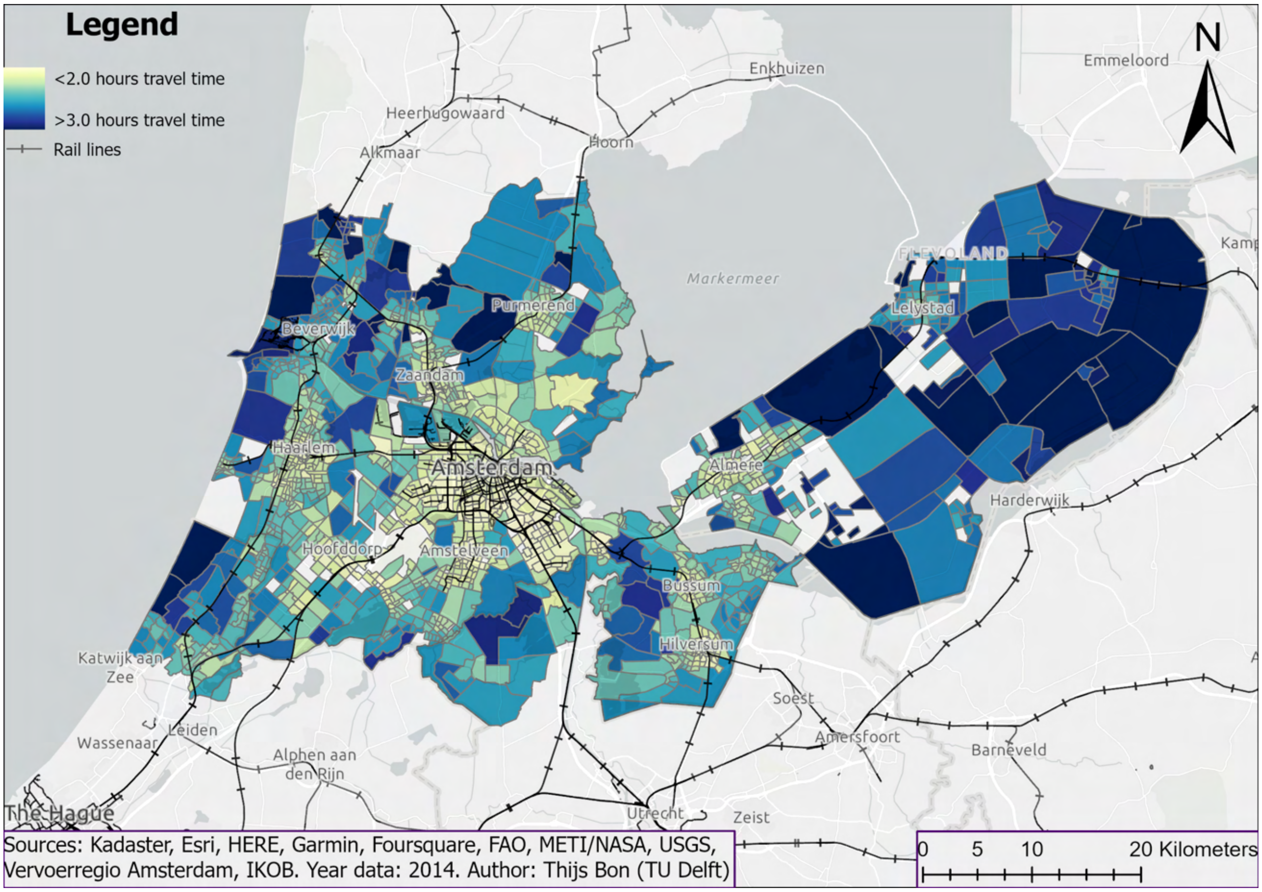

Figure 5: Travel time average public transport.

Secondly, the affordability measures are expressed in a similar manner as the mobility measures, taking the arithmetic average of the zone-to-zone single-trip travel costs for each study area zone. Again for all zones within the study area, averaging over all model zones. Around 13% of the public transport costs skim matrix values equal zero, a substantial amount. For neighboring zones this might be representative of the experienced travel costs, however, values for pairs located geographically far from each other are disproportionally overrepresented here. Upon further inspection, we note that zonal pairs with one or more

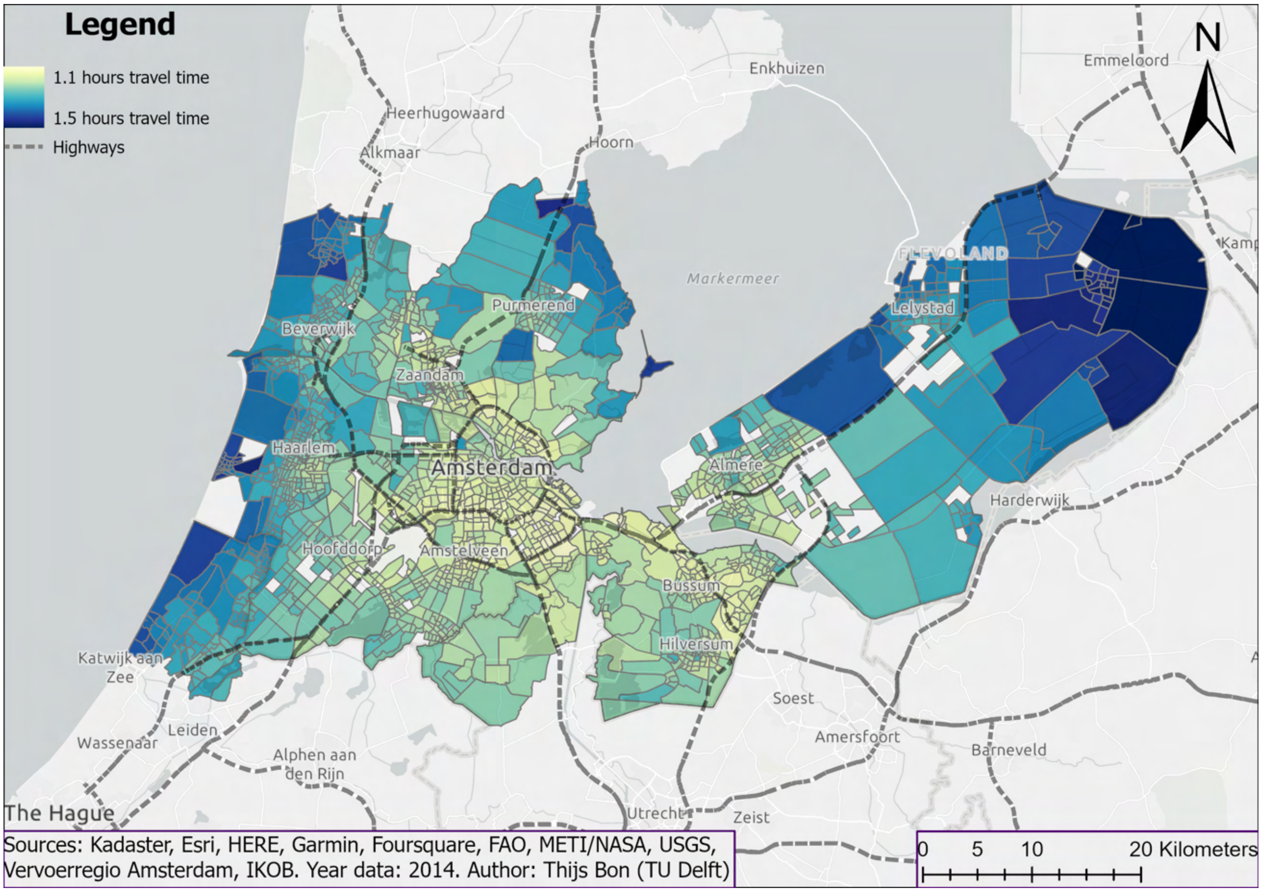

Figure 6: Travel time average car.

zones outside of the Netherlands or with one or more zones in remote areas of the Netherlands comprise most of these unrealistic zero-value pairs. Presumably, the local traffic model is unable to compute public transport travel costs when remote and foreign zones are involved due to a lack of information on service provision or even the impossibility to reach these zones using public transport. The averages are calculated using only valid values. Because the discounted values are expected to have generally higher costs associated with them, the cost averages will be an underestimation and should be interpreted as a lower bound of the public transport costs.

The misrepresentation of the public transport travel costs for certain zones due to the invalid values can be observed in Figure 7, where remote zones exhibit relatively low travel costs. For example, the peninsula zone in the Markermeer lake (northeast of Amsterdam) has the lowest travel costs among all study area zones. This misrepresentation happens because only 10% of the cost pair values of this zone are valid and they are mostly located close to the peninsula. Luckily, as on average 87% of the costs values are valid, most of the study area zones are affected to a much lesser extent and relative differences between zones are still well represented since the invalid values occur mostly in the same manner, i.e., when a zone with many invalid value pairs is involved.

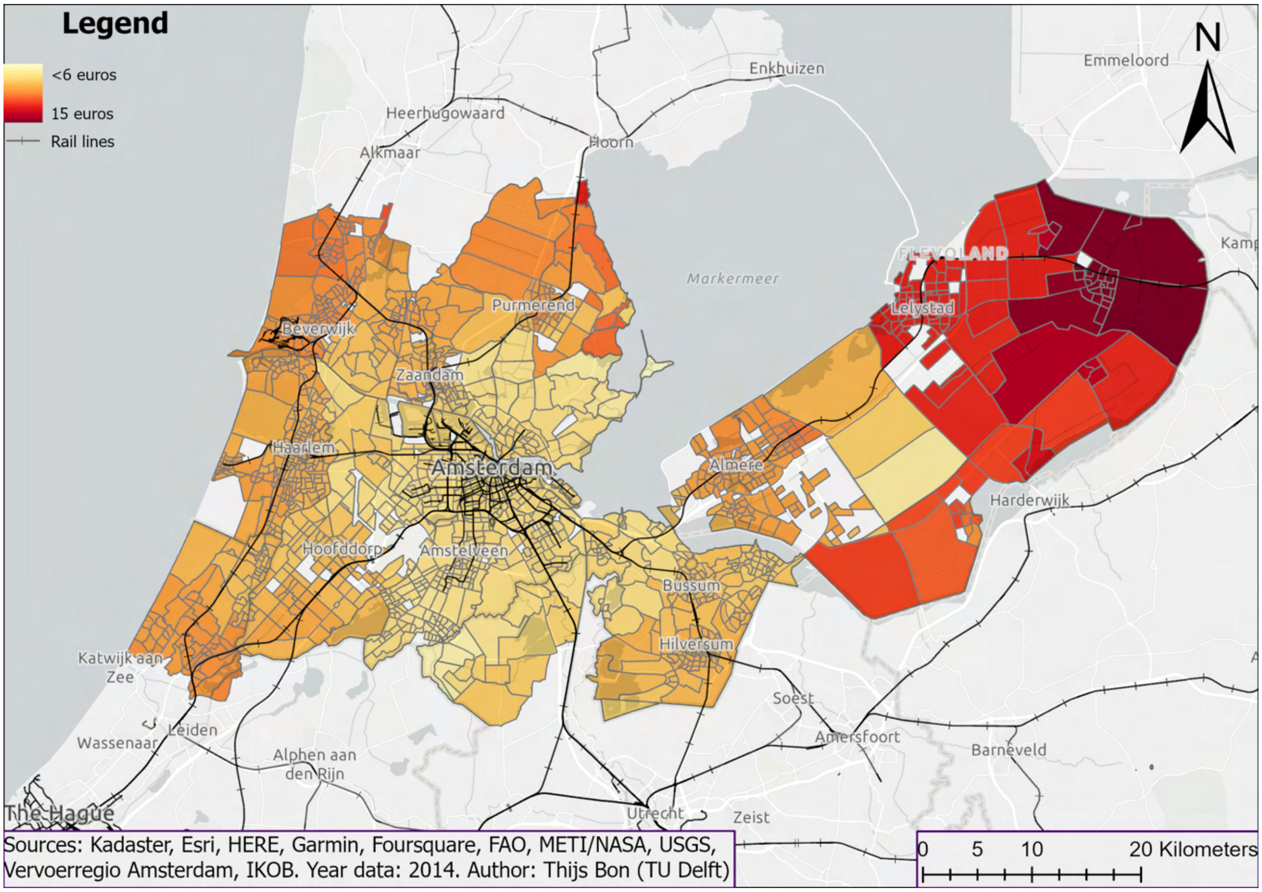

Figure 7: Costs average public transport.

In contrast with the concentration of lower travel times along the rail lines, we conclude from comparing between travel times in Figure 5 and travel costs in Figure 7 that lower public transport travel costs are more dispersed within the study area. This can be ascribed to the fact that travel costs are based on travel distance (Translink, 2021) and rural areas are often solely serviced by relatively slower transportation modes such as the bus. It stands to reason then that centrally located zones, relative to all other model zones, exhibit the lowest travel costs, which is in line with the pattern observed in Figure 7 where the lowest travel times are concentrated around Amsterdam.

Relatively lower travel costs occur in centrally located zones for car transport too, as can be seen in Figure 8. Compared to public transport, the lower car travel costs are located somewhat more towards the south of the study area. A reason for this might be the many invalid values for public transport costs associated with zones outside of the Netherlands or in remote areas, these zones are generally closer in travel distance to the southern regions. Another factor might be that car transport generally follows a more direct line between origin and destination compared to public transport since public transport users are dependent on predetermined routes and car users have the freedom to determine their own route. As a consequence, the lowest car travel costs would be corresponding to the geographical center among all model zones to a greater degree.

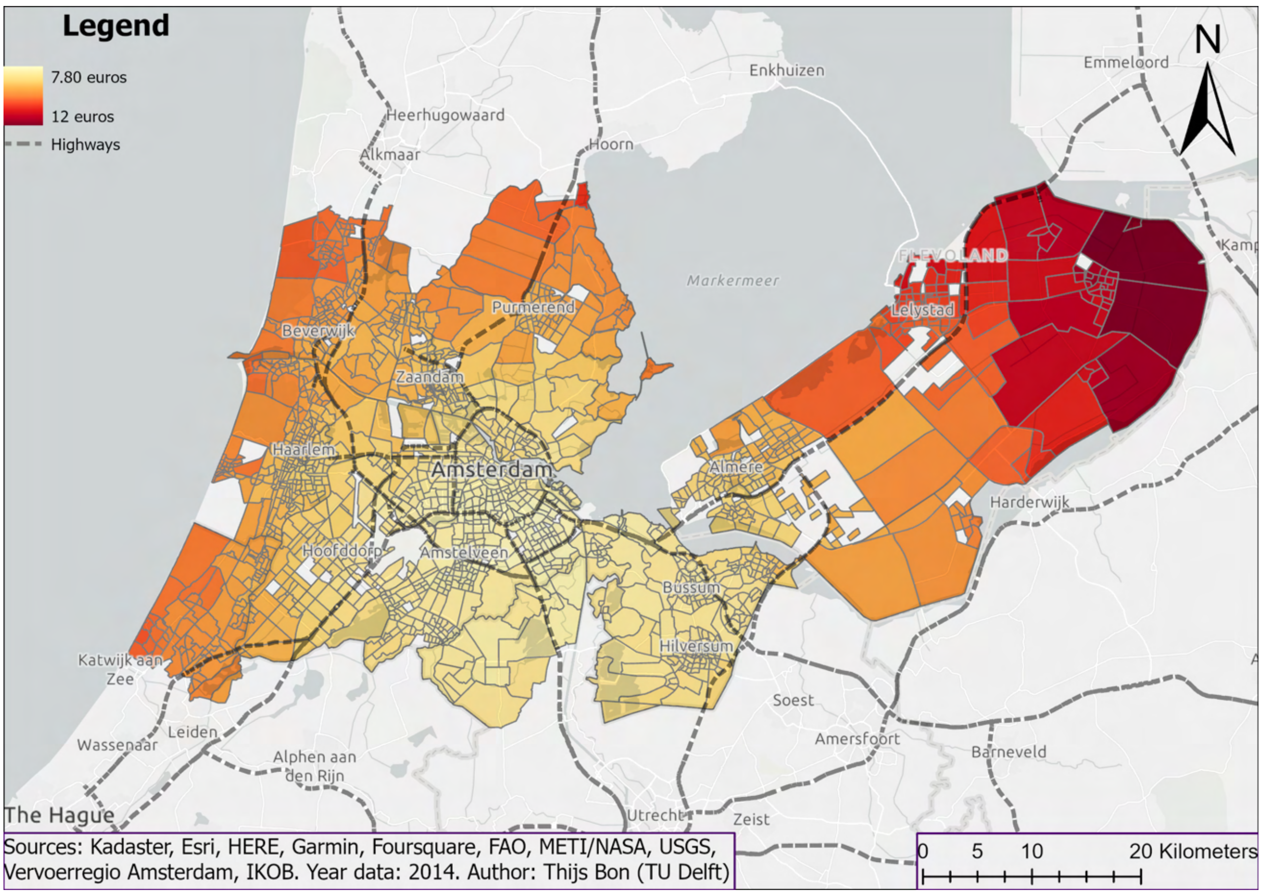

Figure 8: Costs average car.

Finally, as accessibility measures, we consider the number of total jobs within half an hour of travel time. Jobs existing in all model zones are used in computing this job accessibility indicator for the study area zones. This measure has already been used for comparable research within the Netherlands (Susilo & Maat, 2007; Martens & Bastiaanssen, 2019). and it satisfies most of the criteria for a good accessibility measure as laid out by Geurs & van Wee (2004).

However, thirty minutes of travel time is still an arbitrary threshold level—decisions of people on whether a job is considered to be accessible to them will generally not exhibit such a harsh cutoff level. Nevertheless, it will still give a good impression of relative differences among neighborhood zones because a relatively high number of jobs accessible within half an hour of travel time most likely also means a relatively high number of jobs accessible within any specified amount of travel time. Since the validity of our main results is not dependent on an absolute level of job accessibility, the relative job accessibility measure with a thirty-minute threshold level will suffice.

The regions with 100,000 or more jobs accessible within half an hour of public transport travel time are limited to Amsterdam and the surrounding area, as seen in Figure 9. Some zones within cities not far away from Amsterdam are also within the upper part of the spectrum in terms of the number of jobs within half an hour of public transport travel time. Furthermore, many regions have relatively very few jobs considered accessible through public transport. Half of the zones have 13,700 or fewer jobs within half an hour of public transport travel time, see the corresponding median value in Table 1 (Section 4.1). This concentration of comparatively much larger public transport job accessibility to a small portion of the zones inside the study area results in many light-colored zones in Figure 9.

It becomes apparent from Figure 10 that, in contrast with public transport, zones with a relatively high number of jobs accessible within half an hour of car travel time are substantially more dispersed throughout the study area. Especially striking is that many of the zones with relatively many jobs within thirty minutes of travel time are located in the south of the study area. Many jobs are located in the Randstad area, which is the main economic area in the Netherlands of which Amsterdam, Utrecht, The Hague, Rotterdam, and some smaller cities in between are a part. Since the Randstad cities other than Amsterdam are situated roughly to the south of the study area, it is very likely that the jobs from those zones strongly affect the southern region of the study area in terms of jobs accessible within half an hour of car travel time. Hence, the southern region is less dependent on Amsterdam for high job accessibility by car.

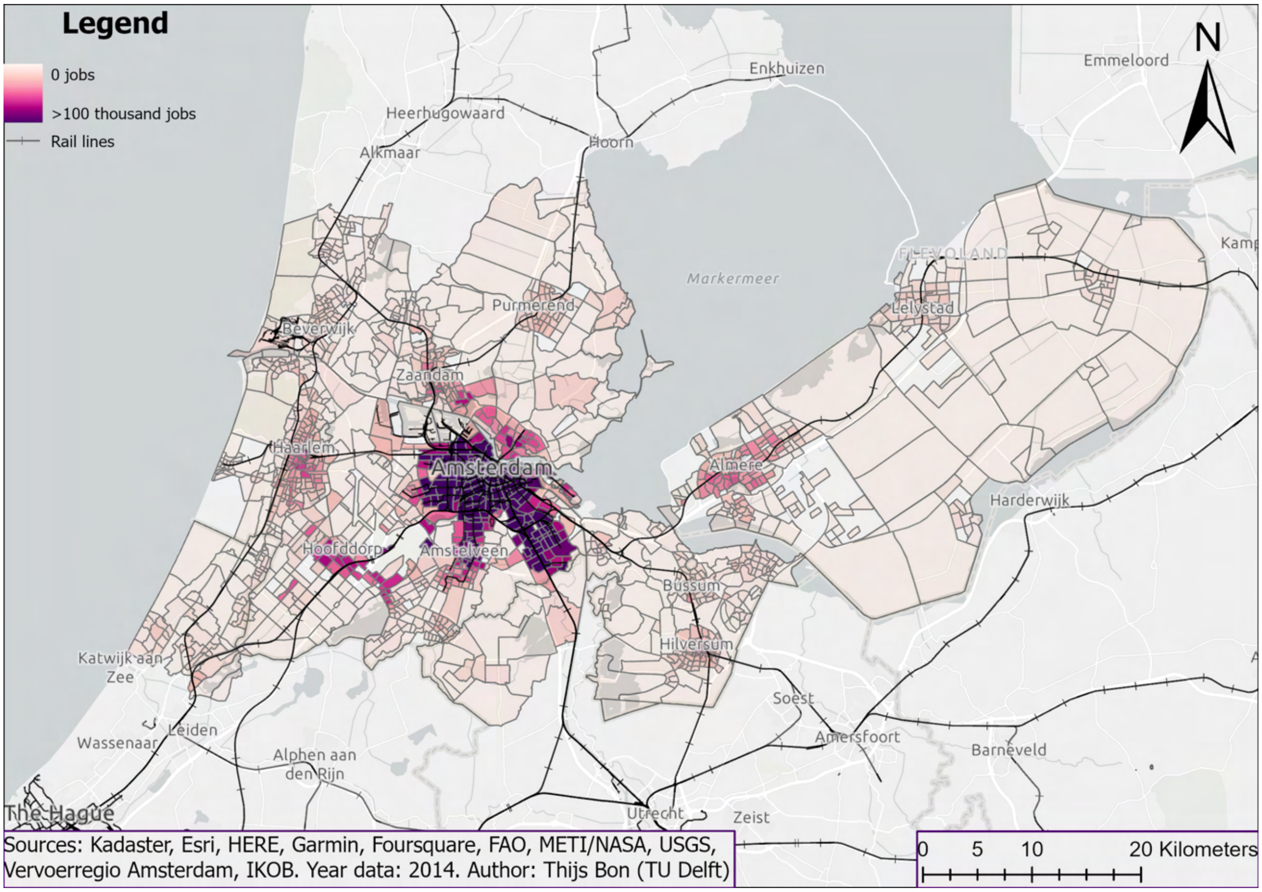

Figure 9: Number of jobs within half an hour of public transport travel time.

Comparing the number of jobs accessible within half an hour of travel time between car and public transport, we notice a very substantial advantage of the car. The number of jobs within thirty minutes of public transport travel time for zones with relatively high public transport job accessibility is around the same level as the number of jobs within half an hour of car travel time for zones exhibiting the lowest amount of job accessibility by car. In Table 1 (Section 4.1) we observe that the minimum number of jobs accessible within thirty minutes of travel time by car among all study area zones is equal to 102,000. This number of jobs would correspond to the darkest color in Figure 9 and thus correspond to the upper end of the number of jobs accessible within half an hour of public transport travel time.

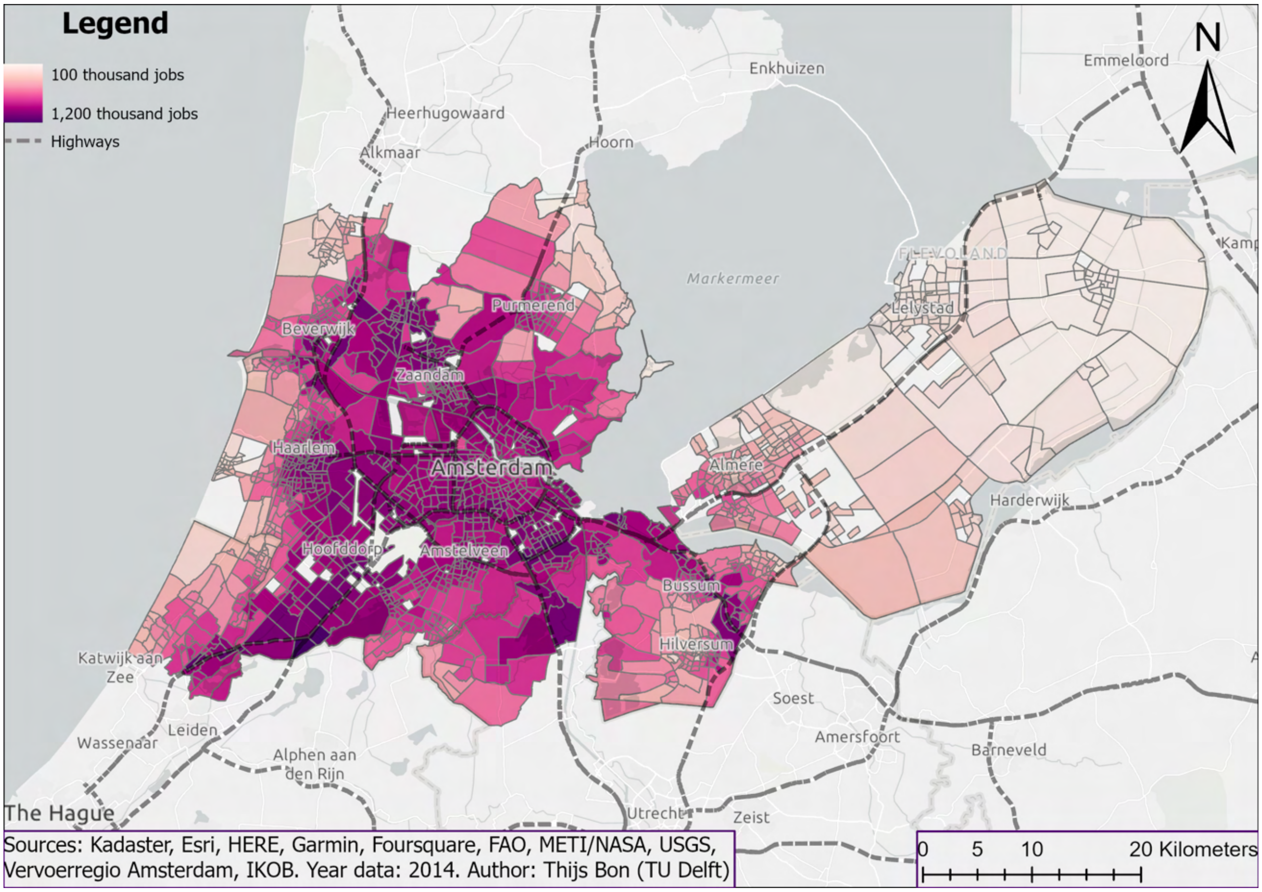

Figure 10: Number of jobs within half an hour of car travel time.

5 Results and Discussion

Since our regression results report only on correlations, we cannot exactly determine the cause of observed relationships and only provide probable reasons that are most in line with the observed correlations based on spatial distributions, logical reasoning, and other research.

The weighted least squares regression results in Table 2 show that mainly the public transport travel times and the number of jobs accessible by public transport within half an hour are relatively strongly correlated with the (estimated linear combination of) socioeconomic and built environment variables; the adjusted R^2 values of their respective regressions are equal to 0.437 and 0.859. The other adjusted R^2 statistics vary from 0.0827 for car-accessible jobs to 0.1176 for car travel costs. All regression results are obtained through weighted least squares because the White tests' null hypotheses of homoskedastic error terms were all rejected and the studentized residuals plots for all regression and respective significant variable pairs did not reveal any model misspecification.

The comparatively low adjusted R^2 statistics for car transport poverty measures indicates that job accessibility, travel times, and travel costs by car exhibit only a weak link with the observed variation in the zonal socioeconomic and built environment characteristics. This weak link suggests that there are only small variations in the car transport poverty environment among people when distinguished by the socioeconomic and built environment characteristics of their residential zone.

By contrast, the strong link between the variation among public transport job accessibility and travel times and the variation among socioeconomic and built environment characteristics suggests that there is a substantial difference in the public transport poverty environment among people when distinguished by the socioeconomic and built environment characteristics of their residential zone. The weak link with the socioeconomic and built environment characteristics of public transport costs compared to the other public transport poverty indicators is presumably related to the fact that fees are distance-based (Translink, 2021). As such public transport costs are mainly determined by whether a connections exists and are independent of the frequency and travel time of the particular connection.

Further evidence for small variations in public transport costs and car transport poverty measures among zones are the geographical distributions shown in Figure 6, Figure 7, Figure 8, and Figure 10 (see Section 4.3). These figures depict that there is substantially less variation among the study area neighborhood zones in terms of (respectively) car travel time, public transport costs, car transport costs, and car job accessibility—when compared to the geographical distributions of public transport travel time and public transport job accessibility, respectively visualized in Figure 5 and Figure 9 (see Section 4.3). Moreover, we observe relatively low variation levels of transportation affordability and car-related transport poverty measures from the descriptive statistics in Table 1 (Section 4.1).

The results thus suggest that the included built environment and socioeconomic characteristics of a neighborhood zone in the Amsterdam Metropolitan Area are quite unrelated to the transport poverty measures for car and public transport costs, whereas these characteristics do play an important role in public transport travel times and for the number of jobs accessible through public transport within thirty minutes.

Table 2: Regressions weighted least squares results summary.

Dependent var.:	Public transport: log(Travel time)		Public transport: Travel costs		Public transport: log(Accessible jobs)		Car: log(Travel time)		Car: log(Travel costs)		Car: Accessible jobs	
	Estimate	t value	Estimate	t value	Estimate	t value	Estimate	t value	Estimate	t value	Estimate	t value
log(Inh. density)	-0.010^{***}	-7.96			0.333^{***}	29.4						
	(0.001)				(0.011)							
log(Job density)	-0.033^{***}	-22.9	-0.087^{***}	-5.11	0.521^{***}	35.7	-0.009^{***}	-12.1	-0.014^{***}	-13.3	35.12^{***}	9.94
	(0.001)		(0.017)		(0.015)		(0.001)		(0.001)		(3.53)	
Household size	0.010^{***}	4.53	0.175^{***}	4.58	-0.077^{***}	-2.69	0.004^{***}	2.60	0.006^{***}	3.18	-16.89^{**}	-2.40
	(0.002)		(0.038)		(0.029)		(0.001)		(0.002)		(7.03)	
Age 0-18			1.816^{***}	4.57								
			(0.398)									
log(Age 18-34)			-0.216^{***}	-5.90								
			(0.037)									
Age 35-54			-2.486^{***}	-6.47								
			(0.384)									
log(Age 55-64)	0.008^{***}	3.67			-0.240^{***}	-12.6	0.003^{**}	2.46			-21.62^{***}	-3.33
	(0.002)				(0.019)		(0.001)				(6.49)	
log(Age 65^{+})			-0.126^{***}	-4.48							17.33^{***}	3.08
			(0.028)								(5.63)	
Male					1.017^{***}	4.97						
					(0.205)							
log(Car ownership)	0.025^{***}	10.9	0.553^{***}	10.2	-0.447^{***}	-15.6	0.008^{***}	3.65	0.008^{***}	4.38	-18.61^{**}	-2.58
	(0.002)		(0.054)		(0.029)		(0.002)		(0.002)		(7.22)	
log(Income)			-0.909^{***}	-6.23	-0.245^{**}	-2.11			-0.038^{***}	-4.26		
			(0.146)		(0.116)				(0.009)			
Constant	0.833^{***}	110	11.99^{***}	22.6	2.197^{***}	4.78	0.178^{***}	37.8	2.304^{***}	70.3	780.5^{***}	35.3
	(0.008)		(0.530)		(0.460)		(0.005)		(0.033)		(22.1)	
Adjusted R^2	0.437		0.098		0.859		0.113		0.1176		0.0827	

Note: *$^{*}p<0.1$; $^{**}p<0.05$; $^{***}p<0.01$.* The total number of observations is equal to 1704. Standard errors are given in brackets underneath the respective coefficient estimate. Var. is the abbreviation for variable and inh. stands for inhabitant.

A potential reason for the observed substantial differences in terms of correlational strength (R^2 statistics) with the socioeconomic and built environment characteristics between car travel time, car transport costs, car job accessibility, and public transport costs on the one hand and public transport travel time and public transport job accessibility on the other is that good public transport provision over large areas is much more difficult to organise than a good environment for car transport. A car user is dependent only on the presence and quality of road infrastructure, disregarding any car availability concerns since our indicators do not account for the extent to which someone has a car to his disposal. By contrast, a public transport user is dependent on frequencies, transfers, predetermined routes, and access and egress options. Therefore, it stands to reason that good public transport provision is focused on areas with relatively higher levels of job and inhabitant density, which is in line with our observations on the relative correlational strength of these built environment characteristics as well as with the adjusted R^2 statistics.

Combining the regression results and normalized data coefficient estimates as visualized respectively in Table 2 and Figure 11, we note that job density almost invariably exhibits the most significant, most precisely determined, and strongest correlation with the transport poverty measures. This is concluded by comparing, respectively, between all the t values, standard errors, and normalized data coefficient estimates related to the independent variables. The direction of the relationships is such that a higher job density is correlated with lower travel times, lower travel costs, and more accessible jobs for both car and public transport. This suggests that favorable transportation conditions lead to the concentration of jobs, or that a higher concentration of jobs leads to transportation infrastructure development (or a combination of both)—and that job density largely dominates inhabitant density and the included socioeconomic characteristics of a zone regarding the connections to the transport poverty indicators.

Inhabitant density displays the same direction of correlation with favorable transportation conditions as inhabitant density—higher inhabitant density is related to lower travel times and more accessible jobs. However, significant estimates of the relationships between inhabitant density and the transport poverty measures were obtained solely for public transport travel time and public transport accessible jobs. Hence, it is suggested that inhabitant den-

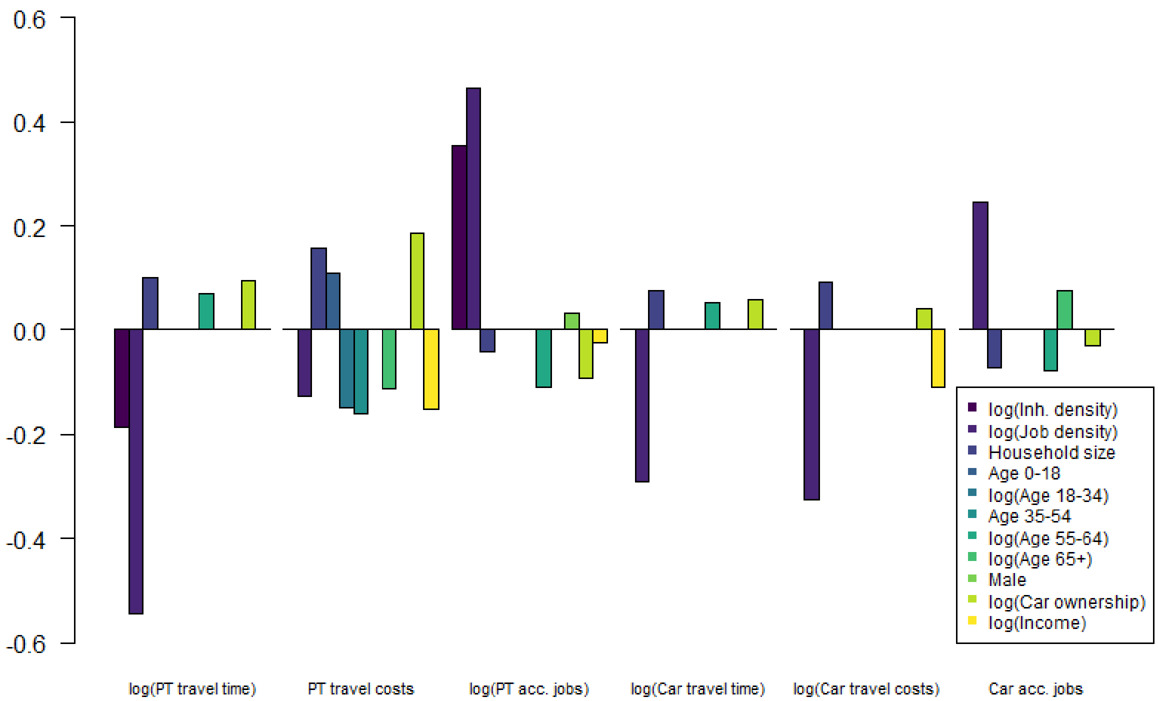

Figure 11: Coefficient values weighted least squares on normalized data. Inhabitant is abbreviated by inh., accessible by acc., and public transport by PT.

sity is an important factor relating to differences in the public transport poverty environment while this does not seem to be the case for the car transport poverty environment.

Somewhat surprisingly, all socioeconomic variables display quite weak connections to the transport poverty indicators; all absolute coefficient values of the regressions on the normalized data are smaller than 0.19, see Figure 11. A coefficient value of 0.19 in the regressions with normalized data means that a change in the corresponding independent variable equal to its standard deviation is related to a change in the dependent variable with 0.19 of the standard deviation of the dependent variable.

The weak correlations of the socioeconomic variables with the transport poverty measures might have to do with the aggregated nature of the data, differences among inhabitants of a zone are averaged out while considering the socioeconomic characteristics of an individual in relation to the transport poverty measures might reveal more information on their connections. Furthermore, our transport poverty measures account only for the resources and opportunities provided by the environment in which the neighborhood zone exists. Since the literature highlights the importance of the socioeconomic characteristics included in our analyses, see our literature review in Section 2, our results suggest that this correlation manifests to a substantially larger degree at a less aggregated level or in relation to transport poverty factors outside of the resources and opportunities existing in the environment.

The socioeconomic variable of household size is correlated with disadvantageous transportation conditions, the direction of the corresponding coefficient estimates is such that larger household size is correlated with higher travel times, higher travel costs, and less accessible jobs for both car and public transport. It is difficult to support a certain explanation for the tendency of bigger households to be situated in zones where disadvantageous transportation conditions exist based on our results. A possible interpretation is that bigger households may share transportation resources among the members of a household and thus are less dependent on the transportation environment of their residential location. It could also be that smaller households generally exhibit a higher demand for transportation since they are arguably more dependent on transport to satisfy social contact needs. Another possible explanation is that bigger households attach a higher value than smaller households to residential location characteristics conflicting with advantageous transportation conditions, such as dwelling space.

The demographic makeup of a zone in terms of age cohorts and gender proportions is mostly insignificantly correlated with the transport poverty measures. This insignificance of demographics can be seen from the large number of respective empty spaces in Table 2. This insignificance suggests that the age and gender of inhabitants are mostly irrelevant in relation to differences in the existing transportation environment.

The sole exception to the insignificance of age and gender coefficient estimates is the proportion of inhabitants in the age cohort of 55 up to and including 64, a higher proportion of inhabitants belonging to this age cohort is correlated with higher travel times and less accessible jobs, both for car and public transport. Seeing that people within this age cohort are still younger than the generally observed age of retirement in the year from which the data stems (around 65 years old in 2014), the association with less favorable transportation conditions might seem surprising. A possible explanation is that parents might move to a generally less accessible location once their children have moved out of the parental home since the demand for accessibility (to e.g., schools and social contacts) arguably becomes lower and household members can rely more on car usage.

Car ownership is found to be significant for all six transport poverty indicators, it is associated with higher travel times, higher travel costs, and fewer jobs accessible within half an hour of travel time. This correlation is less pronounced for car than for public transport, judging from the respective t values in Table 2 and normalized data regression coefficient estimates in Figure 11. Logically, the public transport environment is worse for zones where car ownership is high since bad public transport conditions might induce the necessity of a car to fulfill mobility needs or because car-owning people are less sensitive to the quality of public transport infrastructure regarding residential location choice. The observed correlation between worse car travel conditions and higher car ownership can be explained by car dependency in areas that are less accessible in general because activities are more often located in other areas that exhibit better transportation conditions.

Lastly, income correlations were found to be significant only for public transport-accessible jobs and both costs measures, with higher incomes being related to lower costs and less accessible jobs through public transport. We noted already in Section 4.3 that zones located in the center and the central southern part of the study area generally coincide with the zones associated with lower travel costs, presumably due to their relatively central location with respect to all model zones. The zones in the central and southern central parts of the study area are also zones where average income levels are higher.

5.1 Policy Recommendations

We identified from the results that differences in transport poverty among the built environment and socioeconomic makeups of the neighborhood zones in the Amsterdam Metropolitan Area mainly exist for public transport travel times and public transport job accessibility. Regarding transport poverty for car users, this does not seem to be an issue when compared to public transport poverty. Namely, the highest levels of car transport poverty correspond to the lowest levels of public transport poverty in our analysis. Therefore, the greatest potential to limit transport poverty on the neighborhood zonal level likely exists for public transport poverty aspects.

Although public transport job accessibility displays the strongest correlation with the built environment and socioeconomic characteristics, we focus on public transport travel times. The reason for this is that the strong relationship between job density and job accessibility is somewhat self-evident and largely independent of the transportation system for locations that are very close (as the crow flies) to a lot of jobs. Therefore, travel times will give a better indication of how well the transportation system functions. Moreover, the log-transformed measures of public transport job accessibility and public transport travel time averages are highly correlated (Pearson correlation coefficient of -0.78) and thus are very similar.

Job density and inhabitant density were shown to display respectively the strongest and second strongest relationship with the public transport travel times. Both higher job density and higher inhabitant density are linked to lower public transport travel times. Public transport poor zones with relatively high levels of job or inhabitant density deviate from the established strong relationships and could be interesting areas to improve public transport provision because they are relatively transport-poor despite high levels of job or inhabitant density.

We consequently end up with three focal measures: job density, inhabitant density, and average public transport travel times. For each of the three focal measures, we determine three relative levels: low, medium, and high. Each level contains a third of all study area zones and all levels are relative to all other study area zonal observations. For example, the lowest job density level contains all study area zones that fall into the lowest one third in terms of job density.

It should be noted that job density exhibits a substantially stronger correlation with public transport travel times compared to inhabitant density, suggesting that lower public transport travel times are markedly more important for jobs than for inhabitants. Hence, we suggest attaching more importance to lowering public transport travel times to job locations compared to residential areas when designing policy measures to limit transport poverty at the neighborhood zonal level.

The zones that are most interesting for public transport provision improvement are colored

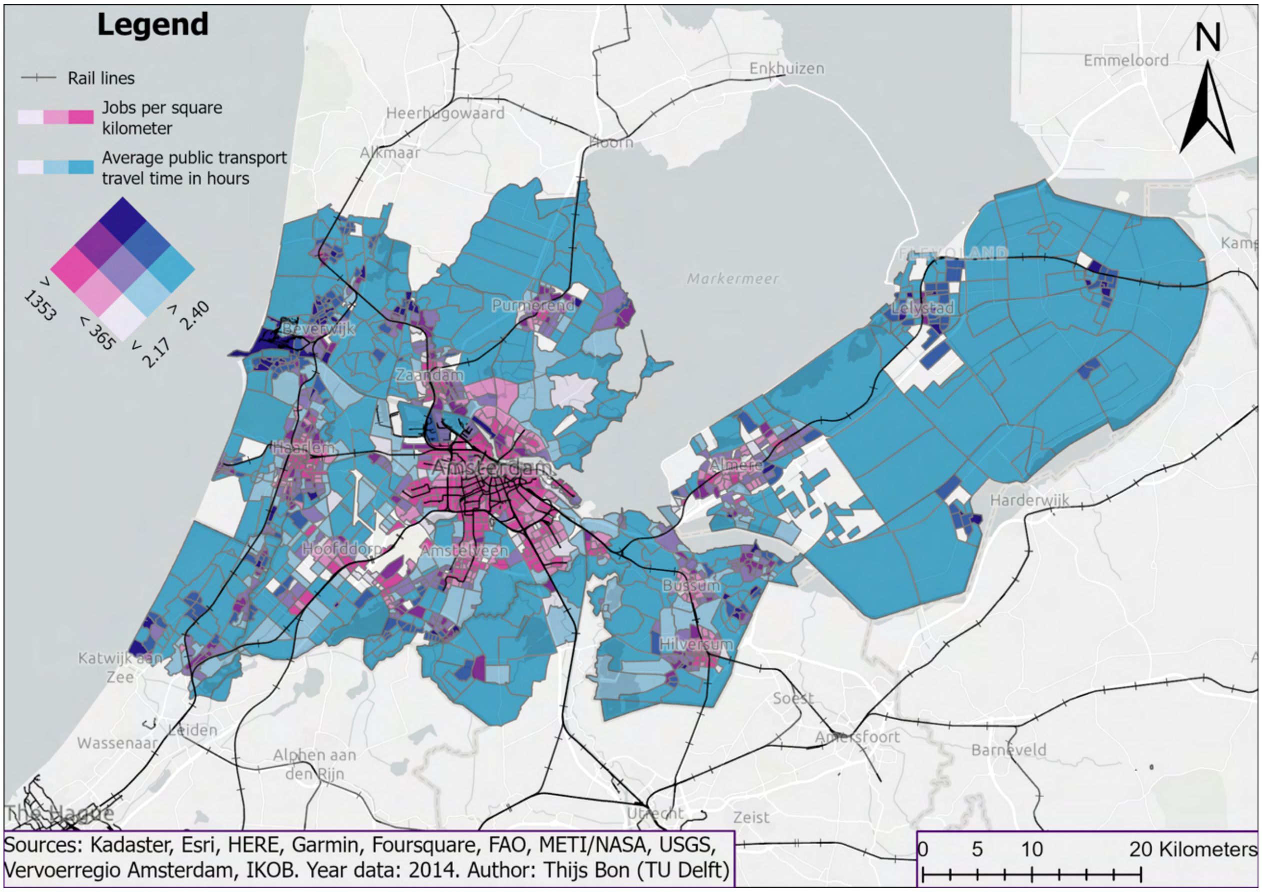

Figure 12: Travel time average public transport and job density.

dark blue in Figure 12 and in Figure 13. These zones exhibit simultaneously relatively high travel times and a high job density, and relatively high travel times and a high inhabitant density, respectively. For instance, the zone to the south-west of Beverwijk which contains the Tata Steel factory has a high level of job density and at the same time a high level of average public transport travel time. The closer the color of the zone is to dark blue, as may be identified using the legend, the more interesting it is to consider public transport provision improvement.

Finally, we find that there is only a very weak correlation between the transport poverty environment and the socioeconomic characteristics of household size, age, gender, car ownership, and income. Nevertheless, it is probably still effective to design certain policy measures targeting the neighborhood transportation environment at locations where inhabitants are at a greater risk of experiencing transport poverty based on these socioeconomic characteristics. For instance, low-income groups often depend on public transport as they cannot afford a car

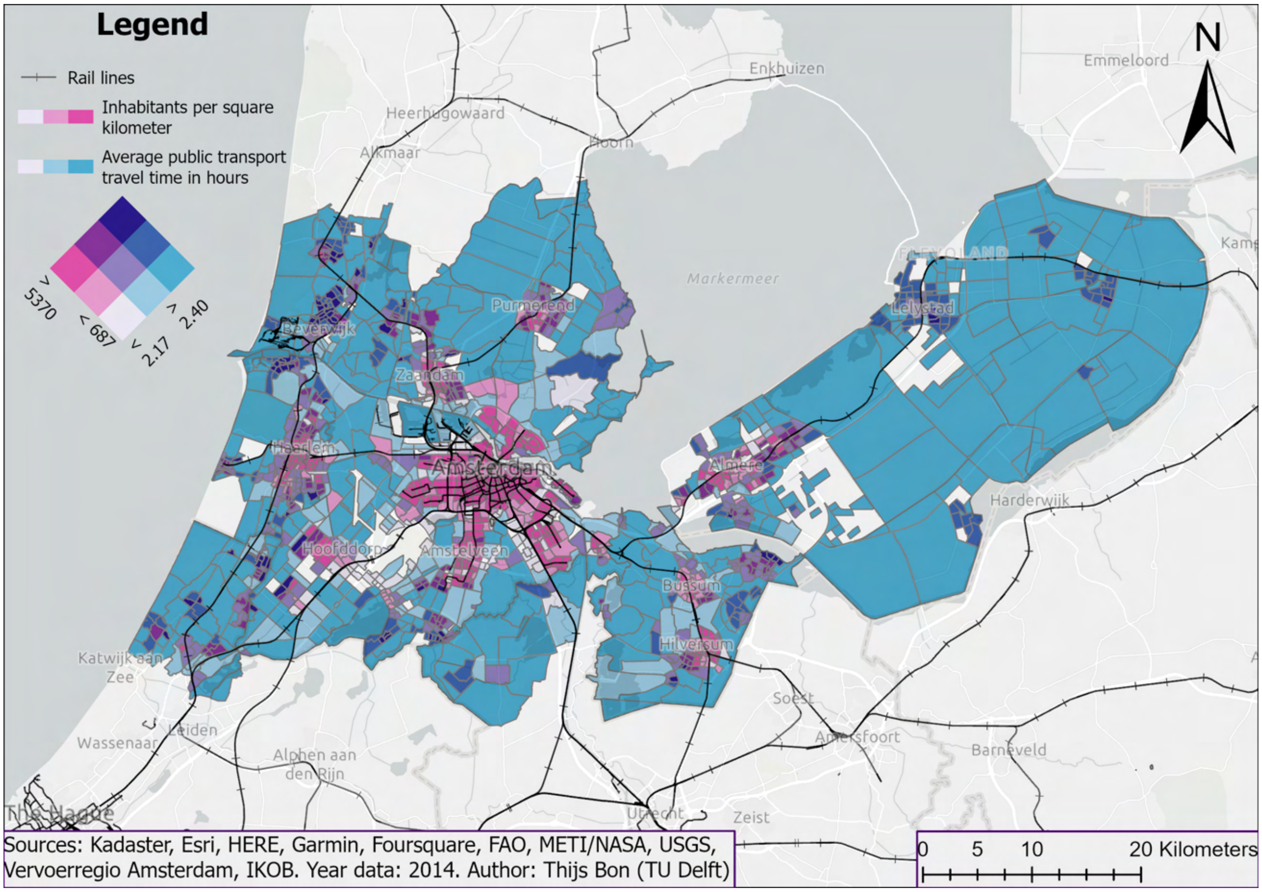

Figure 13: Travel time average public transport and inhabitant density.

or prefer to not own a car for financial reasons. Therefore, policy measures limiting transport poverty for areas with a relatively high concentration of low-income inhabitants are arguably very effective.

6 Conclusion

Transport poverty has been shown to worsen well-being (Awaworyi Churchill & Smyth, 2019; Awaworyi Churchill, 2020; Delbosc & Currie, 2011) and lower employment probability (Bastiaanssen et al., 2020). Since well-being and employment levels influence public policy, it is important to investigate how transport poverty is related to factors shown to influence it or vice versa.

Much of the existing research identifies the importance of socioeconomic and built environment characteristics in relation to transport poverty (Martens et al., 2019; Fan & Huang, 2011; Shay et al., 2016; Jomehpour Chahar Aman & Smith-Colin, 2020; Voerknecht, 2020; Kampert et al., 2019; Lucas et al., 2018). Often, these scholars give careful consideration to the selection of the (combination of) characteristics based on logical reasoning. However, research into how different socioeconomic and built environment factors are related to each of the distinct dimensions of transport poverty in terms of strength, significance, and direction has not yet been conducted.

Our transport poverty measures consist of statistics on travel time, travel costs, and the number of accessible jobs within thirty minutes, for both public transport and car. To decide which built environment and socioeconomic characteristics to include, a literature review is conducted that identifies which characteristics are often related to transport poverty. Regarding the built environment; we include inhabitant density and job density, while the socioeconomic characteristics in our analyses consist of household size, five age cohorts (0-18, 18-34, 35-54, 55-64, and 65^+), gender, car ownership, and income.

We establish the significance, strength, and direction of correlations between transport poverty indicators on the one hand and built environment and socioeconomic variables on the other for neighborhood zones within the Amsterdam Metropolitan Area through a combination of logarithmic variable transformation, insignificant variable elimination, data normalization, and weighted least squares. Additionally, we use spatial distributions as evidence to support explanations of the established correlations and to provide policy recommendations.

We find that public transport travel times and especially public transport job accessibility exhibit high correlations with the estimated linear combination of the built environment and socioeconomic variables, the adjusted R^2 values of their respective regressions are equal to 0.437 and 0.859. The car transport poverty indicators and public transport costs display much lower adjusted R^2 values of around 0.1, which is likely related to much less variation in these transport poverty indicators among the study area zones when compared to the other two transport poverty measures.

Furthermore, higher job density is shown to be associated with favorable transportation conditions for all transport poverty measures and to substantially outweigh inhabitant density and the socioeconomic variables in terms of significance and strength for almost all transport poverty measures, which suggests focusing mainly on job density among the included variables when analyzing the existing differences in the transport poverty environment on the neighborhood level. Still, inhabitant density also seems to play a noteworthy role in the observed differences in transport poverty levels, a higher inhabitant density is also linked to favorable transportation conditions, but only for the public transport poverty measures of travel time and job accessibility.

By contrast, all included socioeconomic variables seem rather inconsequential in relation to the observed differences in levels of environmental transport poverty among the study area zones because correlations are mostly either insignificant or relatively weak. Nonetheless, we report the significant relationships that we found. Household size and car ownership are correlated with a disadvantageous transportation environment for all six transport poverty measures. The correlation with worse transportation conditions applies for the age cohort of 55 up to and including 64 as well, although the coefficient estimates relating to the cost measures were found to be insignificant. Higher average income levels were observed to be associated with lower costs for both car and public transport, which seems mainly caused by the simultaneous occurrence of both higher incomes and lowers costs in the central and southern central part of the study area.

The coefficient estimates for the remaining demographic variables on gender and age cohorts (0-18, 18-34, 35-54, 65^+) proportions of inhabitants were almost exclusively insignificant in relation to the six transport poverty measures, which suggests that they are mostly irrelevant regarding observed differences in the existing transport poverty environment at the neighborhood level.

The established relationships between the built environment and socioeconomic variables, and the transport poverty indicators apply to transport poverty resources and opportunities within the context of neighborhood zonal averages in the Amsterdam Metropolitan Area. Since we use aggregate neighborhood zonal data, we implicitly assume the transportation environment measures to be defined uniformly across the heterogeneous group of people within a zone, whereas travel times and costs might actually differ among these people. Moreover, differences in the socioeconomic characteristics of the zonal inhabitants are averaged out and the average measure is assumed to apply to all inhabitants.

Other than using average measures for potentially heterogeneous groups of people, there are a few more limitations that we would like to mention. First, subjective perceptional measures that influence how people experience transport poverty, such as safety and comfort, are not taken into account. Second, transport poverty outcome measures (e.g., trip frequency, activity participation, and social isolation) are not included in our analyses. Third, externalities like traffic safety, noise pollution, and air pollution, which (partly) result from the transportation environment are not included. These transport externalities might also be included in the definition of transport poverty, their potential relevance may be illustrated using an example; people might opt to live in the vicinity of a busy highway or train station due to the associated favorable accessibility, costs, and travel time levels. Finally, our findings are limited to a somewhat small area within the Netherlands.

Considering the mentioned limitations, we have several recommendations. First and foremost, further studies may study the relationships between socioeconomic and built environment characteristics, and transport poverty at the individual level. Studying the individual level might reveal the strength, significance, and direction of the supposedly existing strong relationships between socioeconomics and transport poverty that we were unable to find. Considering the individual level will, among others, open up the possibility to use (disposable) income measures in determining transportation affordability, enable the inclusion of a personal measure of car availability, and allow for differences among individuals living in the same neighborhood zone.

Moreover, it would be interesting to see whether our findings are more generally applicable, outside of our study area. Therefore, similar analyses for different regions could provide more comprehensive information, and potentially reveal the location-specific nature of the relationships between socioeconomic and built environment characteristics, and transport poverty. Lastly, the extent to which the inclusion of measures of subjective perception, transport poverty outcomes, and transport externalities would impact our results are potentially interesting avenues for further research.

7 Acknowledgement

We want to thank Hans Voerknecht for providing us with the data on which we based our transport poverty, socioeconomic, and built environment characteristics. Additionally, we want to thank Vervoerregio Amsterdam for providing additional data and for fruitful discussions.

References

Awaworyi Churchill, S. (2020). Ethnic diversity and transport poverty. *Transportation Research Part A: Policy and Practice*, *139*, 297–309. doi: 10.1016/J.TRA.2020.07.012

Awaworyi Churchill, S., & Smyth, R. (2019). Transport poverty and subjective wellbeing. *Transportation Research Part A: Policy and Practice*, *124*, 40–54. doi: 10.1016/J.TRA.2019.03.004

Bastiaanssen, J., Johnson, D., & Lucas, K. (2020). Does transport help people to gain employment? A systematic review and meta-analysis of the empirical evidence. *Transport Reviews*, *40*(5), 607–628. doi: 10.1080/01441647.2020.1747569

CBS. (2003). *Huishoudens.* Retrieved from https://www.cbs.nl/nl-nl/achtergrond/2003/14/huishoudens

CBS. (2022). *Huishoudens nu.* Retrieved from https://www.cbs.nl/nl-nl/visualisaties/dashboard-bevolking/woonsituatie/huishoudens-nu

Delbosc, A., & Currie, G. (2011). Exploring the relative influences of transport disadvantage and social exclusion on well-being. *Transport Policy*, *18*(4), 555–562. doi: 10.1016/J.TRANPOL.2011.01.011

Den Dulk, C. J., Van De Stadt, H., & Vliegen, J. M. (1992). *Een nieuwe maatstaf voor stedelijkheid: De omgevingsadressendichtheid [A new measure for degree of urbanisation: the address density of the surrounding area]* (Tech. Rep.). Den Haag: CBS.

Fan, Y., & Huang, A. (2011). *How Affordable is Transportation? A Context-Sensitive Framework* (Tech. Rep.).

Geurs, K. T., & van Wee, B. (2004). Accessibility evaluation of land-use and transport strategies: review and research directions. *Journal of Transport Geography*, *12*(2), 127–140. doi: 10.1016/J.JTRANGEO.2003.10.005

Godfrey, B., Armstong, B., Davison, G., Malan, J. d. V., & Gleeson, B. (2015). *Delivering sustainable urban mobility* (Tech. Rep. No. 8). Australian Council of Learned Academies (ACOLA).

Goudappel. (2022). *Traffic modelling software OmniTRANS Expert | Goudappel.* Retrieved from `https://www.goudappel.nl/en/expertise/data-and-it-solutions/traffic-modelling-software-OmniTRANS-Expert`

Groenemeijer, L., van Leeuwen, G., van der Lelij, M., & Marchal, B. (2022). *SEGs NRM 2022* (Tech. Rep.). Delft: ABF Research.

Jomehpour Chahar Aman, J., & Smith-Colin, J. (2020). Transit Deserts: Equity analysis of public transit accessibility. *Journal of Transport Geography*, *89*, 102869. doi: 10.1016/J.JTRANGEO.2020.102869

Kampert, A., Nijenhuis, J., Nijland, H., Uitbeijerse, G., & Verhoeven, M. (2019). *Indicator risico op vervoersarmoede* (Tech. Rep.). CBS.

Lucas, K., Phillips, I., Mulley, C., & Ma, L. (2018). Is transport poverty socially or environmentally driven? Comparing the travel behaviours of two low-income populations living in central and peripheral locations in the same city. *Transportation Research Part A: Policy and Practice*, *116*, 622–634. doi: 10.1016/J.TRA.2018.07.007

Luz, G., & Portugal, L. (2022). Understanding transport-related social exclusion through the lens of capabilities approach. *Transport Reviews*, *42*(4), 503–525. doi: 10.1080/01441647.2021.2005183

Martens, K. (2013). Role of the Bicycle in the Limitation of Transport Poverty in the Netherlands. *Transportation Research Record*, *2387*(1), 20–25. doi: 10.3141/2387-03

Martens, K., & Bastiaanssen, J. (2019). An index to measure accessibility poverty risk. *Measuring Transport Equity*, 39–55. doi: 10.1016/B978-0-12-814818-1.00003-2

Martens, K., Bastiaanssen, J., & Lucas, K. (2019). Measuring transport equity: Key components, framings and metrics. *Measuring Transport Equity*, 13–36. doi: 10.1016/B978-0-12-814818-1.00002-0

Moreno, C., Allam, Z., Chabaud, D., Gall, C., & Pratlong, F. (2021). Introducing the "15-Minute City": Sustainability, Resilience and Place Identity in Future Post-Pandemic Cities. *Smart Cities 2021, Vol. 4, Pages 93-111*, *4*(1), 93–111. doi: 10.3390/SMARTCITIES4010006

Shay, E., Combs, T. S., Findley, D., Kolosna, C., Madeley, M., & Salvesen, D. (2016). Identifying transportation disadvantage: Mixed-methods analysis combining GIS mapping with qualitative data. *Transport Policy*, *48*, 129–138. doi: 10.1016/J.TRANPOL.2016.03.002

Smits, E.-S. (2011). *Origin-Destination Matrix Estimation in OmniTRANS* (Tech. Rep.). Utrecht: Utrecht University.

Susilo, Y. O., & Maat, K. (2007). The influence of built environment to the trends in commuting journeys in the Netherlands. *Transportation*, *34*(5), 589–609. doi: 10.1007/S11116-007-9129-5/TABLES/5

Translink. (2021). *OV-chipkaart - Kilometertarief.* Retrieved from `https://www.ov-chipkaart.nl/zo-werkt-reizen/tarieven/kilometertarief.htm`

Vervoerregio Amsterdam. (2022). *Regionaal verkeersmodel VENOM - Vervoerregio Amsterdam.* Retrieved from `https://vervoerregio.nl/venom`

Voerknecht, H. (2020). Een nieuwe kijk op bereikbaarheid. In *Nvc 2020.*

White, H. (1980). A Heteroskedasticity-Consistent Covariance Matrix Estimator and a Direct Test for Heteroskedasticity. *Econometrica*, *48*(4), 838. doi: 10.2307/1912934

Zondag, B., & Pieters, M. (2005). Influence of Accessibility on Residential Location Choice. *Journal of the Transportation Research Board*, *1902*(1), 63–70. doi: 10.1177/0361198105190200108

Appendix

A, Study Area and the Netherlands

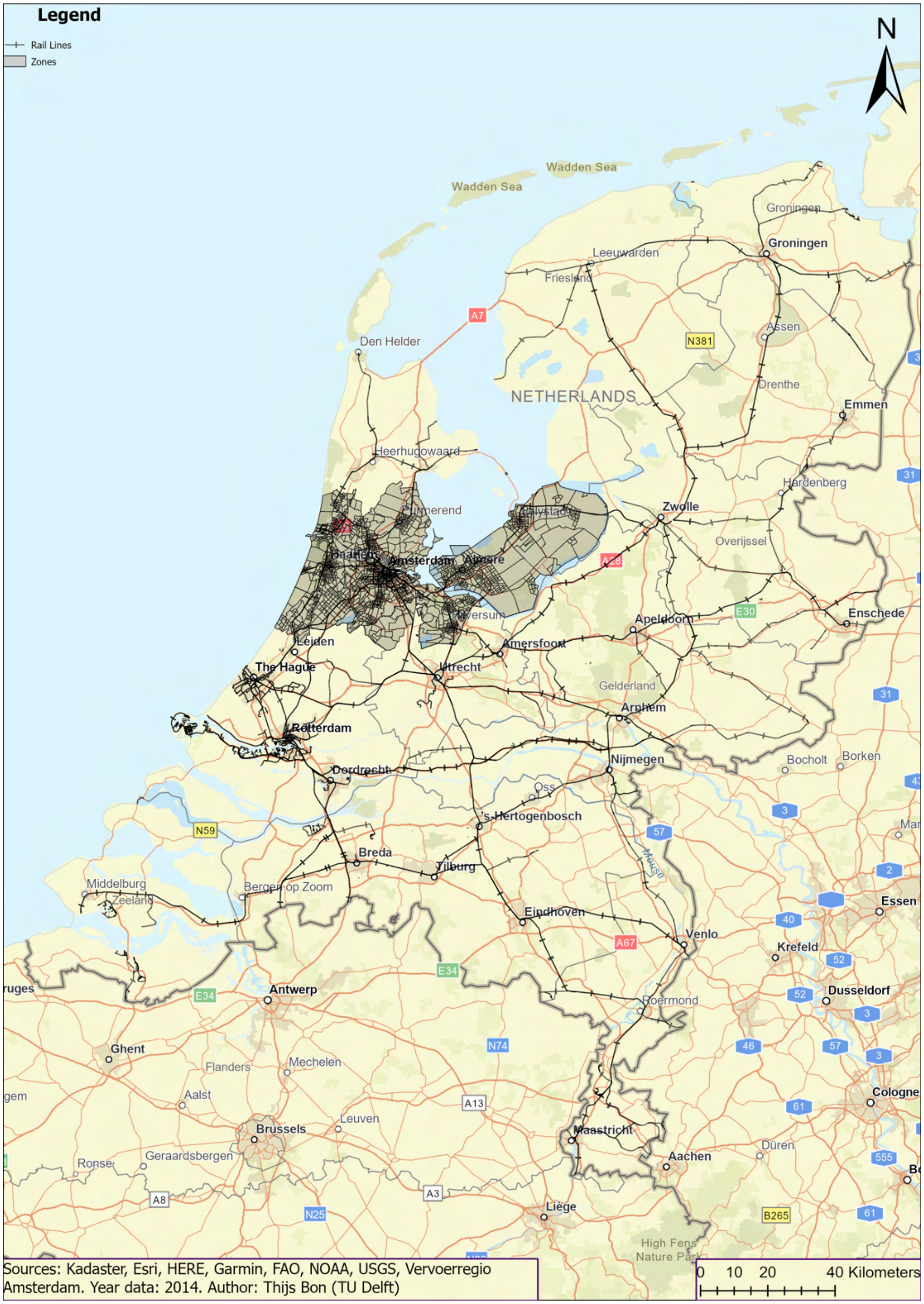

Figure 14: Amsterdam Metropolitan Area within the Netherlands.