3 The path-dependency of low-income neighborhood trajectories: An approach for analyzing neighborhood change

Merle Zwiers, Reinout Kleinhans, Maarten van Ham
Published in *Applied Spatial Analysis and Policy (2017), 10(3): 363-380.*

§ 3.1 Introduction

Socio-spatial polarization is increasing in large cities throughout Europe (Tammaru et al., 2016). Socio-spatial polarization refers to the process where the gap between the rich and the poor is increasing, which is translated into spatial segregation along ethnic or socioeconomic lines. In the European context, this has resulted in distinctive spatial patterns in large cities where the rich are increasingly located in historic city centres, while the poor reside in the more disadvantaged outer-city neighborhoods (cf. Hulchanski, 2010; Van Eijk, 2010). Despite substantial government investments to counteract such socio-spatial polarization over the past few decades, this process seems to be persistent, though it varies over time and between places (Bailey, 2012).

In most of the studies on socio-spatial polarization the continuous dynamic character of neighborhoods is neglected, reducing neighborhood change to comparing two points in time. However, neighborhoods are constantly changing in their population composition as the result of residential mobility and demographic events, thereby changing the aggregate status of neighborhoods. Many studies investigating neighborhood change focus on exceptional cases of gentrifying or declining neighborhoods (Bailey, 2012; Bailey et al., 2013; Bailey & Livingston, 2007; Clark et al., 2006; Finney, 2013; Hochstenbach & Van Gent, 2015; Jivraj, 2013; Van Ham et al., 2013). Although these studies have provided important insight in the drivers behind neighborhood change, they are typically limited to time-specific case-studies in particular cities. As a result, we do not know if neighborhoods with similar characteristics experience similar processes of change over time – or if processes of gentrification or downgrading are the exception to the rule. In addition, we have limited understanding of how processes of gentrification and downgrading affect other neighborhoods. As neighborhoods do not operate in a societal and policy vacuum, changes in one neighborhood are likely to affect other neighborhoods as well. It has, for example, been argued that processes of urban restructuring or gentrification are likely to lead to new concentrations of deprivation in other neighborhoods through the displacement of low-income groups (Bolt et al., 2009). As such, the upgrading of one neighborhood might go hand-in-hand with the deterioration of another neighborhood (Bråmå, 2013; Musterd & Ostendorf, 2005a).

In addition, many studies in this field rely on percentile shifts and point-in-time measures to analyze change, neglecting the possibility that development over time might be more non-linear than linear or need much more time to take effect (see also Van Ham & Manley, 2012). Because the physical structure of neighborhoods hardly changes, neighborhoods can maintain their overall status over longer periods of time (Meen et al., 2013; Tunstall, 2016). However, selective mobility and demographic events lead to a constantly changing population composition (Van Ham et al., 2013). In this paper we argue that to fully understand processes of neighborhood change, the next step in neighborhood research is to focus on detailed neighborhood *trajectories* and to identify typologies of neighborhood change over longer periods of time. Analyzing interrelated neighborhood trajectories and understanding why some neighborhoods are more prone to change than others is therefore highly relevant to the debate on spatial manifestations of inequality and neighborhood development.

In this paper, we present an approach for analyzing neighborhood change by focusing on long-term neighborhood change combined with a detailed analysis of neighborhood trajectories. Focusing on the trajectories of low-income neighborhoods in the Netherlands over the period 1971-2013, we analyze the role of physical characteristics in neighborhoods change. In the Dutch context, neighborhood and housing quality is often related to the debate on neighborhood change, however, few empirical studies try to analyze to what extent physical characteristics are related to today's spatial patterns. Different starting positions in terms of housing quality can have long-lasting effects on neighborhood status through processes of path-dependency (Meen et al., 2013). In addition, because the Dutch government has invested heavily in urban restructuring by changing the share of owner-occupied and social-rented dwellings in particular neighborhoods, we analyze the effect of demolition and construction on the different neighborhood trajectories. Changes to the housing stock generate mobility processes and may thus affect neighborhoods in both direct and indirect ways.

To analyze neighborhood trajectories we use a combination of methods. Sequence analysis allows for the analysis of complete pathways through time and is therefore a promising method for longitudinal neighborhood research. Sequence analysis is gaining popularity in the social sciences and is increasingly used by researchers interested in patterns of socio-spatial inequalities (e.g. Coulter & Van Ham, 2013; Hedman et al., 2015; Van Ham et al., 2014). However, sequence analysis is ultimately a descriptive method and its potential for *explaining* trajectories is limited. Researchers have therefore developed a methodological framework that combines sequence analysis and a tree-structured discrepancy analysis, allowing for the analysis of the relationship between covariates and sequences (Studer et al., 2011). As such, this framework can provide insight in how different covariates affect neighborhood trajectories in different ways. To our knowledge, this paper offers the first empirical application of this combination in the field of urban research, constituting a new approach towards researching neighborhood dynamics and a move towards the visualization and analysis of complex trajectories. In this paper, we only highlight the most important aspects of the combination between sequence analysis and a tree-structured discrepancy analysis. Based on our presentation, researchers should be able to get a basic understanding of both methods (for a full understanding of sequence analysis researchers are referred to Gabadinho et al., 2011; for a tree-structured discrepancy analysis to Studer et al., 2011).

The remainder of this paper is organized as follows. We start with expounding our approach for analyzing neighborhood change. We then move to describe the combination of sequence analysis and the tree-structured discrepancy analysis in more detail. In the data and method section, we elaborate on the structure of the dataset and the methodological choices made. We then discuss the substantive results and reflect on the applicability of the methods for neighborhood research.

§ 3.2 Longitudinal neighborhood change

Time is an important dimension in neighborhood research. There are generally two viewpoints on this: one emphasizes the general stability of neighborhood status over longer periods of time as a result of path-dependency (Dorling et al., 2007; Meen et al., 2013). Another viewpoint argues that neighborhoods are highly dynamic and are constantly experiencing population change (Van Ham et al., 2013). These two views on neighborhood change are however rather complimentary than competing. On the one hand, neighborhoods are indeed very dynamic and are constantly changing in their population composition as a result of residential mobility and demographic events. On the other hand, because the housing stock of neighborhoods is rather static, the overall socioeconomic status of neighborhoods does not change much over time. In other words: because the physical spatial structure of neighborhoods remains broadly unchanged, similar types of residents move in and out of these neighborhoods, thereby maintaining the status quo.

This is not to say that there are no changes in neighborhood status at all: neighborhoods can experience processes of decline or gentrification over time because the population in-situ experiences changes in employment status (Bailey, 2012), or because of selective out- and inflow of different income groups (Van Ham et al., 2013). However, extreme processes of decline or gentrification whereby neighborhoods experience a complete transformation of their population composition and overall status are rare (Cortright & Mahmoudi, 2014; Tunstall, 2016). Moreover, when neighborhoods experience processes of decline or gentrification, the effects of these processes on the urban mosaic are often only visible after longer periods of time (e.g. Hulchanski, 2010).

When such extreme changes do occur, they can often be explained by the physical quality of the neighborhood. Processes of gentrification have been related to the desirable location, high quality, and architectural aesthetics of pre-war or other historic neighborhoods (e.g. Bridge, 2001; Zukin, 1982; 2010). As higher income groups gradually move into these neighborhoods, housing values and prices go up, thereby pushing lower income households out. In a similar vein, many unattractive post-war neighborhoods have experienced processes of extreme neighborhood decline over the past few decades. Researchers have argued that this decline can be explained by the low quality of and technical problems with dwellings and neighborhoods built after the Second World War (Prak & Priemus, 1986; Van Beckhoven et al., 2009).

In the Netherlands, these extreme processes of neighborhood decline in postwar neighborhoods (built between 1945 and 1970) led to the development of large-scale urban restructuring programs. These urban restructuring programs were aimed at creating a social mix in these neighborhoods by demolishing social housing and constructing more upmarket owner-occupied or rental dwellings (Kleinhans, 2004). Researchers have argued that urban restructuring programs have led to minor improvements in the socioeconomic position of these neighborhoods (Kleinhans et al., 2014; Permentier et al., 2013). This can be explained by the fact that while a large number of social rented dwellings has been demolished, the overall share of social housing remained high in most restructuring neighborhoods (Dol & Kleinhans, 2012). Urban restructuring is only effective in reducing sociospatial segregation when a substantial part of the social housing stock in a neighborhood is replaced by owner-occupied dwellings (Bolt et al., 2009). Quite often (part of) the original residents in restructuring neighborhoods moved back to the newly constructed dwellings. This meant that while these neighborhoods have experienced a physical upgrade; the socioeconomic status of the population remained largely unchanged (see e.g. Kleinhans et al., 2014).

There are thus two important, yet related, gaps in the literature on neighborhood change. First, many studies focus on exceptional cases of change involving gentrification, downgrading, or urban restructuring in particular cities or neighborhoods, failing to answer the question if neighborhoods with similar characteristics experience similar changes over time. Second, few studies have analyzed the role of path-dependency of physical characteristics of neighborhoods in processes of change for a large sample of neighborhoods in different cities. As a result, we have little insight into which neighborhoods are more prone to change than others. Analyzing the effect of physical characteristics and/or physical changes on neighborhood trajectories is important for our understanding of why some neighborhoods experience change while others remain stable for longer periods of time and help to answer the question which neighborhood characteristics are predictors of future processes of change.

However, research on neighborhood change is complicated because neighborhoods have different starting positions, may experience different paces and processes of change over time, and the effects of changes in context might be non-linear or might only be visible after longer periods of time (Van Ham & Manley, 2012). To fully capture patterns of neighborhood change, it is therefore necessary to adopt a twofold approach: (1) Change should be analyzed over longer periods of time (20-40 years) to capture the effects of longer term processes; and (2) The focus should be on continuous change of neighborhood trajectories instead of simply comparing two points in time. As such, a dual approach would contribute to the identification of neighborhood change typologies providing insight in (the drivers of) different spatial dynamics.

§ 3.3 Analyzing neighborhood trajectories

The methods for analyzing trajectories are limited: the most common statistical methods treat time as another level (in multilevel models), as dummy variables (in regression models), or as growth curves (time-series models). While all of these models have their advantages and disadvantages for studying change over time, they do not easily allow for the identification of *patterns* of change. Sequence analysis, a method that originates from the biological sciences to map DNA patterns, however, allows for the study of patterns of change and is gaining increasing popularity within the social sciences because of its ability to show complete pathways. Social researchers are using sequence analysis to explore class careers (Halpin & Chan, 1998), labor market patterns (Abbott & Hrycak, 1990; Brzinsky-Fay, 2007; McVicar & Anyadike-Danes, 2002; Pollock et al., 2002), family histories (Elzinga & Liefbroer, 2007) and life-course trajectories (Billari & Piccarreta, 2005; Martin et al., 2008; Wiggins et al., 2007).

The main goal of sequence analysis is to explore trajectories of subjects (individuals, neighborhoods, et cetera) over time and to identify groups of subjects that experience similar trajectories (Gabadinho et al., 2011). Sequences are comprised of different *states* that show the order and duration that the individual subject occupied in each state. Focusing on neighborhood trajectories, a neighborhood can, for example, be in the 6th socioeconomic neighborhood category in 1971, then move up to the 5th category in 1999, and the 4th category in 2000, to end up in the 3rd category in 2013. The neighborhood categories in this example represent the different *states* that a neighborhood can move through. The *sequence* of this particular neighborhood would then look like this: 6th category-5th category-4th category-3rd category. This is an example of the most straightforward state sequence format (STS), however, other sequence representations are also possible (for a detailed understanding of state sequence representations, see Gabadinho et al., 2011). All sequences together can then be visualized as a series of individual neighborhood trajectories, which represents how each neighborhood moves through the different states over time. There are different ways to visualize sequences depending on the objective of the researcher (Gabadinho et al., 2011).

Many researchers are however interested in going a step further and explain variation in sequences. For that reason, sequence analysis is often combined with cluster analysis where similar sequences are clustered together. However, cluster analysis has several disadvantages. First of all, the clusters can be very arbitrary because different algorithms generate different clusters. In addition, cluster membership tends to be unstable and the optimal number of clusters is very difficult to assess (Studer, 2013). Cluster analysis reduces sets of sequences to a number of standard trajectories which are a rather crude approximation that consider deviations from the standard as noise (Studer et al. 2011).

In a few recent papers Studer and colleagues (2010; 2011; 2012; 2013) indicate a tree-structured discrepancy analysis as a valuable alternative to cluster analysis. The advantage of this method over cluster analysis is that a tree-structured discrepancy analysis does not create a number of groups that is supposedly representative for the entire population, instead it shows the effect of different variables on the set of sequences in a stepwise approach. Discrepancy analysis is similar to the analysis of variance (ANOVA)-types of analyses and measures the variability between sequences (Studer et al., 2011). The researcher can select a number of explanatory variables which are hypothesized to be related to the different sequences. Based on these predictor variables, the tree-structured discrepancy analysis will group similar sequences together. This is done by using the pairwise dissimilarities between sequences to compute the discrepancy within groups (Studer et al., 2010; 2011). In practice, this means that two sequences are compared to determine to what extent they are different from one another. This level of mismatch is then quantified by the dissimilarity measure (Studer & Ritschard, 2016). In this paper, we use Optimal Matching distances to quantify dissimilarity. Optimal Matching computes the distance between pairs of sequences using a chosen cost scheme. This cost scheme constitutes of (1) insertion and deletion costs (indel) which capture whether the same state occurs in two sequences, and (2) substitution costs that focus on the timing of states and whether the same state occurs at the same time point in two sequences (Aisenbrey & Fasang, 2010). Here we have set the indel costs to one and we base the substitution costs on the inverse transition frequencies between different states, which is in line with previous studies (e.g. Aassave et al., 2007; Barban, 2013; Kleinepier et al., 2015; Widmer & Ritschard, 2009). This means that we are more focused on distinct trajectories (i.e. a change from the 1st category to the 6th category is considered to be more costly than a change from the 1st category to the 2nd category) than on timing (i.e. we place less importance on differences in neighborhood states at different points in time). We have replicated our results using a different dissimilarity matrix to ensure robustness. We have used Optimal Matching with indel costs of one and substitution costs of two, which is equivalent to the Longest Common Subsequence distance (Studer & Ritschard, 2016). All of our results remain the same.

There are different ways to measure dissimilarity and the choice of dissimilarity algorithm has been subject to debate for many years (see Abbot & Tsay, 2000; Aisenbrey & Fasang, 2010; Gabadinho et al., 2011). Different dissimilarity measures focus on different aspects of the trajectories; researchers interested in change are advised to use one of the Optimal Matching algorithms; researchers focused on timing should employ one of the Hamming algorithms; while researchers interested in duration are recommended to use algorithms such as the Longest Common Subsequence or Chi^2 or Euclidian distances (for an excellent overview, see Studer & Ritchard, 2016). Optimal Matching remains the most popular dissimilarity matrix used in the social sciences because of its flexibility and can generally be used to understand the 'common narrative' between trajectories (Elzinga & Studer, 2015).

The tree-structured discrepancy analysis visualizes the relationship between predictor variables and the sequences trajectories. The tree starts with all sequences in an initial group. The tree-structured discrepancy analysis then selects the most important (significant) predictor and its most important values to split the group into two distinctly different groups using the dissimilarity measure and a pseudo R^2 and a pseudo F test. Significance is assessed through permutation tests (5,000 permutations are sufficient to assess the results at the 1% significance level, see Studer et al., 2011). Looking at, for example, the share of social housing, the model identifies the threshold value at which the sequences differ most, resulting in two significantly different groups of sequences that show different trajectories below and above the threshold value. In practice, this could mean that the model illustrates the trajectories for a group of neighborhoods with low shares of social housing and a group of neighborhoods with high shares of social housing. For each of the newly created groups, the discrepancy analysis splits the groups into two again, using the second most important predictor and its values (for that group) for which the highest pseudo R^2 is found. Using our example, for the group of neighborhoods with high shares of social housing, the model then shows the effect of a different variable, again creating two groups that show distinctly different trajectories. The process is repeated until a stopping criterion is reached or when a non-significant F for the selected split is encountered (Studer et al., 2010). The overall quality of the model can be assessed through the pseudo F test and the pseudo R^2 that provide information on the statistical significance of the tree and the part of the total discrepancy explained, respectively (Studer et al., 2010).

A tree-structured discrepancy analysis can be seen as the next step in sequence analysis and contributes to the creation of meaningful groups of sequences (Studer et al., 2011). In this paper, we adopt an exploratory approach and use the tree-structured discrepancy analysis to understand how variation in neighborhood sequences can be explained by the physical characteristics of neighborhoods.

§ 3.4 Data and methods

§ 3.4.1 Data and measures

Research on neighborhood change ideally requires individual-level georeferenced data at short-time intervals over a longer period of time. Unfortunately, in many countries, such longitudinal data are unavailable or inconsistent through time. Researchers are therefore confronted with a trade-off between data quality and data availability. This paper used longitudinal register data from the System of Social statistical Datasets (SSD) from Statistics Netherlands. For 1999 to 2013, we have data for the full Dutch population. Historic neighborhood-level data from before the 1990s is extremely scare in the Netherlands due to the move from a census based system to a register based system. The last Dutch census was conducted in 1971, and the alternative country-wide individual-level registration system was ionstalled by 1995. Data on neighborhood income levels is however only available from 1999 onwards, hence our focus on 1999 to 2013. Combining the recent register data with the last census from 1971 allowed us to analyze long-term neighborhood change, however, this meant that there was a 28-year time-gap in our dataset. Nevertheless, the inclusion of 1971 data provides a unique viewpoint on long-term neighborhood change in the Dutch context.

Our definition of a neighborhood is based on 500 by 500 meter grids. The use of 500 by 500 meter grids enabled the comparability of geographical units over time (as other administrative definitions of neighborhoods have changed drastically over the last 40 years) and allowed for a detailed analysis on a relatively low level of aggregation. We focused on the 31 largest cities of the Netherlands, resulting in a total of 8,917 500 by 500 meter grids (including newly constructed neighborhoods in the period 1971-2013). The choice for including the 31 largest cities in the Netherlands is related to the scale of urban restructuring programs over the past few decades and can therefore be understood as a political construct. To ensure the stability of spatial boundaries over time, we use the city boundaries of 2013. Because of the high density of these cities, the average grid consists of 900 residents. For privacy reasons, grids with less than 10 residents have been excluded from the analyses.

We analyzed changes in the share of low-income households in neighborhoods over time. Low-income households are defined as the bottom 20%, which in 1971 included households with an income below 8,000 guilders and in 2013 households with an income below 17,167 euros. Neighborhoods have been categorized according to their share of low-income households into deciles. Because there were few neighborhoods with more than 50% low-income households, the last four deciles have been grouped together.

To examine the role of the physical characteristics of neighborhoods on their trajectories over time, we have included several control variables. We first included a dummy variable for the four largest cities in the Netherlands (Amsterdam, Rotterdam, Utrecht, and the Hague) because we expect more dynamics in big cities. To analyze the path-dependency of neighborhood quality, we included the share of social housing and the share of post-war housing in 1971. We included the change in the share of owner-occupied dwellings between 1971 and 2013 as an indicator for high-quality construction. To assess the effect of changes to the physical structure, we analyzed the effect of demolition, defined as the cumulative number of demolished postwar rental dwellings over the period 1999 to 2013. We have no information on demolition in 1971, however, as many postwar dwellings were still relatively new in 1971 and as large-scale urban restructuring of postwar areas started in the 1990s, it is highly unlikely that the demolition of postwar rental dwellings in 1971 was more than incidental. A summary of all the variables used in the analyses is presented in Table 3.1.

§ 3.4.2 Methods

To provide a detailed illustration of long-term neighborhood change, we first zoomed in on Amsterdam and Rotterdam. We visualized how the spatial distribution of low-income neighborhoods has changed in Amsterdam and Rotterdam between 1971 and 2013. Amsterdam and Rotterdam are the two largest cities in the Netherlands, but have experienced different neighborhood trajectories over time. The economy of Amsterdam is characterized by a strong service sector, while Rotterdam's economy remains tied to the harbor (Burgers & Musterd, 2002; Hochstenbach & Van Gent, 2015). The average income level of the population is therefore higher in Amsterdam (Hochstenbach & Van Gent, 2015). Amsterdam has experienced strong gentrification in the past few decades, which is often ascribed to the historic architecture of inner-city neighborhoods. Although some neighborhoods in Rotterdam have also experienced processes of gentrification, the dominant process in Rotterdam has been neighborhood downgrading since the 1970s (Hochstenbach & Van Gent, 2015).

To come to a better understanding of patterns of neighborhood change, we next focused on neighborhood trajectories of the 31 largest cities using a combination of sequence analysis and a tree-structured discrepancy analysis. We have first conducted a multifactor discrepancy analysis to assess the raw effects of the variables on the sequences trajectories (see Table 3.3). The multifactor approach offers insight in which covariates are significantly associated with the neighborhood trajectories and provides information on the significance of the variables (using permutation tests) and the strength of the model using a pseudo F and a pseudo R^2 (see also Studer et al., 2011).

We then combined sequence analysis and a tree-structured discrepancy analysis to analyze variation in neighborhood trajectories. Sequence analysis is used for the visualization of neighborhood trajectories showing the neighborhood status at each point in time using a color scheme. Each neighborhood category is assigned a different color where the red to blue scheme represents the low to high neighborhood status scale. There are different ways to visualize sequences (for an overview, see Gabadinho et al., 2011). In this paper, we used a sequence distribution plot showing the overall neighborhood distribution instead of individual sequences. Importantly, this means that we are focused on the general pattern of neighborhood trajectories rather than individual neighborhoods. The tree-structured discrepancy analysis then visualized how our control variables affect the trajectories in a tree-structured sequence plot (Studer et al., 2011).

TABLE 3.1 Summary of the dataset

	MIN	MAX	MEAN	SD
Neighborhood category:				
1971	1	6	2.21	1.78
1999	1	6	2.07	1.15
2000	1	6	2.09	1.13
2001	1	6	2.12	1.16
2002	1	6	2.13	1.16
2003	1	6	2.15	1.16
2004	1	6	2.20	1.19
2005	1	6	2.21	1.19
2006	1	6	2.28	1.22
2007	1	6	2.36	1.29
2008	1	6	2.39	1.28
2009	1	6	2.38	1.27
2010	1	6	2.36	1.26
2011	1	6	2.37	1.27
2012	1	6	2.38	1.27
2013	1	6	2.49	1.32
Four largest cities	0	1	0.20	0.40
Percentage social housing 1971	0	100	12.77	27.47
Percentage postwar dwellings 1971	0	100	28.24	39.55
Change percentage owner-occupied dwellings 1971-2013	-97.70	100	6.24	26.66
Total number of demolished dwellings 1999-2013	0	1,536	16.15	67.15

Source: System of Social statistical Datasets (SSD)

We have used the default stopping criteria of a p-value of 1% for the F test (R = 5,000), fixing the minimal group size at N = 446 (5% of the total N = 8,917), and allowing for the creation of five levels (see also Studer et al., 2011). The analyses were conducted in R version 3.2.1 ('World-Famous Astronaut') using the TraMineR package (Gabadinho et al., 2011).

§ 3.5 Results

We first zoom in on Amsterdam and Rotterdam in Figure 3.1 and 3.2. Table 3.2 tabulates the neighborhood categories in 1971 and 2013 for each city. Both illustrate a process of increasing poverty concentration in these cities. Table 3.2 shows that the share of low-income neighborhoods in the last two categories has remained stable over 40 years: the share of neighborhoods with more than 40% low-income households has not increased. However, the spatial distribution of these neighborhoods is characterized by increased spatial concentration as shown in Figure 3.1 and 3.2. While both cities were characterized by a large share of high-income neighborhoods (category 1) in 1971, they show more variation in the neighborhood income distribution by 2013.

The maps show the distribution of low-income households in 1971 and 2013. Figure 3.1 illustrates how inner-city neighborhoods in Amsterdam have maintained their high status over time, while the postwar neighborhoods at the outskirts of the city have experienced downgrading. Low-income neighborhoods in Amsterdam are now increasingly concentrated outside the city centre (cf. Van Gent, 2013). Figure 3.2 shows significant downgrading of large parts of Rotterdam over the last 40 years. Contrary to Amsterdam, Rotterdam's inner city neighborhoods have experienced downgrading, while the high-status neighborhoods in the northern part of the city have maintained their status (cf. Hochstenbach & Van Gent, 2015).

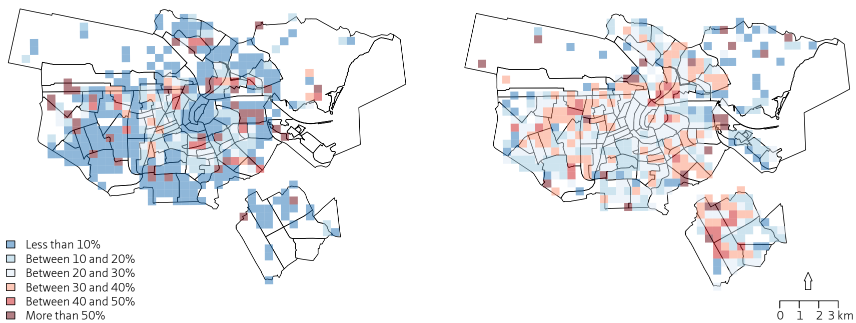

FIGURE 3.1 Percentage low-income households in Amsterdam, 1971 and 2013
Source: System of Social statistical Datasets (SSD)

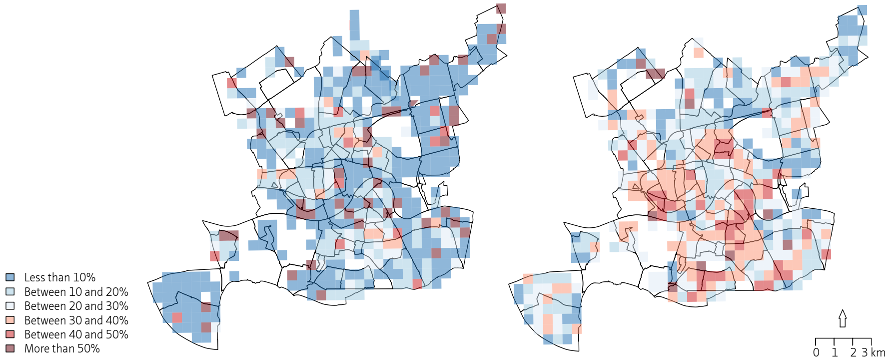

FIGURE 3.2 Percentage low-income households in Rotterdam, 1971 and 2013
Source: System of Social statistical Datasets (SSD)

TABLE 3.2 Distribution of low-income households in neighborhood deciles in Amsterdam and Rotterdam, 1971 and 2013

	PERCENTAGE NEIGHBORHOODS AMSTERDAM		PERCENTAGE NEIGHBORHOODS ROTTERDAM	
	1971	2013	1971	2013
Percentage low-income households:				
<10	57.1	11.3	57.0	19.0
10-20	18.2	23.5	21.0	25.5
20-30	7.8	34.2	7.4	23.2
30-40	3.3	21.7	3.1	21.1
40-50	4.0	4.6	2.4	8.8
>50	9.7	4.6	9.2	2.3
N	**424**	**497**	**458**	**478**

Source: System of Social statistical Datasets (SSD)

We are interested in the neighborhood trajectories underlying the patterns described above and how these trajectories are related to a set of predictors. We are particularly interested how the physical characteristics of neighborhoods are associated with neighborhood trajectories over time. As mentioned earlier, we have first conducted a multi-factor discrepancy analysis to assess the raw effect of our variables on the neighborhood sequences. The results are shown in Table 3.3.

The global statistics show that the model is significant (F = 43.58, R = 5,000) with an R^2 of 14.4%, meaning that our set of variables provides overall significant information about the diversity of neighborhood trajectories. All variables are significant at the 1% level (assessed through 5,000 permutations), with the exception of our dummy variable for the four largest cities. The share of social housing in 1971 and the number of demolished dwellings appear to be the most important predictors of neighborhood trajectories between 1971 and 2013.

TABLE 3.3 Multifactor discrepancy analysis

	PSEUDO-F	PSEUDO-R^2
Four largest cities	1.428	0.001
Percentage social housing 1971	117.701**	0.078
Percentage postwar dwellings 1971	43.201**	0.029
Change percentage owner-occupied dwellings 1971-2013	20.874**	0.014
Total number of demolished dwellings 1999-2013	45.316**	0.030
Overall model	**43.584****	**0.144**

Note: significance is assessed through permutations (R = 5,000).
Source: System of Social statistical Datasets (SSD)

Figure 3.3 shows the tree-structured discrepancy analysis for the neighborhood trajectories in the 31 largest Dutch cities. The initial node shows the distribution of neighborhood states by year (box 1). Overall, the 31 largest cities are characterized by a more or less even distribution of neighborhoods. Over time, the share of high-income neighborhoods is decreasing while the share of low-income neighborhoods is increasing. In the tree, the most significant variables and their most significant values are used in respective order. For each group, we see how the selected variable (and the threshold values of the variable) affects the neighborhood trajectories, showing the group size, the within-discrepancy, and the R^2 for that split. Our overall model has an R^2 of 19.5%, which is higher than the R^2 from the multifactor discrepancy analysis, meaning that the tree has better explanatory power, which can be explained by the fact that the tree automatically accounts for interaction effects (Studer et al., 2011). Our neighborhood characteristics explain 19.5% of the variability in neighborhood trajectories. We have forced the model to use the dummy variable for the four largest cities – Amsterdam, Rotterdam, the Hague, and Utrecht – for its first split because we were interested to see how the trajectories of neighborhoods in these large cities differ from the trajectories in the other cities. We find that neighborhoods in the four largest cities (box 3) are characterized by more neighborhood dynamics than the other cities (box 2).

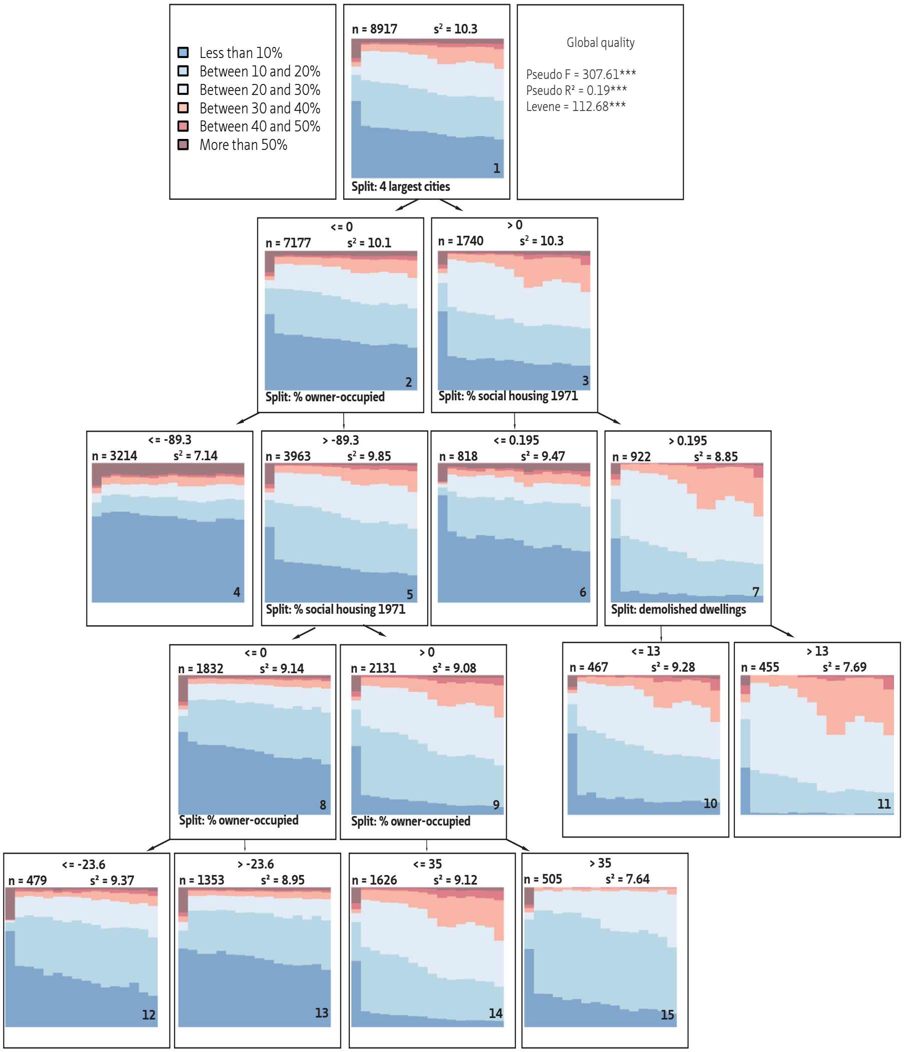

FIGURE 3.3 Tree-structured discrepancy analysis of neighborhood trajectories, 1971-2013
Source: System of Social statistical Datasets (SSD)

Since 1971, the four largest cities have experienced a substantial decrease in their share of high-income neighborhoods and an increase in low-income neighborhoods. The model shows that the share of social housing in 1971 is the most important indicator in explaining variance in neighborhood trajectories in the four largest cities (box 6 and 7). The neighborhoods with hardly any social housing in 1971 are characterized by high-income trajectories, while the neighborhoods with higher shares of social housing show more downward trajectories.

For this latter group, the number of demolished dwellings between 1999 and 2013 seems to matter (box 10 and 11). Demolition took place in neighborhoods that were experiencing downgrading (box 11). These processes of decline were the reason for the Dutch government to target these neighborhoods for urban renewal through the demolition of low-quality social-rented dwellings (Kleinhans, 2004).

The left side of the tree shows that changes in the share of owner-occupied dwellings between 1971 and 2013 is the most important predictor for neighborhood trajectories in the other 27 cities (box 4 and 5). Box 4 consists almost solely of newly constructed neighborhoods with high shares of owner-occupied dwellings since 1999. These neighborhoods are characterized by more neighborhood stability. Existing neighborhoods that have seen increases in their share of owner-occupied dwellings are characterized by more downward trajectories (box 5). Here the share of owner-occupied dwellings interacts with the share of social housing in 1971. For those neighborhoods that have seen an increase in the share of owner-occupied dwellings (box 5), the share of social housing seems to matter. Higher shares of social housing in 1971 are associated with more downward trajectories (box 9). For this latter group, higher increases in the share of owner-occupied dwellings are associated with more high-income trajectories (box 15). This interaction between the share of social housing in 1971 and changes in the share of owner-occupied dwellings captures the Dutch policy of social mixing by changing the tenure composition in neighborhoods.

§ 3.6 Discussion and conclusion

Especially in the four largest Dutch cities, our results show an increase in the share of low-income neighborhoods since 1971. Amsterdam and Rotterdam, in particular, have experienced increasing poverty concentrations in specific neighborhoods. Most of these neighborhoods were built after the Second World War and were characterized by concentrations of social housing. The Netherlands, historically, had a large social housing sector with relatively high-quality housing. Contrary to many other countries, social-rented dwellings were inhabited by mix of socioeconomic groups, not just low-income households (Van Kempen & Priemus, 2002). In 1971, many postwar neighborhoods were still relatively new and were considered to be high-status neighborhoods (Van Beckhoven et al., 2009). By 2013, these postwar neighborhoods have experienced significant downgrading and are characterized by concentrations of poverty as is shown in Figure 3.1 and 3.2. The downgrading of these neighborhoods can be explained by their physical characteristics, in particular, the low-quality housing and its multiple technical and physical problems. This, combined with relative downgrading due to new housing construction elsewhere, fuelled processes of neighborhood decline (Kleinhans, 2004; Prak & Priemus, 1986). At the same time, this process led to the residualization of the social housing stock in the Netherlands, where the social housing sector increasingly became the domain of low-income households (Van Kempen & Priemus, 2002).

In the 1990s, the Dutch government launched large-scale urban restructuring programs to target the most disadvantaged neighborhoods. In practice, this meant that many low-quality postwar social-rented dwellings were demolished to make room for more expensive privately rented or owner-occupied dwellings (Kleinhans, 2004). Figure 3.3 captures this process very well: we see that demolition took place in downgrading neighborhoods with relatively high shares of postwar rental dwellings in the four largest cities. At the same time, we see that the changes in the share of owner-occupied dwellings interacts with the share of social housing in 1971 in the other 27 cities. If we interpret a rising share of owner-occupied dwellings in these neighborhoods as an indicator of the Dutch policy of mixing tenure, it then seems to be most effective in neighborhoods that have experienced substantial increases in the share of owner-occupied dwellings, thereby contributing to more high-income trajectories (see also Bolt et al., 2009). The question however remains if such changes to the housing stock will lead to significant neighborhood upgrading and to what extent these effects will be temporary or long-lasting (Tunstall, 2016; Van Ham & Manley, 2012; Zwiers et al., 2016)

Our analyses seem to indicate a high degree of path-dependency as the initial quality of dwellings and neighborhoods was found to be associated with neighborhood trajectories over time. While the four largest cities generally show a change towards a more equal neighborhood distribution, there is some indication of increasing poverty concentration. Especially neighborhoods with high shares of social housing in 1971 have experienced strong processes of neighborhood decline. Zooming in on Amsterdam and Rotterdam in Table 3.2 and Figures 3.1 and 3.2, we see that both cities were characterized by high shares of high-income neighborhoods in 1971, but show more variation in neighborhood income groups by 2013, albeit with more poverty concentration in many postwar neighborhoods.

The main contribution of this paper is the introduction of a new method for exploring neighborhood trajectories. Our empirical exercise confirms the need for an approach that incorporates both long-term neighborhood changes and a more detailed analysis of neighborhood trajectories, because neighborhoods are extremely dynamic but the effects of downgrading and upgrading on neighborhoods are only visible after longer periods of time. A focus on neighborhood trajectories lends itself for the identification of different patterns of change over time. The combination of sequence analysis and a tree-structured discrepancy analysis contributes to an understanding of how changes in a particular group of neighborhoods are related to the trajectories of other neighborhoods. As such, these methods provide an integrated approach towards neighborhood change, by focusing on trajectories and by identifying factors that contribute to changing trajectories over time. The analyses show how specific levels of change function as thresholds for a different direction of neighborhood trajectories. It is however unclear to what extent these thresholds can be used as more than cut-off points. Future research should aim to explore the meaning of these thresholds for the identification of risk factors for neighborhood change and its implications for spatial policy.

A tree-structured discrepancy analysis can be seen as the next step in sequence analysis, providing a new way of researching neighborhood dynamics. The combination between sequence analysis and a tree-structured discrepancy analysis has proven to be a powerful tool to visualize and understand complex, contextualized patterns of change over time. These methods could contribute to an understanding of 'when' or 'under what circumstances' neighborhood trajectories diverge in a particular direction, instead of 'if'. Such research is necessary, because the time-period, frequency, and composition of mechanisms that influence neighborhood trajectories may be non-linear, can be temporary or long-lasting, may vary over time, and might be conditional on other factors (Galster, 2012; Van Ham & Manley, 2012).