



Investigating Social Frontiers in the Netherlands: A Data Analytics Approach

Eleanor Bale

A thesis submitted in partial fulfilment of the requirements for the degree of
Doctor of Philosophy

The University of Sheffield
Sheffield Methods Institute
School of Education

June 2026

Acknowledgements

This thesis would not have been possible without the help and support of my supervisors at the University of Sheffield Gwilym Pryce and Aneta Piekut. I am also grateful for the support of TU Delft with facilitating and funding access to the data, and specifically Ana Petrović and Maarten van Ham for their advice. I would like to thank Duncan Lee and Maria Petrillo for their guidance and help.

Thank you to my family and friends for always being supportive, listening and believing in me, and to Dougal, Florence, Bella, Lily, and Arthur for being the best possible distractions.

The results of this thesis are based on analysis of non-public microdata from Statistics Netherlands (**University of Sheffield - project number 9197**).

Abstract

Introduction: The impact and presence of segregation has long been a topic of interest in social science (Massey and Duncan, 1985). Yet transitions between communities are vastly underrepresented (Legewie and Schaeffer, 2016; Dean *et al.*, 2019), with segregation measures often bypassing the steep transitions. Dean *et al.* (2019) coined the term ‘social frontiers’ to describe these steep boundaries, highlighting that they could have large consequences on society through increasing crime and conflict, impacting household mobility (Olnier *et al.*, 2024), reducing access to resources, and heightening feelings of relative deprivation (Iyer and Pryce, 2024).

Objectives: This thesis aims to address some of the gaps in the relatively new social frontier literature, making large contributions to the conceptual, methodological, and empirical findings. The thesis especially focuses on the frontiers’ prevalence, evolution, and impact in the Netherlands.

Methodology: This thesis implements a range of techniques. I apply four different identification methods including the established Lee and Mitchell (2013) method. Additionally, the thesis uses multiple measures to describe the prevalence of social frontiers and discuss their evolution. Finally, applying causal matching methods to evaluate the impact of living close to social frontiers.

Results: The thesis has several key findings which form an important contribution to the literature: the establishment of the Partitioning Around Medoid (PAM) clustering method for identifying frontiers; growing rates of social frontiers in Amsterdam and the Hague; the contrasting prevalence of single dimension and intersectional social frontiers; and the negative association between living close to social frontiers and socioeconomic outcomes.

Conclusion: The thesis makes significant methodological, empirical, and theoretical contributions to the literature. It highlights the importance of social frontiers and the need for further research into them. The findings of all the chapters have the potential to drive future research, as well as having large policy implications for the Netherlands.

Contents

Acknowledgements	2
Abstract	3
Contents	4
Table of Figures	7
Table of Tables	8
Glossary	10
Chapter 1 Introduction	12
Research Context	12
Research Contribution and Aims	16
Data and Methodology	21
Thesis Structure	27
Chapter 2 Segregation and the Netherlands	35
Segregation Factors and Patterns in the Netherlands	35
A Brief Summary of the Migration History of the Netherlands	37
Migration Background Classification and Assimilation Policies	40
Social Housing in the Netherlands	42
Conclusion	42
Chapter 3 Detecting Spatial Discontinuities in Amsterdam	44
Abstract	44
Introduction	46
Literature	48
Data and Methodology	57
Simulation Study Results: Testing Alternative Social Frontier Identification Methods	66
Social Frontier Prevalence Results	74
Excess Embrittlement in Amsterdam	81
Conclusion	82
Chapter 4 The Lifecycle of Social Frontiers	87
Abstract	87
Introduction	88
Literature	91
Study Context: Dimensions of Segregation in the Netherlands	99

Data and Methodology	108
Results	111
Conclusion	124
Chapter 5 Exploring the Intersectional Nature of Social Frontiers in the Netherlands	128
Abstract	128
Introduction	130
Literature	132
Data and Methodology	139
Results	140
Conclusion	149
Chapter 6 The Impact of Living near Country of Birth Social Frontiers during Childhood on Socioeconomic Outcomes in Adulthood	153
Abstract	153
Introduction	155
Literature	156
Data	161
Methodology	170
Results	176
Conclusion	179
Chapter 7 Conclusion	183
Overview	183
Key Findings and Conceptual Contributions	186
Limitations and Future Research	191
Concluding Remarks	193
References	194
R Packages and Programs	222
Appendices	224
Appendix A: The Boroughs of Amsterdam	224
Appendix B: Absolute Contiguous Difference Threshold Choice	225
Appendix C: Simulation Study Results Table for 2019	226
Appendix D: Simulation Study Scores for 2010	227
Appendix E: Municipalities in COROP Region Definitions	231
Appendix F: Income Cut-offs	232

Appendix G: Lee and Mitchell (2013) Frontier Prevalence for Chapter 4	233
Appendix H: PAM Clustering Social Frontiers for Chapter 4	234
Appendix I: Frontier Lifecycle Plots	238
Appendix J: Descriptive Statistics for Intersectional Frontiers in Amsterdam and the Hague	239
Appendix K: Intersectional Social Frontier Prevalence	241
Appendix L: Aggregated Education Levels	245
Appendix M: Matching Methods Descriptions	246
Appendix N: Covariate Balancing Scores	247
Appendix O: Sensitivity Analysis Results	253

Table of Figures

Figure 1.1: Thesis structure flowchart.	26
Figure 2.1: Graph showing the number of immigrants arriving to the Netherlands by continent between 1995 and 2022 (CBS, 2023b).....	39
Figure 3.1: Spatial arrangement of non-Westerners in the city of Amsterdam in 2010 and 2019.	56
Figure 3.2: Flowchart representing Chapter 3's structure - a section of Figure 1.1 (the thesis flowchart).	58
Figure 3.3: True social frontiers in 2019.	63
Figure 3.4: TNR scores jitter plot for 2019.....	68
Figure 3.5: TPR scores jitter plot for 2019.....	70
Figure 3.6: Social frontiers identified in the City of Amsterdam in 2010 and 2019 identified using the Lee and Mitchell (A) and PAM clustering (B) methods.	77
Figure 4.1: A hypothesised social frontier lifecycle plot and simulated frontier formation stages.....	96
Figure 4.2: Plots showing the proportion of the variables of interest.	107
Figure 4.3: Frontier lifecycle plots.	114
Figure 4.4: Entrenched frontier lifecycle.	115
Figure 4.5: Percentage of social frontiers in Amsterdam and the Hague.	116
Figure 4.6: Average social frontier length in Amsterdam and the Hague.	117
Figure 4.7: Average social frontier strength in Amsterdam and the Hague.	118
Figure 5.1: Intersectional social frontier diagrams.....	134
Figure 5.2: Comparing the percentage of coterminous intersectional social frontiers and individual social frontiers.	142
Figure 5.3: Comparing the percentage of multidimensional intersectional social frontiers and individual social frontiers.	143
Figure 6.1: Region of common support.....	167
Figure 6.2: Covariate balancing love plot for all the matching methods	173
Figure 6.3: Covariate balancing plot for three chosen methods	175
Figure 7.1: Key thesis findings	185
Figure A.1: City of Amsterdam with the boroughs labelled. Created using shapefiles from the Municipality of Amsterdam - Research, Information and Statistics (2022) and OpenStreetMap (2023).....	224
Figure D.1: The true social frontiers for 2010 data	227
Figure D.2: Jitter plots showing the accuracy scores (TNR - a) and (TPR - b) for the 2010 simulation study for the four identification methods.....	229
Figure I.1: Frontier lifecycle plot showing the evolution of the strength (median difference) of the 5-year long social frontiers identified using the Lee and Mitchell clustering method in Amsterdam and the Hague between 1999 and 2021.....	238
Figure I.2: Frontier lifecycle plot showing the evolution of the strength (median difference) of the 10-year long social frontiers identified using the PAM clustering Lee and Mitchell (2013) method in Amsterdam and the Hague between 1999 and 2021	238

Table of Tables

Table 1.1: Regions and grid cell sizes.....	23
Table 3.1: Population statistics in the City of Amsterdam.....	53
Table 3.2: Median accuracy scores from the simulation study.	66
Table 3.3: Processing time of the four identification methods when applied to Amsterdam 2010 and 2019 data.	72
Table 3.4: Social frontier prevalence results in Amsterdam in 2010 and 2019.....	75
Table 3.5: Lee and Mitchell (2013) country of birth social frontiers per borough in the city of Amsterdam.....	78
Table 3.6: Excess embrittlement level in Amsterdam in 2010 and 2019.	81
Table 4.1: Socioeconomic descriptive statistics in Amsterdam and the Hague in 1999 and 2021.	100
Table 4.2: Social frontier dimension descriptions.	109
Table 4.3: Frontier prevalence measure methodology.	111
Table 4.4: Average frontier duration and re-emergence rate.....	112
Table 5.1: Summary of intersectional frontier prevalence results.....	150
Table 6.1: Descriptive statistics for imputed dataset.	165
Table 6.2: Descriptive statistics for outcome variables - Mean (SD)	168
Table 6.3: Missing data	170
Table 6.4: Sample size for matching methods.	172
Table 6.5: Average Treatment Effect on the Treated (ATT) for income.....	176
Table 6.6: Average Treatment Effect on the Treated (ATT) for income for alternative exposure length.....	177
Table 6.7: Average Treatment Effect on the Treated (ATT) for education years... ..	177
Table 6.8: Average Treatment Effect on the Treated (ATT) for education years for alternative exposure length.	178
Table B.1: 2010 Threshold choices for absolute difference social frontier identifier simulation accuracy results.	225
Table B.2: 2019 Threshold choices for absolute difference social frontier identifier simulation accuracy results.	225
Table C.1: Accuracy scores (TPR and TNR) for 2019 simulation study.....	226
Table D.1: Accuracy scores (TPR and TNR) for 2010 simulation study.	230
Table E.1: Municipalities included within COROP region definitions.....	231
Table F.1: Cut-offs for low- and high-income population in the Netherlands between 1999 and 2021.....	232
Table G.1: Social frontier prevalence table for Lee and Mitchell (2013) method ...	233
Table H.1: Social frontier prevalence table for PAM clustering.	237
Table J.1: Amsterdam descriptive statistics	239
Table J.2: The Hague descriptive statistics.....	240
Table K.1: Intersectional social frontier percentage and median length (% , m) in the Hague.....	241
Table K.2: Intersectional social frontier percentage and median length (% , m) in Amsterdam.....	242

Table L.1: Aggregated education levels.	245
Table M.1: Description of matching methods.....	246
Table N.1: Covariate balance for matching at the age of 10 - values over 0.1 are highlighted.	247
Table O.1: Average Treatment Effect on the Treated (ATT) for income using entropy balancing	253
Table O.2: Average Treatment Effect on the Treated (ATT) for education years using entropy balancing	253
Table O.3: Average Treatment Effect on the Treated (ATT) for income using IPW	254
Table O.4: Average Treatment Effect on the Treated (ATT) for education years using IPW	254
Table O.5: Average Treatment Effect on the Treated (ATT) for income for common support	255
Table O.6: Average Treatment Effect on the Treated (ATT) for education years for common support.	255

Glossary

Social Frontiers: “Boundaries between adjacent neighbourhoods where there are steep differences in the racial, ethnic, religious, cultural or social characteristics” (Dean *et al.*, 2019, p. 272).

Coterminous Intersectional Social Frontiers: “Boundaries between adjacent communities where there exists an overlap between multiple single dimension social frontiers, whether that dimension be ethnicity, economic, or cultural.” (Chapter 5).

Economic Social Frontiers: Sharp transitions between communities based on income or employment status (Chapter 4).

Excess Embrittlement: A measure of how far the identified percentage of social frontiers differs from the expected level if an area was arranged randomly (Manley *et al.*, 2023).

Group Faultlines: “Hypothetical dividing lines that may split a group into subgroups based on one or more attributes” (Lau and Murnighan, 1998, p. 328).

Index of Dissimilarity (D): A segregation measure which measures evenness. The score “represents the proportion of minority[/majority] members that would have to change their area of residence to achieve an even distribution” (Massey and Denton, 1988, p. 284).

Intersectionality: “The interconnected nature of social categorizations such as race, class, and gender, regarded as creating overlapping and interdependent systems of discrimination or disadvantage” (‘Intersectionality’, 2023).

Multidimensional Intersectional Social Frontiers: “Identified if adjacent communities have a steep difference in terms of a combination of population attributes, e.g. the proportion of low-income non-Westerners” (Chapter 5).

Non-Westerner: “Person originating from a country in Africa, South America or Asia (excl. Indonesia and Japan) or from Turkey” (CBS, 2025a).

Social Boundaries: “Any contiguous zone of contrasting density, rapid transition, or separation between internally connected clusters of population and/or activity” (Tilly, 2004, p. 214).

Westerner: “Person originating from a country in Europe (excluding Turkey), North America and Oceania, or from Indonesia or Japan” (CBS, 2025b). This term (and the non-Westerner term) is undeniably outdated with strong racial and ethnic connotations with clear links to the Netherland’s colonial past. The problematic

nature of the terms is discussed further in Chapter 2 – Segregation and the Netherlands.

Chapter 1 Introduction

The Oxford English Dictionary (OED) defines segregation as “the separation or isolation of a portion of a community or a body of persons from the rest” (‘Segregation’, 2024). Social science literature has often focused on the spatial arrangement of communities (van Ham and Tammaru, 2016), with many studies exploring how to measure (Duncan and Duncan, 1955; Massey and Denton, 1988) and model (Schelling, 1971) segregation. Along with noting how segregation has links with social inequality, political factors, and shapes social networks (Piekut, Pryce and van Gent, 2019). However, despite extensive segregation research, there has been an absence of literature addressing “what happens at places where groups border” (Legewie and Schaeffer, 2016, p. 131). These transitions between communities have the potential to have large impacts upon society, most of which are almost completely unexplored empirically (Dean *et al.*, 2019; Iyer and Pryce, 2024). This thesis studies these boundaries - specifically focusing on ‘social frontiers’ - exploring why their inclusion in the literature is of great importance in segregation literature.

In this thesis I identify social frontiers in the Netherlands. The Netherlands is a country with a vast history of immigration and a strong welfare system making both ethnic and economic segregation an area of great interest. Research has explored multiple facets of segregation including the impact of spatial scale, housing and urban renewal policies, and various segregation indices (Hartog and Zorlu, 2009; Musterd and Ostendorf, 2009; Bolt and van Kempen, 2010; Sleutjes, de Valk and Ooijevaar, 2018). Yet very few studies have identified and studied the effect of social frontiers in the Netherlands (Olnier *et al.*, 2024). This is likely due to the social frontier being relatively new meaning that they are unstudied in the majority of countries. The thesis makes use of the extensive microdata from Statistics Netherlands to explore both country of birth and economic frontiers. The level of the data available “open[s] up ever richer ways to measure segregation and its impacts” (Piekut, Pryce and van Gent, 2019, p. 226), allowing a thorough picture of the spatial discontinuities in the country to be created.

Research Context

Dean *et al.* (2019) coined the term social frontiers to describe the “boundaries between adjacent neighbourhoods where there are steep differences in the racial, ethnic, religious, or social” composition (p. 272). If the transitions are steep, they can be described as ‘cliff edges’, potentially representing the “epicentres of pent-up social tensions” (Dean *et al.*, 2019, pp. 271–272). These sharp boundaries have the potential to threaten neighbourhood cohesion, increase conflict and crime, along with limiting positive contact, opportunities, and the socioeconomic outcomes of the individuals who grew up near them (Dean *et al.*, 2019; Iyer and Pryce, 2024). These effects may arise due to the frontiers becoming entrenched between communities in

such a way that they become barriers, reducing the opportunity for relationships to form, along with limiting individuals' access to visible role models or resources, and increasing feelings of relative deprivation (Legewie and Schaeffer, 2016; Dean *et al.*, 2019; Iyer and Pryce, 2024). Additionally, the close proximity of the groups means that there is an increased chance of territorial and defensive behaviour as communities compete over the interspatial zones, as well as an absence of bridge-builders meaning communities become increasingly isolated (Legewie and Schaeffer, 2016; Dean *et al.*, 2019). Furthermore, communities can experience spillover effects from their neighbours, with higher income inequality in gentrified areas (Christafore and Leguizamon, 2019), and an association identified between relative deprivation and the poor health of individuals specifically in relation to coronary heart disease (CHD) mortality rates (Allender *et al.*, 2012).

The term social frontier is an extension of the social and symbolic boundaries literature, distinct in its spatial requirements with the boundaries required to be occurring at a specific geographical location (Dean *et al.*, 2019, p. 274). The spatial element of the frontier is crucial as it specifies that we are studying the geographical transition between neighbourhoods. These transitions impact upon the way that neighbourhoods fit and co-exist alongside each other. However, despite the spatial requirement and them existing at spatial boundaries, social frontiers do not consider the spatial arrangement of residents within the grid cells. Frontiers consider the spatial arrangement of how people live and transitions in relation to geographical administrative boundaries (grid cells in this study), however, not how the within neighbourhood factors impact the boundaries. The location of residents within the cells is likely to have an impact on the location of the social frontiers. Iyer and Pryce (2024) highlight the concept of peripherality with residents living further from the centre of the community often feeling more excluded. If the social frontiers exist on the edge of the community, then these resident's feelings could be exacerbated resulting in the frontiers being more impactful. Equally this could reduce territorial attitudes as residents may feel less of an affiliation with their community. The spatial arrangement of residents within the grid cells is not considered within this thesis as individuals are aggregated to the grid cells – however, a future study could consider the impact of the exact location, potentially in relation to distance to community centre as suggested by Iyer and Pryce (2024). This type of study would be possible using the data from Statistics Netherlands due to the point-based data.

Additionally, the frontiers are forced to lie along the standardised grid cells which are the fixed boundaries used by Statistics Netherlands. This makes their location influenced by choices made by the researcher, with the frontiers limited in where they can exist based on these pre-determined boundaries. They are to some extent a data artefact created by the researcher as a result of numerous choices involving the spatial scale, dimension, and identification method. The location of boundaries and neighbourhoods is a debated topic in segregation and neighbourhood literature - Spielman and Logan (2013) suggest there is a consensus across researchers that

the “spatial organization of cities is multidimensional” (p. 2). There are physical boundaries such as motorways or rivers etc. which lead to neighbourhoods being clearly demarcated. Equally, there are boundaries which are likely to exist despite not being physically marked e.g. political boundaries which establish electoral constituencies. However, Rich (2009) stated that “what a neighborhood truly is exists in the minds of social actors, its residents and neighborhood outsiders” (p. 829). This is difficult to define and quantify particularly in a quantitative study and thus Dean *et al.* (2019) suggest that the identification of social frontiers is actually a process of checking whether administrative boundaries coincide with “step changes across space in the social makeup of neighbourhoods” (p. 279). It is important to consider the impact of grid imposition when discussing the location and impact of the frontiers, as they may not be identified exactly where residents deem there to be a neighbourhood boundary. However, Zhang *et al.*’s (2024) qualitative study found that residents recognised the steepest boundaries, those likely to be defined as social frontiers, as local community boundaries in Rotherham, indicating that the steepest of data manufactured boundaries are still likely to be present and consequential on their neighbourhoods.

There is a tendency in the social frontier literature to use the terms social frontier and spatial discontinuity interchangeably (Dean *et al.*, 2019; Iyer and Pryce, 2024). The process of identifying spatial discontinuities is described as a way of creating a “picture of the geography of ethnic separations” (Harris, 2014, p. 2244). The idea being that similar to social frontiers, studies are looking to identify extremes between neighbouring areas (Harris, 2014; Pavlačka *et al.*, 2023). There are many different ways to describe social frontier: spatial discontinuities, sharp transitions or steep boundaries etc., all of which capture their abrupt nature but not necessarily their impact nor dynamic existence. Iyer and Pryce (2024) suggested that ‘social frontiers’ are potentially an extension of spatial discontinuities as a result of their heightened territorial impact which is likely to have large behavioural impacts (p. 795). Their territorial nature results in social frontiers becoming a “social/psychological/political phenomenon” (Iyer and Pryce, 2024, p. 795). This requirement for social frontiers to become a phenomenon means that they are often framed negatively, particularly due to their territorial nature resulting in them indirectly increasing crime and conflict. Social frontier research is limited in that very few studies have noted their potential positive impact and even fewer have studied these impacts empirically. There is a chance that the increased contact between communities leads to the formation of relationships across boundaries. Conversely communities on either side of the frontier can become stronger as they are united by their similarities (Dean *et al.*, 2019). Additionally, a key aspect of the frontier definition, which is not specified with spatial discontinuities, is their dynamic nature which is a headlining element of Chapter 4. The frontiers are prone to emerging, strengthening, declining, and eventually disappearing in various forms across their lifecycle. The use of the term ‘boundary’ may suggest that the division is permanent, whereas frontiers are flexible, with them often temporary and fluctuating.

Despite their potential impact and complicated nature, social frontiers remain under-researched in social science literature, with the majority of studies failing to consider what occurs when communities meet (Legewie and Schaeffer, 2016). Segregation measures often bypass the identification of steep boundaries due to their aspatial nature meaning that they evaluate each unit (neighbourhood) independently (Jakubs, 1981; Tselios *et al.*, 2015; Yao *et al.*, 2019). Thus, they fail to analyse the transitions between communities. Massey and Denton's (1988) widely cited segregation paper evaluated twenty measures of residential segregation classifying them into five dimensions: evenness, exposure, concentration, centralisation, and clustering. However, none of these can fully capture changes at boundaries, instead they describe the level of segregation of an aggregate area missing the individual level nature of social frontiers. For a comprehensive description of the level of segregation in an area, and to understand the spatial dynamics and inter-group relations, we need to employ a combination of the indexes (Massey and Denton, 1988), evaluating segregation as a multidimensional phenomenon. However, even with the combination of the dimensions, social frontiers are still not identified. Therefore, we require an additional dimension of segregation, one which describes the level of sharp transitions between communities in an area and that helps to explain the extent of communication and cohesion between groups. The inability of existing mainstream segregation measures to identify social frontiers is important because of the large role they play in shaping our perspectives on the nature and level of segregation in an area (Iyer and Pryce, 2024), this is due to them forming the basis of the majority of quantitative segregation research (Iyer and Pryce, 2024).

Although Dean *et al.* (2019) was the first paper to use the term 'social frontiers', it was not the first to highlight the importance of contested and fuzzy boundaries. Legewie and Schaeffer (2016) proposed that it is poorly defined boundaries which have the greatest impact, with the lack of defined territory resulting in conflict due to individuals being unsure about the ownership of interstitial areas. Dean *et al.* (2019, p. 272) suggest that these contested boundaries are a stage in the lifecycle of social frontiers. An unstable boundary can result in increased conflict between neighbours as the ownership of these areas becomes of greater importance and community relations are unsettled (Bannister, Kintrea and Pickering, 2013; Legewie and Schaeffer, 2016). However, it is not solely emerging or contested frontiers which are consequential for society; long-term well-defined frontiers can also have enduring ramifications such as the peace-walls in Belfast (Dean *et al.*, 2019, p. 272). In Belfast, these walls have become entrenched within the communities with ramifications for the mental health of those who live in close proximity (Maguire, French and O'Reilly, 2016). In addition, there was clear evidence of sectarian segregation with members unlikely to enter what they perceive as outgroup areas (Dixon *et al.*, 2020). If a frontier is consistently present over an extended period of time, then there are likely individuals who have been exposed for a significant time.

This long-term (and potentially lifelong) exposure could be of greater consequence for individuals compared to momentary exposure (Iyer and Pryce, 2024).

This is an exciting field of study, with the limited literature meaning there exist large gaps for future research, particularly in the empirical exploration of the social frontier effects and dynamics. Longitudinal geocoded data, such as the Dutch microdata used in this thesis, offers researchers the opportunity to capture the effect of living in close proximity to social frontiers on an individual's life outcomes and wellbeing, and on community stability and conflict levels. If living near social frontiers results in reduced access to civic resources or upward mobility, then there is an imperative for governments to address social frontiers. This could be through policies aimed at reducing gentrification or implementing cohesion strategies such as investing in more community spaces with the aim of building trust between contrasting communities. Additionally, they could consider frontiers when making housing decisions especially in relation to the allocation of social housing and construction, creating integration strategies or redesigning neighbourhoods with the aim of reducing segregation.

This thesis is firmly situated in social frontier research, tasked with addressing some of the gaps in the existing literature. Each of the chapters attempts to build on the growing social frontier literature, generating findings that deepen our understanding of segregation especially in the Netherlands.

Research Contribution and Aims

This thesis makes a number of novel and innovative conceptual, empirical, and methodological contributions to the social frontier literature. Despite each chapter focusing on a different aspect of social frontiers in the Netherlands, they all have the same overarching goals – to understand the pattern and prevalence of social frontiers in the Netherlands and use this to discuss segregation patterns in the Netherlands. Additionally, they all highlight the importance of expanding the study of social frontiers.

The main conceptual contribution comes from the definition of the intersectional social frontiers based on sociology and social frontier literature. The empirical analysis builds upon the theoretical discussions in both the thesis and social frontier literature, especially in relation to frontier dynamics, the identification and prevalence of intersectional frontiers, and the effect of living in close proximity to frontiers on outcomes in adulthood. These empirical insights are made possible through methodological advances which include the introduction of new frontier identification and prevalence measures, and the implementation of matching causal effects.. These contributions are not only beneficial for developing social frontier literature, but they also advance Dutch segregation literature, particularly our understanding of how communities intersect or live alongside each other.

This thesis has the following key aims:

- Developing our understanding of sharp divisions in the Netherlands, particularly how their prevalence varied over regions, time periods, and dimensions.
- Introducing different methods and practices with the aim of increasing the scope of future social frontier research.
- Comparing the expected and observed level of social frontiers to create the excess embrittlement level. Then using this to explore the socioeconomic characteristics of the city which effect the spatial arrangement ultimately impacting the rate of social frontiers.
- Proposing two intersectional social frontier definitions and identifying them in the Hague and Amsterdam.
- Estimating the impact of living near country of birth social frontiers during childhood on outcomes in adulthood.

Developing Dutch segregation research

All of the chapters within this thesis concentrate on spatial divisions in the Netherlands. Chapter 2 situates the empirical chapters in the context of migration and segregation literature in the Netherlands. Each of the empirical chapters (Chapters 3 to 6) engage with literature on Dutch segregation, noting the similarities between my results and existing studies, and providing potential explanations for my findings. The Dutch literature includes studies which focus on gentrification (Musterd *et al.*, 2016; Hochstenbach and Musterd, 2021), the link between economic and ethnic segregation (Bolt, van Kempen and van Weesep, 2009; Feijten and van Ham, 2009), and the effect of the neighbourhood on income and education outcomes (van der Laan Bouma-Doff, 2007; Nieuwenhuis *et al.*, 2017; Troost, van Ham and Manley, 2023). Whilst these studies are well-established and often far-reaching, none of them touch upon the existence or impact of social frontiers.

The only study to focus on social frontiers in the Netherlands is Olnier *et al.* (2024), which examined the impact of country of birth social frontiers on household mobility in Rotterdam. However, even with this study we still have very little understanding of the pattern of frontiers in the Netherlands and how they evolve over time. Within the chapters in this thesis, I discuss multiple aspects of social frontiers including areas that are particularly prone to frontiers, the difference in the frontier evolution depending on the region and dimension, and why these results may vary, discussing the impact of socioeconomic and regional context. Iyer and Pryce (2024) highlighted the importance of including multiple different facets to measure social frontier prevalence, also noting that the majority of social frontier research is cross-sectional - studying a specific city at a moment in time.

Within this thesis I identify the frontiers on multiple dimensions and geographical scales over an extended period of time, resulting in a more substantive discussion about the pattern of social frontiers along with a greater understanding of their potential impact. These findings can help to explore whether certain regions or areas are particularly susceptible to social frontiers and why this may be the case.

Expanding the scope of future social frontier research

One of this thesis' main methodological contributions is the identification of the Partitioning Around Medoid (PAM) clustering method as an alternative social frontier identification method to the existing Lee and Mitchell (2013) method. This is of particular importance for the advancement of social frontier studies as the existing Lee and Mitchell (2013) method, adapted specifically for social frontiers in Dean *et al.* (2019), is complex and time-consuming, especially when applied to large datasets. The computational demands of the Lee and Mitchell (2013) method become a problem when researchers want to extend their study beyond single cities or identify social frontiers using finer spatial intervals such as 100m by 100m grid cells rather than 1km by 1km grid cells which can dramatically increase the size of the dataset. I test alternative social frontier identification methods, with the aim of developing an approach which is quicker but equally as accurate as the Lee and Mitchell (2013) method, ultimately identifying PAM clustering.

The introduction of the PAM clustering method as a substitute identification method in Chapter 3, sped up the identification of the frontiers in Chapter 6 where I identified frontiers nationwide using 500m by 500m grid cells over a 20-year period. An extended study such as this is very uncommon in neighbourhood effects literature, with findings often cross-sectional due to difficulties obtaining the required datasets (de Vuijst and van Ham, 2019, p. 17). In addition, in Chapter 3 I implement a shuffling method to calculate the excess embrittlement level which is a time-consuming process. As I only calculated the excess embrittlement in two years in Chapter 3, it was feasible for me to identify the social frontiers using the Lee and Mitchell (2013) method. However, if studies were to calculate the excess embrittlement level across multiple cities and time periods, then the PAM clustering method would make this process considerably quicker which could deepen our understanding of how and why cities differ from random.

Comparing the prevalence of social frontiers in different regions, time periods, and dimensions

Previous social frontier literature has lacked comparisons between regions, time periods, and dimensions (Dean *et al.*, 2019; Iyer and Pryce, 2024). Without including aggregate prevalence measures it is difficult for researchers to describe the level of social frontiers in an area. Previous studies have included only visual representations of social frontiers through maps; however, it can be difficult to compare cities and time periods based purely on a visual inspection of these maps. To comprehensively compare regions and study frontier evolution, I include multiple

prevalence measures: the frontier percentage, and the average length, duration, and strength as proposed by Iyer and Pryce (2024). The inclusion of these measures facilitates a discussion about the evolution of social frontiers, meaning that changes in the frontier strength and effect can be modelled over the frontier lifecycle.

By studying the duration and strength of frontiers, and thus their dynamic evolution, we are able to evaluate whether frontiers are consistent in their development. Is there a uniform trajectory across different dimensions and regions, or do frontiers emerge and decline at different rates depending on neighbourhood characteristics or the frontier dimension? Additionally, this investigation opens up opportunities for future research to evaluate how the frontier effect varies depending on what lifecycle stage the frontier is at or the length of time an individual is exposed to a social frontier.

Furthermore, there are few social frontier studies which have included a comparison of the frontier prevalence in multiple cities. This could be missing from the literature due to the time-consuming nature of identifying social frontiers in multiple cities which is a computationally taxing activity that requires a large amount of computer power. Additionally, the data used in this thesis from Statistics Netherlands is incredibly detailed both spatially and in relation to the information available, this means that a meaningful comparison can be created at a fine spatial scale. This level of precise data may not be available from other sources (countries). Spienburg *et al.* (2025) identified social frontiers in 520 cities in Europe but did not compare the rate of frontiers in the cities, focusing instead on their association with urban boundaries. Through the inclusion of numerical prevalence measures I am able to create a more detailed, precise analysis, describing and comparing multiple different measures such as length and duration, and using these to compare the level of divides in regions. Comparing different regions allows a discussion about the potential social factors or context that are likely to result in frontiers being prolific in one area but not in another (Iyer and Pryce, 2024). This leads to policy discussions with regard to whether there are strategies that could be implemented to decrease the number of homogeneous neighbourhoods potentially through housing planning decisions or increasing the number of shared spaces. In addition, by identifying areas or dimensions where the frontiers are particularly prolific, the policies can become more targeted.

It is not only regions and time periods that can be compared but also different frontier dimensions. This thesis includes the first identification of economic social frontiers with previous papers focusing on country of birth or ethnicity frontiers (Dean *et al.*, 2019; Křížková *et al.*, 2021; Olnér *et al.*, 2024). I identify and compare the prevalence of four types of social frontiers: low- and high-income, unemployment, and country of birth. I explore how these frontiers have evolved over time and whether this varies between dimensions. The comparison of the different dimensions, regions, and time periods is made considerably more detailed and

validated through the inclusion of the frontier prevalence measures as we can make meaningful conclusions about how the frequency and substantialness of the frontiers vary.

Intersectional social frontiers

One of this thesis' main conceptual contributions is the introduction of the intersectional social frontier definitions. Iyer and Pryce (2024) described intersectional social frontiers as boundaries which are "co-occurring across multiple socio-demographic domains" (p. 784). Even with this description, there is not a clear process or definition for how we identify these boundaries as there has only been incredibly limited discussion about them. Most of the discussions have focused on the potential variation in their effect compared to individual dimension frontiers (Dean *et al.*, 2019; Iyer and Pryce, 2024).

In Chapter 5 I define two types of intersectional social frontiers: coterminous and multidimensional. Despite both falling under the principle of intersectional social frontiers they are two separate interpretations of the concept and thus differ in their prevalence and likely also in their effect. Within this study, the focus is on defining, identifying, and comparing the prevalence of intersectional frontiers rather than their effect. As this is the first study to specifically focus on defining and identifying intersectional social frontiers, it acts as an introduction to them: discussing how they exist in society, especially in comparison to single dimension frontiers. Future studies can build upon the findings to study the variation in the effect of the two definitions, how they differ in prevalence in different regions and how their effect differs to that of single dimension social frontiers.

Understanding the impact of living in close proximity to social frontiers in childhood on outcomes in adulthood

As social frontiers are a relatively new topic in the literature, a large amount of the discussions on their impact has been theoretical – discussing the potential effects of the frontiers on health, educational outcomes, and neighbourhood cohesion (Dean *et al.*, 2019; Iyer and Pryce, 2024). Those who have empirically studied the frontier effects have often focused on conflict and crime (Legewie and Schaeffer, 2016; Dean *et al.*, 2019; Křížková *et al.*, 2021), household mobility (Olnér *et al.*, 2024), or mental health (Maguire, French and O'Reilly, 2016; Lee, Pryce and Ramos, 2025). There is an absence in the literature of studies that explore the effect of the frontier on socioeconomic factors such as income and educational attainment, despite neighbourhood effects and theoretical social frontier literature highlighting the impact of the neighbourhood on life outcomes (Musterd *et al.*, 2012; Troost, van Ham and Manley, 2023; Iyer and Pryce, 2024). Within Chapter 6 I estimate whether living near a country of birth frontier during childhood has an impact on income and educational attainment in adulthood. Neighbourhoods play a crucial role in childhood due to the extended period of time that children spend in their local area whether that be socialising, in education or simply experiencing their local area (Sykes and Kuyper,

2009). Thus, children may be more likely to be impacted by the frontiers as they are exposed on a greater scale compared to adults who can more easily leave their neighbourhood resulting in them often having a larger social network which goes beyond their local community. I apply a causal inference matching method in Chapter 6 to explore whether there is a difference in the income and education outcomes of those who lived near a social frontier compared to those who did not during their childhood.

Data and Methodology

Data

All the chapters within this thesis make use of the extensive registry data from the Centraal Bureau voor de Statistiek (CBS) also known as Statistics Netherlands (<https://www.cbs.nl/en-gb/our-services/customised-services-microdata/microdata-conducting-your-own-research>). The data is an agglomeration of administrative data from a collection of different registers. There is a vast range of data available about every individual who has lived in the Netherlands since 1999 including details about their income, household location, education level, and health. Due to the depth of information available in the datasets, it is possible to provide the complete housing, education, and income history of an individual using the unique identification code which makes it possible to link individuals and households. Additionally, households and individuals can be tracked over time through houses being geo-coded to grid cells of various sizes. The richness and granularity of the data means there is large scope for potential research including longitudinal studies which may not be possible using census or open-access data as they are unlikely to track individuals in every year. The microdata is only accessible under strict conditions with permission required for access. Prior to any results being exported from the remote access, they have to be checked to ensure there are no disclosure risks.

In all four chapters I identify country of birth social frontiers classified using the binary definition from Statistics Netherlands: Western or non-Western.¹ Moreover, Chapters 4 and 5 extend the study of social frontiers by including economic dimensions: high/low-income and unemployment. The income classifications are based on percentiles: the low-income group is the proportion of the population who were in the bottom 20 percentile of earners in the Netherlands and the high-income group the top 10 percentile. The high-income level is similar to the highest tax band in the Netherlands for each year (Belastingdienst, no date), whilst the low-income threshold is close to the poverty threshold used by Statistics Netherlands (CBS, 2024a). The unemployed population were those who did not work for at least one month within the year. The economic social frontiers were identified only using individuals who were of working age (between 20 and 64). Additionally, students were excluded. This

¹ A detailed description of this classification can be found within Chapter 2 (Segregation and the Netherlands) of the thesis.

is because students are often not in full-time employment and their income is unlikely to be representative of their standing, potentially skewing the results.

Spatial scale

In this thesis I identify social frontiers across a range of different region and grid cell sizes. The different regions and grid cells used throughout the thesis are listed in Table 1.1, including the average grid cell population and range. The size of the region studied increases throughout the thesis. Chapter 3 studies frontiers in a single city (Amsterdam), this is extended in the subsequent two chapters to compare two regions and in Chapter 6 the frontiers are identified on a nationwide scale. By expanding the size of the study, the aim is to understand how the level of frontiers varies across regions. In Chapter 6 studying the whole country enhances the strength of the causal study with a greater statistical power and makes the findings more generalisable.

I choose in particular to study the Amsterdam and the Hague regions in Chapters 4 and 5 as larger regions/cities tend to be both the most ethnically and economically heterogeneous areas in the Netherlands (Hartog and Zorlu, 2009; Sleutjes, Ooijevaar and de Valk, 2019), and the Hague has been reported to have high levels of segregation (Aalbers, 2007, p. 6). Additionally, the majority of individuals with a non-Western background lived in larger urban areas particularly in the Randstad region which includes Amsterdam, Rotterdam, The Hague and Utrecht (Sleutjes, de Valk and Ooijevaar, 2018). There are clear signs of spatial discontinuities within the largest cities with the cities often having an uneven spatial distribution of different ethnicities (Karsten *et al.*, 2006). These factors make large cities an interesting area to study, specifically for social frontiers as these particular cities are untouched in social frontier research yet relatively well studied in segregation literature. Additionally, the regions differ in their demographics and spatial structure creating an interesting study. Currently, social frontiers have only been identified in Rotterdam in the Netherlands (Olnier *et al.*, 2024), leaving a large gap in our understanding of segregation in the Netherlands. Increasing the region size throughout the study aims for each chapter to build upon each other, with the thesis as a whole ultimately an extensive study into social frontiers in the Netherlands resulting in a comprehensive picture of them being created.

The frontiers in this thesis are identified at various spatial scales, in Chapters 3 and 6 I identified frontiers using 500m by 500m grid cells whilst in Chapters 4 and 5 I used the 100m by 100m grid cells. In terms of the characteristics of the varying grid cell sizes there is unsurprisingly larger variation within the population of the 500m by 500m grid cells, the average range in the population between the largest and smallest grid cells is approximately 7,000 residents across the years. Additionally, it is unsurprising that the average grid cell population is smaller in Chapter 6 when the whole of the Netherlands is included compared to just the city of Amsterdam, as the

more rural areas of the country are included where the grid cells are less densely populated, despite this the range remains relatively consistent.

Table 1.1: Regions and grid cell sizes

Chapter	Region Size	Grid Cell Size	Average Grid Cell Population	Average Grid Cell Range
3	City of Amsterdam	500m by 500m	2010: 1,329 2019: 1,457	2010: 6,948 2019: 7,194
4	Amsterdam and the Hague regions	100m by 100m	Ams: 95.43 Hague: 95.6	Ams: 1106 Hague: 706
5	Amsterdam and the Hague regions	100m by 100m	Ams: 95.43 Hague: 95.6	Ams: 1106 Hague: 706
6	Netherlands	500m by 500m	172.83	7,363

These statistics were created using the non-public microdata from Statistics Netherlands.

The 100m by 100m grid cells do pose an issue with the proportions appearing more extreme, with differences magnified by the model in way that may not be representative of the actual community. For the calculation of the proportion when the grid cell population is small the denominator will also be small meaning that the rate is easily impacted by even small increases potentially skewing the results. Additionally, the grid cells can be unrepresentative of actual neighbourhoods, Sleutjes, de Valk and Ooijevaar (2018) stated that the average neighbourhood size in Amsterdam was 8,228 residents (p. 203), much larger than the average population of the grid cells used in this study. This can be linked to the previous discussion on quantifying neighbourhoods and highlights that there is likely variation between the estimated grid cells and an individual's experience of a neighbourhood.

The choice of spatial scale for each chapter was made based on multiple factors: the purpose of the study and computational complexity. Chapter 3 was the first application of several of the social frontier identification methods in the literature meaning I was unsure of their computational complexity, thus began the study using the larger scale. Additionally, a crucial part of the study was the analysis of the pattern of social frontiers in Amsterdam, which for the maps to be clear and interpretable was not possible using the 100m by 100m grid cells. A larger spatial scale can be used to make more comprehensive statements about the impact of broader policies potentially making the findings more generalisable and the study more replicable as other countries may not have data available at such a fine spatial scale. Chapter 4 compares the prevalence pattern of the city of Amsterdam at 100m by 100m grid cells with Chapter 3's results about the region of Amsterdam measured using 500m by 500m grid cells. By comparing outcomes measured at different scales it in part addresses the modified areal unit problem (MAUP) which is the bias which can occur when point-based data (in our case the location of a household) is

aggregated into larger geographical areas – larger areas are more likely to smooth over spatial discontinuities which can explain why social frontiers may be less prevalent at larger spatial scales (Owen *et al.*, 2021, p. 212). For Chapter 6 the choice was largely down to computational complexity, measuring social frontiers in the whole country using the 100m by 100m grid cells would have massively increased the size of the study and thus would have been almost untenable to achieve within a realistic time scale and within the capability of the computer. However, Iyer and Pryce (2024) also note that for frontiers to be significant and have a ‘potent symbolic meaning’ they will need to be of ‘significant length’ (p. 784), by identifying frontiers using the 500m by 500m grid cells I am guaranteeing that individuals live close to frontiers which are at least 500m long which are more likely to have a substantial impact than the shorter 100m frontiers.

Methodology

The process of identifying social frontiers forms a large part of the thesis particularly in the first quantitative chapter: Chapter 3, in which I assess and discuss identification methods. The Lee and Mitchell (2013) and PAM clustering identification methods are applied in all four analytical chapters (Chapters 3 to 6). These methods are used to identify the various different types of frontiers both single dimension and intersectional, along with being used to create the excess embrittlement level. An explanation of PAM clustering is included in Chapter 3. However, the key concept behind the method is that it is a simpler, less computationally challenging approach to identifying social frontiers than the established Lee and Mitchell (2013) method. The clustering works by grouping grid cells into two groups and defining the boundary between them as a social frontier if the bordering grid cells were clustered into different groups.

Lee and Mitchell (2013) method (Adapted in Dean *et al.* (2019))²

The Lee and Mitchell (2013) method was adapted in the introductory social frontier paper by Dean *et al.* (2019) specifically to identify social frontiers. The method was then applied in the subsequent social frontier studies (Křížková *et al.*, 2021; Olnér *et al.*, 2024).

The Lee and Mitchell (2013) method uses an iterative estimation algorithm set within a Bayesian hierarchical model (Lee and Mitchell, 2013). The overall model is based on a binomial logistic regression model, where the value in each grid square is estimated with a random effect. The set of spatial random effects for all grid squares is modelled by a conditional autoregressive (CAR) model whose correlation structure is defined by a spatial weight matrix W . The jk^{th} element of this matrix equals one if grid squares (j, k) share a common boundary (not a vertex), and a value of zero otherwise. This matrix in conjunction with the CAR prior induces partial correlation

² Full explanations of the method can be found in Lee and Mitchell (2013) and Dean *et al.* (2019).

between random effects in neighbouring grid squares, while non-neighbouring grid squares are modelled as conditionally independent.

To identify statistically significant step changes (social frontiers), this model treats the elements in W as elements to be estimated, rather than being fixed at one if the cells are adjacent to each other. If the random effects in two neighbouring grid squares are estimated as being significantly different (their 95% credible intervals do not overlap), then the corresponding element in W is estimated as zero in the next iteration of the algorithm, otherwise the element of W is estimated as one. Once the new estimate of W is created, the spatial CAR model is re-fitted with the updated matrix to get updated estimates of the random effects. This process concludes when the estimate of W (and hence the social frontiers) does not change from one iteration to the next.

Dean *et al.* (2019) extended this to a double-hurdle approach by adding a post-processing step to the algorithm, making it specifically aimed at identifying social frontiers. There is a requirement for $w_{jk}^* = 0$ (an element of the matrix) which indicates a statistically significant step change. Additionally, the absolute difference between the fitted random effects of the two contiguous areas is required to be larger than the threshold value, calculated as in Equation 1.

$$Threshold\ Value = mean(d_{u_{j\sim k}}) + 1 * sd(d_{u_{j\sim k}}) \quad (1)$$

Where:

$$d_{u_{j\sim k}} = |u_j - u_k| \quad (2)$$

and u_j is the model fitted random effects at area A_j (Dean *et al.*, 2019, supplementary material).³ Although effective in capturing localised spatial autocorrelation and open frontiers, this method can be computationally challenging especially when applied to large areas or multiple iterations. It also runs into an issue with small grid cells where there are no individuals in the group of interest, for example grid cells where there are no non-Westerners, the method can fail and thus these cells' boundaries would have to be edited separately – adding to the processing time of the method. Additionally, the method is limited in that it only identifies social frontiers between two population groups meaning it can simplify the segregation level of an area.

³ A_j and A_k are two areal units which share a common boundary.

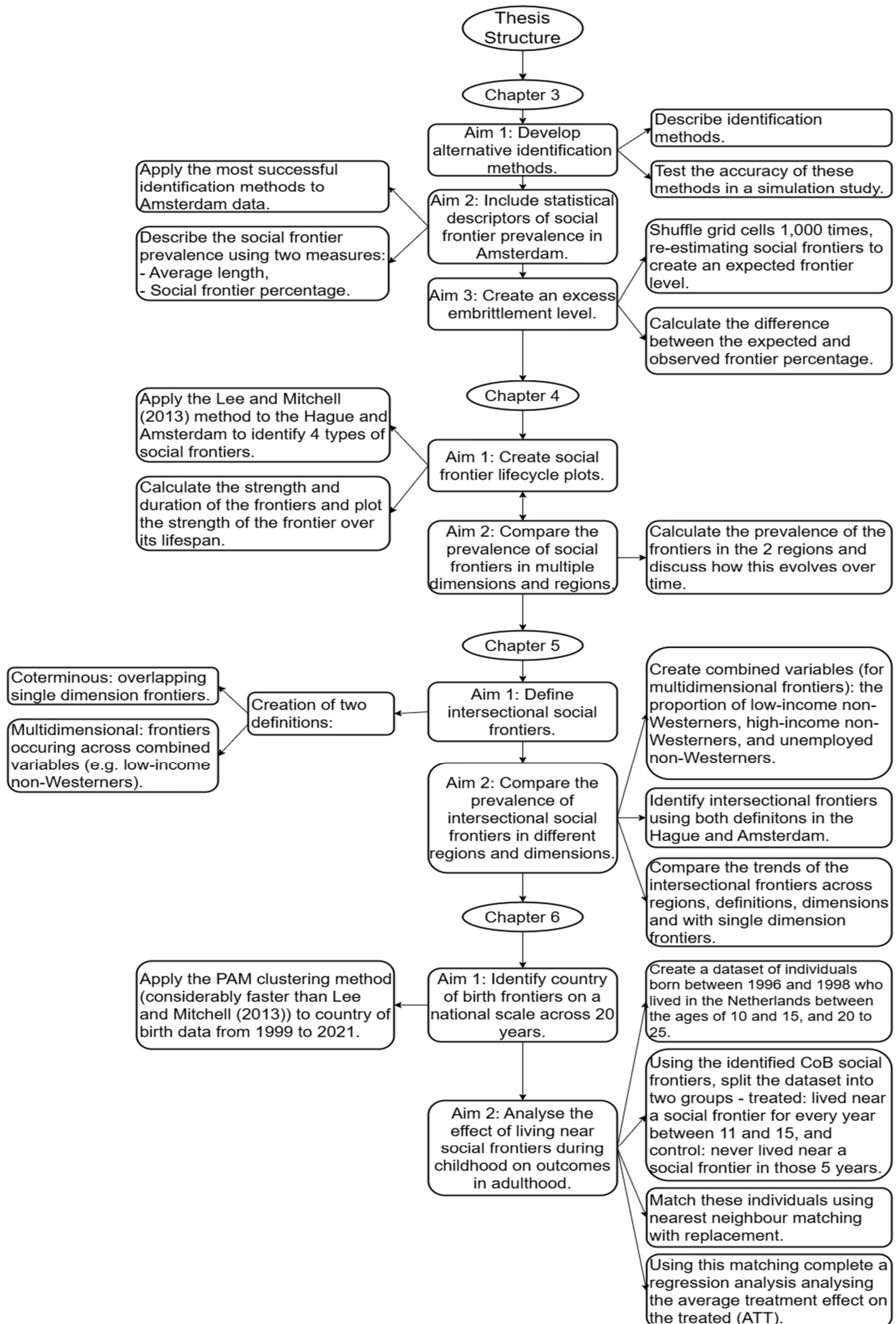


Figure 1.1: Thesis structure flowchart.

Thesis Structure

This thesis is composed of seven chapters: a main introduction (Chapter 1), a background section (Chapter 2), four quantitative chapters (Chapters 3 to 6), and a main conclusion (Chapter 7). The quantitative chapters are all standalone, with each focusing on a different gap within the current social frontier research in the Netherlands. The chapters link to each other through their methodology, findings, and study of the Netherlands. My thesis ends with Chapter 7, an overview of the thesis, along with a discussion on the key findings and results, and suggestions for future studies and potential policy implications. In the following section, I will provide an overview of the content of my four standalone quantitative chapters, discussing their key aims - the structure of these chapters is represented in Figure 1.1.

Chapter 2 – Segregation and the Netherlands

Chapter 2's main contribution to the thesis is that it provides an introduction to the migration history, migration background classification, segregation patterns, assimilation, and social housing policies of the Netherlands. This information can provide an insight and potentially explain segregation patterns identified throughout the thesis as well as explaining any potential effect identified in Chapter 6. Additionally, it discusses the outdated nature of the classification as well as suggesting potential alternatives for future studies.

Chapter 3 - Detecting Spatial Discontinuities in Amsterdam: Comparing social frontier identification methods and the prevalence of sharp transitions in the city of Amsterdam.

As the first standalone quantitative chapter within the thesis, Chapter 3 makes a number of methodological and empirical contributions to the social frontier literature. It does this by establishing an alternative identification method, including multiple prevalence measures, and using a shuffling algorithm to create the excess embrittlement level. Social frontiers are a relatively limited research topic in published literature and there remain large aspects of them that are unexplored. This includes the study of multiple time periods and regions (Iyer and Pryce, 2024), along with the creation of an expected level of frontiers for a given index of dissimilarity, identifying whether certain cities are particularly prone to social frontiers (Manley *et al.*, 2023). The introduction of alternative identification methods aims to improve the computational efficiency of the social frontier identification process, thus increasing the potential scope of future studies by making it easier to identify frontiers in multiple cities and over an extended time period in a shorter processing time. The accuracy of these methods is assessed using a simulation study to identify the best performing method. Not solely methodological, the chapter situates itself within Dutch segregation literature, comparing the prevalence of social frontiers in Amsterdam in 2010 and 2019 through aggregate prevalence measures and visualisations. Along with creating an excess embrittlement level and discussing its evolution over time and comparing it to the index of dissimilarity. The double

analysis, the inclusion of both prevalence measures and maps, means that as well as studying how frontiers change over time numerically, I can also identify areas within the city which are particularly prone to social frontiers.

This chapter's main contributions to the literature are:

1. The introduction and testing of alternative social frontier identification methods to provide a less computationally complex alternative to the established Lee and Mitchell (2013) method.
2. The comparison of the percentage and average length of country of birth social frontiers in Amsterdam in 2010 and 2019, noting the differences in the prevalence levels over time. Additionally, I use social frontier maps to identify areas in the city which have high levels of social frontiers, discussing why this may be the case.
3. A calculation of the excess embrittlement level for the city of Amsterdam and comparing this to the index of dissimilarity and the 2010 and 2019 results.

An alternative identification method is identified through the use of a simulation study to test the accuracy of three alternative social frontier methods along with the existing Lee and Mitchell (2013) method. The simulation study implements a binomial distribution to create 100 datasets under nine different scenarios, adapting the population size and strength of the social frontier to assess how the methods perform under different conditions. Country of birth social frontiers were then identified in the city of Amsterdam using PAM clustering (the method identified as best performing in the simulation study) and the Lee and Mitchell (2013) method. Their prevalence was measured through their average length and the percentage of frontiers in the area comparing the results in 2010 and 2019. The excess embrittlement level was calculated by shuffling the grid cells and creating a median (expected) level of frontiers.

Chapter 3's main contribution to the literature is the advancement in the social frontier identification process by the identification of the PAM clustering method as an alternative to the Lee and Mitchell (2013) method. The PAM clustering method has a faster processing time, identifying relatively similar frontiers as the Lee and Mitchell (2013) method for our study of Amsterdam. Although the difference in the processing time for this study is relatively minimal, I found that it increased relative to the size of the dataset. The simulation study finds that the results of the methods slightly contrast each other, the Lee and Mitchell (2013) method is relatively conservative, with a deliberately high requirement for defining social frontiers. Whereas the PAM clustering is more likely to overestimate the prevalence, with a lesser requirement for the boundary transition, but rarely misses a frontier. The conservative nature of the Lee and Mitchell (2013) method is a deliberate decision to ensure that the social frontiers identified are only the boundaries with the steepest transitions with a high threshold, thanks to the 2-step process, required to define a

boundary as a social frontier as to ensure that the frontiers identified are likely to be of significance.

As a result of the simulation study and processing time, I suggest that the Lee and Mitchell (2013) method should be used in the majority of studies, especially those exploring frontier prevalence and dynamics as only the strongest frontiers would be studied meaning the findings would be more cautious. The PAM clustering method plays a role in the social frontier literature when the study is sizeable, it is the first piece of research as it can provide an introduction for researchers on a larger scale, or when designing policies as it ensures that individuals are not missed. Additionally, Chapters 4 and 5 include the PAM clustering method in the appendix as a sensitivity test for the Lee and Mitchell (2013) method to provide a robustness to the study. This introduction of the PAM clustering method influences the methodology of the subsequent chapters in the thesis, making it feasible for frontiers to be identified across the whole of the Netherlands over a 20-year period in Chapter 6 within a reasonable time scale. In terms of the prevalence findings, the identification methods produced slightly varying trends. However, when I consider these results alongside the simulation study, it is likely that there was very little change in the prevalence of the country of birth social frontiers in Amsterdam when comparing 2010 and 2019. In relation to the excess embrittlement level, the number of social frontiers would be expected to be decreasing in the city.

Chapter 4 - The Lifecycle of Social Frontiers: A study into the dynamics and evolution of steep boundaries in the Netherlands.

Chapter 4 investigates how social frontiers evolve over their lifecycle, exploring how their strength changes as they develop. This is the first piece of empirical social frontier literature to graph social frontier dynamics, enhancing our understanding of how social frontiers behave. This chapter calculates the strength of the country of birth and economic frontiers over a 20-year period, exploring whether they behave as theorised by Iyer and Pryce (2024) in their lifecycle plot. This inclusion means that future literature can be extended to analyse whether the impact of the frontier varies depending on the length of exposure or the lifecycle stage.

Although its primary aim is to study frontier dynamics, Chapter 4 also makes a crucial contribution by being the first social frontier study to compare their prevalence in multiple cities and over an extended period of time. Additionally, it is the first to include multiple dimensions, identifying low- and high-income, unemployment, and country of birth frontiers. The chapter extends on Chapter 3 by identifying the four frontier dimensions in two regions: Amsterdam and the Hague between 1999 and 2021. Segregation is by its nature multidimensional (Mauritti *et al.*, 2016; Boterman, Musterd and Manting, 2021), with large overlaps existing between socioeconomic factors meaning that they can reflect each other (Ziller and Spörlein, 2020). By identifying frontiers in multiple dimensions, I am able to create a more comprehensive picture about the frontier prevalence and segregation patterns within

the regions and explore whether there is any variation in the results which is dependent on the region and dimension studied. If there is a clear difference in the results of the regions, then there could be additional factors which result in areas becoming prone to social frontiers such as housing or assimilation policies.

The prevalence of these frontiers is studied not only using the two prevalence measures introduced in Chapter 3, but also with additional strength, average duration, and re-emergence measurements, some of which were suggested by Iyer and Pryce (2024) as a way to create a clearer picture of segregation levels in an area. These measures make it easier to understand whether frontiers are likely to have a meaningful impact upon communities as we can have a better understanding of the extent to which individuals are exposed to frontiers.

This chapter's main contributions to the literature are:

1. The creation of social frontier lifecycle plots to assess how frontier strength changes over their existence and comparing them to the theorised lifecycle by Iyer and Pryce (2024), along with a discussion about whether the trends identified are consistent across regions, dimensions, and duration.
2. The identification of social frontiers in multiple dimensions including the first identification of economic frontiers.
3. The first comparison of social frontier prevalence in multiple regions, allowing a discussion about whether the frontier level and pattern is consistent across cities.

The results discussed in the main section of the chapter are based on the frontiers identified by the Lee and Mitchell (2013) method whilst the PAM clustering results are included in the appendix as a robustness check. The frontier prevalence is described using the percentage, average length, duration, and strength measures.

A large section of Chapter 4's results are focused on the evolution of social frontier strength over their lifespan with consistent findings across the regions and dimensions. The strength of the frontiers evolved almost identically to that proposed by Iyer and Pryce (2024). They steeply increased prior to their formation, stabilised during the middle of their life, before decreasing prior to their demise. What was also of interest was the high rate of frontiers re-emerging after their initial decline: over 50% reappeared, whilst the lifecycle plots showed they began to grow in strength quickly after their decline. This highlights the potential long-term nature of the divisions with communities often not fully merging in their neighbourhood. In terms of the frontier prevalence there were several key findings: country of birth frontiers were the most prevalent of all the dimensions in Amsterdam and the Hague; bar the high-income frontiers, the rate of frontiers grew between 1999 and 2021; and the Hague had a higher percentage of frontiers than Amsterdam in 2021 for all but the high-income frontiers.

Chapter 5 - Exploring the Intersectional Nature of Social Frontiers in the Netherlands: Identifying and studying the patterns of co-occurring steep spatial divides in Amsterdam and the Hague regions.

Chapter 5 makes several conceptual and empirical contributions to the social frontier literature as the first formal discussion and identification of intersectional social frontiers. The chapter combines sociology and social frontier literature to create two separate definitions of intersectional social frontiers. I describe these intersectional social frontiers using two different terms: coterminous and multidimensional. Coterminous intersectional social frontiers are identified where there is an overlap in the single dimension frontiers. This is closely related to Iyer and Pryce's (2024) description of intersectional frontiers: boundaries which are "co-occurring across multiple socio-demographic domains" (Iyer and Pryce, 2024, p. 784) and acts as representation of the intersection on a neighbourhood level. The second type of intersectional frontiers - multidimensional - are more closely related to Crenshaw's intersectionality framework (Crenshaw, 1989; Crenshaw, 1991). These frontiers are identified if there exists a steep difference between adjacent communities in the proportion of a combined variable, such as the proportion of low-income non-Westerners.

The definition of the multidimensional intersectional frontiers is more closely related to the sociological theory of intersectionality, where individuals are part of multiple groups and the frontiers represent zones of compounded difference. If these two groups are both what we would consider marginalised, then there could be an issue of double discrimination (Beale, 1979, cited in King, 1988; Crenshaw, 1991; Berdahl and Moore, 2006), with the close contact potentially increasing feelings of relative deprivation on multiple scales. Theoretical and empirical discussions about intersectional social frontiers have been incredibly limited and have only touched upon the potential variation between their effect and single dimension frontiers. I apply both of these intersectional social frontiers definitions to Amsterdam and the Hague over a period of 20 years, studying how results vary depending on the method, dimension, and in comparison to single dimension frontiers.

This chapter's main contributions to the literature are:

1. A conceptual discussion about intersectional social frontiers including defining them using existing social frontier and sociology literature, resulting in two definitions.
2. A study of the pattern of multiple different intersectional social frontiers in Amsterdam and the Hague regions.
3. A comparison of the pattern of single dimension and intersectional frontiers. This has potential policy implications as policies which are targeted at specific groups of individuals could bypass those who fall into multiple groups due to a lack of understanding of how the effect of the intersection could vary from the single dimension impact.

These two intersectional social frontiers definitions are applied to Amsterdam and the Hague regions, identifying low-income non-Western, high-income non-Western, and unemployment non-Western frontiers between 1999 and 2021. Coterminous intersectional frontiers are identified based on where single dimension Lee and Mitchell (2013) frontiers overlap. The multidimensional intersectional frontiers are identified by applying the Lee and Mitchell (2013) method in regard to the proportion of the combined variables: low-income non-Westerners, high-income non-Westerners, and unemployed non-Westerners in each grid cell.

Chapter 5 finds that the different intersectional social frontier definitions produce contrasting prevalence results. However, both definitions identified the high-income non-Western frontiers as the least common with Amsterdam and the Hague similarly divided. For the multidimensional intersectional frontiers, both the low-income non-Western and unemployment non-Western frontiers steadily increased, whilst the coterminous frontiers had a less consistent pattern but ultimately there was a higher percentage of frontiers in 2021 compared to 1999.

Of particular interest is the variation between the rate and pattern of single and intersectional social frontiers; there were periods when the intersectional frontiers occurred more frequently than the single dimension ones. Future researchers may be interested in whether this difference impacts upon the frontier effect and why certain groups of the population are more likely to become divided than others. This has potential policy implications, especially in relation to policies or targeted interventions as the current attempts to reduce divides might be bypassing those experiencing segregation on multiple levels due to their problems potentially being a result of the unification of multiple aspects (Crenshaw, 1989; Crenshaw, 1991).

Chapter 6 - The Impact of Living near Country of Birth Social Frontiers during Childhood on Socioeconomic Outcomes in Adulthood. A study which implements matching methods to investigate whether living near sharp boundaries between the ages of 11 and 15 has adverse effects on income and education in adulthood.

All of the previous research chapters (Chapters 3 to 5) have included a discussion about the potential effects of social frontiers on society and how these impacts highlight the importance of future research to deepen our understanding of social frontiers. These discussions have been based on the empirical findings of existing social frontier literature (Legewie and Schaeffer, 2016; Maguire, French and O'Reilly, 2016; Dean *et al.*, 2019; Křížková *et al.*, 2021; Olnér *et al.*, 2024; Piekut *et al.*, 2024; Lee, Pryce and Ramos, 2025), and theoretical discussions (Iyer and Pryce, 2024). Chapter 6 is the first study to assess whether living near social frontiers during childhood has a long-lasting impact on outcomes in adulthood. This extensive study is possible due to the detailed data which I am able to access through Statistics Netherlands, this data means that a strong covariate balance can be achieved. Based on the literature, I hypothesise that children will be negatively affected by

living close to social frontiers and this will result in lower levels of income and educational attainment in adulthood (Dean *et al.*, 2019; Iyer and Pryce, 2024). Neighbourhood effects literature has found a range of results: Troost, van Ham and Manley's (2023) results showed that living near affluence had a stronger effect than poverty on educational attainment, however, in contrast Nieuwenhuis *et al.* (2017) identified that increased feelings of relative deprivation amongst young people can impact adulthood outcomes, along with mental health and interaction levels. Additionally, several studies have identified the link between adult outcomes and crime and conflict (Bowen and Bowen, 1999; Damm and Dustmann, 2014), children's access to resources (Galster, 2012), and the potential positive impact of more diverse social networks (Bezzo and Vanhuysse, 2024). Yet none of these have studied whether these impacts are magnified when an individual lives near a steep boundary.

This chapter's main contributions to the literature are:

1. An analysis into the effect of living in close proximity to social frontiers during childhood on income and education outcomes between the ages of 20 and 25.
2. The identification of social frontiers on a national scale. This is far quicker with the implementation of the PAM clustering method proposed in Chapter 3.

The study uses matching and stratification methods to create two comparable groups - treated and control - which are as similar as possible based on observable characteristics. Achieving this balance means that I am able to compare income and educational outcomes whilst minimising selection bias. The treated group is composed of those who lived near a social frontier between the ages of 11 and 15, whilst the control group never lived near a social frontier between these ages. As an additional level of the study, I have also included a discussion about the variation in the treatment effect depending on exposure time to provide a robustness to the study. Following a comparison of covariate balancing, nearest neighbour matching with replacement was identified as the most successful method, with it being capable of maintaining a large sample size whilst minimising the standardised mean difference between the covariates. The method matches individuals by creating propensity scores: the probability of an individual being assigned to the treatment variable (Rosenbaum and Rubin, 1984). The score is created using a large number of covariates including parents' education, household characteristics, and neighbourhood level information. The aim is to create a study as close to a randomised trial as possible, attempting to extract only the impact of living near a social frontier by eliminating alternative explanations for any income or education differences.

Chapter 6's findings suggest that living near a country of birth social frontier during childhood is negatively associated with education and income outcomes in adulthood. Although the difference was not huge - treated individuals at the age of

25 earned on average €697 less a year and had spent on average 0.03 fewer years in education. There was a negative treatment effect at every age. Based on the results of the sensitivity analysis, I must be cautious in making causal claims about my results. However, this does not mean that living near the social frontier is insignificant on the outcomes nor that the effect is non-existent, just that it is difficult to detect a causal relationship. However, I can find an association between a decrease in earnings and education and social frontier proximity in childhood.

Chapter 2 Segregation and the Netherlands

Segregation Factors and Patterns in the Netherlands

There are many factors which influence where people live and thus the pattern of segregation in an area. These include systemic factors resulting from policy decisions or housing availability, neighbourhood sorting due to preferences or historical patterns, and economic factors such as work opportunities, gentrification, or the location of social housing. This chapter will briefly delve into these factors and consider how they present themselves in the Netherlands briefly noting regions or cities where their impact is visible. Spierenburg, van Cranenburgh and Cats' (2024, p. 5) segregation study in the Netherlands discovered an increase in non-Western segregation between 2015 and 2020 in the majority of municipalities in the Netherlands, however, they still displayed clear signs of diversity. There exist neighbourhoods in the Netherlands which are highly concentrated with non-European migrants (Andersson et al., 2018), which was also identified by Boterman, Musterd and Manting (2021) in Amsterdam where there were clusters of both non-Western and Dutch-born households. However, Hartog and Zorlu (2009, p. 24) denied the existence of mono-ethnic neighbourhoods for immigrants, finding that at least 80% of immigrants lived in neighbourhoods where the proportion of migrants from their own country was less than 10%.

The theory of homophily suggests that people are drawn towards those who are similar to themselves, for a range of different reasons. The preference can result in numerous divides within a person's environment, including their neighbourhood (McPherson, Smith-Lovin and Cook, 2001). Schelling's (1971) model of segregation posited that even a mild individual preference amongst individuals can result in a society becoming increasingly segregated. Thus, even if individuals do not have an extreme wish to live in a neighbourhood which is mostly similar to themselves, complete segregation will most often occur (Zhang, 2011, p. 170). Boschman and van Ham (2015, p. 1165) found that in Utrecht non-Westerners were 17 times more likely to choose to live in neighbourhoods which were made up of at least 80% non-Western migrants compared to neighbourhoods which were majority Dutch-born. They identified that individuals had a preference to live amongst their 'own-group': each "percentage point increase in the share of the own group leads to a 7% [$\exp(0.069) = 1.07$] increase in the odds of selecting the neighbourhood" (Boschman and van Ham, 2015, p. 1165). Residential context, particularly migrant concentration, also plays a role in determining whether an immigrant to the Netherlands decides to apply for Dutch nationality (Leclerc, Vink and Schmeets, 2023). Furthermore, Boschman and van Ham found that the ethnic composition of a neighbourhood was more consequential than economic characteristics, such as the housing market, on neighbourhood selection (Boschman and van Ham, 2015, p. 1165).

Homophilious preferences can arise due to a number of reasons including the increased access to resources and services that come as a result of living close to those similar to ourselves, with it often easier to form networks based on similarities, along with a reduced possibility of experiencing prejudiced or discriminatory views (Zorlu and Mulder, 2008). Substantial established communities can act as a large pull factor for individuals towards an area and explain settlement patterns such as the large Surinamese community in the Zuidoost borough of Amsterdam with the community originally established following the independence of Suriname in the 1970s (Mingle, 2018; Zahirović and Sterk, no date). Sleutjes, de Valk and Ooijevaar (2018) found that the Zuidoost borough has a large concentration of individuals with Surinamese origin - the largest concentration in the Netherlands (p. 208). Newer arrivals are likely to congregate towards established areas as it can act as a safety net (Musterd *et al.*, 2016), potentially further entrenching segregation.

It is not only homophilious preferences which can explain segregation patterns in the Netherlands or elsewhere, additional factors can include housing preferences, household income, or government policies (Sleutjes, de Valk and Ooijevaar, 2018). In 2013, migrants earned on average 34% less than non-migrants in Amsterdam (OECD, 2018, p. 37) highlighting the impact of economic factors on housing options. Low-income households particularly are likely to be impacted by the presence of social housing, an area which I discuss in detail later in this chapter in relation to its presence in the Netherlands. Additionally, segregation patterns are likely impacted by gentrification – the process where working class areas become increasingly unaffordable as wealthier individuals move in. The increased price of houses results in lower-income households being forced out of areas and thus clusters of higher-income households becoming visible within the spatial arrangement of the city. Separation levels⁴ were often higher in suburban towns surrounding large urban cores, with segregation in city centres as it becomes richer potentially due to gentrification (Spierenburg, van Cranenburgh and Cats, 2024, p. 5) and clusters of non-Westerners are shifting towards post-war neighbourhoods surrounding the city centre (Sleutjes, de Valk and Ooijevaar, 2018, p. 199).

In relation to location preferences, non-Western individuals are more likely to live in urban areas as these tend to be larger economic and political centres (Sleutjes, de Valk and Ooijevaar, 2018). These factors often vary in importance depending upon the initial reason for migrating. If moving through family reunification schemes, then immigrants' locations were often predetermined by the location of their existing family, with the individuals tending to be younger and more likely to be dependent on their family, whilst the housing market plays a lesser role in determining their location (Zorlu and Mulder, 2008). Family reunification was the most common reason for migration for non-EU or EFTA immigrants between 1999 and 2023 (nearly a third of

⁴ Spierenburg, van Cranenburgh and Cats define separation as “the share of this group [population of interest] living in these regions” (2024, p. 1).

the migrant population) (CBS, 2024b). Even family reunification migrants who initially moved into a mixed neighbourhood on arrival were more likely to leave, moving towards a highly concentrated ethnic neighbourhood (Zorlu and Mulder, 2008). Whilst in contrast, economic migrants' priorities tended to be more aligned with the housing market and the economic demographics of a neighbourhood, with them tending to gravitate towards areas where there are job opportunities or where they have a job lined up (Zorlu and Mulder, 2008). For example, the region Twente in the Netherlands has a large Turkish community due to a large group of workers arriving to work in the textiles industry (Sleutjes, de Valk and Ooijevaar, 2018). If communities are established due to the presence of industries or work opportunities, these are only likely to increase in size as the presence of co-ethnics is often a pull-factor for new arrivals to the Netherlands (Zorlu and Mulder, 2008; Sleutjes, de Valk and Ooijevaar, 2018). Those who arrive for specific jobs often have their location pre-determined by the company that they work for or where work is available (Zorlu and Mulder, 2008, p. 249).

To understand more about the segregation patterns and factors in the Netherlands, it is important to discuss the migration history of the Netherlands to understand historical factors and motivations for migration which can explain the push and pull factors impacting migration.

A Brief Summary of the Migration History of the Netherlands

The Netherlands is a country with a long history of immigration. In 2022 the country had a population of over 17.9 million with 14.5% of the population born outside of the Netherlands (CBS, 2025c) and 25.2% having a migration background (defined as having at least one parent born abroad) (CBS, 2022a). In 2022, over 400,000 people migrated to the Netherlands, with the majority (236,029) coming from another European country and the largest number of migrants from outside Europe coming from Syria (17,042) (CBS, 2025d). In the Netherlands, the largest migrant groups were of Turkish, Moroccan, Surinamese, and Indonesian origin in 2022 (CBS, 2025c) with the Turkish, Moroccan, and Surinamese population the dominant groups in the non-Western classification (Hartog and Zorlu, 2009).

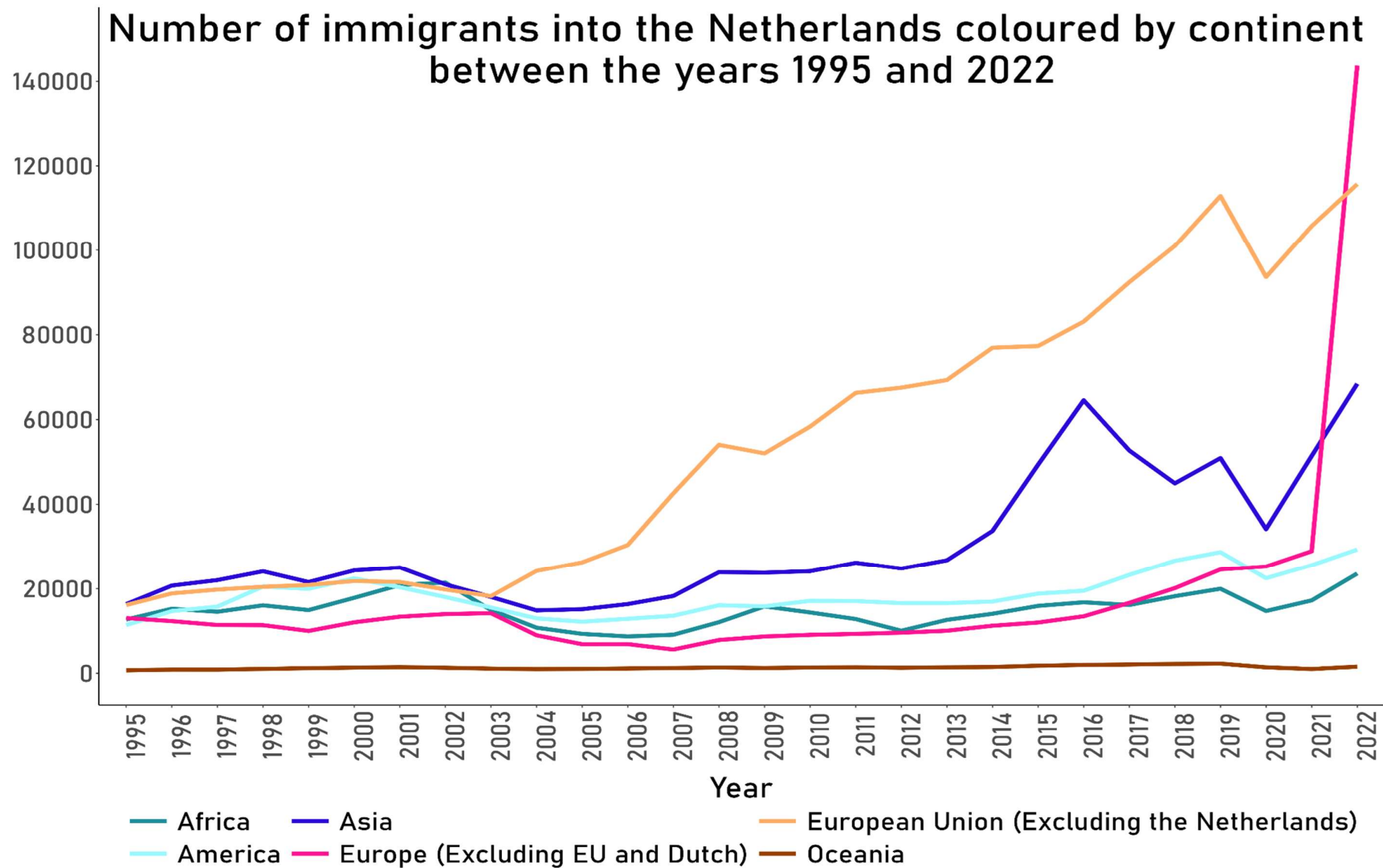
Since World War Two Dutch migration has been through a number of stages and cannot be separated from the country's history as a colonial empire from the 17th to 20th century. Firstly, the end of the Japanese occupation of Indonesia, a part of the Dutch empire prior to its independence in 1949, along with deteriorating relations between Indonesia and the Netherlands resulted in almost an exodus of Dutch individuals from Indonesia (van Amersfoort and van Niekerk, 2006, p. 325; Bosma, 2012a, p. 9). This eventually resulted in around 300,000 individuals arriving from Indonesia between 1945 and 1962, in three main waves (Bosma, 2012a, p. 9). The first was shortly after the end of World War Two, where a large proportion of arrivals were first generation Dutch families returning to the Netherlands (Bosma, 2012a, p.

9). The second and third waves occurred in the 1950s and 60s, again with a large number of repatriates but also Indische Nederlanders who may not have been to the Netherlands previously, and highly educated Indonesians (Bosma, 2012a, p. 9).

In the 70s and 80s immigration from both former and current Dutch colonies continued, with around 180,000 Surinamese migrants arriving in the 1970s (van Amersfoort and van Niekerk, 2006, p. 335). This was due to the declining economic and political situation in Suriname, along with the concerns about the potential loss of Dutch citizenship as a result of independence (Bosma, 2012a, pp. 9-10). The third post-colonial immigration flow came from the Dutch Caribbean (Jennissen *et al.*, 2023, p. 19). Those who arrived in the 1970s and 80s faced issues with growing levels of unemployment (Bosma, 2012a, p. 15). However, post-colonial immigrants were more similar to Dutch born citizens in their social capital than labour migrants, due to higher rates of intermarriages, a more similar schooling system to the Dutch, and a large proportion of them already being Dutch citizens (Bosma, 2012a, pp. 13-15).

In addition to individuals arriving from former colonies, following the Second World War there were a large number of migrant workers arriving from “southern Europe, Turkey and the Maghreb” following the changing economics after the war (Jennissen *et al.*, 2023, p. 19). In the 1960s ‘guest workers’ from Turkey and Morocco were often joined by their family through the family reunification scheme (Sleutjes, de Valk and Ooijevaar, 2018, p. 196) which was the most common reason for immigrating between 1976 and 2005 (Jennissen *et al.*, 2023, p. 20). Labour migration grew again in the 1990s and following the removal of the requirement of work permits for workers from the European Union in 2004 (Jennissen *et al.*, 2023, p. 19) - this increase is visible in Figure 2.1. In 2022 the majority of European migrants migrated to the Netherlands for work opportunities, 46.7%, whilst family reunification was the most common reason for non-European migrants, 41.58% (CBS, 2022b).

More recently, there has been an increase in asylum seekers following conflicts worldwide, with over 40% more asylum applications in 2022 compared to 2021 (CBS, 2023a). In the mid-2010s this included a large number of Syrian and Eritrean asylum seekers (CBS, 2025e), whilst more recently there has been an increase in Ukrainians, who have been relocating following Russia’s invasion in 2022, with over 97,000 individuals arriving in 2022 (CBS, 2025d) - again this increase is visible in Figure 2.1.



Source: CBS

Figure 2.1: Graph showing the number of immigrants arriving to the Netherlands by continent between 1995 and 2022 (CBS, 2023b).

Migration Background Classification and Assimilation Policies

Migration background definition

Within this thesis, individuals are defined as Western or non-Western based on the definition from Statistics Netherlands. Non-Western individuals are defined by Statistics Netherlands as people with origins from any country in Africa, South America, Asia (excluding Indonesia and Japan), or from Türkiye (CBS, 2025a). Western individuals are from the remaining countries and additionally in this study includes those born in the Netherlands. The definition is undoubtedly outdated: it does not reflect the geographical location of an individual's country of birth whilst having strong racial and ethnic connotations (van Schie, 2018). The Western and non-Western definitions reflect the concept of race (Coenders, 2024, p. 706) with groups described as Western when they closely resemble the Dutch born population (Yanow and van der Haar, 2013). There are examples of the grouping being confusing or inconsistent. For example, the majority of Asia is classified as non-Western, however, Indonesia and Japan are defined as Western, due to their "socioeconomic and cultural position" (CBS, 2025a). The terms are undoubtedly linked to the Dutch colonial past, with the categories revealing the power structure and binary classifications that came with Western colonialism (Coenders, 2024, p. 702).

The inconsistency and racial underpinning of the definition is not the only problem. The binary definition collapses multiple different nationalities into two terms, ultimately suggesting that all non-Westerners are exactly the same and the same for Westerners, hence, implying that all individuals in the same group have the same experience and are equally segregated within the Netherlands (Coenders, 2024, p. 706). This is despite Sleutjes, de Valk and Ooijevaar (2018) finding that the four main migrant groups in the Netherlands – the Turkish, Moroccan, Surinamese, and Antilleans – display very different geographical patterns in their spatial arrangement within the Amsterdam region. Additionally, the binary definition indicates that the two groups (Western and non-Western) are opposites suggesting they differ in terms of multiple socioeconomic dimensions (CBS, 2022c).

Dutch assimilation policies and their implications

The use of the non-Western definition has likely shaped the experiences of those who migrate to the Netherlands, with the labelling potentially shaping their image within Dutch society (van Schie, 2018). Fang and van Liempt (2021) found that non-Western migrants faced multiple obstacles in the housing market, including obtaining a mortgage if attempting to purchase a house along with difficulty accessing rental properties due to the perception that there will be differences in their cultural capital, such as not speaking Dutch fluently. Tselios *et al.* (2015, p. 431) suggested that the ability to speak Dutch was strongly linked to an individual's level of social integration

in the Netherlands. Social integration is defined as the level at which individuals are included within the Dutch system, along with their attitudes towards it and the relationships they have formed (Bosswick and Heckmann, 2006, p. 2; Tselios *et al.*, 2015, p. 416). Additionally, Zorlu and Mulder (2008, p. 248) argue that the spatial assimilation theory suggests that when migrants have a similar human capital, including religion, language skills, and cultural background, there exists a smaller social distance between them and those in the native country. A large proportion of first generation Moroccan and Turkish migrants - some of the largest non-Western groups - tend to become “stuck at the bottom of the Dutch labour market” (Witteveen and Alba, 2018, p. 97), which to some extent is also reflected among second generation migrants (Belfi *et al.*, 2022).

There have been policies, specifically the New Civic Integration Act 2021, which are directly aimed at encouraging those who move to the Netherlands to “participate in Dutch society”, aiming to increase their knowledge of Dutch culture, language and encourage employment (Government of the Netherlands, no date-a, para. 1). Even within this policy, there are exceptions: citizens of EU countries, along with those from Liechtenstein, Norway, Iceland, and Switzerland, are not required to partake (Government of the Netherlands, no date-a). Blankvoort *et al.* (2021) suggested that this can be seen to be dividing those arriving to the Netherlands by constructing an ‘other’ group which may be described as “undesirable migrants” who have clear differences from Dutch citizens (p. 3512). There is an expectation that migrants need to distance themselves from their own culture and through doing this they will become more assimilated within Dutch society (Ghorashi, 2014, p. 102).

Changes in migration definition

Statistics Netherlands has identified the controversial nature of the definitions, noting the out-of-date terminology and difficulty in interpretation (CBS, 2022c). This, along with the racial connotations and the inconsistency in the definitions, highlights the problematic nature. As a result, they are currently being phased out of datasets, with the Western and non-Western categories replaced by zones describing continents or specific countries which tend to have a strong migration history with the Netherlands such as Suriname or Türkiye (CBS, 2022c).

However, in the case of my thesis these were the terms used when my initial datasets were constructed and thus are used throughout the thesis for consistency. There is also an argument that the classification reflects a historical divide in the Netherlands (Bosma, 2012b; Coenders, 2024) and the terms having been used within studies and the country for an extended period of time has resulted in them continuing to influence politics and attitudes within the country (Coenders, 2024). The use of these terms within this thesis means it can act as an introduction to social frontier research in the Netherlands, with the opportunity for future research to include multiple different ethnic groups. The study of different groups will show

whether there is a difference in the experience of segregation for those from various countries and how this has impacted on the segregation levels in the area.

Social Housing in the Netherlands

The Netherlands has a long history of social housing and welfare support. In 2023 around 66% of rental housing was owned by non-profit housing associations, whilst housing association properties made up 28% of total housing nationwide and nearly 40% in Amsterdam (CBS, 2025f). In the Netherlands social housing is provided by non-profit housing associations, with rents controlled and a maximum monthly rent in 2024 of €879.66 (Government of Netherlands, no date-b). In 2024 a minimum of 92.5% of social housing was leased to those who earned less than €47,699 for one-person households and €52,671 in multi-person households (Government of the Netherlands, no date-b). The housing associations are not-for-profit with a requirement to invest any profits into improving the quality of housing and new developments. Despite social housing improving conditions for renters, there has been a shift towards rent liberalisation and reducing the share of social housing in the housing market (Howard, Hochstenbach and Ronald, 2023). However, the volatility of the housing market and high prices means that private renting or house ownership is unachievable or incredibly difficult to access for the majority of middle-income households, resulting in them remaining in social housing, hence the demand being greater than the supply (van Gent, 2013, p. 515).

Recently, there has been a decrease in social housing availability. Housing construction in 2010 was at its lowest rate since World War One (Hochstenbach, 2024), and housing associations have been selling off a significant amount of social housing in the city centre (Hochstenbach and Musterd, 2017). This reduction adversely affects lower-income households resulting in greater income inequality (Boterman, Musterd and Manting, 2021), with lower-income households frequently priced out (Hochstenbach and Musterd, 2017). The limited social housing availability combined with the reduction in benefits unfavourably affects disadvantaged groups (e.g. those who are unemployed or on low-incomes), often further reducing housing choices (Marcuse and Van Kempen, 2000).

Conclusion

This chapter acts as an introduction to segregation and migration in the Netherlands. It does this by firstly discussing segregation factors considering how these manifest themselves in the spatial economic and country of birth patterns in the Netherlands. The neighbourhoods which I identify in this chapter as having a large migrant population, being prone to gentrification or which have experience large growth into the suburbs, can be used to explain the patterns of social frontiers which are identified in the forthcoming chapters. This is particularly the case for Chapter 3, where the location of the frontiers in Amsterdam is discussed in relation to gentrification and established communities. The social frontier patterns are also

discussed in relation to social housing policies and availability, noting how a reduction in social housing resources has resulted in lower-income households being priced out of the city centre potentially explaining the social frontier patterns identified in the chapter. Additionally, to understand the patterns of segregation and social frontiers which are discussed both in this paper and the literature it is useful to have an insight into the migration history in the Netherlands as this can explain why some migrant groups are considered to have assimilated more easily than others.

A key part of the chapter is the description of the migration background classification which is an ever-present concept used throughout the thesis. Its inclusion is partly a limitation of the thesis due to the problematic nature of the definition – its relation to the Dutch colonial past and the enforcement of a binary classification. Thus, this chapter plays a key role in explaining why I still used these terms within my thesis along with suggestions for how future studies could include alternative classifications to improve the social frontier literature in the Netherlands.

This chapter provides a crucial insight and introduction into segregation patterns, the migration history, terminology, and segregation policies of the Netherlands. The topics discussed within it will re-occur throughout the thesis, whether that be to explain social frontier patterns, provide reasoning for method or scale choices, and to generate suggestions for future studies.

Chapter 3 Detecting Spatial Discontinuities in Amsterdam

Comparing social frontier identification methods and the prevalence of sharp transitions in the city of Amsterdam.

Abstract

Introduction: The identification and effect of steep boundaries between adjacent communities, defined as ‘social frontiers’, is an under-researched topic in social science studies. This is despite them having the potential to heighten territorial attitudes, increase crime rates and negatively influence life-outcomes (Dean *et al.*, 2019; Iyer and Pryce, 2024). To extend the study of social frontiers, researchers need to be able to identify frontiers using large datasets. This would be made considerably easier with the development of an alternative, more computationally efficient identification method along with the inclusion of prevalence measures to facilitate comparisons.

Objectives: This chapter introduces and tests the accuracy of four social frontier identification methods: the established Lee and Mitchell (2013) method and three alternatives: Partitioning Around Medoid (PAM) clustering, k-means clustering, and an absolute contiguous difference method. In addition, it describes the level of social frontiers in Amsterdam through two prevalence measures along with computing the level of excess embrittlement.

Methodology: The chapter implements a simulation study to assess the accuracy of the identification methods when applied in a controlled data environment. These identification methods are the first application of non-spatial clustering, seemingly counterintuitive with the spatial requirement of the boundaries, but allows abrupt neighbourhood boundaries to be identified efficiently without imposing closure with frontiers. Spatial clustering methods may result in frontiers being forced to lie along the full perimeter of a neighbourhood. In addition to clustering methods, the chapter implements an absolute contiguous difference method along with the existing Lee and Mitchell (2013) method (Lee and Mitchell, 2013; Dean *et al.*, 2019). To assess the social frontier prevalence, I implement two measures: the percentage of social frontiers and their average length, comparing these to an expected level created through the randomisation of grid cells.

Results: The PAM clustering method correctly identified 95.8% of social frontiers whilst also correctly identifying 96.3% of non-social frontiers on average. It outperforms the other methods in its ability to correctly identify a frontier, but is more likely to identify a false positive, classifying a border as a social frontier when it is not, with a false positive rate of 3.7%. In terms of our study the PAM clustering operated quicker than the established Lee and Mitchell (2013) method, but this does come at a cost of the robust nature of the Lee and Mitchell (2013) method. The prevalence of social frontiers in Amsterdam was discussed using both the Lee and

Mitchell (2013) and PAM clustering methods, with little change between the prevalence in 2010 and 2019. Whilst the pattern for the expected level corresponded with the index of dissimilarity.

Conclusion: The chapter concludes that the PAM clustering method is a strong alternative to the Lee and Mitchell (2013) method as the frontiers identified are relatively similar and the processing time was 50 seconds shorter in this particular study. This finding is especially consequential for studies identifying frontiers on a large scale as the difference in the processing time can make large studies which would be almost unfeasible with the Lee and Mitchell (2013) method possible. It suggests that there are scenarios in which each method is suitable with the Lee and Mitchell (2013) method identifying stronger and likely more significant frontiers. Whilst the methods differed slightly in their results, the social frontier prevalence in Amsterdam differed slightly in 2010 and 2019, there was not a notable change in the overall prevalence. However, the expected prevalence under randomness decreased as the index of dissimilarity decreased, with the city consistently having fewer divides than if it was arranged randomly.

Introduction

Sharp transitions between communities can result in intergroup tension due to them inhibiting interaction and evoking territorial attitudes (Iyer and Pryce, 2024). Equally, such boundaries provide an opportunity for positive contact, with the close contact potentially promoting social harmony (Dean *et al.*, 2019), resulting in the formation of cross-ethnic relationships and reduced feelings of resentment between groups (Zorlu and Mulder, 2008). Established sharp boundaries and territory zones between communities can foster stability within groups by clearly demarking territory, reducing the ambiguity of land ownership (Gold, 1982; Legewie and Schaeffer, 2016). Yet as noted in my introduction chapter, boundaries are vastly under-researched in social science literature (Legewie and Schaeffer, 2016, p. 131; Kramer, 2017; Dean *et al.*, 2019), meaning we are unsure what circumstances foster particular outcomes or what the most likely effect would be. Sharp boundary identification is bypassed by traditional segregation measures, such as those discussed in Massey and Denton (1988), with neighbourhood transitions often overlooked as indexes tend to ignore the spatial arrangement of a region (Tselios *et al.*, 2015; Dean *et al.*, 2019). This means that measures cannot fully capture how communities live alongside each other, highlighting the importance of including social frontiers as an additional dimension of segregation in research.

The development of social frontier and boundary literature is in some form a response to these gaps in the literature, a growing body of empirical work is slowly emerging, with studies exploring a variety of outcomes associated with contested and sharp neighbourhood boundaries including anti-social behaviour and crime (Legewie and Schaeffer, 2016; Dean *et al.*, 2019; Křížková *et al.*, 2021), geographical mobility (Olnier *et al.*, 2024), and mental health (Maguire, French and O'Reilly, 2016; Lee, Pryce and Ramos, 2025). The introductory study by Dean *et al.* (2019) first coined the term 'social frontiers' for steep transitions, focusing on the empirical identification of steep boundaries by adapting Lee and Mitchell's (2013) localised spatial smoothing model specifically for identifying social frontiers. This has become the most widely used approach in social frontier literature (Dean *et al.*, 2019; Křížková *et al.*, 2021; Olnier *et al.*, 2024; Myatt *et al.*, 2026).

However, the method is not without its flaws: it is computationally complex meaning the process of identifying social frontiers in a large area is time consuming along with being difficult to interpret with a complex mathematical basis. For this study in particular, the data analysis was executed in a remote access due to the highly confidential nature of the data, there was a limit on power and memory and thus to apply this method on a region especially multiple times could take several days if not weeks to process. This study aims to address this by introducing and testing the accuracy of three alternative frontier identification methods: two clustering methods and an absolute contiguous difference method. The aim is to identify a method which identifies similar frontiers at a comparable rate to those identified by the Lee and Mitchell (2013) method whilst having a shorter processing time. Thus, reducing the

complexity of future studies by making it easier to study large areas, repeat the identification process across multiple time periods, and create an excess embrittlement level using a shuffling algorithm. This inclusion is important for research as it means that researchers can consider whether regional characteristics, such as natural or physical boundaries, impact upon the level of social frontiers by comparing multiple regions and studying changes over time. Along with a discussion about whether there are particular policies or housing planning decisions that can explain changes in social frontier prevalence or why some areas are prone to steep transitions.

Despite advances in the theoretical and empirical understanding of social frontiers, the question of the difference in social frontier prevalence in different regions remains unexplored. As far as I am aware, the only previous social science steep boundaries paper to include a statistical prevalence description measure is Harris' (2014) study of average discontinuity in local authorities in the UK. However, Harris' paper (2014) does not specifically measure social frontiers, with areas not required to boundary each other but only to be within 100m. Thus, it does not specifically focus on the close contact of the contrasting groups, nor does it consider alternative prevalence descriptors such as average length. Iyer and Pryce (2024) theorised that for a frontier to have a significant impact upon communities it is required to be of substantial length. In terms of measuring the steepness of boundaries, Legewie and Schaeffer (2016) and Wickes *et al.* (2025) included an edge intensity measure to describe the severity of individual boundaries, investigating the relationship between edge intensity and conflict levels in neighbourhoods. Previous social frontier papers have included visual representations of frontiers (Dean *et al.*, 2019; Křížková *et al.*, 2021). However, none of the papers have included a statistical description of the level of frontiers in an area nor compared the rate of social frontiers over time, meaning that they were unable to form concrete empirical conclusions about how the social frontiers evolved over time. I aim to address this gap by including two prevalence measures to evaluate whether the country of birth social frontiers increased in Amsterdam. The prevalence measures will facilitate temporal and spatial comparisons, inviting a discussion about why social frontiers proliferate in some areas but are absent from others. In this chapter I identify frontiers in two non-consecutive years, 2010 and 2019, making it the first social frontier study to use aggregate measures to compare prevalence in a social science environment.

This chapter not only includes the percentage and length prevalence measures, but also the level of excess embrittlement. This term has been defined by Manley *et al.* (2023) as the degree to which the observed level of frontiers differs from what we would expect based on the region's index of dissimilarity. This inclusion makes it possible to study how far the spatial arrangement of a city has contributed to the social frontier prevalence. The embrittlement level is calculated through a shuffling algorithm which randomises the proportions of the group of interest across the different neighbourhoods (Manley *et al.*, 2023). The measure explores the

relationship between the social frontier prevalence and the traditional segregation index: the index of dissimilarity (D) which measures how evenly a particular population group is distributed across areal units, the larger the score the greater the level of segregation in an area (Duncan and Duncan, 1955; Massey and Denton, 1988). The randomisation of the grid cells maintains the overall distribution of group proportions across neighbourhoods, keeping D constant, but means an average social frontier prevalence level can be calculated for various arrangements of the city for a given D. The excess embrittlement measures the extent to which this expected level created through randomness differs from the actual observed level. There is an expectation that if there is a significantly greater level of frontiers than forecasted (based on the index of dissimilarity), then there are underlying processes impacting upon the divisions within areas (Manley *et al.*, 2023).

In the remainder of this chapter, I argue that social frontiers are the missing dimension of segregation research, discussing this along with their causal effects and the level of ethnic segregation in Amsterdam. Following the exploration of the literature, I will discuss the methods and data I employ. In the findings section I discuss the success of the alternative identification methods based on the results of the simulation study, before examining the frontier prevalence and excess embrittlement level in Amsterdam. The chapter concludes with a summary of the findings, and a discussion about the limitations of the study and suggestions for future studies.

Literature

The missing dimension of segregation research

Boundaries between contrasting neighbourhoods are an area of research which has frequently been overlooked by segregation researchers (Legewie and Schaeffer, 2016, p. 131). Tselios *et al.* (2015) highlighted that neighbourhood and social interaction studies often do not control for the characteristics of adjacent neighbourhoods and for interaction between them, with very little known about the spillover effect between neighbourhoods (Dekker, Rijnks and Mierau, 2021). Relationships between contrasting socioeconomic groups are likely to be impacted by their level of interaction, and the potential for this interaction is likely to be related to how neighbourhoods are arranged within a region. Communities who live alongside each other have an increased opportunity for contact which can result in the formation of relationships, however, for this contact to be positive Allport (1954) - social psychologist, the author of the 'contact hypothesis' - suggested it has to occur under certain conditions including equality of groups and respect for authority.

Social frontiers differ from social boundaries as the latter term is often used to refer to non-spatial barriers between groups, whereas social frontiers specifically require a spatial element. Communities must be directly bordering each other with the severity

of the frontier measured in terms of how the residential composition of either side of the neighbourhood boundary contrasts with each other (Dean *et al.*, 2019). As such, they diverge from administrative delineations of geographical boundaries due to their social requirement despite due to the methodology being forced to lie on the administrative boundaries (Dean *et al.*, 2019). Social boundaries are defined as “any contiguous zone of contrasting density, rapid transition, or separation between internally connected clusters of population and/or activity” (Tilly, 2004, p. 214). Examples include gender-boundaries in the workplace and the unequal access to resources or opportunities (Lamont and Molnár, 2002, p. 168). The spatial nature of social frontiers is crucial as it is the spatial arrangement and the distribution of neighbourhoods that shapes the level of interaction and thus the experiences of those who live near these boundaries. I discussed in the introduction (Chapter 1) how social frontiers can be thought of as partially spatial, with the arrangement within the grid cells not considered, however, what is important about the grid cells is that they border as this increases the likelihood of contact between communities and likely increases territorial attitudes.

The absence of literature studying transitions between different communities has partly been a result of segregation measures failing to identify sharp boundaries. Massey and Denton (1988) - whose work laid the foundations for segregation research - classified segregation indexes into five dimensions: evenness, exposure, concentration, centralisation, and clustering. To understand the segregation level of an area, we need to use a combination of these segregation dimensions (Massey and Denton, 1988, p. 309). However, even if all these indexes are included in a study, they cannot capture the transitions between neighbourhoods. These indexes all share the common issue of being aspatial, meaning that they view each neighbourhood as individual, and thus independent of each other (Tselios *et al.*, 2015; Yao *et al.*, 2019), with the location of the individual units not impacting the segregation score (Jakubs, 1981), hence boundaries are ignored. For example, two cities could have the same index of dissimilarity (D), a measure of evenness, but the geographical arrangement of neighbourhoods in the region is completely different where one city can have steep and sharp neighbourhood boundaries, whilst the other has smooth transitions. If we solely measured segregation based on the index of dissimilarity, then we may expect individuals to have the same experiences in both neighbourhoods, yet that is unlikely to be true. The effect of the neighbourhood on individuals is not restricted to solely their grid cell. There exist spillover effects between communities, with social interaction and crossover in movements likely between individuals living in close proximity to boundaries (Kramer, 2017; Dekker, Rijnks and Mierau, 2021, p. 2). These methods have dominated the segregation field despite recent research highlighting the importance of the spatial arrangements or spillover effect of neighbourhoods on social inequality and aspects of society (Kramer, 2017, p. 82; Legewie, 2018, p. 1974). To advance segregation studies, we should consider social frontiers as an additional dimension of segregation,

deepening our understanding of how the spatial arrangement of neighbourhoods impacts upon society.

Studies in a variety of research topics have employed the use of boundary detection algorithms. These include cancer detection (Lu and Carlin, 2005), respiratory disease risk (Lee and Mitchell, 2013), and brain tumour detection (Kalaiselvi, Sriramakrishnan and Vasanthi, 2015). Within the limited social science studies that identify sharp geographical boundaries, there have been a variety of methods implemented. Edge detection algorithms (Legewie and Schaeffer, 2016) and areal wombling techniques (Legewie, 2018) have been used to detect racial boundaries - where adjacent neighbourhoods are declared to share a contentious boundary when the difference in the proportion of the groups exceeds a predetermined threshold. For many of these approaches there is a requirement to set a threshold as to when a boundary is severe enough to be significant (Lu and Carlin, 2005; Legewie, 2018) or defined as a social frontier (Dean *et al.*, 2019). Setting a threshold potentially introduces researcher bias, with researchers holding a greater power as they ultimately choose when a boundary is determined to be steep enough to be defined as a social frontier. If these thresholds are not consistent across studies, then it is not possible to fairly compare the results for regions and time periods as there is no set frontier definition.

Additionally, some of the existing approaches make use of smoothing algorithms with the aim of reducing random noise and errors from measurements (Legewie and Schaeffer, 2016). However, in the process of identifying social frontiers the use of smoothing algorithms could potentially mask the discontinuities which we are attempting to identify, reducing the reliability of the model, and distorting the overall prevalence of social frontiers. To resolve the issue of smoothing and open frontiers (allowing cells to have a social frontier on one side but a smooth transition on the other), Dean *et al.* (2019) adapted the Lee and Mitchell (2013) detection model. They added the additional requirement of the threshold – described additional step to the algorithm in order to identify a statistically significant change - this is described in the methodology section in Chapter 1. As far as I know this is the only method used to identify social frontiers. It was applied in Dean *et al.* (2019), Křížková *et al.* (2021), Olnér *et al.* (2024), and Myatt, Lindgren and Osland (2026) on UK, Czech, Dutch, and Norwegian data respectively. The method works at a high degree of accuracy; however, it is time consuming and computationally complex particularly when applied to large datasets.

This chapter acknowledges that the complexity of the Lee and Mitchell (2013) method can make it difficult to apply on a large dataset. Thus, attempts to build upon the methodological advances of Dean *et al.* (2019) through the application and testing of three alternative identification methods, evaluating their ability to accurately identify frontiers through a simulation study.

Causal effects of social frontiers

Empirical findings

To understand the importance of studying social frontiers, it is crucial to discuss their potential consequences on society. There is relatively little empirical research examining how social frontiers or contested boundaries impact on society, with the majority of research focusing on their consequences for neighbourhood crime and conflict levels (Legewie and Schaeffer, 2016; Legewie, 2018; Dean *et al.*, 2019; Křížková *et al.*, 2021; Wickes *et al.*, 2025). Steep boundaries can increase competition over territory, resulting in hostile relationships and eventually leading to an increase in crime rates. Mixed results have been identified, with Dean *et al.* (2019, p. 282) finding higher adjusted crime counts close to ethnic or country of birth social frontiers compared to administrative boundaries. Whilst, Křížková *et al.* (2021, p. 297) found a statistically significant negative relationship between crime and foreigner frontiers, indicating crime rates were lower within 100m of a social frontier. Future research could consider the different circumstances which influence the impact of the social frontiers, whether that be the cultural distance between the contrasting groups, the effect studied, or the social context with the country or region studied.

Dean *et al.* (2019) emphasised how we should view social frontiers as dynamic contested spaces whose impact is likely to evolve over time. Despite this, the majority of studies have been cross-sectional, studying a single city at a moment in time (Iyer and Pryce, 2024). When frontiers are developing or boundaries are being contested then territory is likely to be less well-defined which results in higher levels of conflict and disagreement (Legewie and Schaeffer, 2016). Steep boundaries often “lack the social control and cohesion of adjacent homogeneous areas” (Legewie, 2018, p. 1960), with interface areas becoming unstable and conflict arising over territory (Bannister, Kintrea and Pickering, 2013, p. 487). If the intensity of the boundary stabilises, relationships between the groups may improve as living in close residential proximity to different ethnicities has been found to enhance cross-ethnic relationships, reducing tension and easing social mobility (Zorlu and Mulder, 2008, p. 247). The formation of relationships across boundaries can reduce the strength of the frontier, blur the edges, and smooth the transition between communities (Dean *et al.*, 2019; Iyer and Pryce, 2024). To explore frontier dynamics and the variation in their effects, social frontiers need to be identified in multiple time periods and regions within a single study. For this to be feasible within research, there is an imperative for developing consistent, computationally efficient, and accurate social frontier identification methods which can identify frontiers on a larger scale in a short space of time.

Theorised effects

There are a number of additional theorised impacts of social frontiers, most of which are yet to be explored empirically. These include the mooted impacts on wellbeing including mental health and feelings of self-worth, negative implications for life course outcomes, and reduced social mobility. These effects can be explained by the ability of the social frontiers to potentially reduce geographical mobility, increase feelings of relative deprivation, and indirectly result in higher levels of crime and conflict (Iyer and Pryce, 2024). Lee, Pryce and Ramos (2025) found strong and consistent evidence of a negative association between the prevalence of mental health problems and the intensity of ethnic and religious social frontiers. However, this is an ecological correlation – not measured at an individual level - and does not necessarily imply a causal link.

There have also been potential positive effects of social frontiers theorised. Social frontiers can become sites of intergroup dialogue and alliance-building due to the similarities between the residents resulting in the existence of stronger networks (Zorlu and Mulder, 2008). The close contact provides groups an opportunity to learn about each other and experience positive contact. This is only as long as the contact is under constructive circumstances where groups consider themselves equal, cooperate towards common goals and respect authority (Allport, 1954). These conditions facilitate the increased contact that comes with proximity resulting in positive outcomes with lower levels of prejudice beliefs about contrasting groups (Pettigrew and Tropp, 2005, p. 271).

Both the theorised and empirical effects highlight the importance of developing a robust methodological foundation for social frontier research. This includes having a choice in identification methods depending on the requirement of the study and being able to describe the social frontier prevalence in an area, both of which are addressed in this chapter.

Ethnic segregation in Amsterdam

In addition to making methodological contributions, this chapter studies the prevalence of country of birth social frontiers in the city of Amsterdam, the capital city of the Netherlands, in 2010 and 2019. Social frontiers research in the Netherlands has been limited, Olnier *et al.* (2024) identified country of birth social frontiers in Rotterdam but the rest of the country remains unexplored. Amsterdam is the largest city in the Netherlands with a population of over 800,000 and over 35% of the population having a non-Western background in 2019 as visible in Table 3.1. The city experienced an over 1 percentage point growth in its proportion of non-Westerners between 2010 and 2019.

Table 3.1: Population statistics in the City of Amsterdam.

Year	Population	Non-Western Population	Non-Western Percentage	Index of Dissimilarity
2010	789,607	274,434	34.75%	0.3955
2019	882,710	316,253	35.83%	0.3526
These statistics were created using the non-public microdata from Statistics Netherlands.				

Figure 3.1 shows the spatial arrangement of non-Westerners in Amsterdam with clusters clearly visible. There is a large concentration of non-Western individuals particularly on the outskirts of the city centre, with the clusters especially visible in the north-east (Noord district), the area to the west of the city centre (Nieuw-West district) and the Zuidoost borough of Amsterdam (an exclave of the city due to it not sharing a border with any of the other boroughs)⁵ (OECD, 2018). The Nieuw-West borough is a post-war construction whose non-Western population grew rapidly between 2000 and 2010 (Entzinger, 2019), with a particularly large Turkish and Moroccan population (Sleutjes, de Valk and Ooijevaar, 2018).

The Zuidoost borough has a large non-Western population. It is home to over 130 different nationalities (Zahirović and Sterk, no date) with especially large groups of individuals originating from Suriname, the Caribbean and West Africa (Smets and den Uyl, 2008). One of the largest neighbourhoods within the borough is the Bijlmermeer, a 1960s high-rise development aimed at Dutch middle-class families. However, a missing metro line connecting the area to the centre of Amsterdam, along with a change in housing style preference, led to the area becoming unpopular amongst the Dutch middle-class (Zahirović and Sterk, no date). This coincided with the arrival of a large number of Surinamese immigrants in the early 1970s, discussed in greater detail in Chapter 2. These migrants were often forced into poor quality housing on arrival due to negative political beliefs spreading surrounding immigration and the population becoming increasingly separated from the Dutch born population (van Gaalen, 2025, p. 11). As a result of this, in 1974 around 90 families took action by squatting in empty apartments in the Bijlmermeer. The protests and arguments which followed eventually resulted in the residents being given leases, leading to the establishment of a long-standing Surinamese population in the Bijlmermeer (van Gaalen, 2025). These densely populated clusters are visible in Figure 3.1 at the top of the Zuidoost borough.

Clearly visible in Figure 3.1 is the higher percentage of non-Westerners in the outskirts of the city compared to the city centre. Amsterdam has clear signs of gentrification, where the upward social mobility of one group, typically the middle-class, has resulted in the displacement of the lower status residents (Musterd *et al.*,

⁵ A map showing the outlines of the boroughs in Amsterdam is visible in Figure A.1 for reference when describing the location of the frontiers.

2016, pp. 230-231). This displacement is more likely to occur where the social distance between groups is greater,⁶ as the odds of leaving the neighbourhood increases (Musterd *et al.*, 2016, p. 239). The high house prices in the centre of Amsterdam have resulted in it becoming unaffordable for the majority of individuals, especially low-income households (Hochstenbach and Musterd, 2021). This, along with the selling off of social housing (Hochstenbach and Musterd, 2021, p. 482), has resulted in 'exclusionary' displacement. This is where even if homes are becoming vacant, lower-socioeconomic groups are excluded as they cannot afford the prices or there are additional factors making the region inaccessible (Marcuse, 1985; van Gent, 2013, p. 516). Furthermore, the centre has increasingly become a business hub for large companies with deindustrialisation frequently attracting foreign highly-skilled workers who have similar housing preferences to the highly-skilled native Dutch (Burgers and Musterd, 2002, pp. 408-409; Sleutjes and Boterman, 2016, pp. 165-166; Beckers and Boschman, 2019). Whereas non-Westerners were more likely when choosing a house to prioritise "low housing costs, proximity to work and having a private garden" (Sleutjes and Boterman, 2016, p. 165).

Gentrification and housing policies are a large influence on the housing location choice of migrants, but they work alongside several other factors which include the large presence of social housing in the Netherlands (Zorlu and Mulder, 2008; Hartog and Zorlu, 2009). Social housing strategies are discussed within Chapter 2 with their importance particularly relevant in Amsterdam due to housing association properties making up nearly 40% of housing stock in 2023 (CBS, 2025e). In Amsterdam there is on average a 13 year wait for social housing, with those who earn just above the social housing limit often unable to afford to purchase a house or the high price of private rental accommodation (van Deursen, 2023). Changes in social housing availability is potentially consequential for the spatial arrangement of non-Westerners in the Netherlands as migrant households were "significantly less likely to own their own home" across all the age groups (CBS, 2024c, p. 47) and thus more likely to live in rental accommodation (CBS, 2024c, p. 8). Additionally, the average income of non-Western migrants was 34% lower than that of the non-migrant population in Amsterdam in 2013 (OECD, 2018), likely restricting their options for housing.

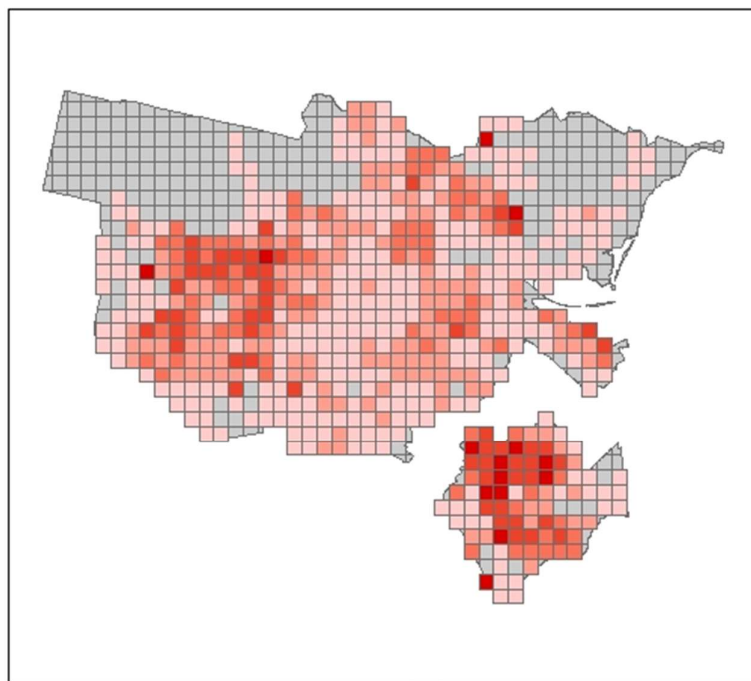
The combination of these socioeconomic factors, along with the homophilious preferences and family reunification schemes discussed in Chapter 2, are likely to influence the location of social frontiers in Amsterdam and importantly make it a fascinating city to consider. Where the non-Western clusters are visible on the map, then it is likely that social frontiers are also present, separating the diverse areas. The aggregate measures in this chapter allow and encourage research which includes different regions or countries. If, as we suspect, the socioeconomic factors in the Netherlands such as social housing, gentrification, and high levels of welfare

⁶ Social distance is related to the "sum total of psychological difference" (Schnittker, 2004, p. 219). It is a measure of the difference between individuals or groups within society.

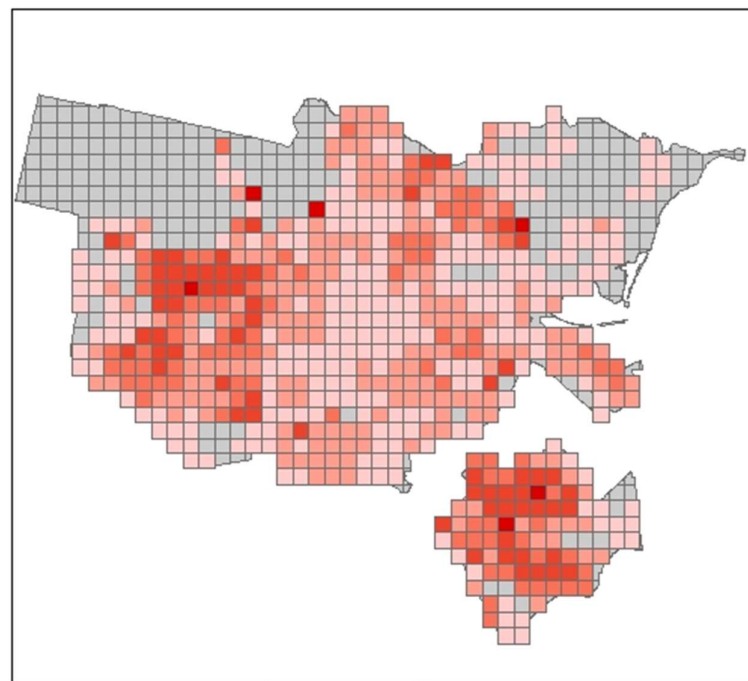
support (Musterd *et al.*, 2016) are linked to social frontiers patterns, then it is of interest to study whether these trends are replicated in other Dutch cities. Studies including multiple cities would be made easier with the creation of a less computationally challenging identification measure to speed up the process and the use of prevalence measures to compare cities and time periods.

Proportion of Non-Western individuals in the City of Amsterdam in 2010 and 2019

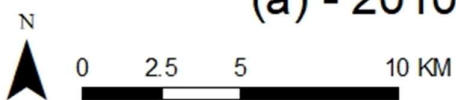
Proportion of Non-Western population



(a) - 2010



(b) - 2019



Data Source: CBS

Figure 3.1: Spatial arrangement of non-Westerners in the city of Amsterdam in 2010 and 2019.

Data and Methodology

Data

The chapter utilises registry data from Statistics Netherlands (CBS) geocoded to 500m by 500m grid cells. This study uses the larger 500m by 500m grid cells as an introduction to both social frontiers in Amsterdam and several of the identification methods. The 500m by 500m grid cells allow the creation of visualisations which can result in a meaningful discussion about the location and pattern of the social frontiers. As well as providing the first opportunity to explore the computational complexity of the methods.

Amsterdam is split into 925 grid cells with a median population of 785 and 995 and a mean of 1329 and 1457 in 2010 and 2019 respectively. In this chapter, empty grid cells or grid cells without at least 5 neighbours were removed prior to the identification methods being applied as they either cannot have a social frontier boundary or they pose issues to the identification methods. These requirements resulted in the removal of 331 and 319 cells (approximately 35% of total grid cells) in 2010 and 2019 respectively. The social frontiers in this chapter are identified using the proportion of Western and non-Western individuals in the adjacent grid cells.⁷

Social frontier identification methods

This chapter has three main aims to advance the study of social frontiers: (i) develop alternative identification methods, (ii) include statistical descriptions of social frontier prevalence, and (iii) measure the difference between the observed and expected frontier prevalence for a given index of dissimilarity, D . The stages of this chapter are represented graphically in Figure 3.2 (which is a snippet of the thesis flowchart in Figure 1.1). For three of the methods this is the first time they have been used in social frontier identification, therefore, to form an argument about method suitability for identifying frontiers, I include a simulation study to assess the accuracy of the methods. As a robustness test, I alter aspects of the population in the simulation study, checking how this impacts upon the methods' success in identifying frontiers. It is the simulation study alongside the speed of the process that means I can create a shortlist of successful identification methods. Faster methods could make it easier to extend studies to include multiple or larger regions and more time periods, however, it is crucial that speed does not come at the cost of accuracy. The Lee and Mitchell (2013) method is established and deliberately cautious in the frontier identification process, which these potential methods have not been adjusted to do. These developments encourage broader discussions about the dynamics and evolution of social frontiers, the role of contextual factors, and ultimately the development of a taxonomy of social frontiers.

⁷ The Western and non-Western definitions can be found in Chapter 2 (Segregation and the Netherlands).

This chapter builds upon the Lee and Mitchell (2013) method by evaluating three alternative identification methods: an absolute contiguous difference method and two non-spatial clustering methods: k-means and PAM clustering. These will have the ability to identify open social frontiers whilst not smoothing any discontinuities, as with the Lee and Mitchell (2013) method. However, they are less computationally complex and thus potentially quicker than the Lee and Mitchell (2013) method. Below, I describe the methodology of the three proposed methods. The methodology of the original method developed by Lee and Mitchell (2013) and adapted in Dean *et al.* (2019) is described in detail in Chapter 1.

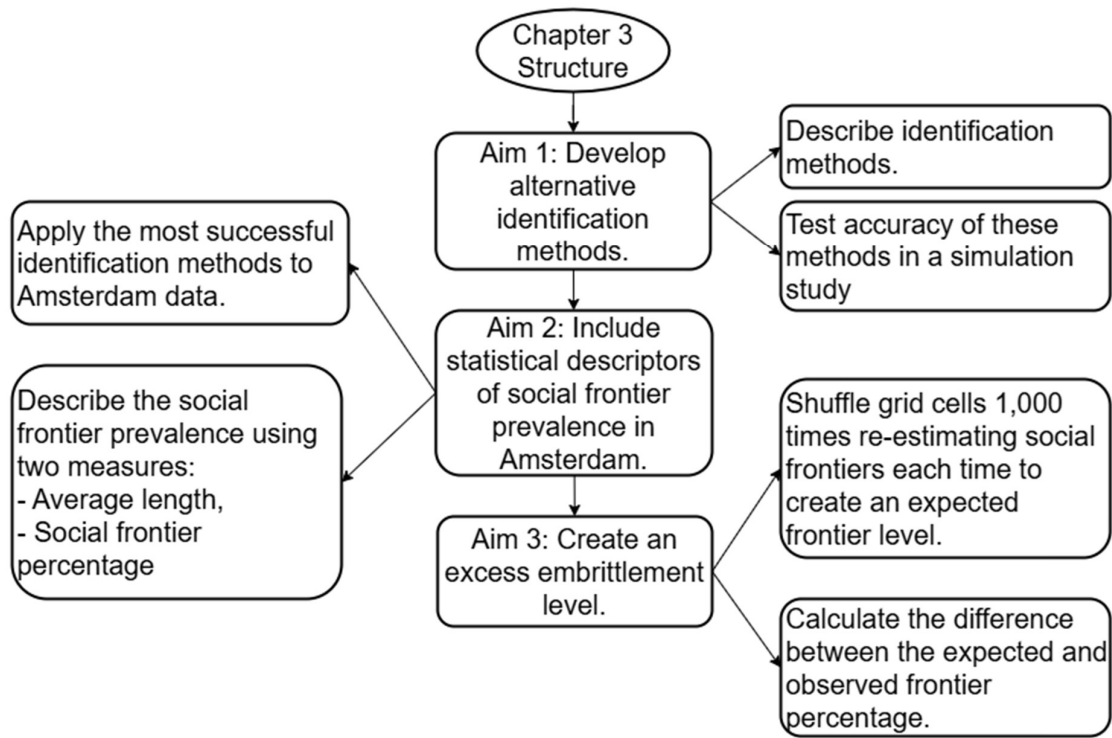


Figure 3.2: Flowchart representing Chapter 3's structure - a section of Figure 1.1 (the thesis flowchart).

Absolute contiguous difference

A simple and computationally convenient method for identifying social frontiers is to define a boundary as a social frontier when:

$$|Prop_j - Prop_k| > \theta \quad (3)$$

Where $Prop_j$ is the proportion of non-Western individuals in grid cell j , θ is a set threshold, and j and k are adjacent grid cells. Kramer (2017) takes a similar approach to map the “rate of change in local racial demographics” (p. 82) across boundaries in Philadelphia, determining a high slope to exist when there was

approximately a 20 percent difference in the non-White percentage on either side of the barrier (p. 88). Although simple and quick to implement, the main shortcoming of this method is that it does not quantify the uncertainty associated with frontier estimation. This is problematic, particularly when the number of residents within an area is small, as small differences in the number of the individuals in the group of interest can result in a large difference appearing in the proportion. This can result in the method indicating a social frontier on the basis of a small number of observations. A second difficulty with this method arises in the selection of a threshold which determines when a boundary is steep enough to be defined as a social frontier. In this chapter, the threshold was chosen using a simulation study. The 75th percentile was identified as producing the most accurate estimation of the frontier on average, correctly classifying 92.3% and 92.9% of boundaries in 2010 and 2019 respectively, when compared with the 'true frontiers'.⁸

K-mean and k-medoid clustering

This chapter tests how successfully the two non-spatial clustering algorithms: k-means and a k-medoids method (Partitioning around Medoids (PAM)) identify social frontiers. Clustering groups objects based on a property or properties, aiming to find 'similar' objects. In this chapter, the objects (grid cells) are clustered into two groups based on the proportion of non-Western residents. One group contains the cells with the highest proportions of non-Westerners and the other includes the cells with the lowest. Once the clusters have been created, any pair of contiguous grid squares that are in different groups are defined as having a social frontier between them. If they are in the same group, then they are determined to have a smooth transition between them and thus, no social frontier. In contrast to the absolute difference method, clustering methods are unsupervised: the researcher does not select a threshold and patterns can be identified within the data without intervention, simplifying the model and reducing the researcher's influence. There is still an element of threshold choice with the choice of centroid/medoid, however, the researcher does not decide it.

Both clustering methods are required to satisfy two principles for: $C_1, C_2, \dots, C_k$ where C_k denotes the set of observations in each cluster:

Principle 1: $C_1 \cup C_2 \cup \dots \cup C_k = \{1, \dots, n\}$. It is a requirement for every datapoint to belong to at least one of the k clusters.

Principle 2: $C_k \cap C_{k'} = \emptyset$ (the empty set). The clusters do not overlap – no observation can belong to more than one cluster. (James *et al.*, 2023, p. 515)

K-means

K-means clustering works by generating k (in this case 2) random centroids and then assigning each datapoint to its closest centroid by calculating the Euclidean

⁸ A discussion surrounding the choice of thresholds is included in Appendix B.

distance. The algorithm then computes the mean non-Western proportion of each group. This replaces the centroid and the cells are reassigned to their closest centroid, with the process repeated until the clusters do not change between iterations. The aim of the method is to minimise within-cluster variance which measures the degree to which the observations within the clusters vary from each other, in this case the distance is measured using the Euclidean distance. This evaluates the sum of the squares distance between the datapoints and the centroid of the clusters. The aim of the study is to find the clusters which minimise the within cluster sum of squares:

$$\min_{C_1, C_2, \dots, C_k} \sum_{i=1}^k \sum_{x \in C_i} \|x - \mu_i\|^2$$

where $C_1, C_2, \dots, C_k$ are the k clusters, $x \in C_i$ is the observations in the cluster i , μ_i is the centroid (the mean) of the cluster, and $\|x - \mu_i\|^2$ the squared Euclidean distance (Schubert and Rousseeuw, 2019, p. 172).

K-medoids (PAM clustering)

In addition to the k-means clustering I have also tested the success of PAM clustering. PAM clustering applies the same logic as k-means but selects two random data points for the initial medoids and assigns each datapoint to the closest medoid. These medoids are then swapped, testing pairs of observations to attempt to improve the quality of the clustering. This is continued until the average dissimilarity, across all the observations, is minimised. The aim is to find the medoid of the cluster C such that:

$$\text{medoid}(C) = \arg \min_{x_i \in C} \sum_{x_j \in C} d(x_i, x_j)$$

(Schubert and Rousseeuw, 2019, p. 172).

Measuring similarity based on average dissimilarity rather than sum of squares makes the clustering less vulnerable to outliers, additionally, the method, unlike k-means clustering, is unimpacted by the initial choice of medoid (Kaufman and Rousseeuw, 1990, p. 117). Using an actual data point rather than the mean reduces the likelihood of being influenced by any variation in the data (Madhulatha, 2012, p. 721). However, it is more computationally complex than k-means, meaning the algorithm takes longer to operate (Xu and Tian, 2015, p. 184).

Limitations

There are several limitations of the clustering methods and statistical trade-offs that could come as the price for the decreased processing time and simpler methodology that the clustering methods offer. Although the most commonly used clustering method (Xu and Tian, 2015), k-means is susceptible to outliers and noise due to the use of the mean, especially when there is large variation in the data. Extreme values are likely to heavily skew the mean and thus the allocation of observations to the

clusters. This can pose issues with both grid cell sizes, for the 500m by 500m grid cells the range between the grid cell size is larger (Table 1.1). However, for the 100m by 100m grid cells the use of the proportion means that as the denominator (grid cell population) will be small it will be impacted by small changes in the non-Western population, potentially making relatively minor differences seem more extreme. For this study the clusters created for the city of Amsterdam were relatively similar sized (there are 402 and 204 observations in each k-means cluster in 2019), suggesting this is less likely to be an issue here.

Despite the PAM clustering method being significantly less susceptible to outliers compared to the k-means clustering due to the use of an actual data point, there is a potential issue with both clustering methods in that it is obligatory that every observation is grouped within a cluster (James *et al.*, 2023, p. 529). This means that at least one frontier will always be identified even if the spatial arrangement of non-Westerners across the city is completely smooth.

For the k-means method the use of the mean as the centroid could result in clusters becoming distorted, resulting in the creation of very small clusters (van der Laan, Pollard and Bryan, 2003, p. 576). This weakens the quality of the clustering, meaning we cannot be as confident in the frontiers identified. Despite this not being an issue as much for PAM clustering as the medoid will be unaffected by outliers, the choice of the medoids/centroids could lead to some inconsistencies in the frontier identification. For example, if there are two adjacent grid cells which are 20% and 50% non-Westerners then these two clusters would likely be grouped into different clusters, assuming that the medoids were around 15% and 55%, meaning a social frontier would be identified between them. However, two adjacent cells with proportions of 50% and 100% would not have a frontier between them, according to this classification, despite having a large difference in the non-Western proportion, there is likely inconsistency in the frontiers identified. The experience of those who live in a grid cell where there are no non-Westerners bordering a grid cell where 20% of the residents are non-Westerners likely differs of those who live in a grid cell with 50% and 70% - the increase for the former may feel more dramatic and significant despite being equal. The inconsistency will likely always pose an issue with the empirical identification of social frontiers. Additionally, as both clustering methods will always cluster the observations into two groups

Furthermore, the k-means method stops when it identifies the 'local optimum' and thus the results depend largely on the initial choice of centroid (James *et al.*, 2023, p. 517). To address this, when I applied k-means in this study I included '*nstarts* = 5' in the code which runs the method using 5 different random starting positions and only returns the most successful clustering.

Simulation study⁹

The main objective of this study is to introduce alternative frontier identification methods which are less computationally complex than the existing Lee and Mitchell (2013) method. This study evaluates the performance of the four social frontier identification methods through a simulation study. Using both the 2010 and 2019 Amsterdam data I simulated 100 datasets using a Binomial distribution which models success in a number of trials (n) – in this case the grid cell population - where the probability of success is constant for each trial (p), in this case the potential outcomes are ‘social frontier’ or ‘non-social frontier’. For our study these datasets are simulated under nine different scenarios which are created based on a different combination of the number of trials and success probabilities.

The value of n for each simulated data simulation (number of trials) uses either:

- The true population n – this is the actual population of each of the grid cells used to model the distribution.
- A large population $2n$ – this is twice the actual grid cell population.
- A small population $\frac{n}{2}$ – this is half the actual grid cell population.

The probability of success for each trial – described as the true proportion surface - is based on either:

- A large difference in the proportion of non-Westerners between the bordering grid cells (0.2 and 0.8).
- A medium difference in the proportion of non-Westerners between the bordering grid cells (0.3 and 0.7).
- A small difference in the proportion of non-Westerners between the bordering grid cells (0.4 and 0.6).

The true proportion surface is a flat proportion surface created using the ‘true boundaries’ visible in Figure 3.3. These boundaries are identified by clustering the Amsterdam grid cells into either a high or low cluster based on the proportion of non-Westerners in the cells using a k-means algorithm. Multiple true proportion surfaces are created in the simulation study based on the level of difference between the cells. An example of how they are created is described below.

Example: Say we are simulating a surface with a small difference between frontiers. The cell A which has been classified into the low cluster (based on the proportion of NW) is allocated the value 0.4 and the cell B which has been classified into the high cluster is allocated the value 0.6. The binomial distribution then generates random deviations, more closely mimicking the real-world data.

⁹ The results for 2010 and 2019 were very similar, thus I only include the 2019 findings in the main chapter with the 2010 results included in Appendix D.

The four identification methods were then applied to each of the 100 simulated datasets for the nine different scenarios. For each of the methods and nine scenarios I computed the True Positive Rate (TPR) - the percentage of the identified social frontiers which matched the 'true frontiers' - and the True Negative Rate (TNR) - the percentage of non-social frontiers which matched the 'true non-frontiers'. The closer the score to one, the better the model performed. By including multiple scenarios, I am assessing whether the methods were consistent in their accuracy across various grid cell populations and differences in proportions. My hypothesis is that the models will perform better when the population and proportion differences are large as this is when the frontier is more pronounced, making it easier to identify the social frontiers due to higher edge intensity and steeper transitions.

True Boundaries in the City of Amsterdam in 2019

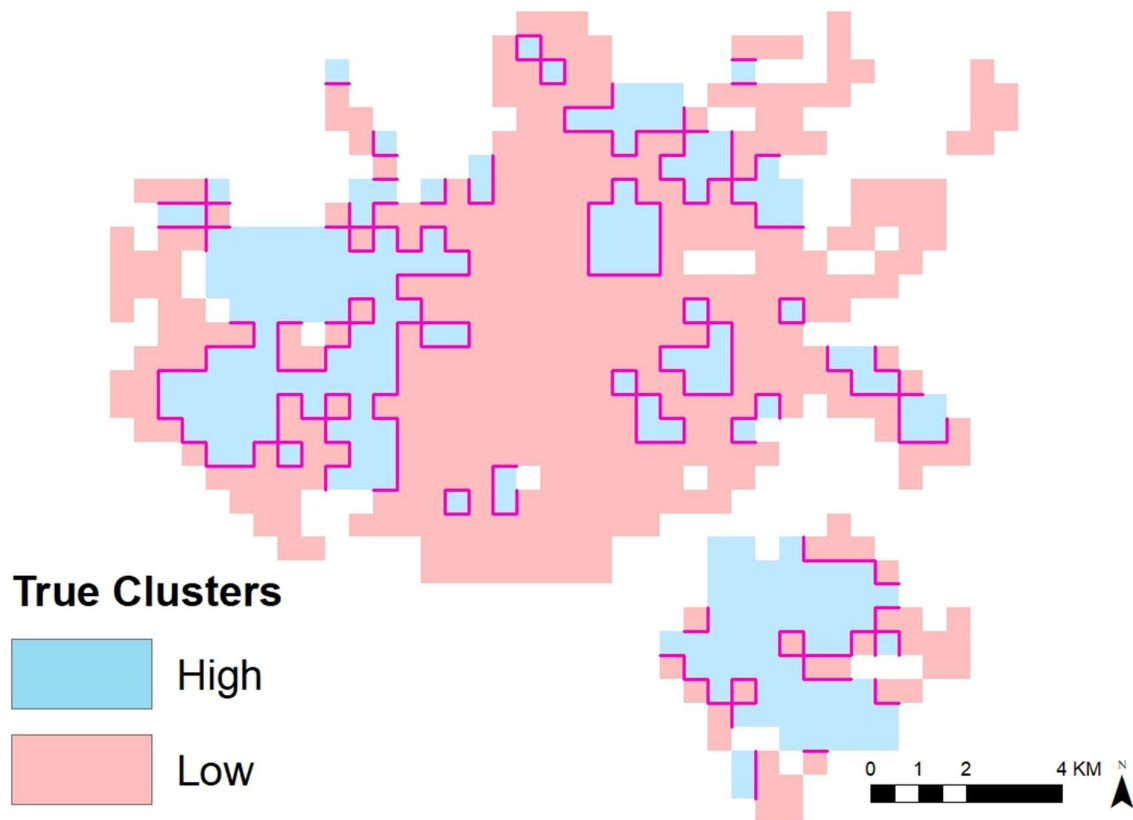


Figure 3.3: True social frontiers in 2019.

In this chapter I only apply the simulation study on the 500m by 500m grid cells which are used throughout the chapter. However, it would be useful for future studies to consider how the frontier identification results vary over the different spatial scales – including the inclusion of grid cells larger than 500m by 500m. This would allow an

exploration into how the different methods perform and how the level of frontiers varies depending on the spatial scale. This could be done using a scale sensitivity analysis (through a simulation study) using a similar method to the simulation study but rather than creating the datasets using the different scenarios based on the combination of frontier strength and population, the same dataset would be used throughout but measured at an increasing spatial scale. By comparing the frontiers identified it will test whether measuring at a larger spatial scale reduces the number of frontiers identified and explore how consistent the frontier locations are. Very few studies have explored/included different spatial scales within social frontier research, however, Sleutjes, de Valk and Ooijevaar (2018) did explore segregation in the Netherlands across different spatial scales finding a large difference in the results. I would theorise, based on existing literature, that frontiers would be more prevalent when measured at the smaller spatial scale as larger areas can appear more homogeneous, with the larger scale blending transitions, resulting in the level of segregation appearing higher when measuring at a smaller scale (Wong, 2004).

Frontier prevalence

This chapter's second aim was to include statistical descriptors of the level of social frontiers in an area to analyse their prevalence in the city of Amsterdam in 2010 and 2019. In previous papers, frontiers have only been visualised, with no statistical description included (Dean *et al.*, 2019; Křížková *et al.*, 2021). These visualisations provide an insight into the location of social frontiers in a city and lead to a discussion about why social frontiers are prolific in some areas but not others, as well as whether the social frontiers are consistent in their location over time. However, they do not make it easy for comparisons to be made between cities or for conclusive statements to be made about changes over time. It is more reliable and easier to compare statistical descriptors rather than subjectively compare maps.

In this chapter the prevalence of frontiers in Amsterdam is measured using two different methods: (1) the percentage of the total boundaries in the city which are social frontiers, and (2) the average (median) length of adjacent social frontiers. The first method provides a basic description of the presence of social frontiers in an area by dividing the number of social frontiers by the total number of boundaries. This measure enables researchers to evaluate the consistency of the frontier percentage over time and compare different areas. Comparing cities using the raw number of social frontiers would be inconclusive as cities vary in size, and even when considering the evolution of a city there often exists variation in the number of boundaries over time.

However, measuring only the percentage of boundaries does not provide a complete picture on how substantial the frontiers are. Therefore, I also measure the average length. Short frontiers are likely to have a reduced territorial effect and have a lesser impact on the community compared to frontiers of a significant length (Iyer and Pryce, 2024), as they occur sporadically and are less likely to form barriers. Longer,

more substantial frontiers have a greater potential to evoke defensive behaviour, with residents potentially feeling more protective over their communities (Iyer and Pryce, 2024, p. 794). If frontiers are visible in only one year when identified over an extended continuous period of time, this could indicate that they are less consequential for individuals than long-standing established frontiers as they are not entrenched within communities. However, equally Legewie and Schaeffer's (2016) contested boundary theory could mean that even if only present for a year, these frontiers forming could result in higher levels of conflict and dispute over unestablished territory. The average length measure was calculated by grouping the frontiers which 'touch' and then computing the average length of the continuous frontiers. Short social frontiers (e.g. those of a single cell length 500m) have a large impact on the average score. As the significance and reliability of these identified single frontiers is unknown I also include a median length measure where the single frontiers have been excluded from the calculation. In their study evaluating the link between housing mobility and social frontiers, Olnier *et al.* (2024) only included 'stable frontiers' in their analysis: frontiers which were consistently present and above the threshold during a five-year period. This ensured that the frontiers identified were not temporary or a result of a random spatial discontinuity.

Excess embrittlement

The excess embrittlement level describes the difference between the observed percentage of frontiers and the expected level (Manley *et al.*, 2023), with the expected level of frontiers calculated by randomisation. The grid cells are randomly redistributed across the city and social frontiers are then re-estimated, with this process repeated 1,000 times. The expected frontier prevalence is calculated by creating a median score for the shuffling results in both 2010 and 2019 in Amsterdam. Due to the aspatial nature of the index of dissimilarity, the index score will remain consistent across the different spatial arrangements of the neighbourhood which are created through the shuffling. The excess embrittlement level compares the level of social frontiers which have been identified for the actual arrangement of an area, with the average level of frontiers identified for exactly the same city but with grid cells randomised (Manley *et al.*, 2023). The excess embrittlement level leads to a discussion about whether there are additional processes at play or certain arrangements of cities which impact upon the level of divides in an area (Manley *et al.*, 2023).

The method is somewhat flawed in that certain neighbourhood arrangements will create almost an unrepresentative image of the actual level of spatial segregation in the region. For example, a checkerboard pattern would result in a region having a large proportion of social frontiers, however, realistically this is likely to be one of the most cohesive neighbourhoods with no large clusters. Whereas a straight divide between the north and south would result in fewer frontiers, only those lying between the north and south, but in reality the communities are far more divided and less likely to interact. This highlights the need to consider the broader spatial

arrangement beyond the local region or city in order to understand the patterns identified and any trends beyond the city. This is an issue with the excess embrittlement measure; however, the method could be used as an entry and starter to the social frontier prevalence discussion. If a city's expected level of frontiers differs greatly from the observed prevalence then future research could consider what aspects of the city or external factors have resulted in this.

Simulation Study Results: Testing Alternative Social Frontier Identification Methods

I use two measures to describe the accuracy of the identification methods: the true negative rate (TNR) and the true positive rate (TPR). I performed the simulation study using both the 2010 and 2019 data, as the results were very similar I have only included those from 2019 in the main text with the results reported in Figures 3.4 and 3.5 and Tables 3.2 and C.1. The 2010 results can be found in Appendix D.

Figures 3.4 and 3.5 show the TNR and TPR scores for all the four identification methods: Lee and Mitchell (2013), k-means clustering, PAM clustering, and absolute contiguous difference, and Table 3.2 shows the median TNR and TPR scores across all combinations. The accuracy is estimated by comparing the identified frontiers for each simulated dataset to the 'true frontiers' in Figure 3.3. Within Figures 3.4 and 3.5 there are nine jitter plots with each one representing a combination of the different levels of grid cell population and difference. The first column in Figures 3.4 and 3.5 is the scores for the smallest population (half the actual grid cell population); the middle column is the actual population; and the third column is the scores of the largest population (twice the actual grid cell population). The rows are based on the frontier strength: the top row is the scores for the frontiers with the smallest difference (0.6 and 0.4); the middle row is the medium difference (0.7 and 0.3); and the bottom row is the largest difference (0.8 and 0.2). The aim of the simulation study is to identify a less computationally challenging potential alternative to the existing Lee and Mitchell (2013) method, whilst maintaining a high degree of accuracy when compared to the 'true frontiers'.

Table 3.2: Median accuracy scores from the simulation study.

	Lee and Mitchell (2013)	PAM	k-means	Absolute Contiguous Difference
True Negative Rate	99.74%	96.26%	96.07%	97.17%
True Positive Rate	83.91%	95.79%	95.79%	90.61%

True Negative Rate (TNR)

The TNR describes the method's ability to correctly identify the non-social frontiers: the boundaries where there is a smooth transition between the neighbourhoods. The score is calculated by identifying the boundaries where both the methods and the true proportion surface identify a non-social frontier.

The Lee and Mitchell (2013) method had the highest TNR on average with 99.7% of non-social frontiers correctly identified when averaging across the different combinations. The worst performing model for identifying non-social frontiers on average was the k-means clustering, with a 96.1% average accuracy score; the PAM clustering returned a very similar score of 96.3%. Whilst these TNR scores were the lowest, there is still a high level of accuracy achieved, indicating that all of the methods were consistently successful in identifying non-social frontiers as they all had an accuracy score of over 96% on average.

The average scores do not show the variation within the method's scores across the different population and strength of the frontiers in terms of their ability to correctly identify non-social frontiers, and so I have also included a jitter plot in Figure 3.4. All the methods performed best when the frontier strength and population were large as the bigger differences made it easier for the methods to differentiate between the social frontiers and non-social frontiers. It is important to note that all the methods performed well (the y-axis scale starts at 0.8) however, there is some variation within the scores. The Lee and Mitchell (2013) method had almost no variation (visible in Figure 3.4) and was incredibly consistent, always scoring above 99%. In contrast, the clustering and absolute difference methods showed variation in their TNR, especially for the simulated datasets where the population and frontier strength were small. The TNR for the absolute difference method dipped to around 87% when the population and frontier was small. The difference in the TNR scores for the different population size and frontier strength was around 10 percentage points for PAM and k-means clustering (maximum 99% for large frontier and population and minimum 89% for small frontier and population). This variation makes the clustering and absolute difference method less reliable than the Lee and Mitchell (2013) method, as they are more likely to identify false positives and thus overestimate the level of frontiers in an area. Overall, I identify the Lee and Mitchell (2013) method as the most successful at identifying non-social frontiers with large levels of consistency across the different combinations. This is likely due to it being more conservative with a deliberately stricter requirement for social frontiers meaning that the boundaries it identifies are stronger and more certain.

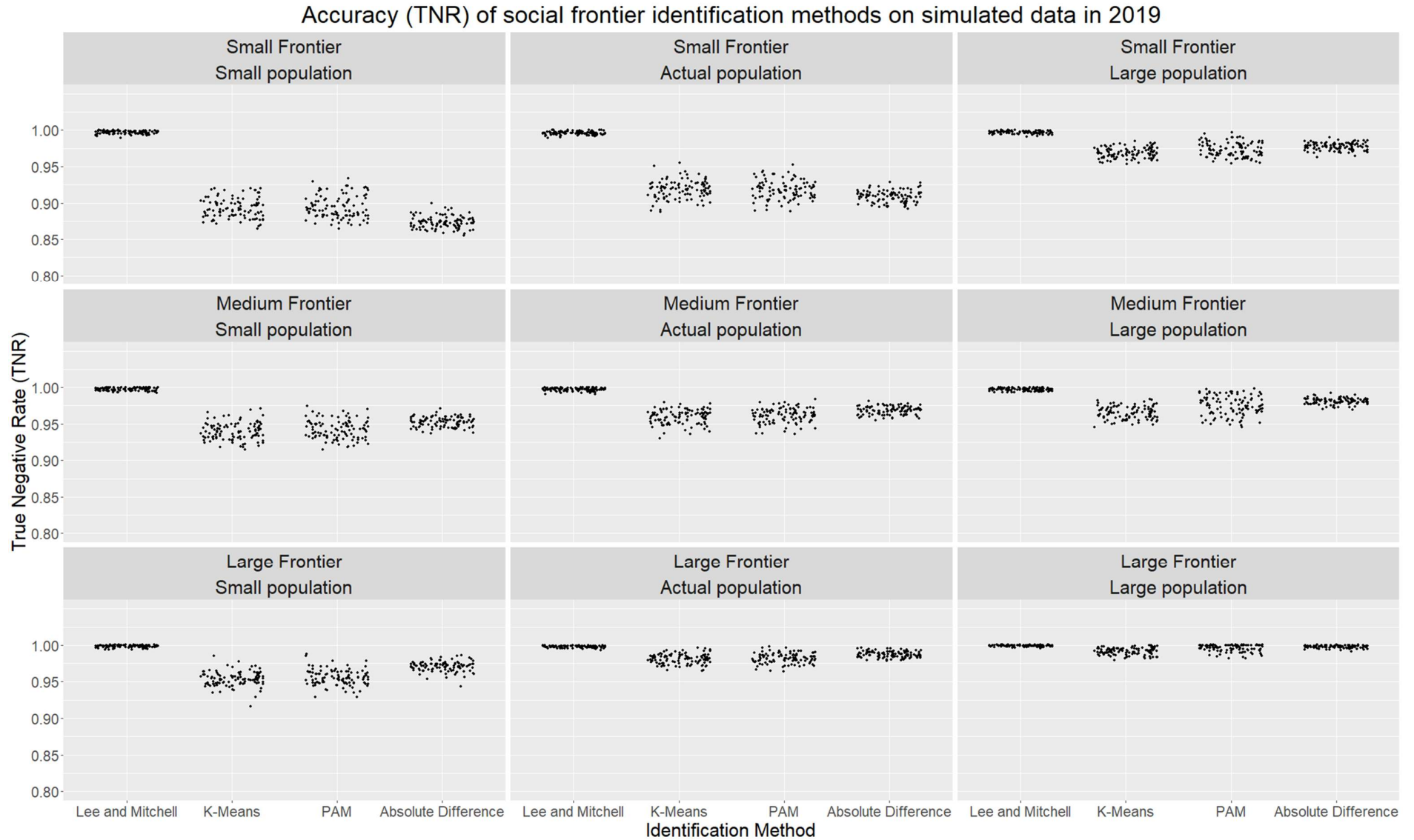


Figure 3.4: TNR scores jitter plot for 2019.

True Positive Rate (TPR)

The TPR describes the method's ability to correctly identify the social frontiers, where there is a sharp change in the proportion of non-Westerners in the adjacent communities. These were identified by comparing the 'true frontiers' to the frontiers identified by the different methods.

The TPR scores have a larger range and are less consistent than the TNR scores as visible in Figure 3.5. In terms of the results for each of the methods, these are the reverse of the TNR scores. The Lee and Mitchell (2013) method returned the lowest score on average - 84%, a significantly smaller percentage than the clustering methods which averaged approximately 95% as shown in Table 3.2. Figure 3.5 shows that the method struggled particularly when the frontier strength and population was small, with a large drop in the scores from approximately 95% for the large frontier and population to 70% of 'true' social frontiers correctly identified when applied to the smallest population and strength using the Lee and Mitchell (2013) method. The decrease in accuracy when frontier strength and population decreased was replicated across all four methods as with the TNR.

In terms of the variation in the scores visible in Figure 3.5, the absolute difference method has the largest difference with approximately a 34 percentage point difference between the TPR scores of the largest and smallest population and grid cell difference combination. The score for the absolute difference method drops to less than 60% of frontiers correctly identified which indicates that a large proportion of the social frontiers would be missed if using the absolute difference method. The clustering methods also showed a range in their scores, although the median TPR score was high at 95.8%. This variation is especially visible within the smaller frontier and population where there is a difference of over 5 percentage points between the maximum and minimum TPR score for the small frontier and population. The clustering methods consistently produce a higher level of accuracy across the different scenarios compared to the Lee and Mitchell (2013) and absolute difference methods. However, there is more crossover between the Lee and Mitchell (2013) and absolute difference methods visible in Figure 3.5, with the Lee and Mitchell (2013) method more consistent and less impacted by the reduction in the frontier strength and population. Despite the Lee and Mitchell (2013) method having a lower TNR on average, it showed relatively little variation in the scores when the population was smaller and had lower levels of disparity in terms of the extremes in the accuracy scores compared to the absolute difference method (Figure 3.5). Overall, the clustering methods are more successful at correctly identifying social frontiers according to this simulation study, with a high level of accuracy identified.

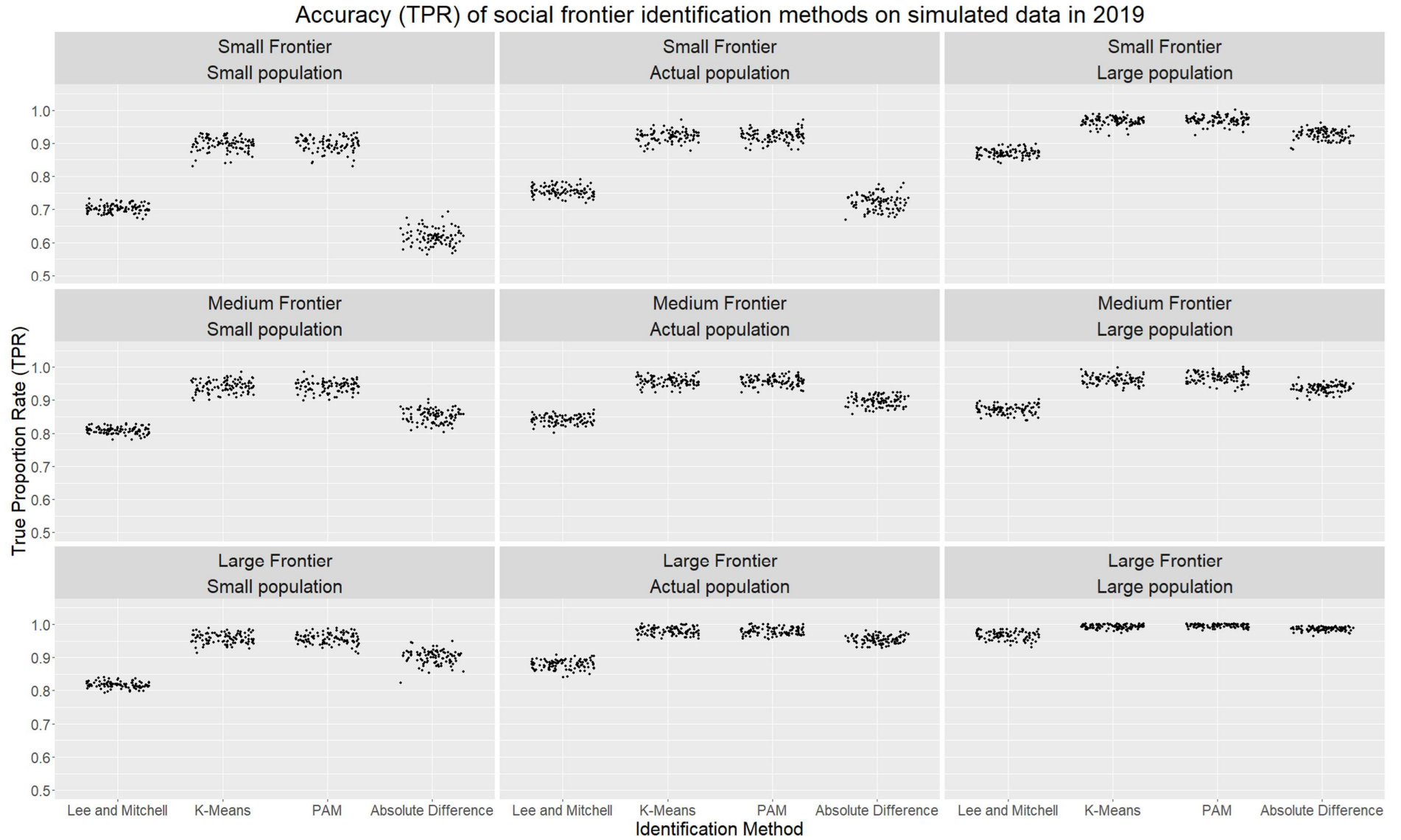


Figure 3.5: TPR scores jitter plot for 2019.

Computational complexity

A crucial reason for introducing new frontier identification methods is the time-consuming nature of the Lee and Mitchell (2013) method. A reduction in the processing time could allow studies to be extended to identify frontiers using large datasets. This could result in studies being conducted that identify frontiers in multiple regions, time periods, or using a finer spatial scale. Table 3.3 shows how the runtime of the methods increases as a function of N , starting at $N = 10$ grid cells and going up to $N = 10,000$. Table 3.3 also includes the runtime for $n = 100,000$ and $n = 1,000,000$ which are estimates based on the calculated processing times. The Lee and Mitchell (2013) method consistently had the longest execution time, taking more than 70 times as long as the PAM clustering when applied to 10,000 grid cells.

When the datasets were small the difference in processing time was almost negligible, not worth any compromises in accuracy. However, for a dataset of 10,000 grid cells the difference in the processing time was 1,485 seconds (equivalent to almost 25 minutes). This difference was also reliant on the method running without encountering an issue – the Lee and Mitchell (2013) method struggles when grid cells have none of the group of interest which can add to the complexity of the process. This difference in processing time could be problematic when applying the method multiple times whether that be to study the evolution of the frontiers over an extended period of time or when randomising the grid cells to create the expected level of frontiers. In Chapter 6 I identify frontiers nationwide applying the method to over 96,000 grid cells, if I had applied the Lee and Mitchell (2013) method then based on the processing time in Table 3.3, this would be expected to take around 88 hours longer than the PAM clustering method assuming that the computer memory and processing power remain consistent. Due to the highly sensitive nature of the data, all the calculations and processing had to be performed in the remote access which limited the computer power potentially slowing the process down further. Particularly in the case of larger studies, the method choice is largely consequential for the feasibility of a study. For smaller areas the Lee and Mitchell (2013) method is not considerably more time consuming than the alternative methods and thus remains a strong identification choice due to its high degree of accuracy and it being well-established in the literature.

Table 3.3: Processing time of the four identification methods when applied to Amsterdam 2010 and 2019 data.

Method	Time Taken (in seconds) for increasing dataset sizes (minute/hour conversions in brackets)					
	N = 10 grid cells ¹⁰	N ² (100 grid cells)	N ³ (1,000 grid cells)	N ⁴ (10,000 grid cells)	N ⁵ (100,000 grid cells) ¹¹	N ⁶ (1,000,000 grid cells) ¹⁰
Lee and Mitchell (2013)	8.12	36.30	124.6 (2.1 minutes)	1,507 (25.1 minutes)	15,035 (4.2 hours)	150,336 (41.7 hours)
k-means	0.049	0.052	0.24	14.09	144 (2.4 minutes)	1,440 (24 minutes)
PAM	0.050	0.049	0.30	21.54	220 (3.6 minutes)	2,206 (37 minutes)
Absolute Contiguous Difference	0.048	0.049	0.26	11.47	117 (1.95 minutes)	1,170 (19.5 minutes)

Simulation results conclusion

Based on the variation in the accuracy scores visible in Table 3.2 and Figures 3.4 and 3.5 along with the processing time discussed in Table 3.3, the PAM clustering method appears the best alternative identification method if it is unfeasible to apply the Lee and Mitchell (2013) method. The method is consistent in its accuracy across the different datasets simulated under the 9 different scenarios, with only slight variation in the TPR scores when the population shrinks. There is very little difference in the results of the two clustering methods, however, the PAM clustering method marginally outperforms k-means with a higher average TNR thus meaning it is less likely to falsely identify boundaries as frontiers (false positive).¹² The TNR is lower than that of the Lee and Mitchell (2013) method which is a slight concern as it suggests that the clustering may misdiagnose non-social frontiers as social frontiers, this would inflate the social frontier prevalence potentially resulting in findings appearing more extreme than they are in reality. The method also has a quicker processing time than the Lee and Mitchell (2013) method, which is especially apparent as the study increases in size – the Lee and Mitchell (2013) method was estimated to take 104 times as long as the PAM clustering method for 100,000 grid cells. Furthermore, the clustering methods remove the researcher bias due to the lack of threshold inclusion. Although the PAM clustering processing time is slightly longer than the k-means clustering, the potential impact of extreme values on k-means can make it unstable and less reliable. The number of grid cells included in

¹⁰ The population of the 10 grid cells was 252, 4,382 for 100 grid cells, 63,969 for 1,000 grid cells and 978,445 for 10,000 grid cells.

¹¹ These are predicted processing times based on the calculated processing times for N, N², N³, N⁴.

¹² This could be a result of the choice of medoid which has altered the clusters resulting in the k-means grouping being slightly less consistent and accurate.

this chapter means that the difference in the processing time is negligible and not largely consequential for the feasibility of the study. However, when studying a larger region or applying the identification process multiple times, the quicker clustering methods could increase the scope of studies.

In terms of how the PAM frontiers differ from those identified by the Lee and Mitchell (2013) method, they had an average step change of 23.85% which is around 6 percentage points smaller than that of the Lee and Mitchell (2013) method. The Lee and Mitchell (2013) method clearly has a stricter and higher requirement for a boundary to be identified as a frontier compared to PAM in Amsterdam. I would expect this to be consistent across different regions, with the TPR indicating that fewer frontiers are likely to be identified compared to PAM, highlighting the cautious nature of the Lee and Mitchell (2013) method. The higher requirement for the Lee and Mitchell (2013) method is likely a result of the two-step approach which means the frontier identification is more stringent. The Lee and Mitchell (2013) method is a more conservative identifier, with a superior TNR, suggesting it is unlikely to incorrectly identify boundaries as social frontiers. Its lower TPR potentially means that it underestimates the level of frontiers as it is more likely to identify social frontiers as non-social frontiers compared to the PAM clustering method.

The other alternative method: absolute contiguous difference, is inconsistent and unlikely to provide a rigorous identification of the frontiers, particularly when the grid cell difference and population are smaller. The large variation in the accuracy scores is likely a result of the choice of threshold; the measure would return a 100% TPR score if the threshold was lower, as visible in Tables B.1 and B.2, as it would identify nearly all the boundaries as social frontiers. However, this would also return a TNR score of 0%. The element of threshold choice in the absolute difference method means that the method relies on human intervention/choice which weakens it significantly and makes it an unreliable measure of identifying frontiers.

The Lee and Mitchell (2013) method is a cautious, rigorous test with a high requirement for social frontiers. Additionally, it is well-established within the literature, providing a comparison for my findings in different studies in terms of investigating the consistency of the impact and reliability of the identified frontiers. This method is best applied when researchers are interested in the pattern of social frontiers or their impact, but likely only when frontiers are studied on a smaller scale due to the computational complexity of the method. A crucial part of the qualitative study by Zhang *et al.* (2024) about how social frontiers are recognised in communities, is that the steeper social frontiers were those that were more likely to be recognised by residents, providing support for the use of the Lee and Mitchell (2013) method due to it having a higher requirement for a social frontier.

In studies where it is not feasible to use the Lee and Mitchell (2013) method, e.g. the dataset is incredibly large or the process is being repeated multiple times vastly

increasing the processing time, then the PAM clustering method can act as a strong alternative. The method is unlikely to miss any social frontiers, with its high TPR, meaning that individuals who live near social frontiers are unlikely to be missed. This makes the method particularly useful when creating targeted policies to reduce social frontiers, as individuals are rarely missed – it is likely better for a policy to over target than under. Additionally, the shorter execution time of the method means that it should be employed when studying a large area over an extended period of time as it significantly reduces the complexity and processing time of a study. The results of the PAM clustering should be discussed and stated cautiously as the nature of the method means it will always identify a social frontier even if the spatial arrangement of the region is almost completely smooth as the method will always create two groups. As a result, the method can be seen as manufacturing frontiers, with a lower true negative rate indicating the presence of false frontiers, meaning that results should be discussed cautiously. As a result of the simulation study, I suggest that the majority of studies should use the Lee and Mitchell (2013) method if possible as its cautious nature means that the frontiers identified are only those which are likely to have the greatest impact. However, there is a role for the PAM clustering, if looking at policies, large regions, or for the creation of the expected level if calculated over an extended time period. I highlight the benefits of PAM clustering in my Chapter 6 which would have been almost unfeasible without the introduction of the PAM clustering method.

Social Frontier Prevalence Results

I analyse the prevalence of country of birth social frontiers in Amsterdam identified using the Lee and Mitchell (2013) and PAM clustering methods. I chose to include the results of both methods in this section as it is the first application of the PAM clustering method and thus I wanted to see how the methods compare both in the location and prevalence of frontiers they identify. Following the identification, I compare the location of the frontiers and their prevalence in 2010 and 2019. I do this firstly, as with the previous social frontier literature, by comparing frontiers visualisations in Amsterdam, before discussing the statistical prevalence through the percentage, length, and single frontier measurements.

Method comparison

There are some key differences in the maps produced by the PAM clustering and Lee and Mitchell (2013) methods visible in Figure 3.6. PAM clustering frontiers tended to be longer and more substantial, which is also visible in the results in Table 3.4, with areas that are enclosed by social frontiers clearly visible in Figure 3.6 (B). The Lee and Mitchell (2013) method in contrast identified shorter, more staccato frontiers with them occurring less frequently than the PAM frontiers (Figure 3.6 (A)). These findings correspond with that of the simulation study which indicated that the Lee and Mitchell (2013) method was more conservative, whilst the PAM clustering method was more likely to overestimate. However, despite the clear differences in

the prevalence measures, areas which are dominated by Lee and Mitchell (2013) frontiers also have a high density of PAM clustering frontiers. In terms of how the locations compare, approximately 69% of the boundaries identified by Lee and Mitchell (2013) as frontiers were also identified as frontiers by the PAM clustering method. Additionally, the methods matched in their definitions for around 78% of the boundaries, suggesting that although the PAM clustering method identified more frontiers the methods were largely consistent in their frontier locations. The consistency across the boundary classifications and the pattern visible in Figure 3.6, suggests that there exist areas of the city which are particularly divided, adding a robustness to the frontier location findings.

Table 3.4: Social frontier prevalence results in Amsterdam in 2010 and 2019.

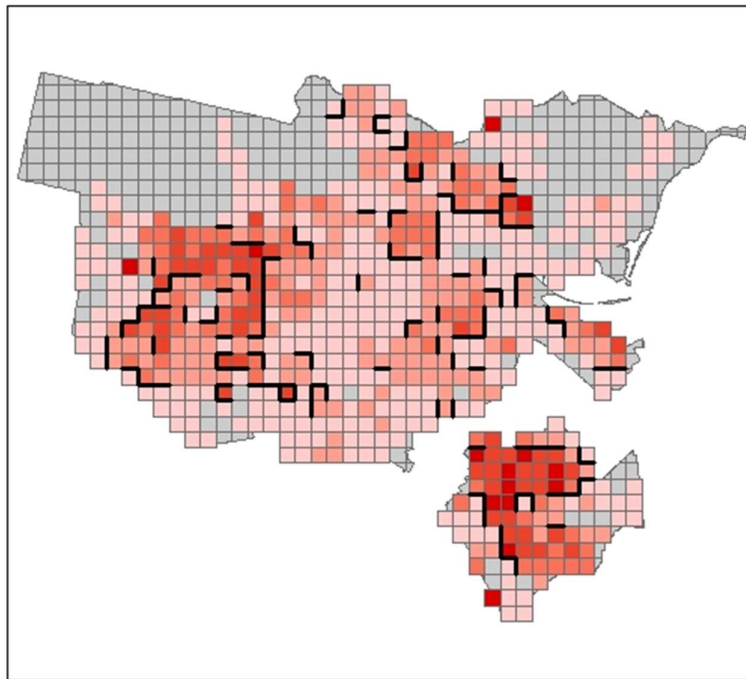
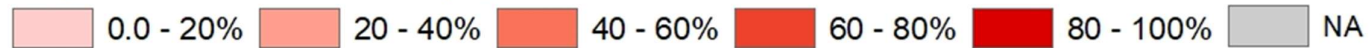
Method	Year	Percentage	Single Frontiers % (number of frontiers)	Median Length	Median Length without singles
Lee and Mitchell (2013)	2010	13.58%	1.97% (20)	1000m	1500m
	2019	13.33%	2.22% (23)	1000m	1500m
PAM Clustering	2010	23.72%	0.39% (4)	2000m	2250m
	2019	25.17%	0.48% (5)	2750m	4000m

Frontier prevalence pattern

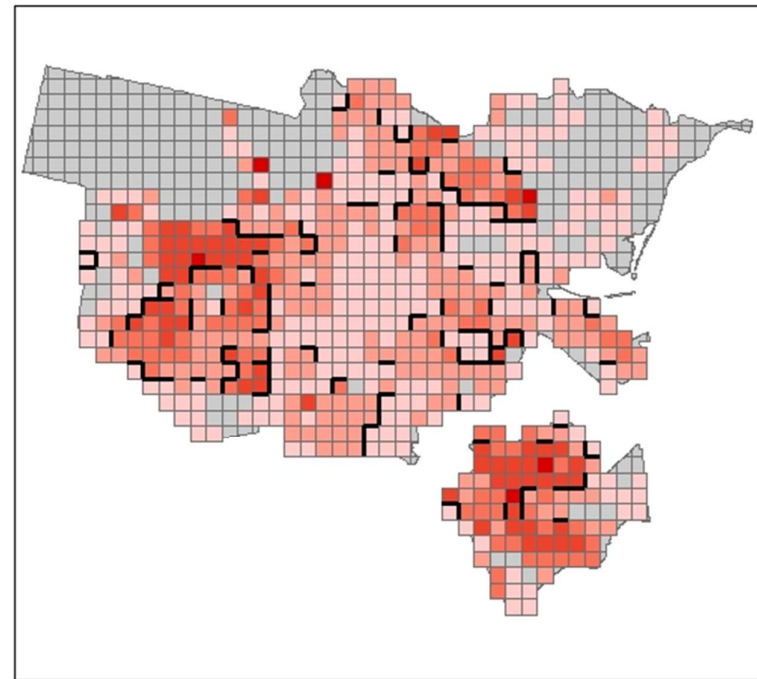
The visualisations in Figure 3.6 show the pattern of social frontiers in the city of Amsterdam which is largely consistent across the methods. In 2010 there were clear clusters of social frontiers visible in the Nieuw-West both within the borough and between the Nieuw-West and city centre. Table 3.5 shows that for the Lee and Mitchell (2013) method the Nieuw-West had the highest proportion of social frontiers in both 2010 and 2019 - nearly 20% of grid cell boundaries were social frontiers. These appear to remain stable - there were two fewer frontiers identified in 2019 compared to 2010 – in the visualisation we can see that particularly the frontiers dividing the outer and inner boroughs of the city remained consistent. The Noord, Oost, and Zuidoost boroughs also had a large frontier presence, around 15% of boundaries in 2010 were identified as social frontiers by the Lee and Mitchell (2013) method, with the social frontier prevalence growing in both the Noord and Oost. The Zuidoost was an area discussed in Chapter 2 and within this chapter's literature section about its established non-Western community particularly those with Surinamese heritage. As an exclave it is already divided from the other boroughs but there also seem to be divisions within the borough.

Social Frontiers in the City of Amsterdam in 2010 and 2019

Proportion of Non-Western population



2010

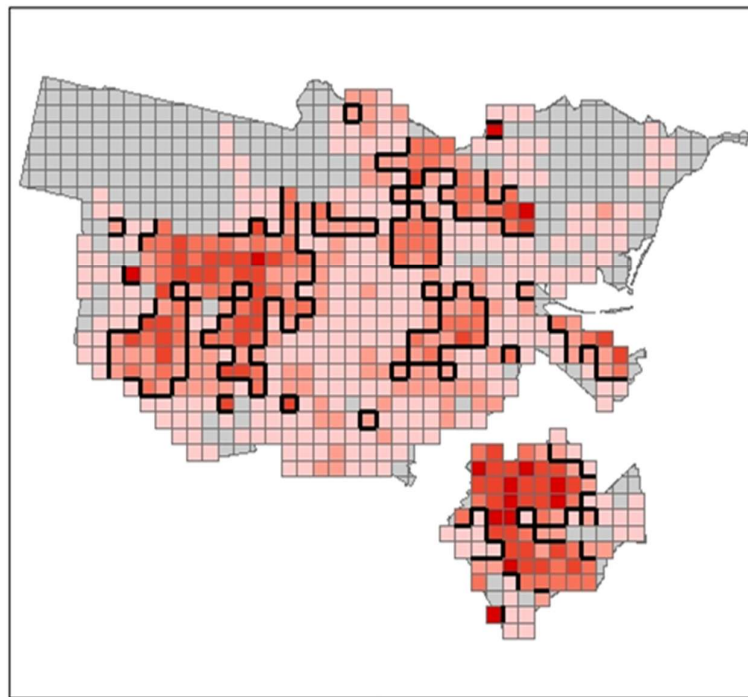
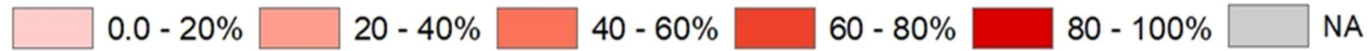


2019

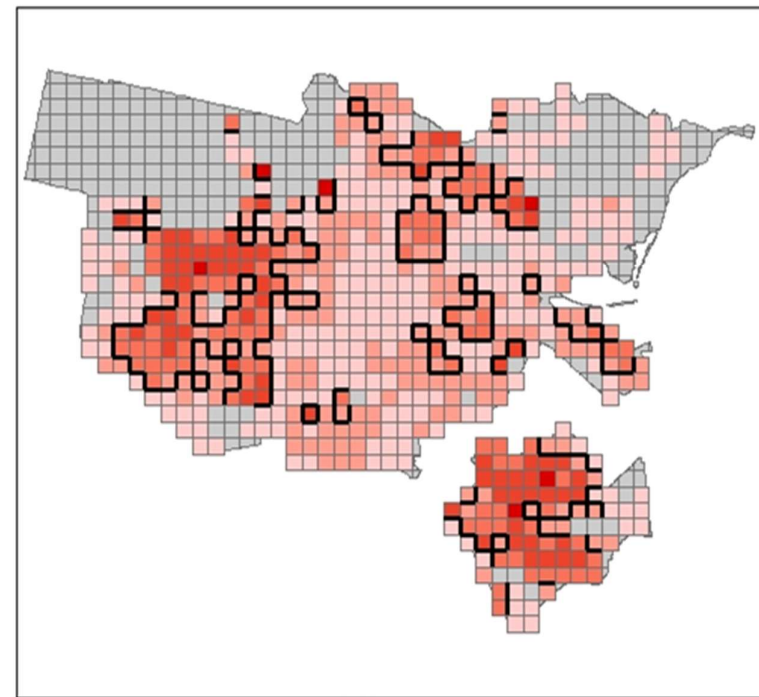
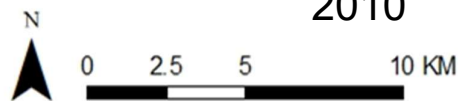
Identified using the Lee and Mitchell Method (2013).
Data Source: CBS

A. Identified using the Lee and Mitchell method.

Proportion of Non-Western population



2010



2019

Identified using the PAM Clustering.
Data Source: CBS

B. Identified using the PAM clustering method.

Figure 3.6: Social frontiers identified in the City of Amsterdam in 2010 and 2019 identified using the Lee and Mitchell (A) and PAM clustering (B) methods.

Additionally, frontiers are visible between the Noord and Oost and Centrum boroughs, particularly surrounding the Centrum borough. These frontiers are likely present due to the gentrification of the city centre resulting in lower-income non-Westerners being forced out of the centre and moving towards the suburbs. For both methods there were very few frontiers identified in the Centrum borough - especially in 2019 with only 2.35% of boundaries in the centre identified by the Lee and Mitchell (2013) method as frontiers. The centre of Amsterdam appears largely ethnically homogeneous (Figure 3.1 and Table 3.5). In actual fact there was a drop in the number of social frontiers in the Centrum borough, which could be explained by the gentrification of the city centre which has impacted upon the location of non-Westerners within the city by forcing lower-income households out of the centre (Zorlu and Mulder, 2008; Hartog and Zorlu, 2009). For both methods there is a relatively stable percentage of frontiers across the Nieuw-West, Noord, Westpoort, and Zuid boroughs in 2010 and 2019. Whilst the change in the frontier percentage in the Oost and West boroughs mostly came as a result of the number of boundaries increasing in the areas, potentially because individuals have moved into previously empty grid cells.

Table 3.5: Lee and Mitchell (2013) country of birth social frontiers per borough in the city of Amsterdam.

	Number of Boundaries		Social Frontiers		Social Frontiers %		Difference in Frontier %
	2010	2019	2010	2019	2010	2019	
Centrum	88	85	10	2	11.36%	2.35%	-9.01%
Nieuw-West	209	209	40	38	19.14%	18.18%	-0.96%
Noord	203	210	28	31	13.79%	14.76%	0.97%
Oost	139	155	21	27	15.11%	17.42%	2.31%
West	72	85	13	13	18.06%	15.29%	-2.76%
Westpoort	85	85	1	1	1.18%	1.18%	0%
Zuid	128	131	12	11	9.38%	8.40%	-0.98%
Zuidoost	138	140	21	15	15.22%	10.71%	-4.50%

The existence of frontiers which emerge or disappear between 2010 and 2019 raises questions about the lifecycle of the social frontier. Iyer and Pryce (2024) suggested that once stabilised, frontiers may attract heterophilous households which will eventually result in the decline of the frontier. Approximately 85% and 79% of boundaries were consistent in their classification across 2010 and 2019 using the Lee and Mitchell (2013) method and PAM clustering respectively. This indicates that the location of the frontiers changed very slightly between the two years, however, it

is difficult without studying the years between 2010 and 2019 to know exactly what has caused the changes in the frontier prevalence.

To build on the discussion surrounding frontier prevalence, this chapter also includes aggregate prevalence descriptors which are used infrequently in social frontier literature to describe the level of social frontiers in an area. Table 3.4 shows that the percentage of boundaries identified as social frontiers by the Lee and Mitchell (2013) method was smaller in 2019 than in 2010, albeit an incredibly minor reduction of 0.25%. This corresponds with the reducing index of dissimilarity reported in Table 3.1, suggesting a reduction in segregation in Amsterdam. Both of these segregation measures reduce despite the proportion of non-Westerners increasing in Amsterdam: 34.75% in 2010 to 35.83% in 2019 (Table 3.1). In contrast, the PAM frontiers increased in prevalence: the 2019 percentage was 1.45 percentage points higher than 2010. Again, as with the Lee and Mitchell (2013) method, the change is relatively small, albeit larger than the Lee and Mitchell (2013) reduction, however, it is of interest that this contrasts with the Lee and Mitchell (2013) results and the index of dissimilarity.

In addition to the social frontier percentage, this chapter also includes an average length measure which is the median length of the consecutive social frontiers in Amsterdam. This measure aims to evaluate how substantial the frontiers are, gauging whether they have the ability to form symbolic boundaries (Iyer and Pryce, 2024). Longer boundaries can be more difficult for individuals to cross, potentially reducing the degree of interaction between contrasting groups and access to resources or services (Wong, 1993). Corresponding with Figure 3.6, the results in Table 3.4 show that the Lee and Mitchell (2013) method identified shorter, often non-consecutive frontiers contrasting to the PAM frontiers. Table 3.4 shows that the PAM frontiers were on average twice as long as Lee and Mitchell (2013) in 2010 and even longer in 2019. In 2019 the average PAM frontier length (2750m) is equivalent to more than five continuous frontiers, suggesting that these frontiers likely have the ability to form substantial divides between communities. When combined with the frontier percentage pattern, the increase in the average frontier length identified by the PAM clustering suggests that the frontiers are increasing in prevalence. The Lee and Mitchell (2013) method produced shorter frontiers on average: 1000m in both 2010 and 2019, which is equivalent to two continuous boundaries. These findings align with the social frontier visualisation in Figure 3.6, where the Lee and Mitchell (2013) visualisation shows a large number of single frontiers especially in the Noord and Oost boroughs corresponding with Table 3.4.

Through the visualisations in Figure 3.6 and the findings in Table 3.4 we can see that the Lee and Mitchell (2013) method identified a large number of single (500m) social frontiers. Whether these single frontiers are meaningful in society is of interest: they could be a result of genuine steep changes in the proportion of non-Westerners or could be due to random noise. Table 3.4 reported that the Lee and Mitchell (2013)

method identified more single frontiers than the PAM clustering, 23 in 2019 (2.22% of the total frontiers), whereas the PAM clustering method identified only 5 (0.48%). Both the methods had a higher percentage of single length social frontiers in 2019 compared to 2010. Potentially suggesting that the frontiers are decreasing in their significance if, as I suspect, single frontiers are less consequential for the cohesion of contrasting communities.

In addition to including the single frontier prevalence measure, I also included a median length measurement without single frontiers in Table 3.4. The average length without the single frontiers was longer than the average length measure, especially for the PAM method in 2019 where the frontier length averaged 4000m, equivalent to eight continuous frontiers, nearly double that of 2010. This suggests that there is a substantial divide present in the city especially in the North and West based on Figure 3.6. Whilst the Lee and Mitchell (2013) frontiers also had a higher average but with a less substantial increase when single frontiers were removed.

We know from the simulation study that the Lee and Mitchell (2013) method is more conservative and has higher requirements for social frontiers, whereas the PAM clustering method is likely to overestimate. As this is the first inclusion of PAM clustering in a social frontier study, I included the results from both methods to provide a robustness to the study, assess how the methods compare to each other in terms of the frontiers identified, and explore whether the pattern detected in the simulation study is replicated in the prevalence results. The Lee and Mitchell (2013) and PAM clustering frontiers varied slightly in the results, thus it is likely that the 'true' level of social frontiers falls between the results of the methods, indicating that the percentage of social frontiers was very similar in 2019 to 2010, whilst their average length may have very slightly increased. This is also visible in the frontier visualisations in Figure 3.6 with many frontiers unchanged and areas remaining quite consistent. In terms of the average length, the methods found that it either stayed the same or increased slightly, indicating that frontiers were not reducing in their substantiveness. Overall, this suggests that the frontier prevalence did not change dramatically between 2010 and 2019. However, they still have a large presence in the city: over $\frac{1}{8}$ and $\frac{1}{4}$ of boundaries were identified as social frontiers in 2019 by the Lee and Mitchell (2013) and PAM methods respectively. Thus, a large proportion of Amsterdam's population likely lived near these sharp transitions, suggesting that their impact is likely to be felt within a large number of communities. Additionally, as I have only studied two non-consecutive years, we cannot accurately describe the changes in the intervening years. Studying multiple continuous years would allow a more accurate description of the fluctuations and patterns over time to be included, with differences potentially being able to be explained by policies or housing decisions implemented in an area.

Excess Embrittlement in Amsterdam

Within this study I compared the social frontier prevalence in two years, however, I also wanted to measure whether the observed level of frontiers is what we would expect for this area. One way to test this is to calculate the expected frontier prevalence had the city been arranged randomly and compare this to the observed frontier prevalence to calculate the excess embrittlement level (Manley *et al.*, 2023). The excess embrittlement level - the difference between the observed percentage of frontiers and the expected based on the index of dissimilarity - can help to discuss several aspects: firstly, whether the excess embrittlement pattern corresponds with the social frontier prevalence trend; secondly, whether there has been an increase in the excess embrittlement level of the city suggesting it has become further removed from random; and thirdly, whether the expected frontier pattern corresponds with the index of dissimilarity scores of 0.3955 and 0.3526 in 2010 and 2019 respectively.

In reality cities are not arranged randomly: “everything is related to everything else, but near things are more related than distant things” (Tobler, 1970, p. 236). Additionally, residents do not choose where they live arbitrarily and “to choose a neighborhood is to choose neighbors” (Schelling, 1971, p. 145). There is always an element of factors such as individual preference, housing policy decisions, and financial factors etc. which impact on housing location. Thus, I hypothesise that the expected prevalence will be higher than the observed value, as a city arranged randomly is unlikely to be arranged in such a way that is more spatially cohesive - and have fewer sharp transitions - than the actual city. My hypothesis is confirmed in Table 3.6. For both methods the expected level was higher than the observed in both 2010 and 2019.

Table 3.6: Excess embrittlement level in Amsterdam in 2010 and 2019.

Year	Lee and Mitchell (2013)			PAM Clustering		
	Expected	Observed	Excess Embrittlement	Expected	Observed	Excess Embrittlement
2010	17.03%	13.6%	3.43%	46.06%	23.7%	22.36%
2019	16.88%	13.33%	3.55%	44.84%	25.17%	19.67%

In terms of the excess embrittlement level, for both methods the expected level of social frontiers was smaller in 2019 compared to 2010. The difference between the expected frontier percentage of 2010 and 2019 for the Lee and Mitchell (2013) method was similar to that of the observed percentage. The lower expected percentage for the PAM clustering method contrasts with the observed findings, suggesting that the observed results were potentially more of an abnormality and therefore a slightly less reliable estimation.

In relation to the pattern of the excess embrittlement level, for the Lee and Mitchell (2013) method the city is becoming very slightly further removed from random, albeit there was only a 0.12% difference. This difference indicates that the city, or at least the grid cells, is becoming arranged in such a way that reduces the level of divisions in the area. Whilst the PAM clustering method is becoming slightly closer to random, which could imply that there are processes at work increasing division levels (Manley *et al.*, 2023).

I cannot make any conclusive statements about the correlation between social frontiers and the index of dissimilarity as we have only considered one city and two non-consecutive years. However, we can see that the reduced expected level of frontiers corresponded with the decrease in the index of dissimilarity between 2010 and 2019 in Amsterdam. The expected level for both methods decreased, dropping by 0.15% and 1.22% for the Lee and Mitchell (2013) and PAM clustering methods respectively, consistent with the decreasing index of dissimilarity. This suggests that the city of Amsterdam is actually arranged in such a way which reduces the social frontier prevalence, when we compare it to what we would have expected based on the index of dissimilarity. This could be linked to the infrastructure, with buildings arranged in such a way that reduces divides which are ignored when the city is shuffled. This may be a result of policies aimed at reducing the divides or due to the presence of social housing which results in residents being placed, limiting housing choice thus reducing the influence of homophilious preference. To investigate whether this is consistent across a continuous period of time and regions, future research could extend my study and evaluate whether there is a significant relationship between the index of dissimilarity and expected frontier prevalence.

Conclusion

Summary of findings

This chapter makes three large empirical and methodological contributions to the social frontier literature: (i) it develops a robust framework comparing the accuracy of social frontier estimation methods and finds the PAM clustering approach as the most computationally efficient alternative to the Lee and Mitchell (2013) method; (ii) it compares frontier prevalence in Amsterdam in 2019 relative to 2010; and (iii) the chapter computes excess embrittlement levels, comparing the expected level of frontiers in 2010 and 2019 along with a discussion about the similarity between the trends of the embrittlement level and index of dissimilarity.

The option of the PAM clustering method as an identification method provides researchers who are looking to identify frontiers on a large scale if they require a reduction in the computational complexity, whilst maintaining a high degree of accuracy. This could lead to frontiers being identified in larger regions (potentially even nationwide), in multiple years, or a finer spatial scale being used. The accuracy

testing of the four methods from the simulation study provides reliable evidence that PAM clustering is an appropriate choice for identifying social frontiers. For this particular study, the findings suggest that the PAM clustering method has a higher true positive rate than the Lee and Mitchell (2013) method – thus is more likely to correctly identify a boundary as a social frontier – whilst also being more likely to overestimate the social frontiers in a region with a lower true negative rate. The chapter recommends scenarios which would suit the use of either the PAM clustering or Lee and Mitchell (2013) method. When considering the pattern of the social frontier prevalence or dynamics the Lee and Mitchell (2013) method is preferred due to its more stringent requirements meaning that the only frontiers identified are the strongest, hence strengthening findings about how frontiers evolve. Thus, in Chapters 4 and 5 the Lee and Mitchell (2013) method is applied. For larger studies, e.g. those that study frontiers nationwide or in multiple regions over an extended time period, or those which are aiming to be used to form policies the PAM clustering method enables this to be possible by reducing the execution time as shown in Table 3.3 and thus being capable of being an introductory method. This explains the choice of the PAM clustering method as the social frontier identification method in Chapter 6.

The pattern of the social frontier prevalence in Amsterdam when comparing 2019 and 2010 differed for each identification method. This difference, although potentially complicating the findings, does highlight the need to be cautious in discussing findings. Neither method identified a large change in frontier percentage nor in the boroughs prone to social frontiers (Figure 3.6 and Table 3.5). The frontiers were particularly prolific in the Noord, Nieuw-West, Oost and Zuidoost boroughs, and surrounding the city centre (a borough outline map can be seen in Figure A.1). The findings demonstrate the need for the inclusion of both visual and statistical descriptor measures within studies to deepen our understanding of how: the frontiers exist in society; the 2010 and 2019 results differ and the difference in the frontier locations. Ideally maps showing social frontiers would be interactive, showing the changes over an extended consecutive period of time. I have addressed this to some extent by discussing the rate at which frontiers remain in the same location. Additionally, for the excess embrittlement level I found that based on the index of dissimilarity we would expect the city to be decreasing in frontier prevalence similar to the pattern of the Lee and Mitchell (2013) method. If a city is becoming further removed from random, then this suggests that there are processes or policies which are resulting in the city becoming less divided. It could be of interest to consider what changes occurred between these years which resulted in the decreasing segregation levels.

Limitations

My approach has some limitations. Firstly, the non-Western classification condenses a large group of different nationalities into one, simplifying any conclusions about segregation levels as it is unlikely that all these nationalities have the same

experiences within the country. Hartog and Zorlu (2009) found that segregation levels varied amongst the largest nationality groups in the non-Western definition. Turkish and Moroccan individuals were more likely to live together, whilst the Surinamese population often lived amongst other immigrants from non-Western countries (Hartog and Zorlu, 2009, p. 21). Additionally, the current definition is based not only on an individual's country of birth but also their parents'. This is despite research showing that second-generation migrants often have higher levels of contact with those born in the Netherlands and a larger knowledge of the Dutch language (Gijsberts and Dagevos, 2007). I was limited in this chapter by the classification used within the data when my datasets were created. However, these classifications are changing to country of origin with four levels based mostly on continent of birth (CBS, 2022c) and thus future studies may be able to adapt the methods to include multiple groups, studying whether patterns are consistent. Additionally, currently the Lee and Mitchell (2013) method is only able to identify social frontiers between two population groups. This is a limitation of the method as it fails to capture the complete picture of the pattern of segregation as people from different groups likely have differing experiences. For example, the different ethnicities within the non-Western classification will have likely experience different levels of segregation based on economical or historical factors.

This chapter begins to consider the presence of single frontiers through their exclusion in the aggregate length measure, however, their significance and the difference in their effect remains unexplored. These single frontiers could be a result of random noise, meaning that they are only visible in one time period or they could be actual steep transitions which are simply smaller in length and still divide communities. Identifying frontiers in more than two non-consecutive time periods could enable researchers to investigate whether these frontiers persist, the external factors that contribute towards their existence (e.g. housing policies such as social housing locations), and allow studies to evaluate whether these frontiers have a significant impact on communities.

By only studying residential social frontiers, thus being 'place-based', this chapter somewhat assumes that if people are not mixing within their communities (specifically where they live), then they have a reduced opportunity for social interaction and social mixing - meaning that an individual's neighbourhood determines their opportunity for interaction when holding everything else constant. Thus, suggesting that any opinions or stereotypes that an individual holds are likely to have developed through interactions within their communities (O'Sullivan and Wong, 2007). However, more recently segregation studies have shifted towards person-based, going beyond residential segregation to consider activity spaces, schools, workplaces, etc. (Piekut, 2021; Dixon *et al.*, 2022; Li *et al.*, 2022). This is due to the diminishing importance of neighbourhoods in relation to the formation of the majority of relationships and individuals' exposure to different cultures; neighbourhoods are no longer the only place where individuals spend their time.

Activity spaces provide an opportunity for contrasting groups to form relationships and thus likely reduce prejudice whilst increasing social mixing (Xu *et al.*, 2019, p. 11). As communication between individuals has become easier with larger social networks and increased transportation (Spielman and Logan, 2013), relationships are more easily built outside of local neighbourhoods. By studying only residential social frontier dynamics, we limit our understanding of segregation.

Suggestions for future research

Following my discussion surrounding the limitation of place-based segregation, person-based segregation research could be incorporated into social frontiers by exploring whether social frontiers have a greater impact on individuals who feel their neighbourhood is more restricted. For example, if an individual works or attends a school within the same neighbourhood as they live, is the frontier effect greater than for an individual who leaves their neighbourhood to work in a less homogeneous, more diverse area? Furthermore, the aggregate measures in this chapter could be extended to calculate the number of people, or proportion of the population, living in close proximity (e.g. within 100m) of a social frontier. A similar technique was implemented by Maguire, French and O'Reilly (2016) when they analysed the relationship between mental health and proximity to physical barriers in Belfast, with proximity calculated based on whether there was a peace wall within the area (on two different scales). Issues could arise here when determining proximity due to the requirement of fine scale spatial data to accurately calculate the distance between the individual's household and the social frontier, however, this could be possible with this extensive Dutch dataset due to the large level of information it contains about an individual's location.

Furthermore, the inclusion of the aggregate description measures facilitates a discussion about frontier dynamics and the variance in their effects depending on the stage of the frontier and their presence in the city. Causal effect studies on social frontiers are currently very limited and they have all focused on a single moment in time, limiting our understanding of their impact. Finally, all four of the identification measures in this chapter used a boolean definition for classifying frontiers with boundaries either defined as social frontiers or not. For the simulation study this made it easier for the accuracy calculations and caused little issue. However, when considering the effects of social frontiers it is likely that their impact varies depending upon their strength. Thus, the inclusion of a steepness (strength) measure would be beneficial to understand how the frontier's strength varies over its lifecycle and to evaluate the effect of the frontier at different lifecycle stages and strength levels.

This chapter's three main aims were (i) to test the success of the three proposed identification methods, (ii) to compare frontier prevalence in Amsterdam in 2010 and 2019, and (iii) to calculate the excess embrittlement level discussing its relationship with the index of dissimilarity. I conclude that researchers can justify using the PAM clustering instead of the Lee and Mitchell (2013) method if the study includes a large

dataset, requires a level of computation that is unrealistic when using the Lee and Mitchell (2013) method, or is an introductory/policy based study which wants to ensure it does not miss anybody who lives near a social frontier. Although PAM clustering produces high accuracy scores, questions could arise over its potential to overestimate the social frontiers which could mean that results are an exaggeration. Thus, studies should consider its clustering patterns when determining its suitability at identifying frontiers within the study. The implementation of the two aggregate measures has made for easier comparison between time periods and regions. By addressing these gaps within the literature, I hope to encourage further social frontier research which has potentially large implications for future policies and segregation studies.

Chapter 4 The Lifecycle of Social Frontiers

A study into the dynamics and evolution of steep boundaries in the Netherlands.

Abstract

Introduction: Research into the prevalence of ‘social frontiers’ – abrupt boundaries between adjacent communities – and their impacts has grown in recent years. However, there remains a notable gap in the literature regarding frontier dynamics, particularly, concerning their evolution over time and any variation between regions (Dean *et al.*, 2019; Iyer and Pryce, 2024). Iyer and Pryce’s (2024) theorised social frontier lifecycle has them emerging as hostility between communities increases, before stabilising and eventually disappearing as residents began to mix.

Objectives: This chapter represents the first detailed study of frontier dynamics. I estimate frontier strength and prevalence in the Amsterdam and the Hague regions to create lifecycle plots showing how frontiers evolve over time. The chapter compares frontier prevalence, calculated using four measures, and explores the difference in the dynamics between regions and the four frontier dimensions: low-income, high-income, unemployment, and country of birth.

Methodology: The chapter makes use of the Lee and Mitchell (2013) method to identify social frontiers across four different dimensions between 1999 and 2021 in Amsterdam and the Hague. The prevalence of these frontiers is evaluated through the inclusion of average length, strength, percentage, and duration measurements.

Results: For all the frontier dimensions and time spans the frontiers evolved as postulated by Iyer and Pryce (2024): their magnitude increased prior to identification, flattened as they became established and then decreased as they dissipated. Even when transitions between communities became smooth, more than 50% of the frontiers re-emerged. The economic frontiers were less prevalent than the non-Western frontiers in Amsterdam and the Hague, whilst the two cities had very similar prevalence levels overall. For the majority of dimensions, the frontiers occurred more frequently in 2021 compared to 1999.

Conclusion: The chapter concludes that frontiers are in general increasing in the regions. In both Amsterdam and the Hague, the country of birth frontiers were the most prominent. Within the lifecycle study, even when frontiers decreased in magnitude (frontier strength) prior to their demise, they frequently re-appeared, suggesting that divides between neighbourhoods are unlikely to fully heal.

Introduction

As we saw in the difference in the frontier prevalence in 2019 compared to 2010 in Chapter 3, frontiers are not a static phenomenon (Iyer and Pryce, 2024, p. 795); they emerge, evolve, and possibly disappear as the population changes, policies are introduced, and neighbourhood sorting occurs. The dynamic nature of segregation has long been established. Massey and Denton (1985) argued that residential mobility should be viewed as a longitudinal process that needs to be studied over an extended period. Additionally, studies have explored the factors that impact the evolution of segregation and neighbourhoods over time and help to explain the spatial pattern of neighbourhoods (Sampson and Sharkey, 2008; Bailey, 2012; Kauppinen and van Ham, 2019). They conclude that segregation is inherently dynamic; it is a process which constantly changes and evolves. However, social frontiers, an important yet under-researched dimension of segregation, have often been studied as a static entity, even in Chapter 3 I identified frontiers in two non-consecutive years. Studies have been mostly cross-sectional (Iyer and Pryce, 2024), where frontiers are identified at a specific moment in time, despite this meaning we are unable to capture frontier prevalence, interaction levels, and evolution (Zorlu and Mulder, 2008; Martinovic, van Tubergen and Maas, 2009; Iyer and Pryce, 2024; Spierenburg, van Cranenburgh and Cats, 2024).

This chapter suggests that we should explore the lifecycle of social frontiers, focusing on how their strength varies when they are forming, existing, and eventually diminishing. This inclusion is crucial to understand how frontiers evolve, and how their effect may differ at each stage of their lifecycle. Currently, we know very little about how frontiers develop and whether this differs based on the social context in which they are identified or the dimension which is studied (Dean *et al.*, 2019). Through the inclusion of duration and strength measures I explore how the social frontier magnitude varies throughout its lifecycle, comparing my empirical findings to the theoretical lifecycle proposed by Iyer and Pryce (2024). This theoretical lifecycle predicted that a frontier's magnitude will strengthen prior to and during its initial formation, plateau as the frontier stabilises, and decline as tensions thaw, bridge-builders move in, and the frontiers disappear (Iyer and Pryce, 2024). The lifecycle development opens up opportunities for future studies to consider whether there are aspects of society which promote frontier growth, along with whether frontiers are more consequential at certain stages of their lifecycle.

The dynamic nature of these frontiers can result in their impact varying depending on lifecycle stages (Legewie and Schaeffer, 2016; Dean *et al.*, 2019; Iyer and Pryce, 2024). Sharp neighbourhood transitions and the arrangement of communities within regions have the ability to greatly influence the experience and opportunities of those who live near them. There exist a large number of both empirically proven and theoretically proposed social frontier effects, including increased crime and territorial conflicts (Dean *et al.*, 2019; Iyer and Pryce, 2024), a reduction in cross-ethnic

relationships (Zorlu and Mulder, 2008), a lack of shared community spaces resulting in inter-group tensions over existing infrastructure and limited social mixing (Piekut *et al.*, 2024), and increased feelings of relative deprivation (Iyer and Pryce, 2024). However, social frontiers also provide an opportunity for positive contact with an increased chance of interaction (Dean *et al.*, 2019), which brings with it an opening for differing communities to form relationships, especially if they share communal spaces. Sharing public/community spaces results in increased contact meaning that individuals are more likely to form their opinions based on actual interactions rather than stereotypical or prejudiced beliefs (Peters and de Haan, 2011). If neighbourhoods experience an extended period of conflict they are often targeted by peace-building initiatives aiming to transform spaces that were formally zones of conflict into peaceful spaces such as those seen in Belfast (Mitchell and Kelly, 2011).

In Chapter 3, I proposed that social frontiers should be viewed as an additional dimension of segregation alongside the existing dimensions conceptualised by Massey and Denton (1988): evenness, exposure, centralisation, concentration, and clustering. These dimensions should be used in conjunction with each other to provide a comprehensive description of the segregation level of an area (Owen *et al.*, 2023) and to understand the effects of these frontiers on society. However, how these effects might vary depending on exposure length or the substantiveness of the social frontier has not been studied as far as I know. By advancing the study of social frontier dynamics, I hope that future studies will include frontier evolution measures into their effect studies.

I specifically study frontier dynamics and prevalence in Amsterdam and the Hague, building upon Chapter 3 by identifying social frontiers in consecutive years over an extended period of time, between 1999 and 2021. I define magnitude (strength) as the absolute difference between the two contiguous grid cells in the social frontier dimension. The majority of social frontier studies have focused on a specific city, Spierenburg *et al.* (2025) identified social frontiers in 520 cities but did not compare their prevalence across these different cities, instead studying how the social frontiers aligned with physical boundaries. By comparing the social frontier prevalence in the two regions, I can investigate whether there are region specific factors which impact the frontier level. Or if the results are consistent across regions, then I can identify the national policies or events which have occurred and influenced the prevalence. This chapter studies two large urban regions: Amsterdam and the Hague. Westerners living in large cities have a higher probability of living amongst non-Westerners compared to the country-wide average in the Netherlands as urban areas often have a more diverse population (Bolt, van Kempen and van Ham, 2008; Hartog and Zorlu, 2009; Musterd and Van Kempen, 2009; Glaeser, 2011). Additionally, cities tend to be representative of segregation rates and patterns across the country (Tammaru *et al.*, 2016; Spierenburg, van Cranenburgh and Cats, 2024), meaning they likely encapsulate the divides that are present on a national scale making results more generalisable (Boal and Livingstone, 1984, p. 175).

I not only study frontiers in two regions over an extended period of time but also in multiple dimensions: country of birth, unemployment, and income. To my knowledge, economic frontiers have not been included in social frontier research. Social frontier studies, along with the majority of segregation research, have tended to focus on ethnicity or country of birth, with these perceived as the main factor in community segregation (Tammaru *et al.*, 2016, p. 1; Boterman, Musterd and Manting, 2021). Income frontiers have been overlooked in social frontier research. This is despite research highlighting how income segregation often limits opportunity in society (Tammaru *et al.*, 2019), with lower labour market performance identified amongst those in deprived neighbourhoods (van Ham and Manley, 2010),

The inclusion of multiple social frontier dimensions creates a more comprehensive framework for understanding how different societal factors impact the geographical arrangement of neighbourhoods and the relationship between society and space in regions (Boterman, Musterd and Manting, 2021, p. 497). Social inequalities are multidimensional with these dimensions often overlapping, potentially resulting in them increasing or negating each other's impact (Mauritti *et al.*, 2016). This is especially apparent with the large intersection between income and ethnicity segregation in the Netherlands, meaning that "ethnic segregation also reflects socio-economic inequalities" (Ziller and Spörlein, 2020, p. 5). The identification of frontiers on multiple dimensions naturally leads to the study of intersectional frontiers - those which exist simultaneously in multiple domains (Iyer and Pryce, 2024, p. 784) – which Chapter 5 focuses on, along with enhancing our understanding of segregation levels in the Netherlands. This is in addition to being able to explore whether the lifecycle development and frontier prevalence level is consistent across dimensions.

In this chapter, I first review the social frontier and segregation literature, focusing on the lifecycle, prevalence measures, and segregation dimensions in the Netherlands. I then discuss the data and methodology. This chapter uses the Lee and Mitchell (2013) method to identify the frontiers rather than the PAM clustering method based on the arguments in Chapter 3. In Chapter 3 I noted the conservative nature of the Lee and Mitchell (2013) method, highlighting that it has a higher requirement for a border to be defined as a social frontier compared to the PAM clustering method. As I am modelling the lifecycle of the frontier I wanted to ensure that I only analysed the strongest of the frontiers, which I can be very confident in their identification, to ensure that any pattern I identify is reliable and robust. The data used both in this chapter and Chapter 5 is not large enough to make the process time consuming enough to justify using PAM clustering. It was computationally feasible to implement the Lee and Mitchell (2013) method as the dataset was not incredibly large meaning that it did not take an extensive period of time nor computer memory, although the process was still relatively time consuming as frontiers were identified in every year between 1999 and 2021. I solidify the study by including the PAM clustering results in Appendix H, which acts as a robustness and reliability check for the patterns and traits identified in the chapter. In the results section I compare the frontier prevalence

across the regions and dimensions over twenty years. Finally, I conclude with a summary of the results, the limitations of the chapter, and a discussion of recommendations for future studies.

Literature

Frontier lifecycle stages

The processes and cycles that drive neighbourhood segregation are of great interest in ethnic segregation research (Bailey, 2012, p. 708), in Chapter 2 I discussed the factors which impact and can explain patterns of segregation. To understand how segregation evolves we must consider the factors that impact an individual's neighbourhood selection (Krysan and Crowder, 2017, pp. 41–42). This is in addition to mechanisms such as selective intraurban and international migration, the local neighbourhood context including historical and housing elements, and spatial processes such as gentrification which I discussed in Chapter 2 (Sampson and Sharkey, 2008; Bailey, 2012; Kauppinen and van Ham, 2019; Piekut, Pryce and van Gent, 2019). Boundaries “are historically contingent processes, which emerge, exist for some time and disappear” (Newman and Passi, 1998, p. 201). Additionally, by their nature, political boundaries are also dynamic, actively evolving and impacting on how communities work alongside each other (Newman, 2003, p. 128). The importance of the dynamic nature of spatial segregation and boundaries is established within the field of urban geography, however, research into the dynamics and lifecycle of social frontiers is relatively limited, specifically in relation to empirical research.

The dynamic nature of international and inter-community boundaries highlights that frontiers are not static. Kristof's (1959) work on extracting separate theoretical understandings of the terms ‘frontiers’ and ‘boundaries’ conceptualised frontiers as the transition between two different lives where the contrasting forces are “neither fully assimilated to nor satisfied with either” (p. 273). The terms boundary and frontier are often used interchangeably in the literature, but Kristof (1959) suggested that frontiers are often produced randomly representing zones of transition, whilst boundaries are defined centrally, often “fixed and enforced” (p. 281). In the introduction chapter (Chapter 1) I discussed the creation of the term social frontiers and the importance of their fluid nature in interpreting and understanding their existence and impact, along with differentiating between them and spatial discontinuities. Legewie and Schaeffer's (2016, p. 126) study of contested boundaries found that boundaries are at their most contentious when they are fuzzy as the territory is not yet fully established. Dean *et al.* (2019, p. 272) proposed that these contested boundaries should be thought of as a phase of frontier evolution. When boundaries are unstable, this will also likely present itself in the instability of territory, meaning that communities are fragile and the interspace between them is often contested resulting in high levels of conflict (Legewie and Schaeffer, 2016), the

territorial nature of social frontiers was highlighted by Iyer and Pryce (2024) as a crucial differentiator between them and other sharp transitions.

Iyer and Pryce (2024, p. 793) proposed a stylised hypothetical social frontier lifecycle and reflected on likely differences in frontier impacts depending on the stage of the lifecycle. Figure 4.1 is an adapted version of this hypothetical lifecycle, depicting the variation in the frontier strength over time. As an additional element to Iyer and Pryce's (2024) lifecycle Figure 4.1 includes a further lifecycle line representing the frontiers that have become ingrained in society whose strength has been maintained over an extended period of time. Iyer and Pryce (2024) speculated that growing hostilities between contrasting groups results in the initial development of the social frontier. Once interspatial zones are established, the frontiers stabilise as communities and territory becomes settled. Following this period of stabilisation, frontiers either become entrenched in society where they maintain their strength over an extended period of time or they disappear, diminishing in strength before eventually dissolving as contrasting communities merge together.

All of the current social frontier lifecycle research proposed and discussed within this chapter is theoretical. Newman and Paasi (1998) proposed an agenda for the study of political boundaries, noting the importance of exploring the dynamic nature of boundaries, and considering them collectively rather than independently. Boundaries by their nature are changeable, they go beyond politics to become multidimensional, impacting multiple aspects of society for those who live alongside them (Newman, 2003, p. 123). Empirical social frontier research is very limited especially that which explores how the frontier strength or effect varies over an extended period of time. Both Legewie and Schaeffer (2016) and Wickes *et al.* (2025) included an edge intensity measurement in their contested boundary studies. However, they did not consider how this changed over time nor was this studied directly in connection to social frontiers. In general, social frontier literature has been cross-sectional (Iyer and Pryce, 2024), studying a single city at a moment in time. Consequently, the variation in frontier strength over its lifecycle has been bypassed, along with changes in the potency and direction of frontier impacts on crime, anti-social behaviour, wellbeing, geographical mobility etc.

Empirical research into the lifecycle of a social frontier would therefore open up new avenues of inquiry into the evolution of frontiers and their impacts, exploring whether the frontiers behave as we would expect over their lifespan. The inclusion of multiple dimensions and regions means we can assess consistency across multiple factors. Prior to my empirical analysis, I include a theoretical discussion about the lifecycle stages of social frontiers based on Figure 4.1, discussing their dynamic nature and potential consequences on communities.

Theorising the social frontier lifecycle

Stage one: Formation

This is depicted in the upswing phase of Figure 4.1. Dean *et al.* (2019) proposed that social frontiers developed for two reasons: an “aversion to living at the interface of communities” or as a “product of benign forces” (p. 274). I proposed that some of the frontiers identified in Chapter 3 arose as a result of gentrification in the city of Amsterdam, with them often identified surrounding the city centre where lower socioeconomic groups have been forced out of the increasingly expensive centre. Frontier locations can be affiliated with physical boundaries, such as the ‘Peace Walls’ in Belfast which act as a physical division between Catholic and Protestant communities. These create a clear division of territory and further divide communities.¹³ Dixon *et al.* (2025) found that “residents overwhelmingly remain within the zones associated with their own community” (p. 16), with residents often avoiding areas that they designate to not belong to them. However, Spierenburg *et al.* (2025) found no “systematic alignment” between social frontiers identified based on the proportion of non-EU migrants and physical divides such as railways or roads in Europe (p. 3). Regardless of the cause, the frontier is formed when there is an upswing or downswing in either one or two groups resulting in a large difference between adjacent neighbourhoods, ultimately leading to lower levels of mixing and neighbourhoods becoming separated. Legewie and Schaeffer (2016) found that contested boundaries between racial and ethnic groups, which Dean *et al.* (2019) theorised could be a stage of social frontiers are the most impactful on conflict. This could be extended to be that the boundaries are most likely to be contested when they are forming as this is when territory is least established and stable.

Stage two: Stability

Once the frontier is established, Iyer and Pryce (2024) assume that its stable nature results in reduced contact between groups due to the frontier acting as a barrier. In the case of Belfast, where territory has been established and cemented through the use of ‘Peace Walls’ there are street networks and facilities which have become almost exclusively used by one community. This has further enforced the sectarian divide between the populations (Dixon *et al.*, 2020, p. 463) and highlighted how territory has been inherently accepted by communities. Legewie and Schaeffer (2016) found that well-defined boundaries, as we suspect stable social frontiers to be, resulted in lower levels of conflict than fuzzy, unstable boundaries. The sharp divide results in established zones of territory with lower levels of contact between the groups ultimately making the boundary more manageable (Newman, 2006; Legewie and Schaeffer, 2016). However, contrasting communities which share a social frontier will by their nature have an increased opportunity of contact due to higher exposure levels. This is potentially more damaging for relationships: exposure

¹³ A limitation of this thesis which I discuss in Chapter 7 is the lack of exploration into how social frontiers coincide or are impacted by physical boundaries.

without contact can have greater negative consequences than no exposure, with prejudicial attitudes reinforced, reducing trust and relations between neighbourhoods (Stolle, Soroka and Johnston, 2008; Dean *et al.*, 2019). If contact can be established between groups, then this can vastly improve relationships between groups which ultimately could lead to the demise of the social frontier (Dixon *et al.*, 2025). However, Allport's (1954) contact hypothesis proposed several conditions for this contact to be positive including equal status and common goals, without which communication between groups can be limited or negative.

Stage three: Demise or entrenchment

Following a period of stability, there are likely two options for the social frontier trajectory: firstly, it becomes established, continuing to exist for an extended period of time. Alternatively, relations thaw between communities, resulting in increased mixing and the frontier disappearing.

Entrenched frontiers can become well-established and longstanding divides between communities. This does not necessarily mean that they are a heightened space of tension, rather, there is a possibility that when the territory and zones are established these divides can reduce the chance of violence and crime, potentially making those who live near them feel safer. The most obvious example is the 'Peace Walls' in Belfast established following the Troubles, which act as a divider between Catholic and Protestant communities and were originally introduced as a method of maintaining safety and reducing violence (Dixon *et al.*, 2020, p. 459). Despite being built as a temporary structure, they continue to exist more than fifty years after the beginning of the Troubles.

In contrast, following a period of stabilisation the frontiers could decline as the transition between the contrasting groups smooths. A social frontier could decline for numerous reasons, including urban reconstruction policies aimed at reducing segregation and increasing social mixing. The demolition of housing followed by the construction of higher-cost dwellings aimed at attracting middle-class households is an attempt to encourage social mixing within communities (Bolt and van Kempen, 2011). Bolt, van Kempen and van Weesep (2009) suspected that these policies often result in low-income communities dispersing, potentially indicating a decline in social frontiers. However, these clusters often reappear elsewhere, suggesting that frontiers are likely to re-emerge following their disappearance but in a slightly different location with the overall proportion unchanged. There exist projects such as Startblok (<https://startblokriekervan.nl/en/>) in the Netherlands or the Belfast Interface Project (<https://www.belfastinterfaceproject.org/>) that aim in the case of Startblok to form diverse and cohesive neighbourhoods by housing refugees with young Dutch citizens, or for the interface project, using creative methods to regenerate areas surrounding the Peace Walls. Both of these attempt to reduce segregation levels and increase the cohesion between communities, based on the

theory and proven concept that increased use of shared spaces can help build relationships between groups, reducing conflict near social frontiers (Piekut *et al.*, 2024).

Furthermore, it could be the long-term close proximity of the communities which ultimately results in relationships forming, especially following a period of stability. Positive contact can result in the emergence of cross-ethnic relationships reducing tension and increasing social mobility which is likely to reduce feelings of relative deprivation and thus improve relations (Zorlu and Mulder, 2008; Dixon *et al.*, 2025). In addition, the presence of bridge-builders - heterophilous households attracted to the contested boundary - will result in boundaries becoming sites of residential mixing as the frontiers are bridged between where individuals interact and live alongside each other (Newman, 2006, p. 150; Iyer and Pryce, 2024, p. 786).

If social frontiers cease to exist or physical barriers are removed then crossing into adjacent regions is likely made easier. However, the existing ethnic ties and segregation effects are frequently retained as there still exists the social distance between communities, even those who are co-located within neighbourhoods (Newman, 2006; Dixon *et al.*, 2020; Dixon *et al.*, 2025). This, along with Bolt, van Kempen and van Weesep's (2009) suggestion that clusters of groups are likely to re-emerge in society, suggests that frontiers are likely to reappear even if they initially declined. Long-standing firmly entrenched frontiers have the potential to have a longer-term impact as individuals are exposed to them over a greater period of time.

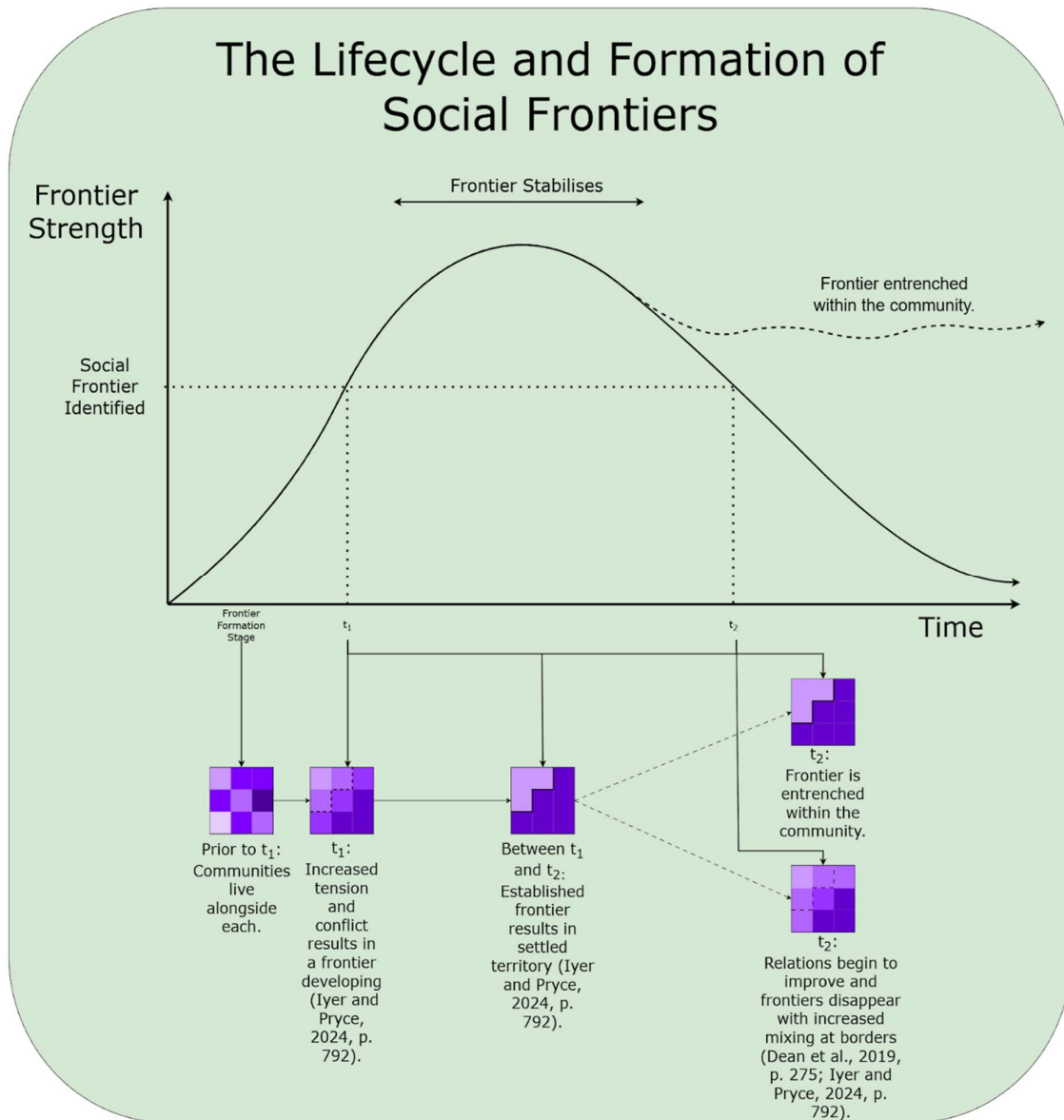


Figure 4.1: A hypothesised social frontier lifecycle plot and simulated frontier formation stages. Adapted from Iyer and Pryce's (2024, p. 793) illustration of frontier effect over time. It additionally includes a split where the frontier either becomes established and entrenched between communities or relations thaw and the frontier dissipates.

Frontier prevalence measures

Chapter 3 highlighted the importance of including multiple social frontier prevalence measures to develop a more complete picture of how ubiquitous frontiers are. Such measures facilitate the comparison of changes in regions over time. Additionally, without the inclusion of frontier prevalence measures it is impossible to empirically explore frontier dynamics. The lack of empirical studies comparing regions or cities

means that policymakers lack concrete evidence about whether the frontier effects vary depending on the region, nor are they able to identify the causes of the divisions and thus cannot target policies (Spierenburg, van Cranenburgh and Cats, 2024). The inclusion of statistical descriptors to describe the level of frontiers in an area in Chapter 3 is unique for a social frontier study. Harris' (2014) study included an aggregate measure which described the prevalence of steep transitions between neighbouring areas in UK local authorities, however, there were weaknesses in both Harris' paper and Chapter 3 in terms of their ability to identify and describe boundaries. Harris (2014) did not focus specifically on the location or dynamics of specific social frontiers: the identified areas were not required to border each other merely to be within a certain distance.¹⁴ Individuals in bordering grid cells are more likely to have close contact and thus, potentially higher levels of tension between groups.

Chapter 3 described the frontier prevalence in Amsterdam using two measures: percentage of social frontiers and average length. These partially capture the level of divisions in an area, however, they cannot describe the consistency of the divides, nor whether the frontier effect varies depending on the steepness of the boundary. To address this, I include two further prevalence measures within this chapter: frontier duration and strength as proposed by Iyer and Pryce (2024) in their theoretical paper. These not only address some of the limitations discussed in Chapter 3 about whether there are frontiers that re-emerge or exist for an extended period by calculating the rate of re-emergence, they also provide the opportunity to engage with the innovative study of frontier dynamics, incorporating both of the measures to explore the strength of the frontier over its lifecycle.

The strength measure is an edge intensity measurement. The hypothesis is that the larger the difference between the bordering grid cells, the steeper the transition, hence the greater the social frontier's impact and the more likely the frontier is to have "potent symbolic meaning" (Iyer and Pryce, 2024, p. 784). Iyer and Pryce (2024, p. 787) also proposed that social frontiers have large territorial consequences, and thus the stronger the frontier the more magnified their impact with a higher likelihood of defensive behaviour within communities. Equally, the presence of strong frontiers may reduce any appeal for bridge-builders whose appearance is likely to diminish the strength of frontiers as there are fewer members of the outgroup on either side of the frontier (Iyer and Pryce, 2024), making it more difficult for groups to gradually merge together (Dean *et al.*, 2019). Dean *et al.* (2019, p. 286) discussed the strength of frontiers in relation to fuzzy contested boundaries (regarding Legewie and Schaeffer's (2016) contested boundary paper) stating that their frontiers had a difference of around 8%, a relatively small difference. Fuzzy boundaries have the

¹⁴ This not necessarily a flaw, in fact it would be useful in a social frontier study to consider a buffer zone or the exact location of individuals within the grid cell to calculate exactly how far apart contrasting resident groups live from each other.

potential for being more consequential on society than established steeper boundaries as there is likely to be heightened conflict over the unstable territory (Legewie and Schaeffer, 2016). In this sense, weaker frontiers - where the transitions are slightly smoother - could be more consequential due to the instability of the boundaries. This does not mean that social frontiers ingrained in society are inconsequential, nor that their presence enhances cohesion between communities, but that their impact is potentially more stable with fewer flare ups in conflict or territory disputes.

For frontiers to have a notable sustained impact on communities, they are likely to have existed for a significant period of time (Iyer and Pryce, 2024), with their extended presence likely to entrench and deepen tensions and divisions between groups. For these long-standing frontiers to be identified, we need to study social frontiers across multiple years. As I identify frontiers across over 20 years I am able to create a duration measure, noting both how long the frontier continuously existed and also the rate at which they re-emerge if they previously disappeared. This leads to a discussion about whether certain regions have more long-lasting frontiers and how the frontier strength varies over its lifespan.

Although there is a potential for variation in the frontier impact depending on its duration, this is unexplored in social frontier research. In diversity-related faultlines - those which “divide a group's members on the basis of one or more attributes” (Lau and Murnighan, 1998, p. 325) - there is an expectation that the faultlines only continue to play a role in creating subgroups if they are constantly active (Lau and Murnighan, 1998, p. 334). Faultlines in organisations can lead to members unconsciously categorising themselves into unspoken groups resulting in limited communication which ultimately increases divides (Lau and Murnighan, 1988, p. 328). If these faultlines are inactive with the groups merging, faultlines tend to diminish in strength resulting in groups becoming increasingly familiar with each other and the faultline reducing in significance (Lau and Murnighan, 1988, p. 333). The length of time that social frontiers exist is likely linked to how established the divisions are between communities. Equally, communities which have long established roots in an area where residents have developed an emotional attachment can lead to the exhibition of territorial defensive behaviour as residents want to protect their home from “outsiders” (De Backer, 2022, p. 1). When strong attachments are formed within communities, this can lead to individuals on the outside of the neighbourhood feeling isolated from opportunities (Bannister, Kintrea and Pickering, 2013).

As discussed in Chapter 3 and in segregation literature, segregation indices and these social frontier prevalence measures should be used in conjunction with each other (Massey and Denton, 1988, p. 309; Owen *et al.*, 2023). The strength and duration measures should be included alongside the length and percentage measures used in Chapter 3 to describe and compare the level of divisions in a

region. It is the combination of these measures (strength, percentage, length, and duration) which can explain the extent to which social frontiers can have a symbolic presence and consequential effect on the lives of residents (Iyer and Pryce, 2024).

As well as being able to describe the difference in the frontier prevalence across regions and dimensions over their lifecycle. Future studies could use the duration measurement proposed here, to explore the difference in the impact of a frontier at each stage of its lifecycle. This facilitates an exploration of the variation of frontier effects depending on strength and age (Dean *et al.*, 2019).

Study Context: Dimensions of Segregation in the Netherlands

Although there has been an increase in empirical social frontier research, these have tended to focus on ethnicity or country of birth divisions (Dean *et al.*, 2019; Křížková *et al.*, 2021; Oliner *et al.*, 2024; Spierenburg *et al.*, 2025). Even in Chapter 3 I focused on country of birth social frontiers, despite discussing the importance of gentrification and social housing in explaining income segregation and thus frontier patterns. The absence of the economic dimension is consistent with inequality and segregation research, especially in Amsterdam where most studies have concentrated on the arrangement of migrant and non-migrant populations (Tammaru *et al.*, 2016, p. 1; Boterman, Musterd and Manting, 2021). The bypassing of income and educational divisions means that the level of segregation in an area is not fully understood. To develop an accurate and thorough description we need to consider a range of neighbourhood characteristics (Vervoort, Flap and Dagevos, 2011; Tselios *et al.*, 2015, p. 417). The economic dimensions play an important role in understanding and describing the interaction level and spatial pattern of neighbourhoods.

In this chapter I identify four different dimensions of social frontiers: low-income, high-income, unemployment, and country of birth, with the existence and pattern of economic frontiers currently unknown. A description of the causes and patterns of ethnic segregation in the Netherlands is included in Chapter 2. This section focuses on economic segregation literature, and patterns of both economic and country of birth segregation particularly in the Amsterdam and Hague regions. I noted their differences and similarities, discussing why these may exist along with the potential for my findings to be generalised beyond these regions.

In Chapter 3 I identified areas within the city of Amsterdam - the North-East, the Western borough and the periphery of the city centre - where the frontiers were particularly prolific in 2010 and 2019. This is likely due to the city centre becoming increasingly gentrified, resulting in it being unaffordable for the majority of lower-income individuals (Hochstenbach and Musterd, 2021). Chapter 3 identified, using the Lee and Mitchell (2013) method, a very slight reduction in the percentage of country of birth social frontiers in the city of Amsterdam between 2010 and 2019.

This chapter extends Chapter 3 by studying the trend of the social frontier prevalence over a 20-year period rather than non-consecutive years (2010 and 2019), enabling a description about the year-on-year differences and thus pattern of segregation.

Table 4.1: Socioeconomic descriptive statistics in Amsterdam and the Hague in 1999 and 2021.

	Amsterdam		The Hague	
	1999	2021	1999	2021
Population	1,144,066	1,381,190	736,655	891,998
Non-Western Population	260,496 (22.77%)	416,540 (30.16%)	150,359 (20.41%)	269,596 (30.22%)
Unemployed Population	194,216 (27.7% of working pop)	164,402 (20.6% of working pop)	118,048 (27.66% of working pop)	113,045 (23.03% of working pop)
High-income Population	67,007 (9.57% of working pop)	99,257 (12.46% of working pop)	42,734 (10% of working pop)	51,004 (10.39% of working pop)
Low-income Population	112,501 (16.06% of working pop)	153,429 (19.26% of working pop)	67,102 (15.72% of working pop)	98,675 (20.11% of working pop)
Work Age Population (20 - 64)	700,417	796,594	426,762	490,797
Average cell size	67.97	72.77	76.15	79.44
Non-Western Index of Dissimilarity (D)	0.526	0.424	0.556	0.456
Unemployed (D)	0.232	0.296	0.245	0.313
High-income (D)	0.449	0.392	0.477	0.480
Low-income (D)	0.203	0.287	0.223	0.332
Median Work Income	€25,464	€42,791	€25,996	€43,704
Median Total Income	€22,926	€38,332	€23,994	€40,098

These statistics are created using the non-public microdata from Statistics Netherlands.

Economic segregation

Levels of economic segregation and income inequalities have been heavily impacted in European countries by their transformation “from industrial to post-industrial societies” (Tammaru *et al.*, 2016, p. 2). The level of economic segregation in an area is impacted by a range of factors including the housing market, welfare state, and the economic structure of a region. A large amount of neighbourhood effects literature has focused on the impact of economic segregation on socioeconomic status (van Ham and Manley, 2010; Galster and Sharkey, 2017), educational attainment (Chetty, Hendren and Katz, 2016; Nieuwenhuis and Hooimeijer, 2016; Troost, van Ham and Manley, 2023), and mental health (Bezzo and Vanhuyse, 2024).

It has been argued that there are both positives and negatives to economic segregation. Putnam (2007) and Cheshire (2009) argued that living in an equally economically deprived neighbourhood can lead to the formation of stronger networks, resulting in increased opportunities for jobs, greater support networks, and knowledge and skills being transferred within the community. However, there are also multiple explanations as to why economic segregation can have negative consequences on society: they can reduce access to role-models and facilities thus lessening opportunities, limit the formation of social networks, and increase feelings of relative deprivation (Bezzo and Vanhuyse, 2024). Particularly in the context of social frontiers, the severity of the boundary can make it seem almost impossible to cross, this along with the absence of bridgebuilders can mean that residents can feel increasingly isolated amplifying the effects of economic segregation. This feeling of isolation and inability to access resources can reduce social mobility with the frontiers seemingly impenetrable (van der Laan Bouma-Doff, 2007; Dean *et al.*, 2019; Iyer and Pryce, 2024). Additionally, Granovetter (1973) stressed the importance of ‘weak ties’ (casual acquaintances), as individuals with multiple, potentially weaker, connections are likely to find these act as stronger bridges between communities and therefore provide more opportunities, such as in securing employment, in comparison to individuals with few very strong connections (Henning and Lieberg, 1996). If individuals feel isolated within their communities, then they are less likely to enjoy the benefits of these ‘weak ties’ and thus can feel limited within the labour market.

The traditionally high-level of social housing, along with the large welfare state in the Netherlands, has often softened economic segregation with the level tending to be lower than other European countries as individuals are cushioned from potential deprivation (Musterd *et al.*, 2016). In Chapter 2 I discussed how the reducing rate of social housing in the Netherlands adversely impacts low-income and unemployed individuals, often restricting their housing choice based on the location and availability of an increasingly reducing social housing sector. I expect that there exist concentrated areas of low-income households within the regions which have developed due to high levels of social housing or lower cost housing being clustered within the regions. In addition, there is also likely to be clusters of higher-income

individuals due to them voluntarily choosing to split from society (Musterd *et al.*, 2016, p. 231). Even in income diverse neighbourhoods there is no indication that this diversity results in higher rates of social interaction, with residents often living alongside rather than amongst each other (Bolt, van Kempen and van Ham, 2008). The Netherlands Institute for Social Research (2024, p. 4) found that in the Netherlands lower-income individuals were less likely to encounter high-income people, not only in their neighbourhood but also in their workplace, friendship circles, and educational settings. Higher-income Dutch individuals are becoming increasingly segregated from lower or middle-income households with a reduced opportunity for contact (The Netherlands Institute for Social Research, 2024). For this study this is likely to manifest itself in both high- and low-income social frontiers surrounding groups of grid cells and these areas likely to have been recognised in prior studies as being economically homogeneous.

Segregation in Amsterdam and the Hague

Descriptive statistics

The descriptive statistics in Table 4.1 show that Amsterdam and the Hague are very similar in relation to their non-Western population and economic statistics. Additionally, when exploring the level of segregation using segregation indexes – in this case the Index of Dissimilarity (D) to measure evenness – the scores are very similar. The Hague's score is consistently higher than Amsterdam's albeit only by a minor amount. The literature reflects the similarity in these findings with the studies often switching in what they argue: Bolt, Hooimeijer and van Kempen (2002, p. 215) noted that the Hague has frequently been referred to as the most ethnically segregated city in the Netherlands, with noticeably stronger rates of segregation than Amsterdam. However, Boterman and van Gent (2023) described Amsterdam as a minority-majority city, meaning that more than half of the population were either born abroad or have one parent who was, reflecting a high level of diversity (Aalbers and Deurloo, 2003).

According to the index, in 1999 52.6% and 55.6% of the non-Western population in Amsterdam and the Hague would have to relocate to another neighbourhood for the region to have an equal distribution of Westerners and non-Westerners, a relatively high score. The non-Western index was lower in 2021 than 1999 in both regions indicating that non-Westerners are becoming more evenly spatially distributed within the regions. This decrease in D is despite the increase in the non-Western population meaning that they make up over 30% of the population in both regions (Table 4.1). The index is not directly affected by the changes in the share of non-Westerners, only changing when the residential pattern has altered (Kauppinen and van Ham, 2019). Thus, it seems that some of the non-Westerners are moving into previously homogeneous Western neighbourhoods, leading to different groups becoming more evenly spread across the region and country of birth segregation decreasing. In Figure 4.2 we can see the consistent increase in the proportion of

non-Westerners in both regions, with the Hague experiencing a larger growth resulting in it having a very slightly higher proportion of non-Westerners in 2021.

In terms of the descriptive statistics in Table 4.1, we can see that the percentage of both high- and low-income individuals increased between 1999 and 2021 in both Amsterdam and the Hague. This is especially so in Amsterdam, where the high-income percentage dramatically increased (see Figure 4.2), with the proportion of high-income individuals in 2021 1.3 times that of 1999. For the low-income population, the increase was larger in the Hague compared to Amsterdam, with the proportion in 2021 equal to 1.28 times that of 1999. Both regions had a higher proportion of low-income individuals than high-income, nearly twice as many for the Hague in 2021. In Figure 4.2 we can see a spike in the percentage of low-income individuals in Amsterdam in 2007, this could potentially be related to the financial crash. Although the plot shows a dramatic spike, the difference is relatively small. In contrast to the non-Western and income populations, the unemployed population fell between 1999 and 2021 in both Amsterdam and the Hague, with the decrease more pronounced in Amsterdam. In Figure 4.2 a slight increase following the Covid-19 pandemic in 2020 is visible, however, rates reversed after the end of pandemic regulations (OECD, 2023, p. 30; CBS, 2024d). Unemployment rates in the Netherlands have decreased in part due to policies aimed at providing career advice, training, and support to the unemployed population (European Commission, 2024, p. 6). Additionally, areas which were traditionally working-class neighbourhoods, where there was often a high economically inactive population, experienced an influx in their migration population (Boterman and van Gent, 2023, pp. 57–58). In terms of the economic segregation patterns noted in Table 4.1, the unemployed and low-income scores showed a similar pattern in Amsterdam and the Hague. They were both higher in 2021 than 1999 and were alike in their score as we would expect due to many low-income individuals not being in paid employment (The Netherlands Institute for Social Research, 2024). The high-income scores differed in their patterns with Amsterdam decreasing in segregation whilst the Hague saw a very small increase.

Patterns of segregation

I discussed the spatial pattern of non-Westerners in the city of Amsterdam in Chapter 3, noting that gentrification, reducing levels of social housing, and the existence of established communities can likely explain the suburbanisation pattern. There are very few visible clusters for non-Westerners living in the central borough. It is plausible to assume that this pattern will also be reflected in the region.

Additionally, through studying low-income frontiers we are also able to see how the dimensions overlap - a topic I look at in greater detail in Chapter 5. The spatial distribution patterns of non-Westerners and lower-income households are likely to be similar within the regions due to the correlation between the dimensions (Hartog and

Zorlu, 2009; Sleutjes, de Valk and Ooijevaar, 2018) with migrants earning on average 34% less than non-migrants in Amsterdam (OECD, 2018).

Both regions have clear extremes in their housing market, a house in the most expensive neighbourhood would cost at most 400 times that of the cheapest neighbourhood in 2021 in both Amsterdam and the Hague (NL Times, 2021). Yet, they do vary in their average house prices: €630,621 in Amsterdam and €465,418 in the Hague in 2025 (CBS, 2026), highlighting the more expensive nature of Amsterdam.

The region's primary industries vary in that Amsterdam is more tourism focused whilst the Hague is the political and administrative centre. This has impacted the housing market within the regions with Amsterdam finding an increasing number of private investors purchasing properties with the intention of using them for short-term rentals whilst the Hague is experiencing a wave of previous ministerial buildings being converted into residential housing (Hekwolter of Hekhuis, Nijskens and Heeringa, 2017, p. 9). The regions also differ in their spatial economic pattern: the centre of the Hague tends to be poorer with the surrounding areas more affluent, whilst the city centre of Amsterdam has become increasingly expensive with the presence of affluent groups vastly increasing (Musterd, Hochstenbach and Boterman, 2020).

The economy of Amsterdam evolved between 1999 and 2021 due to economic growth and the rise of the middle class in the late 90s which resulted in the increasing construction of private housing (Fainstein, 1997, p. 299; van Duijne and Ronald, 2018, p. 639). Musterd, Hochstenbach and Boterman (2020) described Amsterdam as a "multi-layered and global city" (p. 3), noting the city's well-connected nature due to Amsterdam Schiphol airport and the recent business developments resulting in the city becoming a hub for financial and creative industries. The financial crisis in 2008 led to a decline in the construction of new housing which has eventually resulted in a housing shortage and therefore increased pressure on the housing market explaining the high prices (van Gent and Boterman, 2019). Within the city centre there was a shift in the 2000s towards a preference for owner-occupancy housing, with a lesser focus on equality in house availability (van Duijne and Ronald, 2018, p. 645). These changes have resulted in the centre of Amsterdam becoming younger and more affluent and the proportion of low-income households declining (Musterd, Hochstenbach and Boterman, 2020; Boterman and van Gent, 2023), especially as a large proportion of social housing in the centre has been converted into private housing. All these factors have led to the expansion of Amsterdam into the periphery areas with the development of new social housing on former industrial land and post-war estates (Hochstenbach and Musterd, 2017; van Gent and Boterman, 2019). The gentrification of the city centre has resulted in the displacement of lower-income households towards the outskirts of the city, with the remote surrounding areas dominated by lower-income households, shifting poverty

towards the suburbs (Hochstenbach and Musterd, 2017; Musterd, Hochstenbach and Boterman, 2020). Musterd, Hochstenbach and Boterman (2020) described the city centre of Amsterdam as “polarised”, with the proportion of high-income households decreasing and the proportion of low-income households increasing as the distance grows from the city centre (p. 8).

The Hague is the bureaucratic capital of the Netherlands, home to both the Dutch parliament and the International Court of Justice, thus it often attracts an international workforce of professionals (Musterd, Hochstenbach and Boterman, 2020). In contrast to Amsterdam, the Hague has been naturally divided into two sections: an inland side and a seaside with clear patterns of economic segregation visible along the divide. This is likely an area where we would expect to identify social frontiers, similar to the north south divide I discussed in the excess embrittlement section in Chapter 3. Traditionally, the seaside has been occupied by more affluent households, whereas the inland side is mostly blue-collar workers (Meijers *et al.*, 2014). Bailey, van Gent and Musterd (2017) identified a marked income divide in the Hague, noting the extremes were becoming more distinct. However, Musterd, Hochstenbach and Boterman (2020, p. 10) found that recent trends have seen the city centre becoming poorer whilst affluence has increased in the suburbs. Whereas Amsterdam has experienced expansion in recent years, with increased buildings in the suburbs, the Hague is limited in its ability to expand due to the coast forming a natural barrier (Hekwolter of Hekhuis, Nijskens and Heeringa, 2017, p. 9). Despite the importance of the public sector workforce in the city the Hague has an older average population, partly due to the lack of university in the city and the administrative importance of the city (Bailey, van Gent and Musterd, 2017). Families often remain in the central neighbourhoods of the Hague, whilst in Amsterdam families are more likely to move to the surrounding areas (suburbs) of Amsterdam (Hekwolter of Hekhuis, Nijskens and Heeringa, 2017, p. 9).

Identifying social frontiers in multiple regions allows a comparison of frontier prevalence. Any extreme differences in trends can potentially be explained by contrasting policies or specific aspects of the arrangement of the neighbourhoods which may influence frontier prevalence. As discussed in the excess embrittlement section in Chapter 3, this difference can be more extreme whilst not necessarily identified by the social frontiers. Additionally, the inclusion of multiple regions increases the generalisability of the study. Although both regions are in the Netherlands and thus there is consistency in the national policies, they vary due to factors such as their differing main economies, their housing markets, and historical or geographical factors which can explain differences in their spatial arrangements. In relation to generalisability, the colonial history of Netherlands makes its immigration patterns – the large proportion of immigrants arriving from former colonies - and thus segregation and social frontier trends likely to be that similar to that of other countries who had large empires such as the UK and France. Gentrification and housing market trends in Amsterdam make it likely to have

experienced similar trends to Berlin or Vienna which have experienced a similarly high level of demand on their housing markets and implemented comparable strategies to increase the proportion of affordable housing (Wijburg, 2021). Whilst the growing importance of the political and diplomacy industries in the Hague makes it similar to cities such as Geneva or Brussels both of which are also key political and bureaucratic centres within their country and Europe, thus findings from the Hague may be similar to that of those cities. Studying frontiers in different cities within the same country means that national policies can be controlled for, allowing neighbourhood-level factors to be isolated. These factors could include individual regional policies, differences in the economic spatial arrangement and housing market, or natural barriers in the regions impacting the social frontier prevalence.

Percentage of Socioeconomic dimensions in the Hague and Amsterdam

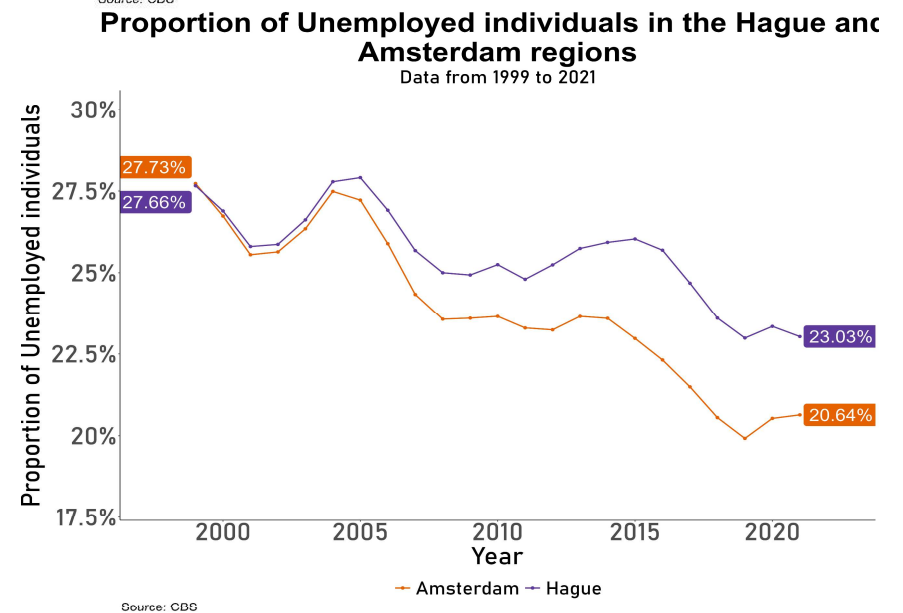
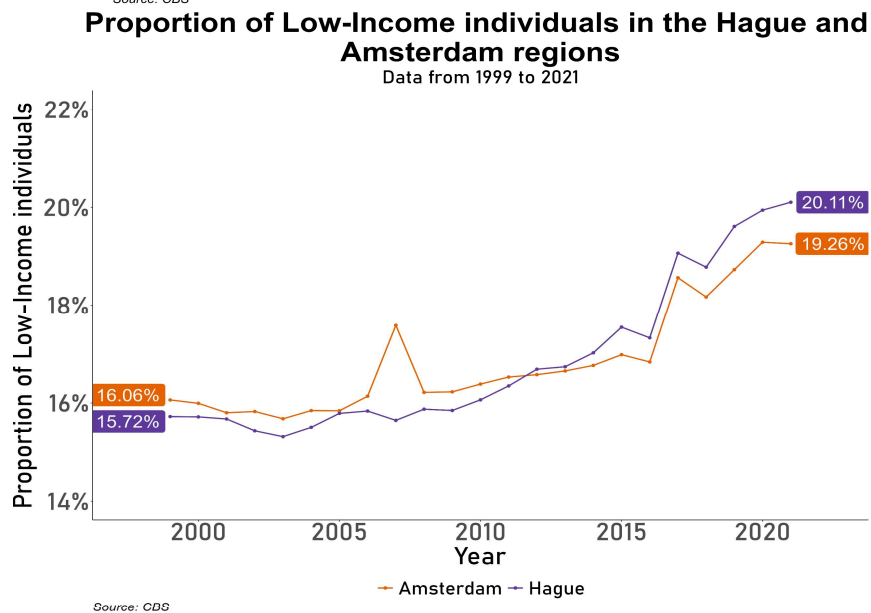
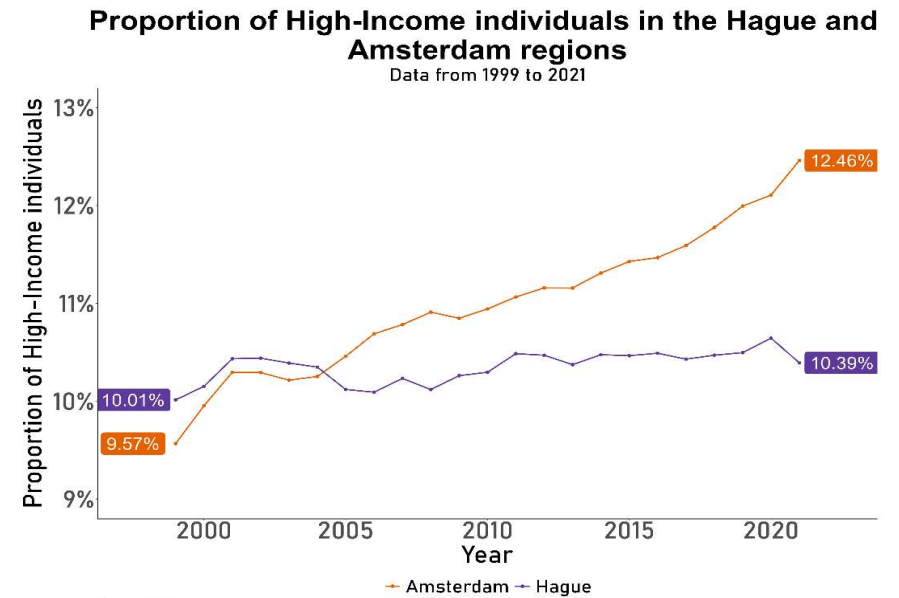
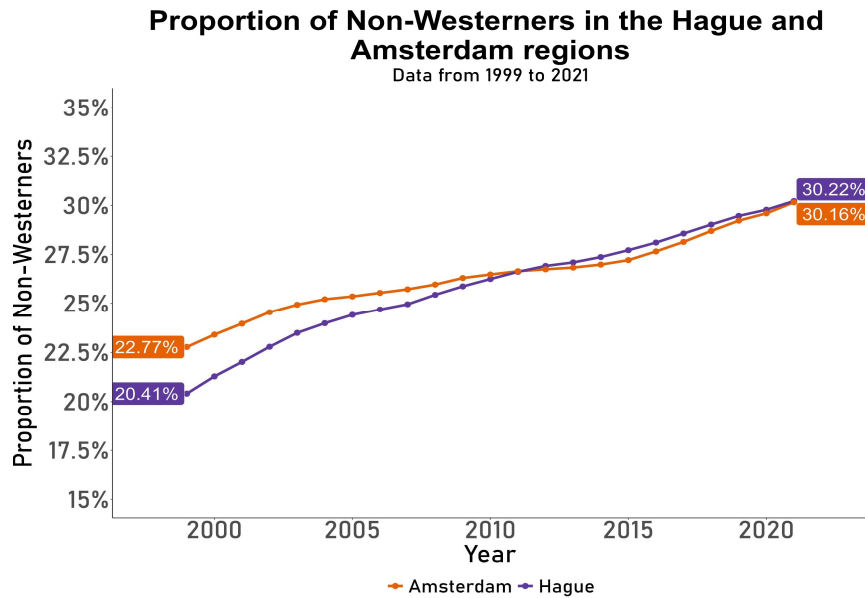


Figure 4.2: Plots showing the proportion of the variables of interest.

Data and Methodology

Data processing

As with Chapter 3, this chapter utilises the registry data from Statistics Netherlands. However, it expands upon the study by identifying unemployment, income (high and low), and country of birth social frontiers in the Amsterdam and Hague regions using the 100m by 100m grid cells.¹⁵ I used the 100m by 100m grid cell size to compare how the prevalence pattern varies from that identified in Chapter 3 where the 500m by 500m grid cells were used, investigating whether the grid cell choice impacts the determined level of segregation. Fowler (2015) argued that we should consider segregation to be continuous across different scales, and thus there is not one single scale which correctly measures all segregation. Additionally, Olnier *et al.* (2024) noted how using a smaller scale reduces the possibility of missing variation which can prove to be important when identifying social frontiers. Furthermore, this chapter is focused on the dynamics of the frontiers and thus a smaller grid cell does not limit our findings due to us being less concerned about the individuals impacted and more so on the patterns, as I use the Lee and Mitchell (2013) method I identify only the strongest of the frontiers.

Economic social frontiers are under researched in the literature despite the large level of research on relative deprivation and economic segregation. Both social frontier and segregation studies tend to focus on ethnicity. Furthermore, the majority of studies have only considered social frontiers in a single city, rather than comparing the prevalence across two regions.¹⁶ These limitations could be in part due to access to data, aggregated data such as census data is unlikely to have the highly detailed level of information about income that is available from Statistics Netherlands. The study is more likely to offer a valuable contribution to the literature if the frontiers in the two cities/regions can be studied at a finer spatial scale or at multiple scales which may not be available in the data.

Amsterdam and the Hague regions are two of the largest in the Netherlands with approximately 1.3 million and 900,000 inhabitants in 2021 respectively. Amsterdam is the capital city of the Netherlands whilst the Hague is the administrative centre and home of the government. Segregation patterns in cities occur as a result of a range of different factors as discussed in Chapter 2 and in this chapter's study context, including individual preferences, particularly homogeneous preferences; housing availability which includes the crucial aspect of the housing market structure; the

¹⁵ The definition of the regions is based on the COROP definition used by Statistics Netherlands (CBS, 2023c) where they are described as Groot-Amsterdam and Agglomeratie's-Gravenhage - Table E.1 includes a list of the municipalities within each region.

¹⁶ Spierenburg *et al.* (2025) identified social frontiers in 520 cities in multiple European countries, however, the prevalence in these cities was not studied individually. The paper focused on the relationship between social frontiers and urban boundaries.

local economy; and government or local policies. The frontiers are measured in a larger area than in Chapter 3. Chapter 3 identified frontiers in the city of Amsterdam, whereas this chapter identifies the frontiers in the region which also includes the more rural outer municipalities of Amsterdam. These outer regions have seen an increase in lower-paid workers with some offering cheaper housing and tending to be poorer on average than the city centre (Boterman and van Gent, 2023, pp. 62–63), meaning there exists a polarisation between the central region and the surrounding neighbourhoods (Boterman and van Gent, 2023, p. 61).

Social frontier dimension definitions

Very few social frontier studies have identified frontiers in more than one dimension - Dean *et al.* (2019) identified country of birth and ethnicity frontiers whilst Thomas (2023) also included religion – however, economic frontiers are relatively untouched. The definitions and population for the dimensions – low-income, high-income, unemployment, and non-Western, are described in Table 4.2.

Table 4.2: Social frontier dimension descriptions.

Dimensions	Description	‘Eligible Population’
Low-Income	Proportion of the population who are in the bottom 20 percentile of earners in the Netherlands. ¹⁷	Work Age population (20 to 64 years old) excluding students. ¹⁸
High-Income	Proportion of the population that are in the top 10 percentile of earners in the Netherlands. ¹⁹	Work Age population (20 to 64 years old) excluding students.
Unemployment	Proportion of the population that have not worked for at least one month in the year.	Work age population (aged 20 to 64 excluding students).
Country of Birth (Split into Western and non-Western).	Non-Westerners are defined as people with origins from any country in Africa, South America, or Asia (excluding Indonesia and Japan) or from Türkiye (CBS, 2025a). The Western category includes all the Dutch-born population along with Western migrants.	Whole region.

¹⁷ Cut-offs for the income can be found in Table F.1.

¹⁸ Students’ income is often inconsistent and based on student finance received during studies which can result in students appearing in the lowest group when they may be receiving alternative support which is not included. They also frequently do not work or do not work for the whole year, impacting their average income. Retired individuals (or those over 65) are also excluded as their income is unlikely to correctly represent their socioeconomic status.

¹⁹ Cut-offs for the income are included in Table F.1.

I chose to use a binary definition for all four dimensions – this was already implemented in Chapter 3 for the non-Western definition based on the data availability. For the income dimensions, using the binary definitions rather than the average income reduces the issue of outliers which can distort the average (Troost, van Ham and Manley, 2023). Furthermore, the Lee and Mitchell (2013) method currently only identifies social frontiers between two groups.

Frontier identification

This chapter uses the Lee and Mitchell (2013) method which was adapted by Dean *et al.* (2019) to identify social frontiers. This is the established social frontier identification method, utilised in Křížková *et al.* (2021) and Olnér *et al.* (2024), which I discussed in Chapter 3 as deliberately cautious in its nature, using a two-step method to only identify the sharpest of boundary transitions. In Chapter 3 it was identified as one of the two most accurate social frontier identification methods by the simulation study.²⁰ The Partitioning Around Medoid (PAM) clustering method was proposed in Chapter 3 as an alternative, quicker identification method to Lee and Mitchell (2013) and is included as a robustness check with the results included in Appendix H. However, the method tends to overestimate the number of social frontiers and will always identify a social frontier even if the surface is completely smooth which could be misleading for the dynamics study as only the strongest – ‘true’ frontiers should be examined. For this chapter I wanted to ensure that when discussing the prevalence of frontiers and exploring how they evolve over time I could be confident the boundaries identified were social frontiers and thus likely to be experienced by residents. Therefore, as it is feasible in this study - the datasets are not beyond the computational capability of the method - the results are based on the frontiers identified by the Lee and Mitchell (2013) method which has more basis and is established in social frontier literature (Dean *et al.*, 2019; Křížková *et al.*, 2021; Olnér *et al.*, 2024). The methodology of the Lee and Mitchell (2013) method is explained in Chapter 1.

Frontier prevalence measurements

Prior to Chapter 3, social frontier papers have only included visualisations of frontiers, with no aggregate descriptive measures incorporated. This chapter builds upon Chapter 3 by incorporating additional prevalence measures: the strength, re-emergence rate, and average frontier duration. Table 4.3 includes a description of the four prevalence measures.

²⁰ I also applied the PAM clustering method to identify the social frontiers in Amsterdam and the Hague with the results included in Appendix H as a comparison and robustness check for the Lee and Mitchell (2013) method. The methods complement each other to provide a stronger picture of divisions in the region. Ultimately, the PAM clustering dynamics plots are very similar to that identified by Lee and Mitchell (2013) cementing our findings. Further information about the methods and their accuracy - based on a simulation study - can be found in Chapter 3.

Table 4.3: Frontier prevalence measure methodology.

Measure	Methodology
Strength (%)	The strength measurement reflects the steepness of the social frontier. It is computed by calculating the median absolute difference in the frontier dimension between bordering grid cells where a social frontier exists between them.
Duration (Years)	Average frontier duration: The average length of time that the frontiers existed continuously. Frontier re-emergence rate: The percentage of frontiers which reappeared in the same location after initially declining. This measure is used to explore the evolution of the frontier strength against its lifecycle to evaluate whether the frontier strength evolves as we may expect - initially increasing, stabilising and then decreasing. The average frontier duration is calculated by assessing the number of years the frontier existed continuously and averaging the overall score for each dimension and region. The frontier re-emergence rate assesses what percentage of the frontiers that disappeared, reappeared in the same location after a brief period of absence.
Length (m)	The average length of the social frontiers. This is calculated by grouping the adjoining social frontiers together and computing their length. These lengths are then averaged (mean) over all the sets of grouped frontiers.
Percentage (%)	The proportion of the grid cell boundaries within the region which are social frontiers. Calculated by dividing the number of social frontiers in the region by the number of boundaries in each region.

Results

Our results are split into two main sections: frontier dynamics and a discussion of the social frontier prevalence in Amsterdam and the Hague. The frontier prevalence results are discussed in relation to which dimension is the most prevalent, their prevalence over an extended period of time, and finally the difference between the regions.

Frontier dynamics and evolution

The limited existing social frontier literature that discusses the dynamics of frontiers means that the rate at which they appear, develop, and potentially disappear is unknown. Additionally, it means we cannot conceptualise the difference in their impact at each stage of their lifecycle (Iyer and Pryce, 2024). To address this absence in the empirical literature, this chapter plots the strength of the frontier against its duration during its lifespan as well as including a discussion about the rate at which frontiers re-emerge. Figure 4.3 shows the frontier strength during the lifecycle of frontiers which have existed for 5-years and Figure 4.4 shows those

which were continuously identified as social frontiers.²¹ The time ‘T1’ marks the year the frontier was first identified, ‘T2’ the year the frontier was last identified, and ‘T1-2’ 2 years before the frontier was first identified etc.

The plots in Figure 4.3 showed very similar trends across the different dimensions and regions, which was also consistent across the plots in Appendix I. The frontiers’ strength steeply increased between T1-1 and T1 (the year the frontier was first identified), before often stabilising between T1 and T2 with the strength flattening out or only having relatively small peaks and troughs, before a steep decrease between T2 and T2+1. Many of the frontiers which were classified as non-social frontiers after a period of being defined as a frontier showed a visible surge in their strength in the year T2+2, suggesting that they may later be re-identified as social frontiers. This hypothesis is somewhat supported by Table 4.4, where over 50% of social frontiers re-emerged across all the dimensions. This is particularly the case for high-income frontiers which have a short average continuous duration but had a higher tendency to reappear.

This rate of re-emergence indicates that divided communities often do not fully merge even if the level of difference diminishes; they often increase and reappear in the following years. The high-income and unemployment frontiers were particularly prone to re-emerging, indicating that the communities are unlikely to completely merge together even if the difference does initially reduce.

Table 4.4: Average frontier duration and re-emergence rate.

		Average Frontier Duration (Years)	Average Continuous Frontier Duration (Years)	Rate of re- emergence
Amsterdam	Non-Western	4	3	57.7%
	Low-income	2	2	51.73%
	Unemployment	3	2	61.92%
	High-income	3	2	70.47%
Hague	Non-Western	4	3	55.88%
	Low-income	2	2	51.75%
	Unemployment	3	2	63.44%
	High-income	3	2	72.3%

²¹ The plots showing the 10-year frontiers and the results from the PAM clustering method can be found in Appendix I.

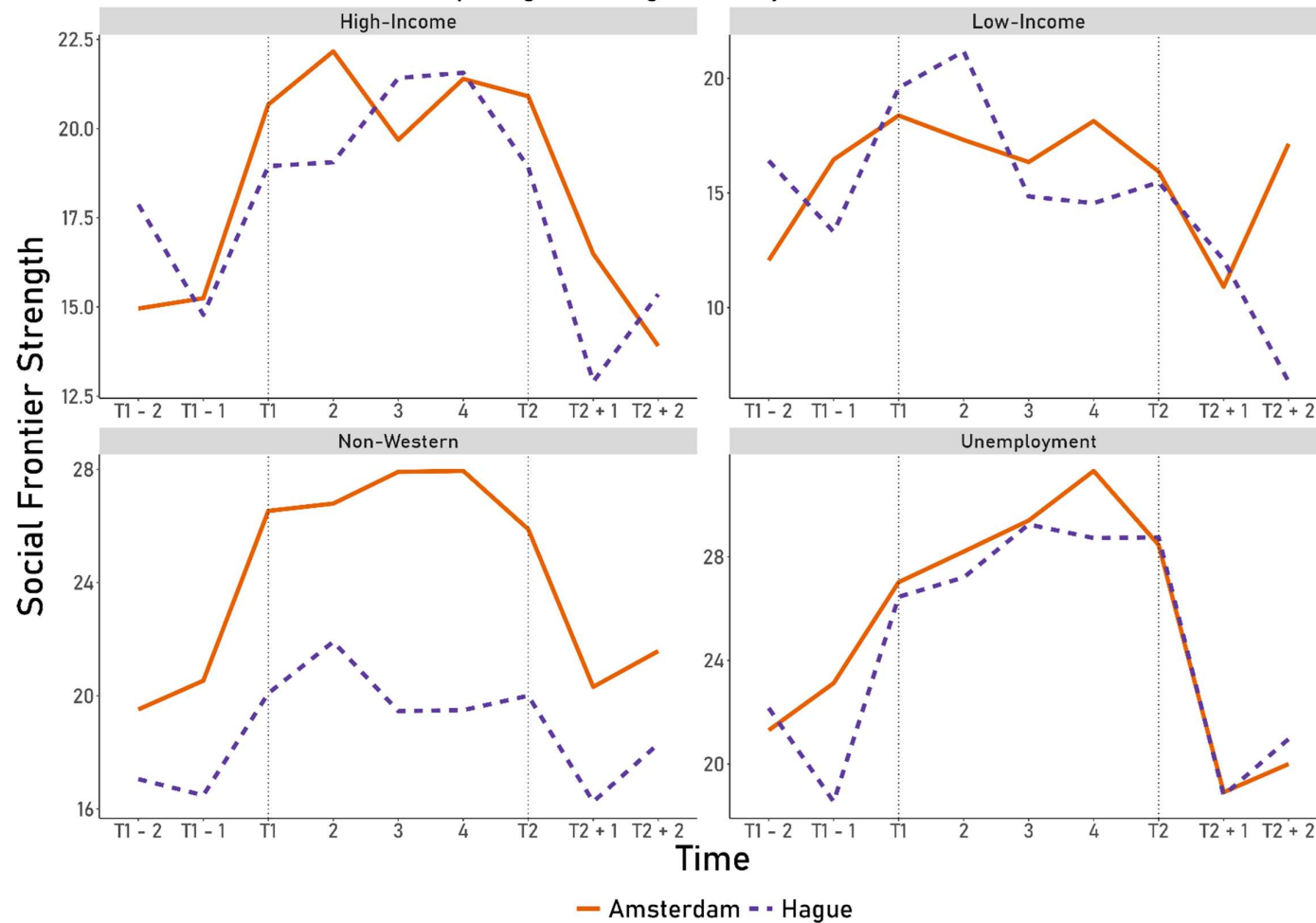
The trends in the strength of the frontiers over their lifecycle are visible in Figure 4.3 and are very similar to the theoretical lifecycle proposed by Iyer and Pryce (2024). Both Dean *et al.* (2019) and Legewie and Schaeffer (2016) suggested that the formation stage of the frontier posed the greatest threat to cohesion between communities and that this is when the frontiers have their greatest impact on neighbourhoods. In our case this is often where we identify the sharpest increase in strength, indicating that this is likely the stage where the presence of the frontiers is most likely to be felt as communities become established.

This is the first time that the trend of the frontier strength has been visualised and the plots almost exactly replicate the hypothesised lifecycle (Iyer and Pryce, 2024). Frontiers that exist for a minimum of five years before diminishing have a consistent average strength during the middle-stage of their life, whether that be the 4 years for a 5-year frontier or 9 years for a 10-year frontier I suspect this means that we would identify a consistent effect during this time period. Additionally, I plotted the social frontier strength in Figure 4.4 for the boundaries which were always identified as social frontiers whilst the boundary existed - which I refer to as entrenched frontiers. Figure 4.4 shows a similar pattern to that theorised in Figure 4.1.²² Once the frontiers have become entrenched and are continuously identified their strength shows some level of stabilisation especially for the non-Western frontiers. The frontier strength never dips lower than that of the time T1 after they have become established. Future studies could explore whether there is an explanation for why some frontiers become entrenched within society and others diminish - considering whether it is potentially based on area characteristics, policy changes, or events which resulted in the frontiers diminishing or becoming further established.

²² In the case of low-income frontiers, none of the frontiers existed as social frontiers for all the 23 years.

The Evolution of Lee and Mitchell Social Frontiers

Exploring the strength of five-year social frontiers

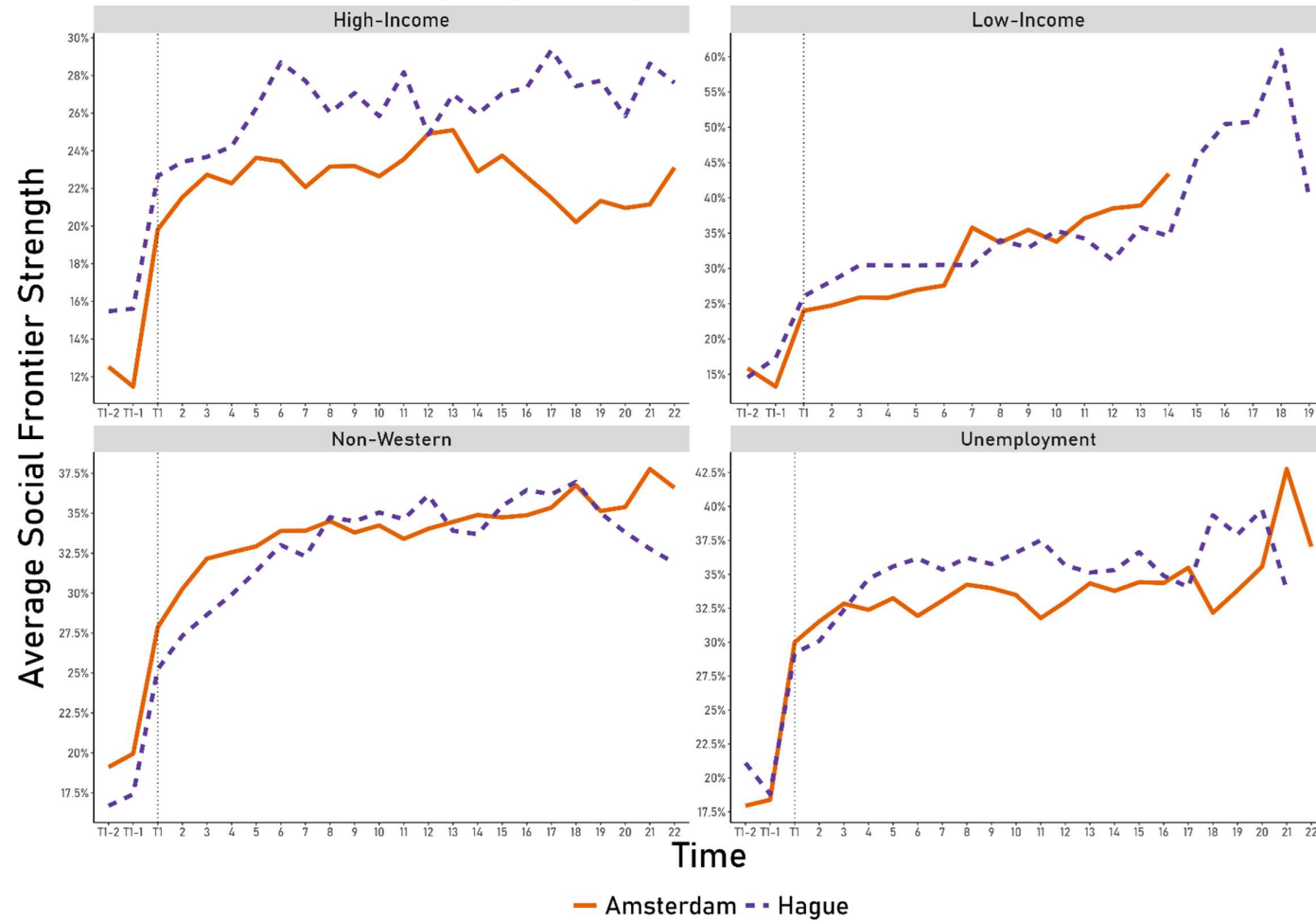


Source: CBS

Figure 4.3: Frontier lifecycle plots.

The Evolution of Entrenched Lee and Mitchell Social Frontiers

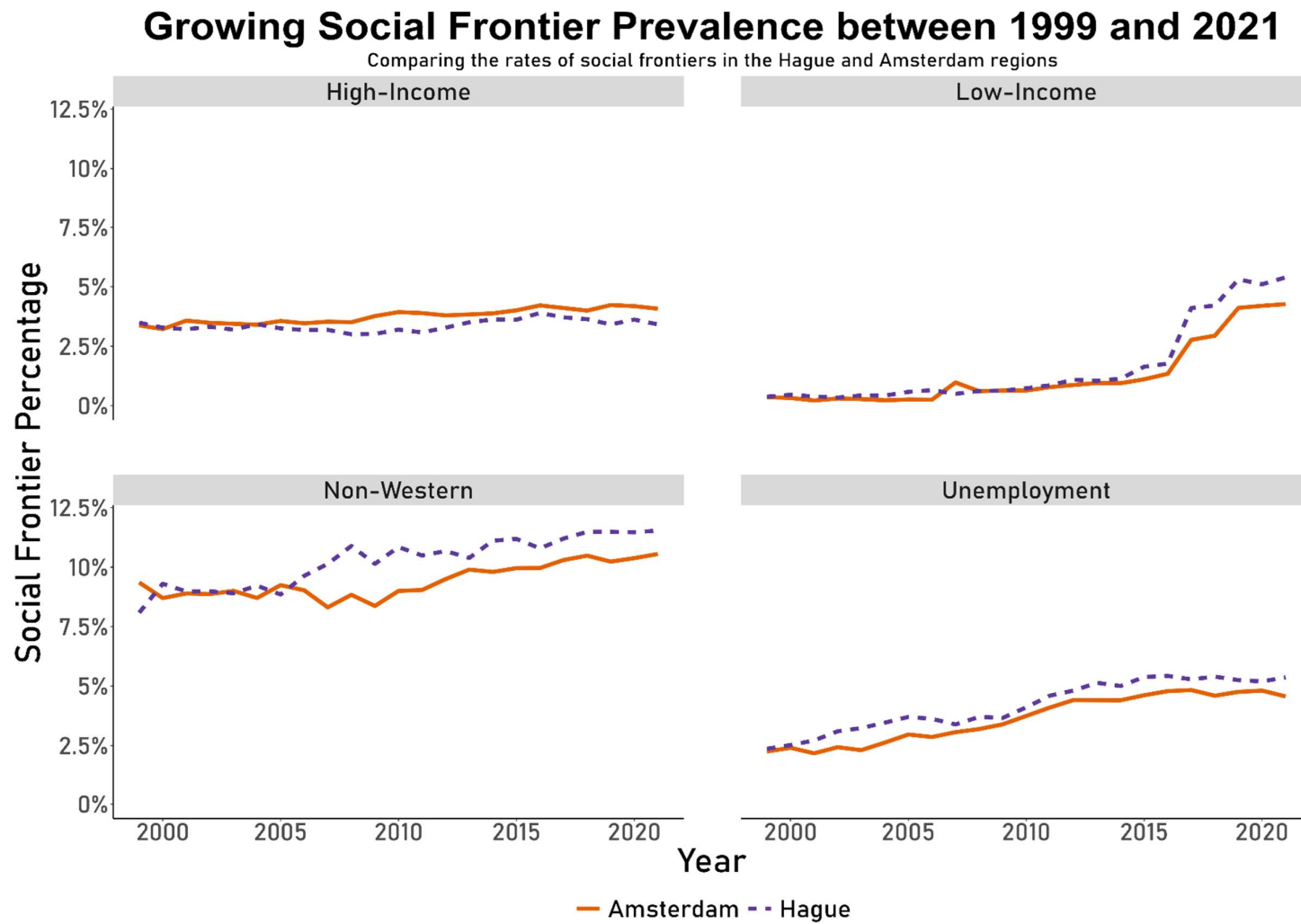
Exploring the strength of entrenched social frontiers



Source: CBS

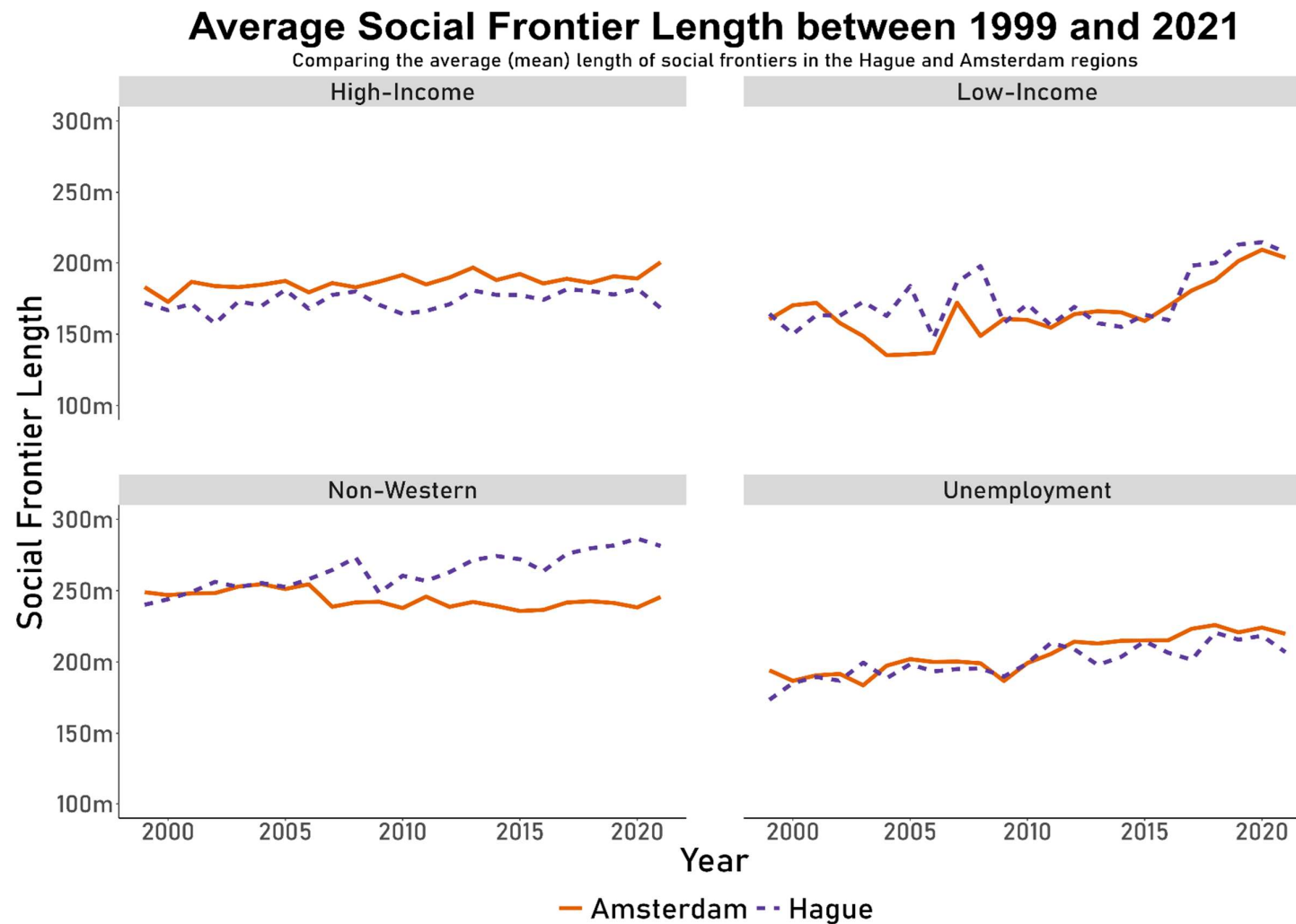
Figure 4.4: Entrenched frontier lifecycle.

Changes in social frontier prevalence between 1999 and 2021



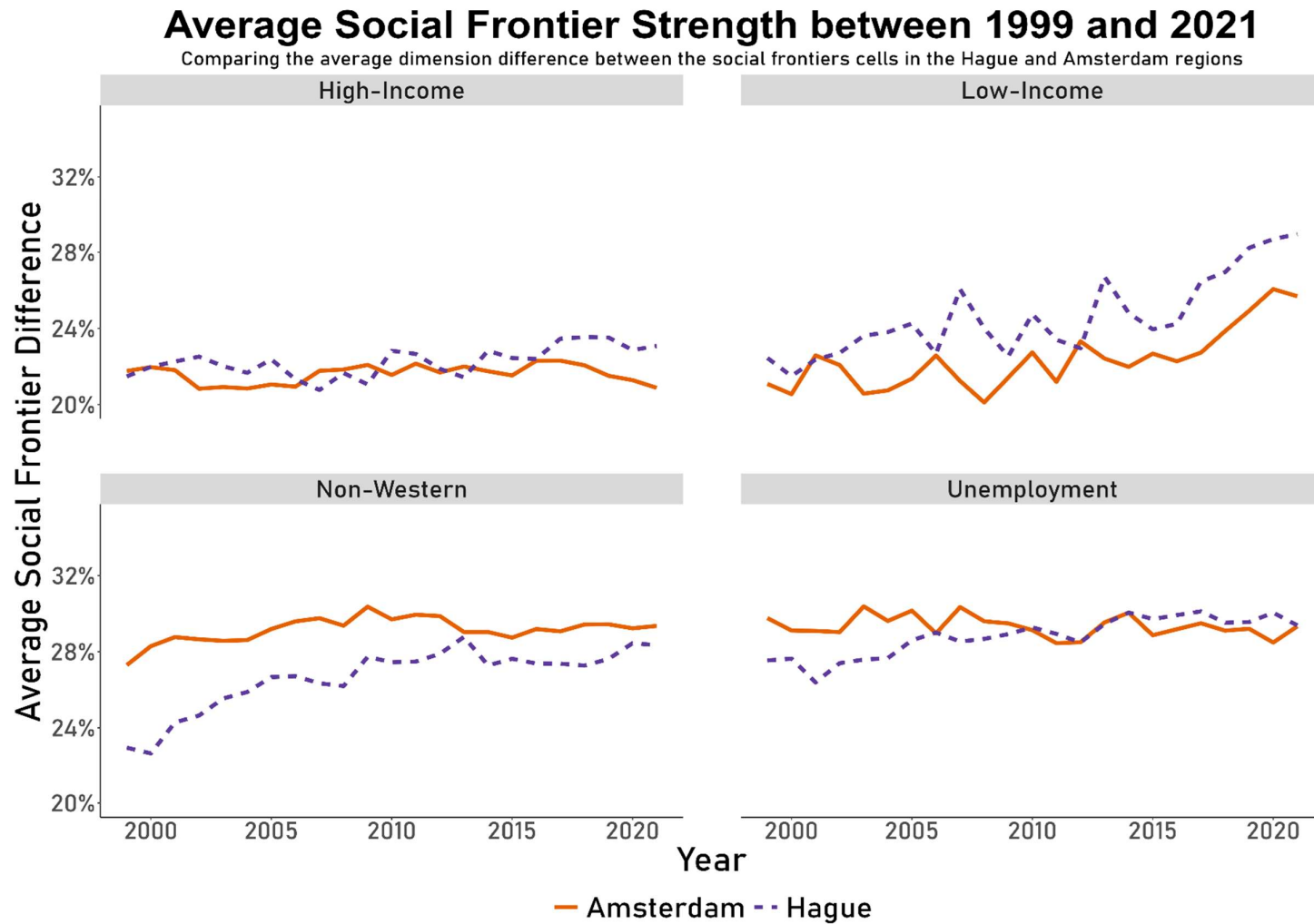
Source: CBS
Data from 1999 to 2021

Figure 4.5: Percentage of social frontiers in Amsterdam and the Hague.



Source: CBS
Data from 1999 to 2021

Figure 4.6: Average social frontier length in Amsterdam and the Hague.



Source: CBS
Data from 1999 to 2021

Figure 4.7: Average social frontier strength in Amsterdam and the Hague.

Most prevalent social frontier dimension

Segregation, when measured by social frontier prevalence, was largest for the non-Western dimension across all the prevalence measures for both Amsterdam and the Hague. This pattern replicates the trends identified in Dutch segregation research (Sleutjes, Ooijevaar and de Valk, 2019; Boterman, Musterd and Manting, 2021; Haandrikman *et al.*, 2023) - with Boterman, Musterd and Manting (2021) noting that “segregation based on income is relatively moderate but segregation on the basis of migration background and educational attainment level is relatively high” (p. 481).

Figure 4.5 shows that non-Western frontiers were identified at least twice as frequently as all the other frontier dimensions. This pattern can potentially be explained by the fact that the location of non-Westerners is often dependent on a combination of factors including economic variables such as housing availability and potentially homogeneous housing preferences (Zorlu and Mulder, 2008). Income and unemployment frontiers, on the other hand, are more likely to appear as a result of income inequality or limited housing availability. Additionally, the large welfare state and support means that the “income differences in the Netherlands are relatively moderate” (Sleutjes, Ooijevaar and de Valk, 2019, p. 361) meaning there are fewer extremes within communities, thus, reducing economic inequality.

The most prevalent economic dimension on average was unemployment - with Sleutjes, Ooijevaar and de Valk (2019) noting that “there is increasing polarisation between groups based on their employment status” (p. 361), especially when compared to the other economic dimensions. The frontiers had a similar average strength suggesting that they have a similar requirement to be defined as a social frontier, as the non-Western frontiers, but occurred less frequently and were on average shorter. The high-income frontiers had the smallest average differences (strength), likely because the high-income dimension was the smallest population group. Hence, there was less range in the high-income populations between the grid cells and smaller proportion extremes than that of the country of birth. However, the frontiers were the second most common in 1999 and in the majority of years were more common than the low-income frontiers until around 2016 in the Hague and 2020 in Amsterdam. Haandrikman *et al.* (2023) found that high-income segregation was often stronger than that of people at risk of poverty, but tended to “take the form of segregated pockets” (p. 20). These pockets are likely what I have identified with the social frontiers with high-income individuals densely concentrated in certain neighbourhoods (Haandrikman *et al.*, 2023, p. 20). In terms of the length, a similar trend is visible in that the high-income frontiers were consistently longer for the majority of the years than the low-income frontiers but not the unemployment frontiers. The low-income social frontier prevalence results were the most volatile across all the measures. For the first 15 years they were the least common and the

shortest, however after 2015 their prevalence increased sharply making them the second most prevalent in the Hague.

Changing social frontier prevalence in Amsterdam and the Hague between 1999 and 2021

In both regions the percentage of social frontiers, with the exception of the high-income frontiers in the Hague, was higher in 2021 than in 1999 (visible in Figure 4.5), indicating that the regions are becoming increasingly segregated when segregation is evaluated using social frontiers. Across the low-income, non-Western, and unemployment frontiers we can see a visible increase in frontier percentage from around 2015, with these dimensions likely linked to each other and thus the frontiers potentially overlapping.

The non-Western increase contrasts with Chapter 3's findings, where the percentage of Lee and Mitchell (2013) country of birth social frontiers decreased in the city of Amsterdam. The difference between the findings of Chapter 3 and this chapter could be due to the larger region and small grid cell studied: the outskirts of the region might display higher levels of segregation. Additionally, the use of the smaller grid cell (100m by 100m) likely results in a greater number of frontiers being identified when using the finer neighbourhood scale as this is likely to increase segregation scores with a greater level of variance identified (Fowler, 2015, p. 3; Bailey, van Gent and Musterd, 2017, p. 14). In this chapter, the grid cells had an average population of 68.5 in 2010, compared to the 785 population average in Chapter 3 meaning that smaller differences appear starker. However, it is of use to consider how the results vary when we identify frontiers using different grid cells, as noted in Chapter 2 and by both Dean *et al.* (2019) and Olnier *et al.* (2024). Studying frontiers using this smaller spatial scale will show the impact of localised policies and reduces the chance of smoothing local spatial discontinuities that could occur when aggregating to a larger spatial scale. In Chapter 3 I proposed gentrification as a potential explanation for non-Western households being more likely to live in the outskirts as they were pushed out of the city centre, hence the increasing rate of social frontiers especially surrounding the city centre. In terms of the year-on-year pattern, the prevalence increased for the majority of years. This is especially true for Amsterdam from 2010 onwards, whilst for the Hague the greatest increases occurred in the first ten years of the study. The increasing social frontier prevalence contradicts the decreasing index of dissimilarity. Segregation indexes showing contrasting trends is not uncommon, in fact it highlights the multidimensional nature of segregation (Massey and Denton, 1988; Feitosa *et al.*, 2007). In this case the index of dissimilarity is aspatial - with the spatial arrangement of the grid cells irrelevant - whilst the arrangement of a region is crucial for determining the prevalence of social frontiers, making it more likely to replicate the pattern of the concentration segregation measure. The decreasing index of dissimilarity tells us that non-Westerners are becoming more evenly distributed across the region, however, our

growing social frontier prevalence suggests that there are still areas which have a high proportion of non-Westerners resulting in the steep transitions. The indexes measure segregation at different scales with the index of dissimilarity an aggregated segregation measure, whilst the social frontier prevalence focuses on the presence of agglomerated communities with the spatial element crucial (Reardon and O'Sullivan, 2004; Feitosa *et al.*, 2007). The Hague does have a higher D than Amsterdam which corresponds with the higher social frontier prevalence, indicating that the Hague is more segregated than Amsterdam on multiple levels.

The percentage of unemployment social frontiers increased gradually between 1999 and 2021. The largest increase between two continuous years was a growth of 0.48 percentage points between 2011 and 2012 in the Hague, highlighting the slow year-on-year growth (Table G.1). In contrast to the non-Western frontiers, the increase in unemployment social frontiers is consistent with the growing index of dissimilarity. The consistency in the measures of segregation supports the concept that unemployment segregation is growing in the Netherlands despite the percentage of unemployed individuals actually decreasing. This indicates that unemployed individuals are becoming more spatially concentrated and less evenly distributed across the regions even though the number of unemployed individuals is falling. Nevertheless, unemployment segregation was still relatively low when considering both the social frontier percentage and the index of dissimilarity. In Figure 4.5 we can see that both regions experienced a peak in their frontier percentage in 2016/2017 before a slight decline, which is also seen in the index of dissimilarity and corresponds with the stabilising unemployment rates identified by the OECD following 2016 (OECD, 2023, p. 39). Sleutjes, Ooijevaar and de Valk (2019, p. 361) theorised that the increase in employment status segregation could be based on the shift in the focus of the welfare system and changing patterns for the labour market.

The most stable and consistent frontiers were the high-income frontiers which changed very little across all the prevalence measures over the 20 years, a pattern which is clearly visible in Figures 4.5, 4.6, and 4.7. The frontier percentage was consistently small (around 3%) in both regions, with only small increases of around 0.2, visible in Figure 4.5 in 2009 and 2015. These small changes could be a result of a bounce back following the global financial crisis which resulted in a greater divide between the rich and middle-class, or the average income of the centre of Amsterdam being greater than the national average (van Gent and Boterman, 2019). Over 20 years the frontiers barely changed in their percentage, length, and strength which is also consistent with the index of dissimilarity, indicating that the level of high-income segregation has remained relatively consistent between 1999 and 2021 in both Amsterdam and the Hague. The stability of the frontier prevalence is likely a result of the clusters of high-income households tending to remain consistent across the years (San Millán, Cottineau-Mugadza and Van Ham, 2025).

For the first 15 years the rate of low-income social frontiers remained very consistent, before a sudden increase in 2015 which is visible in Figure 4.5. The trends for the other prevalence measures were slightly less consistent; multiple peaks and troughs are visible in both Figures 4.6 and 4.7 but ultimately, they did have a higher average strength and length in 2021 compared to 1999. Despite the peaks, the sharp increase in 2015 was also replicated in the average strength and length, with strength about 6 percentage points higher and length about 50m longer on average in both regions (Figures 4.6 and 4.7). These combined with the increase in social frontier percentage show the growing importance and presence of low-income frontiers within the regions. In addition, the increasing social frontier prevalence tallies up with the increasing index of dissimilarity. In Amsterdam there is a possibility that this increase is linked to the construction of new social housing on former industrial land and the housing crisis which occurred following the 2008 financial crash with the disposable income of residents in the centre of Amsterdam notably higher than the outer regions (van Gent and Botterman, 2019). I suspect that the reduction in social housing in the city centre of Amsterdam will have forced the displacement of lower-income households to the surrounding municipalities (Hochstenbach and Musterd, 2021) resulting in the income levels in the centre becoming more homogeneous. The steep increase we see from 2015 onwards in Figure 4.5, could be low-income frontiers developing in the surrounding areas with newer social frontier housing developments arising and lower-income households moving out of the city (Hochstenbach and Musterd, 2017; van Gent and Botterman, 2019). Hochstenbach and Musterd (2017) suggested that the gentrification of the city centre has meant that poverty has shifted towards the suburbs and there is a suggestion that this is also true with social frontiers. However, without frontier visualisations or concrete evidence I can only hypothesise what has caused the increase in social frontiers.

Comparing frontier prevalence in Amsterdam and the Hague

The regions' spatial arrangements, political orientation, and history of migration (Hoekstra, 2015; Hoekstra, 2018) all play a key role when we interpret the difference between the social frontier prevalence of the regions. It can particularly explain the pattern for the prevalence of non-Western frontiers - in 2021 around 11.53% of the boundaries in the Hague were identified as social frontiers whilst it was 10.55% for Amsterdam, the evidence of higher segregation in the Hague is consistent with Dutch segregation literature (Bolt, Hooimeijer and Van Kempen, 2002). This pattern was the reversal of the 1999 findings when Amsterdam had a higher percentage of frontiers. The difference could be explained by the political climate - the Hague's politics is often quite polarised (Hoekstra, 2015) with 14 different political parties within the municipal council including those from both ends of the political spectrum. Whereas Amsterdam has tended to be viewed as more left-wing, with the council mostly made up of Labour and Green party members (Hoekstra, 2015). This

difference could explain why non-Westerners seem to congregate less in Amsterdam, with fewer areas of high concentration.

The location of the economic frontiers are likely to be influenced by the varying socio-spatial structure of the two regions. The Hague is considered to be naturally split in two with a richer seaside, whilst the inland is often more “working-class neighbourhoods” (Kloosterman and Priemus, 2001, p. 170; Meijers *et al.*, 2014; Musterd, Hochstenbach and Boterman, 2020). Whereas Amsterdam has found its city centre to be increasingly gentrified, with lower-income households often moving to the outskirts (Bailey, van Gent and Musterd, 2017, p. 10; Musterd, Hochstenbach and Boterman, 2020, p. 10). The regions were very similarly divided based on low-income frontiers; we can see in Figure 4.5 a steep increase in the percentage of frontiers in the Hague from 2015 with a slightly smaller increase visible for Amsterdam. This resulted in the percentage of low-income frontiers in the Hague being 1 percentage point higher than Amsterdam in 2021 - as with the non-Western frontiers a relatively minimal difference. However, the Hague’s frontiers were also longer and stronger on average by 2021, a trend which was relatively consistent across the 20 years. Although the difference is minute, the higher prevalence in the Hague corresponds with the index of dissimilarity scores and the literature, suggesting that the Hague was the poorest of the Dutch metropolitan regions with a high share of low-income households (Musterd, Hochstenbach and Boterman, 2020, p. 8). Unsurprisingly, the unemployment frontiers exhibit a similar pattern to the low-income frontiers, with the Hague having a slightly higher percentage of unemployment frontiers (0.8). However, Amsterdam had longer average frontiers - by approximately 13m - in 2021. These slight differences in the prevalence are likely to counteract each other (the regions have very similar average differences) resulting in the regions experiencing a similar level of unemployment frontier segregation. This partly corresponds with the index of dissimilarity, which is only slightly higher in the Hague than Amsterdam in 2021. Haandrikman *et al.* (2023, p. 20) noted that lower-income neighbourhoods in Amsterdam are often a result of there being few higher-income households, rather than areas existing which are densely populated with lower-income households.

The trend for the high-income frontiers resembled what we expected based on the literature - Haandrikman *et al.* (2023) found that segregation patterns for high-income individuals were particularly evident in Amsterdam - with Amsterdam having a slightly higher percentage of frontiers than the Hague, albeit only by 0.8 (visible in Table G.1), however, they were also longer on average by around 30m in 2021. We can see in Figures 4.5, 4.6, and 4.7 that the frontier prevalence is fairly consistent, however, the percentage of social frontiers in Amsterdam grew at a slightly faster rate than the Hague based on their similar percentage in 1999. Income gaps have been found to be increasing in the Netherlands, especially in Amsterdam, which is often due to the availability of housing stock being more limited in the larger cities (Modai-Snir and van Ham, 2020).

Conclusion

Summary of findings

Segregation literature has long explored the dynamic and evolving nature of residential segregation (Sampson and Sharkey, 2008; Bailey, 2012; Kauppinen and van Ham, 2019). Research papers have explored how gentrification can intensify segregation by forcing lower-income households out of affordable neighbourhoods (van Gent and Hochstenbach, 2020), creating distinct divisions (Hochstenbach, 2017, pp. 187–188). Papers have also investigated the influence of additional factors impacting population sorting and segregation such as migration, economic factors, structural changes in the neighbourhoods, along with ageing and life course (Bailey, 2012, p. 719). Despite the evidence that these dynamic processes impact segregation, there has been no exploration in social frontier literature as to how these processes may be related to the development and evolution of social frontiers.

This chapter's main contribution to the social frontier literature is the empirical exploration of frontier magnitude over its lifecycle. Previous social frontier literature has theoretically explored their lifecycle (Dean *et al.*, 2019; Iyer and Pryce, 2024) and discussed how the impact of the frontier may change depending on its lifecycle stage (Olnier *et al.*, 2024; Iyer and Pryce, 2024). I found that frontiers evolved similarly to that proposed by Iyer and Pryce (2024), highlighting that social frontiers are not a static concept but a dynamic spatial phenomenon. Prior to their first identification, the average difference between the contiguous cells increased; the strength then stabilised and plateaued before a sharp decrease in the difference prior to the frontier disappearing. This was consistent across both methods, dimensions, and durations (visible in Figures 4.3, I.1, and I.2). I also graphed entrenched frontiers, identifying a very similar pattern to that proposed in the theorised lifecycle in Figure 4.1, once the frontiers became established their strength was relatively consistent. This discussion has the potential to shape future frontier studies through the study of frontier effects at each stage of their lifecycle along with an exploration into the conditions that lead to some frontiers becoming entrenched and others diminishing. Studies could investigate whether stronger frontiers that exist for a brief period of time have a lesser impact than longer lasting, weaker frontiers as residents are exposed over a greater period of time. The study of frontier dynamics brings with it a potential discussion about why social frontiers evolve in such a way and what processes are in operation which impact how they evolve.

The comparison of multiple frontier dimensions has showed that the lifecycle dynamics were relatively similar across the different dimensions, however, they did vary in prevalence over the years. The economic frontiers overall were less prevalent than the non-Western frontiers in both regions, potentially meaning they are of lesser consequence on society. This corresponds with the high-level of welfare support and existing segregation research in Amsterdam where income segregation was quite low, whilst segregation based on migration history was more prominent (Boterman,

Musterd and Manting, 2021). In terms of the economic dimensions, the unemployment frontiers were found to be most prevalent across all measures including a high re-emergence rate and a consistency in their strength and percentage. The income frontiers all occurred at a similar prevalence rate, but the high-income frontiers were more constant, meaning that their impact is likely more uniform. This consistency could be due to higher-income households tending to remain clustered in small zones which remain constant throughout the years (San Millán, Cottineau-Mugadza and van Ham, 2025). The abrupt change in the low-income frontier prevalence could result in the frontiers' impact on society suddenly increasing if they became more noticeable.

This is also the first chapter to compare frontier prevalence across regions. As I suspected, based on the similarity of the descriptive statistics for the regions and them having the same national context, I found that the prevalence results were very similar in Amsterdam and the Hague. However, I did find that the Hague had a higher percentage of frontiers in 2021, except for the high-income dimension, corresponding with the higher index of dissimilarity scores for the Hague. The Hague also saw higher levels of growth in the social frontier percentage compared to Amsterdam between 1999 and 2021: 3.44%, 5.31%, and 3.01% in the non-Western, low-income, and unemployment dimensions respectively. There can be multiple explanations for the difference in the regions. Firstly, the spatial profile of the regions. In Amsterdam the centre tends to have a higher share of middle-class households, with high levels of gentrification in the central neighbourhoods whilst the outskirts tend to have a larger proportion of low-income households (Bailey, van Gent and Musterd, 2017, p. 10; Musterd, Hochstenbach and Boterman, 2020, p. 10). Amsterdam has become a business centre and a crucial part of the global economy, partly due to the airport becoming a transport hub in Europe and the expansion of the financial services, both of which has seen the city centre becoming increasingly affluent (Musterd, Hochstenbach and Boterman, 2020). The Hague has evolved differently, the political and bureaucratic centre of the Netherlands has an increasingly international workforce, but the region is the reverse of Amsterdam with richer suburban areas and a poorer city centre (Musterd, Hochstenbach and Boterman, 2020). The city had a marked divide in income levels running through the middle of the region with evidence of urban renewal and rising incomes along the income divide (Bailey, van Gent and Musterd, 2017, p. 10). Additionally, the regions differ in geographical size, Amsterdam has a larger land area - 165.5km² - compared to the Hague's 82.45km² size (CBS, 2025g). This difference in size obviously impacts the number of boundaries in the area, the Hague has around 7,000 fewer boundaries than Amsterdam. Furthermore, the Hague is more densely populated, likely resulting in more concentrated areas developing, potentially explaining its slightly higher social frontier prevalence. For all the dimensions in Amsterdam and for three of four in the Hague, the social frontier percentage was higher in 2021 compared to 1999. In addition, for the majority of the dimensions the average length and strength of the

frontiers increased, indicating that social frontier segregation is growing in the Netherlands with the regions becoming more divided.

Limitations

This chapter is not without its limitations, the lack of frontier visualisations means we cannot identify any areas within the regions which were particularly prone to social frontiers, so any explanation as to why frontiers rose in prevalence was mostly theoretical. This highlights the need for a combination of statistics and visualisations in social frontier studies to aid a discussion about the processes which increase social frontiers. If I had included this, I potentially could have explained the difference in the results compared to Chapter 3, assessing whether there are areas in the outskirts which have a high level of non-Western frontiers or whether it is due to the inclusion of smaller grid cells. In this case, the inclusion of visualisations was difficult due to the small grid cells (100m by 100m) which when applied to the larger region made visualisations unclear. However, this could have been addressed by focusing on specific parts of the region, discussing the social frontier percentage within each borough, or using a larger grid cell (e.g. 500m by 500m). For example, Spierenburg, van Cranenburgh and Cats (2024) focused their study on the larger scale of municipalities, allowing them to produce visualisations exploring the patterns that their statistics identified.

The Lee and Mitchell (2013) method requires the use of a binary definition for the income dimensions which requires a threshold to be set as to when an individual is defined as low- or high-income. To address this, future studies could consider using a household income variable which may provide a more thorough description of income within an area, or adapting the Lee and Mitchell (2013) method for a continuous variable, thus enabling the use of the median income in a grid cell.

Suggestions for future research

I hope that this chapter will lead to greater discussion surrounding social frontiers, especially the dynamics of the frontiers and the effect of the lifecycle stage on their impact. The study of dynamics opens up the opportunity to explore whether the frontier impact varies depending on the stage of the lifecycle. When exploring contested boundaries, Legewie and Schaeffer's (2016) contested boundary study found that the stage when the territory is not fully established is when the boundary has the highest crime rate. Iyer and Pryce (2024) suggested this contested boundary could be a lifecycle stage of the social frontier. However, currently no empirical social frontier study exists which has researched whether the frontier effect varies depending how long the frontier has existed or how long residents have been exposed to it. Studies could compare crime rates near a frontier at its early stage to rates when the frontier has stabilised. Alternatively, the socioeconomic outcomes of siblings whose exposure length to frontiers during their childhood varied could be compared. In addition, household mobility could be explored by assessing whether the strength of the frontiers impacts upon whether individuals have an increased

likelihood of moving away from the social frontier. By including the dynamics prevalence measures from this chapter in an effects paper, social frontier literature can be extended to study whether we identify the same effects as in Legewie and Schaeffer (2016) or if embedded, established frontiers have a longer lasting impact.

Additionally, if the inclusion of multiple dimensions within social frontier papers becomes more common, then additional comparisons could occur through the inclusion of education level, which has large consequences for social integration levels in the Netherlands (Tselios *et al.*, 2015; Boterman, Musterd and Manting, 2021) or labour market position frontiers (Boterman, Musterd and Manting, 2021). Alternatively, a study could split the country of birth dimensions into separate groups of migrants. Migrants arrive in a country for multiple reasons or differ in their housing preferences on arrival to the Netherlands and thus are likely to have varying experiences of segregation (Zorlu and Mulder, 2008; Sleutjes, de Valk and Ooijevaar, 2018; Spierenburg, van Cranenburgh and Cats, 2024). By including multiple dimensions, I open up the opportunity to consider the intersection between these frontiers, which I explore in Chapter 5. It is often how these dimensions coincide to separate communities and the combination of similarities and differences in sociocultural factors which can impact relationships between communities (Iyer and Pryce, 2024). I anticipate that this study opens up a large amount of opportunity for future social frontier research which could have significant consequences for the experiences of residents living in close proximity to contrasting neighbourhoods.

Chapter 5 Exploring the Intersectional Nature of Social Frontiers in the Netherlands

Identifying and studying the patterns of co-occurring steep spatial divides in Amsterdam and the Hague regions.

Abstract

Introduction: Iyer and Pryce (2024) argued that to understand how communities exist alongside each other we should consider how combinations of similarity and difference can affect intergroup relations. So far in social frontier research this has been overlooked with frontiers only identified on a single dimension. The study of *intersectional* social frontiers, those which are “co-occurring across multiple socio-demographic domains” (Iyer and Pryce, 2024, p. 784), opens up an opportunity to understand how these divides coincide or occur concurrently in space.

Objectives: This chapter engages in a conceptual discussion about the definition of intersectional social frontiers, employing two definitions. The first definition is focused on the overlap of single dimension frontiers, described as ‘coterminous’ intersectional frontiers. The second definition is where the frontiers are ‘multidimensional’, with frontiers identified where there are steep differences in the proportion of a combined variable, for example, non-Western low-income individuals. These definitions are employed to identify intersectional frontiers in Amsterdam and the Hague, discussing how their prevalence changes over time.

Methodology: The chapter applies the Lee and Mitchell (2013) frontier identification method to Amsterdam and the Hague regions. Using the single dimension frontiers identified in Chapter 4, the chapter identifies the boundaries where the individual social frontiers overlap for the coterminous intersectional frontiers. For the multidimensional frontiers, the Lee and Mitchell (2013) method is applied on the combined variables: low-income non-Western, high-income non-Western, and unemployed non-Western.

Results: The high-income non-Western frontiers occurred least frequently whilst the low-income non-Western and unemployment non-Western frontiers had a similar prevalence level; this was consistent across both definitions. In terms of the difference in the results of the definitions, the multidimensional social frontiers often showed a sharper increase over time whilst the coterminous frontiers had a smoother pattern. What is of particular interest is the contrast in the results of the multidimensional and single dimension frontiers.

Conclusion: The chapter makes a substantial contribution to social frontier literature both conceptually and empirically through its study of intersectional social frontiers.

The contrasting pattern of the intersectional and single dimension frontiers indicates that specific groups of individuals are becoming more segregated, even if when studied as a single dimension segregation rates appear to be declining. This finding has potential policy implications. If attempting to reduce social frontiers the intersectionality theory would suggest that a targeted strategy is needed to aid the integration of specific individuals (Crenshaw, 1991).

Introduction

Despite the growing interest in social frontiers and contested boundaries (Legewie and Schaeffer, 2016; Dean *et al.*, 2019; Iyer and Pryce, 2024; Piekut *et al.*, 2024; Spierenburg *et al.*, 2025; Wickes *et al.*, 2025), social frontier papers have tended to only consider individual dimension frontiers, failing to identify the places where communities are divided by multiple factors simultaneously. Iyer and Pryce (2024) defined the boundaries where the communities are divided by overlapping frontier dimensions as 'intersectional social frontiers', describing them as "co-occurring across multiple socio-demographic domains" (p. 784).

This chapter aims to instigate a new avenue of research on the existence, prevalence, and impact of intersectional social frontiers by developing a clear definition of intersectional social frontiers, along with identifying and discussing their presence in the Netherlands. It does this by building upon the sociological theory of intersectionality, where individuals' characteristics overlap or intersect resulting in them experiencing discrimination or privilege due to their unique experiences. In terms of intersectional social frontiers, being divided on multiple dimensions or overlapping frontiers can reinforce feelings of isolation and further reduce the possibility for positive contact (Allport, 1954; Beale, 1979, cited in King, 1988). Intersectional frontiers explore how the dimensions are connected and influence each other, and how this influences the experiences and relationships of those living in different parts of society (Hill Collins and Bilge, 2020). The intersectional nature of divisions can shape multiple aspects of communities, including interaction levels, power dynamics, and territorial attitudes.

Due to the lack of empirical intersectional social frontier research, the frontier identification process, prevalence, and their effect on communities is completely unexplored. This, alongside the limited discussion about the definition of intersectional social frontiers, means that they are almost a completely new aspect of the literature. This absence is despite Boterman, Musterd and Manting (2021) highlighting that to describe the level of spatial inequality in an area, we should consider a combination of characteristics. If segregation is measured only on a single dimension, then any description is unlikely to allow a fully comprehensive understanding of the level of divides in an area. Thus, there is large scope for extending future intersectional social frontier research by considering their definition; identification; and the difference in their prevalence in different regions, dimensions and in comparison to individual dimension frontiers.

This chapter focuses on the definition and identification of intersectional social frontiers, but to provide motivation for the study we discuss their potential impact based on existing and theoretical literature. There exist large gaps in the literature, especially in relation to how intersectional frontiers impact on society and how this differs from the effect of single dimension frontiers. Existing social frontier literature

can help to theorise the potential effects of intersectional frontiers highlighting the importance of future research. As discussed in Chapters 1 and 3, social frontiers and steep boundaries have found to be associated with increased crime rates (Legewie and Schaeffer, 2016; Dean *et al.*, 2019), along with having asymmetric effects on geographical mobility (Olnier *et al.*, 2024). In addition, a negative association has been identified between mental health and social frontier intensity (Lee, Pryce and Ramos, 2025), as well as the theorised negative consequences on wellbeing and life outcomes, such as educational attainment and the labour market position, of those living close to social frontiers (Dean *et al.*, 2019; Iyer and Pryce, 2024). Sharp boundaries have this impact due to them limiting positive contact, potentially alienating individuals from the core of their community, and increasing feelings of relative deprivation, tribalism, and loyalty, all of which have the ability to reduce social mobility (Dean *et al.*, 2019; Iyer and Pryce, 2024).

We are unsure how the effect of intersectional frontiers will vary from that of single dimension frontiers. We can look to sociology literature, where the double jeopardy hypothesis (Beale, 1979, cited in King, 1988) and Crenshaw's theory of intersectionality (Crenshaw, 1991) both indicate that the impact of being divided on two dimensions is greater than one. However, this is hypothetical for social frontiers. Intersectional frontiers can result in communities feeling isolated from each other due to being separated on multiple levels from their neighbours. Additionally, the greater the number of dimensions which divide communities the less they have in common, reducing the likelihood of relationships forming. This has the potential to have material impacts on life outcomes, for example, reducing the potential for weak ties which are important for employment opportunities (Granovetter 1973; Rajkumar *et al.*, 2022). However, a community which has one characteristic in common but differs in another (for example, they are similarly economically deprived but differ in their ethnic composition), may experience increased competition over potentially limited resources which is likely to cause conflict within the neighbourhood (Dean *et al.*, 2019).

The unexplored nature of intersectional social frontiers makes them an area of vast interest as their impact upon communities is unknown. However, prior to a discussion about their impact I first must focus on their identification and definition. To do this I utilise my work in Chapter 4, where I studied multiple single dimension frontiers, to identify intersectional social frontiers in Amsterdam and the Hague, offering a framework for defining and identifying intersectional social frontiers. In this chapter, as with Chapter 4, the frontiers are based on economic and country of birth characteristics (unemployment, high/low-income, and non-Western), with this chapter focusing on their overlap between 1999 and 2021. I study the intersectional frontiers over an extended period of time allowing an exploration into their evolution over time, as well as the differences between the regions and dimensions, and the rate of single dimension and intersectional social frontiers.

This chapter has five main aims. The first is conceptual, aiming to form a concrete definition for intersectional social frontiers. I introduce two definitions based on intersectionality and social frontier literature. The following four aims are all based on interpreting the empirical intersectional social frontier prevalence findings in Amsterdam and the Hague, answering the questions:

- (1) Did the rate of intersectional social frontiers increase between 1999 and 2021 in Amsterdam and the Hague?
- (2) What type of intersectional social frontier was the most prevalent in Amsterdam and the Hague?
- (3) Were intersectional social frontiers more prevalent in the Hague or Amsterdam?
- (4) Were intersectional social frontiers more prevalent than individual social frontiers?

This chapter will be structured as follows. Firstly, I will discuss the literature that helps to clarify the definition of intersectional social frontiers, namely: the intersectionality literature from sociology and the theoretical social frontier literature. I will then use the literature to highlight the importance of future research and discuss how the frontier dimensions are likely to interact with each other. The second section of the chapter describes the methodology and data used including a discussion about the two different definitions and how they are applied to identify the intersectional social frontiers. The results section is separated into the four empirical questions exploring the changes over time, regions, and dimensions, along with a brief discussion about the difference in the results depending on the intersectional frontier definition studied. The chapter ends with a conclusion: summarising the findings, limitations, and recommendations for future studies.

Literature

Defining intersectional social frontiers

Intersectionality is defined in sociology as “the interconnected nature of social categorizations such as race, class, and gender, regarded as creating overlapping and interdependent systems of discrimination or disadvantage” (‘Intersectionality’, 2023). Crenshaw (1991) focused on the intersection between race and gender; however, this principle is applicable in multiple fields. In this study I am interested in intersectional social frontiers, a new concept to empirical social frontier research and thus an area of interest to researchers (Dean *et al.*, 2019; Iyer and Pryce, 2024). Iyer and Pryce (2024) defined intersectional social frontiers as “co-occurring [boundaries], across multiple socio-demographic domains” (p. 784) visible where there are sharp divides between communities in multiple dimensions simultaneously. There is a need for researchers to consider not only individual characteristics but also the combination of these and how they interact with each other. Traditional

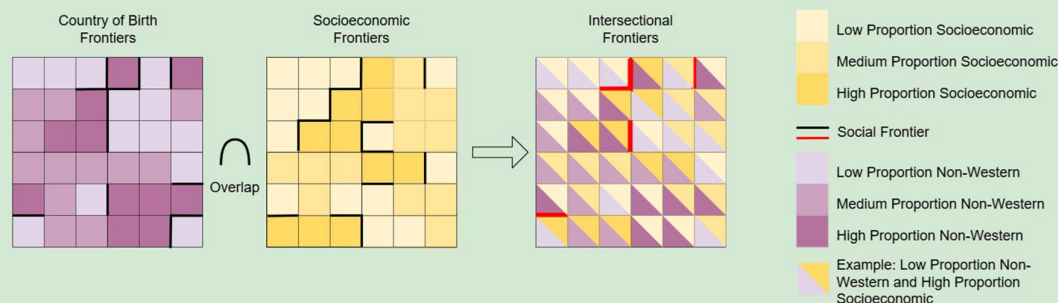
intersectionality literature tends to focus on how aspects of individuals' identities overlap to result in individuals experiencing discrimination or gaining an advantage. However, intersectional social frontiers are not a direct example of intersectionality, instead I use the definition of intersectionality more as a guide for defining intersectional social frontiers. Where intersectional frontiers exist there are sharp spatial divides which are present across multiple dimensions, compounding difference rather than overlapping disadvantages. It is of interest to study what influence this compounded difference is likely to have on society, whether that be the intersection enlarging the effects of the single dimension frontiers - increasing feelings of relative deprivation, conflict over resources, or reducing opportunities for positive contact (Iyer and Pryce, 2024). Or whether, the likelihood of conflict over resources is diminished due to the communities differing in multiple dimensions and the communities are able to comfortably live alongside each other.

As this is the first social frontier chapter to identify and discuss intersectional social frontiers in depth, it is beneficial to introduce a definition along with identifying them. There are multiple ways of understanding and interpreting the concept of intersectional social frontiers. In this chapter I introduce two definitions which are linked to the Iyer and Pryce (2024) description and the sociology definition of intersectionality. I describe these two types of intersectional frontiers as 'coterminous' and 'multidimensional'. The coterminous frontiers are described as the overlap of two individual frontiers, whilst the multidimensional frontiers explore the divides between communities who have contrasting proportions of multiple groups simultaneously. The importance and potential significance of both intersectional social frontiers is unexplored, as is their presence within regions. However, there has been increased importance placed upon multidimensional segregation, with the importance of studying the interconnected nature of multiple dimensions in connecting or dividing communities highlighted in research (Boterman, Musterd and Manting, 2021; Iyer and Pryce, 2024). In Amsterdam, communities are increasingly divided on multiple dimensions, with ethnic segregation becoming more correlated with additional social or cultural factors, thus insinuating that to understand inequalities and community relations we need to include multiple dimensions (Boterman, Musterd and Manting, 2021). The study of multiple dimensions allows a better understanding of the relationship between the communities and the space they occupy (Boterman, Musterd and Manting, 2021).

The difference between the intersectional social frontier definitions is discussed below, and Figure 5.1 provides an insight by visualising the definitions. Coterminous frontiers are determined by the overlap of the single dimension frontiers, Figure 5.1 shows this whilst also demonstrating the potential issue of the inconsistency for the high or low proportions. The multidimensional frontiers are identified based on the combined variable as shown in Figure 5.1.

Intersectional Frontier Definitions

Coterminous intersectional frontiers (Overlapping)



Multidimensional intersectional frontiers (Combined)

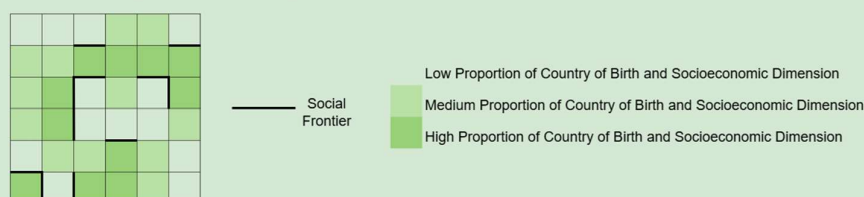


Figure 5.1: Intersectional social frontier diagrams.

Coterminous intersectional social frontiers

The first type of intersectional social frontiers, 'coterminous', occur where there exists an overlap between multiple single dimension social frontiers, whether that dimension be ethnic, economic, or cultural. The term coterminous, interchangeable with conterminous, is defined as the quality of "having the same or coincident boundaries" ('Coterminous', no date). These frontiers are the intersection of existing individual social frontiers, with the frontiers co-occurring in the communities as described by Iyer and Pryce (2024). As a result, communities are divided in multiple economic or racial factors simultaneously. An aspect of these intersectional frontiers is that the communities could be on the high-side for one dimension and low for the other. For example, they could have a high proportion of non-Westerners whilst having a low proportion of low-income individuals (this concept is visible in Figure 5.1). This potential inconsistency is likely to result in the individuals having a different experience to those who live near an intersectional frontier but have a high proportion of both non-Western and low-income individuals. This potentially introduces an aspect of ecological fallacy as we make assumptions about individuals

based on the inferences we have gathered from studying a group (Piantadosi, Byar and Green, 1988), thus, failing to consider the experiences of the individual.

Multidimensional intersectional social frontiers

The ecological fallacy issue is somewhat addressed by my second definition of intersectional social frontiers - described as 'multidimensional'. In this case, a multidimensional intersectional social frontier is identified if adjacent communities have a steep difference in the proportion of a combined variable, e.g. the proportion of low-income non-Westerners. Multidimensional frontiers are more closely related to the sociology definition of intersectionality, where individuals are part of multiple minority groups simultaneously which ultimately shapes their experiences (Crenshaw, 1989; Bauer *et al.*, 2021). If an individual is part of what we would define as two marginalised groups (e.g. they are part of a minority ethnic group whilst also being on a low-income), then if we only considered the experiences of one group, we might ignore the intersectional nature of their experiences (Crenshaw, 1989).

According to Crenshaw's (1989) concept of intersectionality, it is the coinciding oppression in race, gender or class systems which results in a connected system of discrimination. This shapes the experiences of those who experience multiple levels of oppression, especially as we are bypassing the experiences of individuals who experience a combination of discrimination, or a combination of frontiers (Crenshaw, 1991, p. 1278). Thus, these intersectional social frontiers capture the role that these divisions may play in power dynamics, interaction, and even access to facilities or shared spaces within communities (Crenshaw, 1991; Weldon, 2008). In the case of the multidimensional intersectional social frontiers we are not focused on compounded disadvantage, rather compounded differences. However, there is a likelihood that some of those who live close to multidimensional intersectional frontiers may experience disadvantage which may be missed if only single dimension frontiers are identified.

Intersectional social frontier effects

As these intersectional frontiers, and even social frontiers in general, are relatively unexplored in research, increased research on the definition and identification of intersectional frontiers poses a crucial question: what impact do these frontiers have on society cohesion and the experiences of communities? This chapter's role is to measure and define intersectional social frontiers rather than identify their impact, however, to provide motivation for the advancement of intersectional frontier literature I discuss their potential impacts. Along with considering how their impact may vary from that of single dimension frontiers and how the impact of the two types of intersectional frontiers may vary. Particularly of interest is to note how these intersectional frontiers are likely to result in individuals feeling isolated, with the compounded difference between the groups potentially resulting in them feeling to

an even greater extent that the frontiers are impenetrable (van der Laan Bouma-Doff, 2007; Dean *et al.*, 2019; Iyer and Pryce, 2024).

To postulate the effect of multidimensional intersectional social frontiers, I consider sociological research on the double jeopardy hypothesis (Beale, 1979, cited in King, 1988) and Crenshaw's intersectionality (1991). These indicate that the unification of dimensions causes a greater issue as individuals potentially face double discrimination and are often bypassed by policies which are targeted at one group or the other rather than the intersection (Crenshaw, 1991; Berdahl and Moore, 2006). Targeted interventions are often either aimed at a specific attribute of individuals or communities, meaning that those who are part of two stigmatised groups may be missed. Additionally, there is little thought about how any obstacles faced may be a result of the combination of the multiple dimensions (Crenshaw, 1991), a division in one dimension alone may have very little impact but the combination is of great impact. In relation to social frontiers, the double obstacles faced are likely combined with the increased deprivation and higher levels of conflict theorised by Iyer and Pryce (2024), all of which can have consequences on mental and physical health (Greitemeyer and Sagioglou, 2016; Smith and Huo, 2014; Iyer and Pryce, 2024). Communities that are on the high side of the intersectional frontier (e.g. a neighbourhood with a high proportion of non-Westerners and low-income individuals), are likely to experience the same disadvantage as those close to an individual frontier but potentially two-fold, along with any issues that come with being part of two stigmatised groups (Berdahl and Moore, 2006). Equally, the greater the alignment of the attributes, for example when the areas with a greater percentage of non-Westerners also have a higher percentage of low-income individuals, the stronger the 'faultline' becomes, which in the case of management studies was suggested to be linked to increased risk of conflict (Lau and Murnighan, 1998). Hence, individuals living near multidimensional intersectional frontiers could experience greater consequences than those who are separated only by a single dimension frontier (Beale, 1979, cited in King, 1988, p. 46).

Coterminous intersectional frontiers could also have these effects, but they have an additional complication which could alter their effect. The area could have a high proportion of one dimension whilst having a low proportion in the other, as is visible in Figure 5.1. For example, an area A could have a high proportion of non-Westerners and a low proportion of low-income individuals, whilst the bordering area B has a small proportion of non-Westerners but a high proportion of low-income individuals. The combination of similarity and differences may result in the frontier having a different impact to the multidimensional frontiers (Dean *et al.*, 2019; Iyer and Pryce, 2024). The communities still differ from each other, thus there is still the steep transition between them bringing with it the same potential consequences as single frontiers. However, these communities may not face the double-disadvantage of the multidimensional frontiers, in relation to the intersectionality theory.

Two communities which have differing economic deprivation levels and non-Western proportions - with a coterminous intersectional frontier between them - are likely to have little in common potentially making it more difficult for them to form relationships. Whilst two bordering communities with a similar economic deprivation level but a single country of birth frontier between them may form a stronger network based on their economic similarities with feelings of unity based on what they have in common in relation to their human capital (Vervoort, 2012). This could result in the intersectional frontier being of greater consequence than single frontiers (Iyer and Pryce, 2024). However, this is not guaranteed, if the country of birth frontier exists between two communities which are both highly economically deprived then this could lead to increased conflict and tension due to competition for potentially limited resources, with fewer easily accessible opportunities and resources (Lamont and Molnár, 2002; Iyer and Pryce, 2024). This competition for resources is likely to weaken the bonds between the native Dutch and non-Western individuals (Vervoort, 2012), meaning the similarity of the two groups actually results in the communities becoming further removed from each other. If the border is a coterminous intersectional frontier where one community is affluent, then there is less likely to be conflict over resources, meaning that relationships between the groups are potentially less likely to be fraught, improving relations. Additionally, Tselios *et al.* (2015) found that immigrants who live in “areas with high concentrations of immigrants and low levels of economic development” that are surrounded by more prosperous areas with a higher proportion of Dutch born individuals, are likely to show higher levels of social integration, defined by the level of interaction and societal networks formed with Dutch-born individuals, compared to immigrants surrounded by households with similar economic characteristics (p. 437).

As discussed above, intersectional social frontiers also have the potential to have positive impacts, despite the tendency to focus on the negative effects of social frontiers. The close contact between the contrasting groups can promote harmony (Dean *et al.*, 2019), along with encouraging the formation of cross-ethnic relationships (Zorlu and Mulder, 2008). Additionally, as Tselios *et al.* (2015) found, exposure to affluence can be beneficial for those living in economically disadvantaged areas. Having affluent neighbours can have a positive impact on life outcomes with a transmission of resources, whether that be physical access to shared spaces or through connections and networks made from richer neighbourhoods (Sykes and Kuyper, 2009; Toft and Ljunggren, 2016; Troost, van Ham and Manley, 2023). Exposure to affluence during childhood had a greater impact (positive) than exposure to poverty on educational attainment, with individuals gaining skills, access to resources, and having increased educational aspirations (Troost, van Ham and Manley, 2023). This effect could be linked to the school effect: if individuals attend the same school or share community spaces, this can encourage social mixing in a community. Additionally, it could be that the geographic structure of the neighbourhood encourages interaction. In my study I would define these communities as having a social frontier between them, and yet the literature

suggests that the close proximity provides an opportunity for lower-income individuals to form relationships with those from contrasting groups which is of great advantage to them.

I suspect that the effect of intersectional frontiers will depend on multiple aspects: the social, economic or ethnic dimension, the context of the region studied, the strength of the intersectional frontiers, the causal effect dimension studied, as well as the type of intersectional frontier (multidimensional and coterminous) identified. Their potential effect highlights the need for intersectional social frontier research as before we can investigate the effect of these frontiers we first need to be able to define and identify them.

Frontier dimensions

In this chapter I identify the intersectional social frontier on three different dimensions: unemployment and non-Western, low-income and non-Western, and high-income and non-Western. The individual frontier dimensions were identified and their prevalence discussed in Chapter 4. I found that with the exception of high-income frontiers in the Hague the social frontiers grew in prevalence between 1999 and 2021, with the non-Western frontiers the most common. Previous papers that have only identified single dimension frontiers have missed the inherent diversities which exist within these groups; the various subgroups within each dimension are likely to have a different experience of segregation, and thus the impact of the frontiers varies (Spierenburg, van Cranenburgh and Cats, 2024). For example, low-income non-Westerners are likely to have had a dissimilar experience to high-income non-Westerners, based on their ability to choose their housing location and access to resources - this is addressed in this chapter through the multidimensional frontiers. By studying frontiers only in a single dimension we are likely missing the impact that these dimensions have on each other and how they all contribute to where people live (Urria *et al.*, 2025) - this is addressed in part by the coterminous frontiers.

In relation to how we expect these dimensions to interact with each other, there is a large amount of literature exploring the relationship between them, some focusing on ethnicity and income (Hartog and Zorlu, 2009; Tselios *et al.*, 2015; van Elk, Jongen and Koot, 2019), and others employment and ethnicity (Bisschop, ter Weel and Zwetsloot, 2020; Thijssen, Coenders and Lancee, 2021). In the Netherlands, and in the majority of European countries, ethnic minorities are often geographically concentrated in lower-income areas with high levels of social deprivation (Bolt, van Kempen and van Weesep, 2009; Feijten and van Ham, 2009), with a large overlap between ethnic and economic segregation (Ziller and Spörlein, 2020). Areas which have a high concentration of “low-income households also exist in the Netherlands” these often coincide with areas which have a high proportion of non-Westerners, “generally comprising Turks and Moroccans (often sharing the same neighbourhoods), or Surinamese” (Bolt, van Kempen and van Weesep, 2009, p.

502). This relationship occurs as low-income households have fewer choices for their housing and non-Westerners have a lower average position in the labour market (Feijten and van Ham, 2009; Ziller and Spörlein, 2020), with the income per capita decreasing as the percentage of non-Westerners in a neighbourhood increases (Hartog and Zorlu, 2009). Unsurprisingly, a similar relationship was detected between unemployment and ethnicity, non-Western migrants were more likely than the Native Dutch to be unemployed or not be participating in the labour market (Heath, Rethon and Kilpi, 2008). This relationship was still visible even in second generation women with Moroccan, Pakistani, and Turkish ancestry (Heath, Rethon and Kilpi, 2008). A non-Western individual who is absent from the labour market, may have limited levels of social interaction with Dutch born individuals in comparison to those in employment, with reduced opportunities to form strong relationships (Tselios *et al.*, 2015). This is especially true if they live within areas which are ethnically homogeneous (Tselios *et al.*, 2015). The overlap between non-Western and low-income or unemployed individuals indicates that we would expect to identify multidimensional intersectional frontiers within both Amsterdam and the Hague. Additionally, I suspect that there will be crossover between the single frontier dimensions as they are likely to be occurring in similar locations, with this resulting in the identification of a large number of coterminous intersectional frontiers.

For the high-income and non-Western intersectional frontiers the results could vary. I have found less evidence about the relationship between these dimensions and suspect that there are fewer high-income non-Westerners, however, I am unsure how this will present itself in relation to frontiers. In the case of the coterminous intersectional frontiers, there is likely crossover between the single dimension high-income and non-Western frontiers as the single dimension high-income frontiers were mostly the second most prevalent economic frontiers identified in Chapter 4.

Data and Methodology

I identify the intersectional social frontiers by applying the Lee and Mitchell (2013) method, largely for the same reasons listed in Chapter 4. It identifies only the steepest of transitions as social frontiers and thus provides a substantial and reliable description of the level of frontiers in the regions. The coterminous intersectional frontiers are identified as the boundaries where the individual social frontiers overlap, where the individual frontiers were identified using the Lee and Mitchell (2013) method. These coterminous intersectional social frontiers are the boundaries where the non-Western frontiers intersect with either the unemployment, low- or high-income frontiers. The second type of intersectional social frontiers - multidimensional - are identified using a combined variable, either the proportion of high-income non-Westerners, low-income non-Westerners, or unemployed non-Westerners. The Lee and Mitchell (2013) method is applied to these combined population groups to identify the multidimensional intersectional social frontiers. These definitions identify two different types of intersectional social frontiers. The coterminous frontiers look at

the intersection between the frontiers, whilst multidimensional frontiers consider the intersection for the individuals, identifying grid cells where there are steep boundaries in the proportion of those who are part of multiple socioeconomic groups.

The Lee and Mitchell (2013) method is described in Chapter 1. The model has been used in multiple social frontier papers (Dean *et al.*, 2019; Křížková *et al.*, 2021; Olier *et al.*, 2024) and was implemented in Chapters 3 and 4 to identify individual country of birth and economic social frontiers in regions of Amsterdam and the Hague. As in Chapter 4, this chapter uses income, unemployment, and country of birth data geocoded to 100m by 100m grid cells, for reasons similar to that stated in Chapter 4, in the regions of Amsterdam and the Hague, with the regions based on the COROP definition used by Statistics Netherlands (CBS, 2023c).²³

Results

All of the single frontiers discussed in this chapter were identified and their prevalence discussed in Chapter 4.

Methodological comparison

This chapter identifies intersectional frontiers using two definitions: first, the boundaries where single dimension social frontiers overlap - described as 'coterminous frontiers' - and the second where the individuals overlap through their participation in multiple groups - described as 'multidimensional frontiers'.

In terms of the difference in the percentage, there were several key findings visible in Figure 5.2 and 5.3 (as well as Tables K.1 and K.2):

- The percentage of **multidimensional unemployment non-Western** frontiers was higher than the coterminous frontiers in both Amsterdam and the Hague.
- The percentage of **multidimensional low-income non-Western** frontiers was higher than the coterminous percentage in the majority of the years, especially after 2010.
- The percentage of **coterminous high-income non-Western** frontiers was always higher than the multidimensional frontiers in Amsterdam and the Hague.

In the case of the low-income non-Western frontiers, by 2021 the multidimensional frontiers occurred more than twice as frequently as the coterminous frontiers, with the multidimensional definition consistently higher after 2010. The unemployment non-Western frontiers showed a similar pattern with the multidimensional frontier percentage albeit approximately 2 percentage point higher. For the high-income non-

²³ The regions are described as Groot-Amsterdam and Agglomeratie's-Gravenhage - table E.1 in the appendix includes a list of the municipalities within each region.

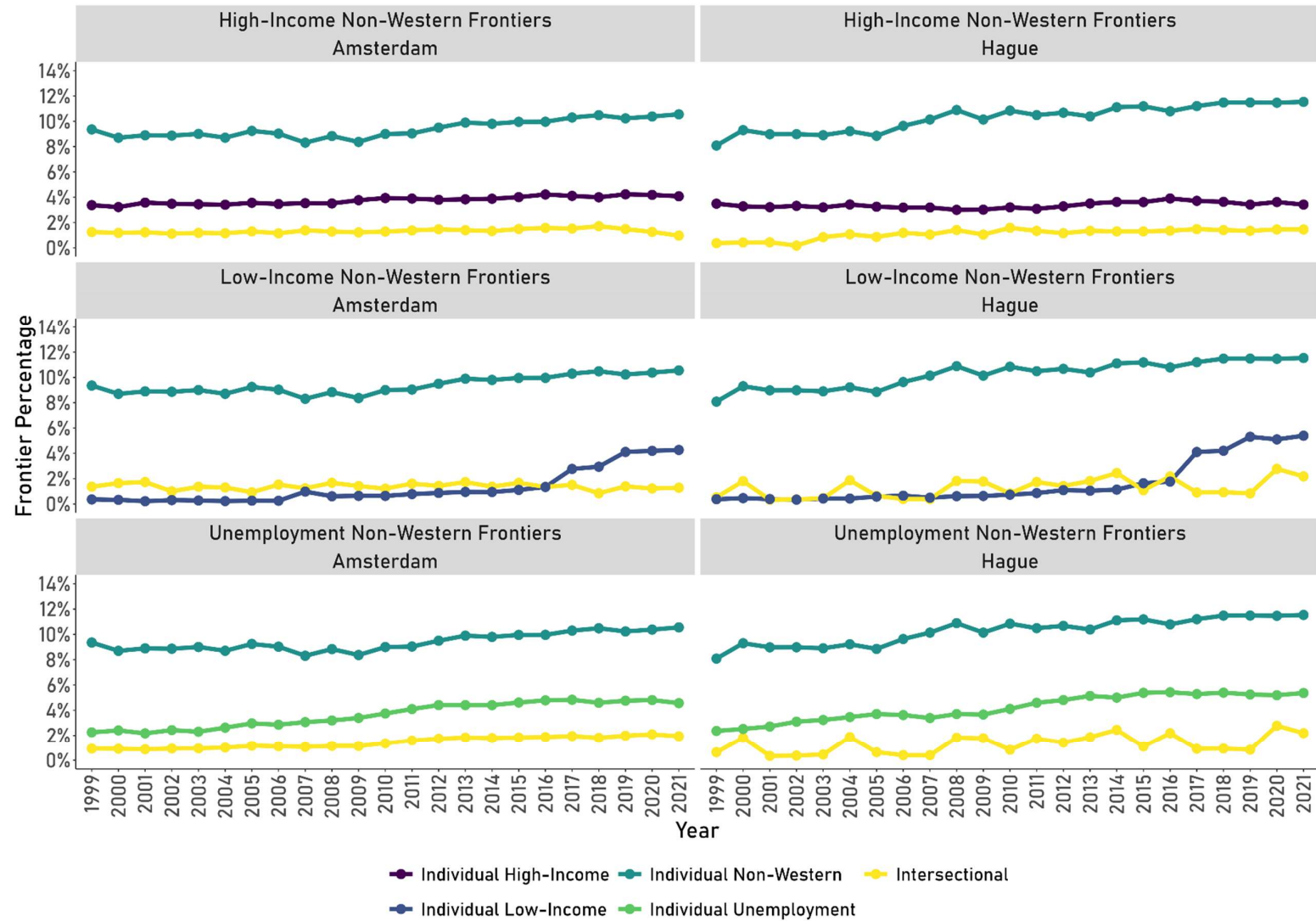
Western frontiers this pattern is reversed, with the coterminous intersectional frontiers significantly more common than the multidimensional frontiers.

This difference in the pattern can partly be explained by the descriptive statistics in Tables J.1 and J.2, with high-income non-Westerners a far smaller group in the population than the other groups. This could explain why the high-income non-Western multidimensional intersectional frontiers were less common and had the opposite pattern to the other dimensions. For the high-income non-Western and unemployment non-Western frontiers, the overall trends tended to be similar with the prevalence constant or showing a slight increase with both methods. However, the coterminous frontiers were more likely to have large peaks and troughs, whilst the multidimensional pattern was smoother.

In terms of the frontier location, the methods concurred the most on the unemployment non-Western frontiers - 28.9% of frontiers identified in Amsterdam were identified by both methods - albeit not a particularly high percentage indicating that the definitions do not often agree. The high-income non-Western frontiers rarely coincided, likely due to the multidimensional method identifying very few frontiers. When multidimensional frontiers were identified they co-occurred with the coterminous frontiers 12.5% of the time. The low-income non-Western frontiers overlapped in around 10% of cases for the methods. Overall, these are relatively low percentages, suggesting that the methods are identifying mostly distinct locations for their frontiers. This is unsurprising as based on the definitions we would suspect that the methods are identifying different types of frontiers. However, they did not vary largely in their prevalence results, for the low-income non-Western and unemployment non-Western, the multidimensional frontiers were slightly more common, whilst this switched for the high-income non-Western frontiers. Both methods had a tendency to display similar patterns, with differences in the results consistent across both regions. This suggests that even if the methods are identifying different frontiers, they are both diagnosing similar trends.

Comparing Coterminous Intersectional Social Frontier Percentages in the Hague and Amsterdam regions

Data from 1999 to 2021

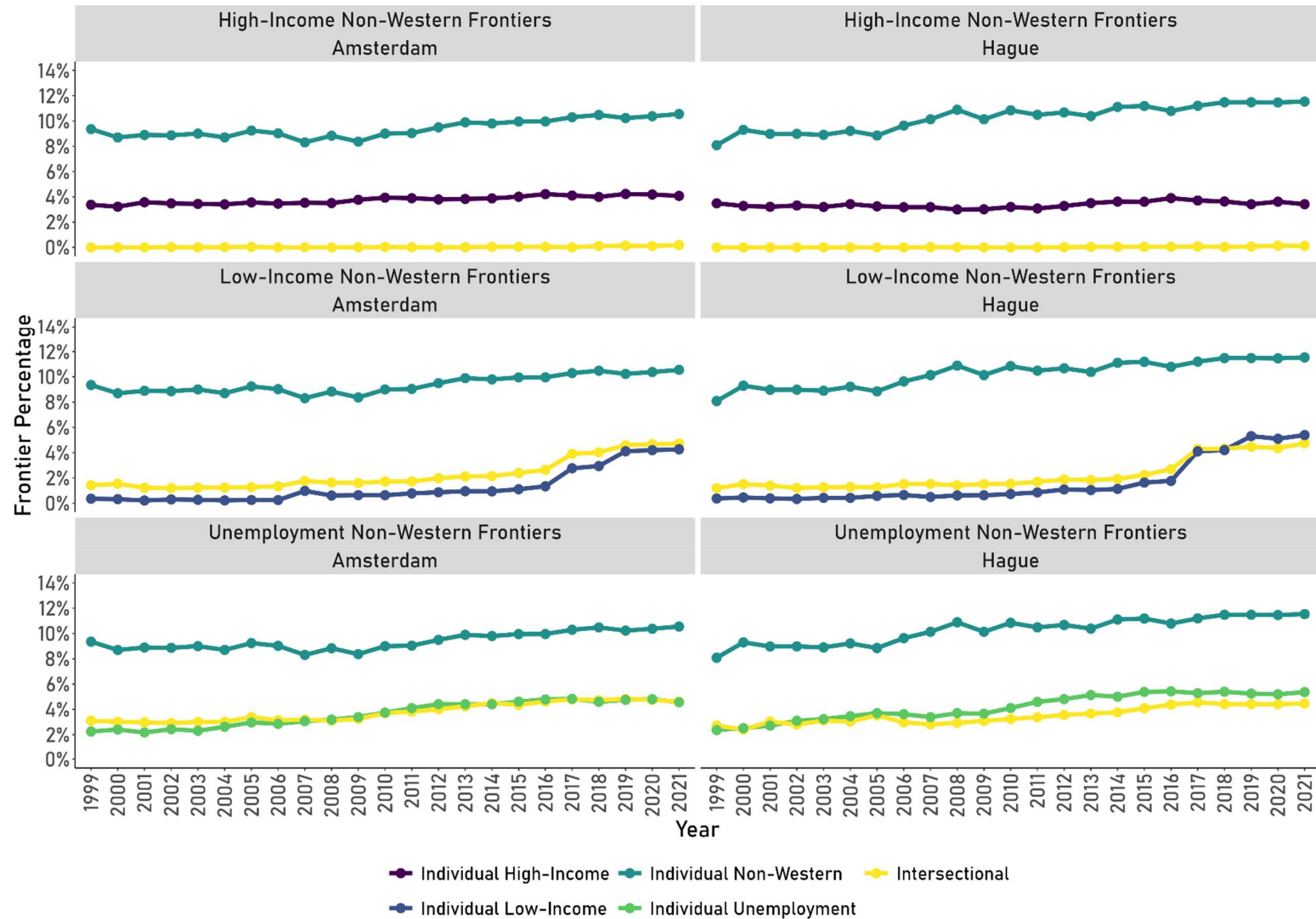


Source: CBS

Figure 5.2: Comparing the percentage of coterminous intersectional social frontiers and individual social frontiers.

Comparing Multidimensional Intersectional Social Frontier Percentages in the Hague and Amsterdam regions

Data from 1999 to 2021



Source: CBS

Figure 5.3: Comparing the percentage of multidimensional intersectional social frontiers and individual social frontiers.

Does the rate of intersectional social frontiers increase between 1999 and 2021 in Amsterdam and the Hague?

Coterminous

Figure 5.2 shows that the trends in the coterminous intersectional frontiers were relatively consistent in both Amsterdam and the Hague. In the Hague, all the frontier dimensions had a greater percentage of frontiers identified in 2021 compared to 1999, indicating increasing divisions in the region. The non-Western high-income frontiers showed a smooth increase with few discontinuities. However, there were more peaks and troughs within the unemployment non-Western and low-income non-Western intersectional frontiers patterns. Amsterdam was less consistent with the non-Western unemployed frontiers showing an increase, but the other two dimensions having a lower percentage of frontiers in 2021. The rate of frontiers in Amsterdam remained quite consistent between 1999 and 2021 with typically 1.5% of boundaries in the region identified as social frontiers. The non-Western unemployment frontiers were slightly higher in 2021 compared to 1999 with a relatively smooth pattern over the 20 years.

The coterminous frontiers were prone to re-appearing after initial decline with over 60% of frontiers re-emerging. This was particularly apparent in the Hague for the low-income non-Western and unemployment non-Western frontiers with approximately 79% of them re-emerging. This suggests a constant presence of intersectional frontiers within the regions and communities rarely fully merging.

Multidimensional

The pattern of all three multidimensional intersectional frontier dimensions in Amsterdam was incredibly similar to that of the Hague as visible in Figure 5.3. The unemployment non-Western and low-income non-Western frontiers steadily increased between 1999 and 2021. Whilst the trend of the unemployment non-Western frontiers was almost completely smooth before slightly flattening out as the frontiers became seemingly established towards 2021. By 2021 the low-income non-Western frontiers occurred twice as frequently as in 1999, whilst the percentage of unemployment non-Western frontiers was 1.5 times that of 1999. The high-income non-Western frontiers show a different story with little change between 1999 and 2021 and rates incredibly low. There were several years where no frontiers were identified. Across all the dimensions, the percentage of frontiers was higher in 2021 than 1999, indicating that divisions on multiple dimensions have increased in both regions.

I also note that the frontiers were particularly prone to re-emerging after their initial decline. In Chapter 4 I discussed the frontier lifecycle, noting that frontiers were prone to increase in their strength after an initial decline. The rate of re-emergence of intersectional frontiers suggest that they are likely to display a similar pattern, with

more than 60% of low-income non-Western and unemployment non-Western frontiers reappearing after initially declining within the 20-year period. The high-income non-Western frontiers had a slightly lower re-emergence rate, likely due to the fewer frontiers identified. The rate of re-emergence suggests that not only are frontiers increasing in prevalence, but they are also likely to either remain constant or have their strength resurge, indicating that divides were likely influential on the regions.

What dimension of intersectional social frontier is the most prevalent in Amsterdam and the Hague?

Coterminous

The pattern for the coterminous intersectional frontiers in Figure 5.2 is much less consistent and clear than the multidimensional intersectional frontiers visible in Figure 5.3. There is no clear hierarchy for the percentage of each dimensions in either Amsterdam or the Hague over the 20 years. For Amsterdam, there is a large amount of crossover and visible oscillations between the most common dimensions. The unemployment non-Western frontiers frequently switched from least to most prevalent, however, the actual percentage of frontiers remained relatively consistent. Additionally, the peaks of the low-income non-Western and unemployment non-Western frontiers tended to match. By 2021 the unemployment non-Western frontiers were the most common.

This trend is also replicated in the Hague. The low-income non-Western and unemployment non-Western frontiers replicated each other almost exactly in the percentage of frontiers, again highlighting the overlap between these frontier dimensions. The high-income non-Western frontiers are a lot more consistent with their growth but remain less prevalent. As with the coterminous frontiers in Amsterdam, the patterns are inconsistent with lots of fluctuations in the most common frontier dimension. However, ultimately in 2021 the unemployment non-Western and low-income non-Western frontiers were more prevalent than the high-income non-Western frontiers in the Hague.

Multidimensional

The multidimensional intersectional social frontiers in Amsterdam and the Hague have a very similar story. The high-income non-Western frontiers are consistently the least frequent, with the divides very rare in both regions. The Hague even had several years where no frontiers were identified, indicating a smooth spatial arrangement of high-income non-Westerners. In comparison the low-income non-Western and unemployment non-Western frontiers were much more common. In 1999 the unemployment non-Western frontiers occurred more frequently, especially in Amsterdam where there were nearly twice as many as the low-income non-Western frontiers. However, by 2021 the frontier prevalence had switched, with the low-income non-Western frontiers experiencing a steep increase resulting in them

being 0.1 percentage point more common than the unemployment non-Western. These patterns are replicated in the Hague, potentially suggesting that they could be visible across the country, with low-income non-Western and unemployment non-Western frontiers appearing at a similar prevalence rate as with the coterminous frontiers.

Discussion

We can see in Figures 5.2 and 5.3 that there is some consistency across the coterminous and multidimensional intersectional frontiers; the high-income non-Western frontiers in 2021 are the least common for both definitions and regions. The order of the dimension prevalence, especially with the small percentage of high-income non-Western frontiers, is potentially unsurprising. Literature has shown that non-Westerners who are not active within the job market are often more segregated (Tselios *et al.*, 2015), potentially explaining the higher frequency of the low-income non-Western and unemployment non-Western divisions. Additionally, income plays an influential role in deciding where individuals live. In terms of the high-income non-Western frontiers, there is a large overlap between high-income neighbourhoods and those with a large Dutch born population, with predominantly Dutch born neighbourhoods typically characterised by higher-incomes and house values (Zwiers, van Ham and Manley, 2018). This factor, along with the significantly smaller number of high-income non-Westerners compared to the other dimensions (Tables J.1 and J.2), potentially explains the fewer frontiers identified. Furthermore, the high-income non-Westerners are more likely to be integrated within Dutch born communities than low-income or unemployed non-Westerners: those who were active participants in the labour market had an increased likelihood of having ties with native Dutch (van der Laan Bouma-Doff, 2007, p. 1012) and the higher the income of non-Western individuals, the less segregated their networks tend to be (van de Water, 2024; Dutch News, 2024). This can be related to the double jeopardy hypothesis and intersectionality theories: high-income non-Western individuals seemingly face only one form of discrimination, which can make it easier for them to live in a predominantly Dutch born community as class plays a large role in defining how well migrants seem to 'fit' within communities (Bonjour and Jan Willem Duyvendak, 2018, p. 897). However, there is potential that they could experience resentment if they are seen to be doing 'well', with anti-migrant attitudes more likely to be present amongst those who have experienced relative deprivation which could be targeted particularly at those who are excelling (Albada, Hansen and Otten, 2021, p. 637).

Are intersectional social frontiers more prevalent in the Hague or Amsterdam?

Coterminous

The coterminous intersectional social frontiers had an ambiguous prevalence pattern and an inconsistent difference when comparing the results of the regions. The regions had a similar percentage of high-income non-Western frontiers, but

Amsterdam had a very slightly smoother pattern with the Hague showing more undulations. However, by 2021 the Hague had a very slightly higher rate of frontiers despite Amsterdam having a higher percentage in all but 5 of the years between 1999 and 2021 and twice as many boundaries being identified as frontiers in 1999 in Amsterdam compared to the Hague. For the low-income non-Western coterminous frontiers, the pattern was very inconsistent. The Hague's pattern was more varied, ultimately higher in 2021, but there were several peaks and troughs over the 20-year period. Amsterdam was also inconsistent with the regions frequently switching in which had a higher prevalence level over the years. Finally, the coterminous unemployment non-Western frontiers: Amsterdam's prevalence level was very consistent, showing a smooth gradual increase over the 20-year period. However, the Hague's fluctuated, often doubling in prevalence between consecutive years, resulting in the Hague having a higher rate than Amsterdam in 2021. The sharp changes in the prevalence rate visible for the Hague makes their presence likely to be experienced differently than in Amsterdam where the frontiers are established and have a constant presence in the neighbourhoods.

Multidimensional

The frontier prevalence patterns for Amsterdam and the Hague were incredibly similar for the multidimensional intersectional frontiers. Amsterdam had a very slightly higher percentage of high-income non-Western and unemployment non-Western frontiers in 2021. However, this difference was almost negligible: a difference of only 0.071 and 0.145 for high-income non-Western and unemployment non-Western frontiers respectively. The rate of low-income non-Western frontiers in Amsterdam and the Hague was almost identical. Amsterdam had a moderately higher percentage in 1999, but by 2021 the frontiers occurred slightly more in the Hague. The similarities between the percentages of the multidimensional intersectional frontiers suggest that the regions are alike in their level of divides and segregation.

Are intersectional social frontiers more prevalent than individual social frontiers?

The identification of intersectional social frontiers provides an opportunity to analyse whether the prevalence rate and pattern is similar to that of the single dimension frontiers. If the intersectional frontiers differ greatly then this could highlight several important features. Firstly, the importance of identifying and researching intersectional frontiers as their presence and effect likely varies from single frontiers. This could lead to a concept similar to ecological fallacy as we have made the assumption that we understand the effect of intersectional social frontiers based on the results of the individual. We cannot identify the effect of intersectional low-income non-Western social frontiers by only individually studying the effect of low-income and non-Western frontiers, we must consider the overlap. Secondly, it could impact any policies which aim to reduce social frontiers as they may be targeting a

wider group of individuals, thus bypassing the experience of those in multiple groups and divided by intersectional frontiers. The difference in the prevalence pattern may be of interest to future researchers, with specific groups of individuals potentially finding themselves more divided than ever and bypassed by intervention strategies.

Coterminous

In Figure 5.2 we can see that the coterminous intersectional frontiers show a slightly different pattern to the single dimension frontiers. The percentage of the intersectional frontiers is always smaller than the percentage of single dimension non-Western frontiers. Additionally, the percentage of high-income non-Western and unemployment non-Western frontier prevalence is also consistently smaller than the single dimension high-income and unemployment frontier percentage. The unemployment non-Western frontier pattern was less consistent than the single unemployment frontiers, particularly in the Hague: it was far more oscillating and had several periods of steeper increases whilst the single frontiers had a smoother increase. For the low-income non-Western frontiers, there is a slightly different pattern, with several years where the percentages of the intersectional and single dimension frontiers overlap, and the sharp increase in single dimension frontiers in 2015 not replicated in the intersectional frontiers. The contrast between the intersectional frontier prevalence and the individual low-income could suggest that although the intersectional divides occurred less frequently than individual divides, the level of intersectional divisions is inconsistent and occasionally increasing, potentially resulting in them having a different impact.

Multidimensional

For all three intersectional frontier dimensions their prevalence pattern, visible in Figure 5.3, closely replicated that of the single economic dimension in both Amsterdam and the Hague, almost exactly replicating their trajectories. The similarity in the patterns suggests that the evolution and trends of the boundaries closely resemble each other. In terms of the difference in the rate of frontiers, the multidimensional intersectional frontiers all occurred less frequently than the single dimension non-Western frontiers across all the dimensions, regions, and years. However, the difference between the intersectional prevalence and the economic dimensions is less consistent with the exception of the high-income non-Western frontiers which were consistently identified at a notably smaller rate than the single dimension high-income frontiers. This difference could in part be explained by the finding that high-income non-Westerners are less likely to have a segregated network (van de Water, 2024; Dutch News, 2024), indicating that those who crossover between the two dimensions may find themselves less divided than those who fall into only one, with the intersection being potentially beneficial for them.

The low-income non-Western and unemployment non-Western intersectional frontiers were more common than the individual low-income and unemployment

frontiers in several years. This is particularly the case for the low-income non-Western frontiers in Amsterdam, with the intersectional prevalence rate higher in every year, whilst for the Hague this was the case until 2018. The periods when the intersectional frontiers were more prevalent than the single dimension could highlight the importance of multidimensional segregation, with individuals who are part of multiple groups experiencing a greater level of division. In addition, Crenshaw's (1991) intersectionality theory suggests that these frontiers can be more difficult to address. Interventions targeted at specific individual frontiers may bypass and therefore be unable to address the intersectional social frontiers, as the policies are not specific enough to deal with the issue.

Conclusion

Summary of findings

This chapter makes several novel conceptual and empirical contributions to the social frontier literature. As the first intersectional social frontier study, it addresses an important conceptual question as to how we define intersectional social frontiers by introducing two interpretations. The first type, coterminous intersectional frontiers study the intersection and overlap of single dimension frontiers, whilst the second, multidimensional intersectional frontiers, focus on the intersection of the individuals, concentrating on those who are part of multiple groups. In terms of the difference in the frontiers identified by the methods, they did not often coincide, highlighting the distinction between the two definitions. However, several of the patterns identified were replicated across both the multidimensional and coterminous intersectional frontiers. In Table 5.1 I have summarised the key findings, noting the most important results for both methods.

The question of which definition to include in future research is one that researchers face. The definitions represent different structural systems and divides between neighbourhoods. The multidimensional frontiers are identified on an individual level focusing on the compounded differences, potentially identifying boundaries where individuals are facing or divided by multiple forms of discrimination or difference. This definition more closely resembles the sociology definition of intersectionality, as both consider how dimensions combine uniquely to impact individuals. However, it varies from the intersectionality framework, in that the multidimensional intersectional frontiers are focused not on overlapping disadvantages but rather on compounded difference and how this demonstrates itself in society. The coterminous intersectional frontiers are identified where single dimension frontiers overlap and intersect. This definition is more of an aggregate descriptor of intersectional social frontiers and the intersection of neighbourhoods. Both definitions offer a valid contribution to the literature and hold a role in social frontier research depending on the context of the study.

Table 5.1: Summary of intersectional frontier prevalence results.

	Coterminous	Multidimensional
Frontier prevalence and trend	Frontiers occurred quite infrequently. All the frontiers occurred more frequently in 2021 compared to 1999, but the change was minimal.	High-income non-Western frontiers showed very little change, whereas the other two dimensions increased in social frontier prevalence.
Dimension comparison	The unemployment non-Western frontiers were the most common in 2021 but the low-income non-Western and unemployment non-Western had very similar trends.	The high-income non-Western frontiers occurred very infrequently. The unemployment non-Western and low-income non-Western frontiers had very similar prevalence patterns.
Regional comparison	The frontiers occurred at a very similar rate in the Hague and Amsterdam, but the Hague had slightly more variation.	The frontiers occurred at a similar rate in the Hague and Amsterdam.
Difference between single and intersectional frontier prevalence	Mostly replicated the pattern of the single dimension frontiers. However, the low-income non-Western frontiers were more common than the low-income single dimension frontiers for an extended period of time.	There were several years when the low-income non-Western frontiers and unemployment non-Western frontiers were more common than the single dimension economic frontiers.

Through the empirical application of my definitions, the identification of intersectional social frontiers in Amsterdam and the Hague introduces an interesting discussion about the difference in the prevalence of the intersectional and single dimension frontiers. There are periods when the intersectional frontier prevalence is greater or differs in its trend to the single economic frontiers. This variation suggests that even when divides are decreasing on a single dimension, specific groups of individuals are becoming increasingly isolated and segregated. This may suggest that there is a need for more targeted policies when addressing the social frontiers, which is linked to Crenshaw's (1991) intersectionality research which found that when we target a strategy on a large group of people, this often disregards the needs of specific individuals within the large groups. It is often the individuals who are in multiple stigmatised groups, such as a non-Westerner who is also on a low-income, who are likely to experience these frontiers on a greater scale than an individual who is part of only one of the groups (Berdahl and Moore, 2006). The difference in the

prevalence is something that could also be analysed in future research, considering different dimensions and regions to see if these results are consistent.

Limitations

This chapter is not without its limitations and weaknesses in relation to the methods and the extent of the study. Firstly, I only considered the intersection between economic characteristics and country of birth. Future research could identify additional intersectional social frontiers to evaluate whether there is any consistency or differences in the results. Boterman, Musterd and Manting (2021) found the highest factors of segregation in Amsterdam were related to education level and migration background, suggesting that the intersection of these frontiers may be more prevalent and consequential. By identifying more intersectional social frontier dimensions, we will form a better picture of divisions in the Netherlands, deepening our understanding of segregation.

Secondly, although I implement two different approaches and definitions of intersectional social frontiers, I do not consider any alternative identification methods. Frontiers could be identified using clustering on multiple dimensions, with the difference in this definition and the results from this method potentially explored in future research. Finally, within the results section I could have included more frontier prevalence measures to capture the difference in the presence of the frontiers within the neighbourhood and how this alters their impact upon residents.

Suggestions for future research

This chapter has focused on the definition and measurement of intersectional social frontiers. However, as noted in the literature review, discussions on the measurement of intersectional social frontiers are motivated by their potential impact on life outcomes. Having provided a robust framework for measuring intersectional frontiers, this chapter facilitates the development of a new avenue of research exploring their impacts. Future studies could explore the effect of intersectional frontiers on socioeconomic factors, crime and conflict, and life course trajectories, including a discussion about whether this effect varies depending on the intersectional social frontier dimensions or definition. Additionally, for the coterminous intersectional social frontiers we can study whether there is a difference in the effect depending on whether the individual lives in the high or low side of the neighbourhood, along with the difference in the prevalence of these frontiers when studied individually.²⁴

In addition to studying frontier effects, future research could explore different intersectional frontier dimensions. Along with investigating whether the difference in the frontier prevalence compared to the single dimension frontiers is replicated

²⁴ Olnier *et al.* (2024) applied a similar study in their research on housing mobility – identifying an asymmetric effect depending on the side of the boundary.

across different regions or whether there are specific factors which result in certain groups becoming more divided. In this study I found Amsterdam and the Hague to be quite similar in their prevalence pattern, however, we know that the cities themselves are relatively similar (Tables J.1 and J.2) despite having a different spatial arrangement as discussed in Chapter 4. Future studies could assess whether this is consistent across the Netherlands, even in cities which vary in their descriptive statistics or in more rural areas, and, if so, it would be of use to understand what aspect of the city results in intersectional frontiers being particularly prevalent.

Chapter 6 The Impact of Living near Country of Birth Social Frontiers during Childhood on Socioeconomic Outcomes in Adulthood

A study which implements matching methods to investigate whether living near sharp boundaries between the ages of 11 and 15 has adverse effects on income and education in adulthood.

Abstract

Introduction: Despite the large body of research on neighbourhood effects that has emerged over the past half-century (Durlauf, 2004; Musterd *et al.*, 2012; Chetty, Hendren and Katz, 2016; Galster and Sharkey, 2017; Chyn and Katz, 2021), the impact of steep neighbourhood boundaries - social frontiers - remains largely unexplored. These boundaries have been theorised to have potentially far-reaching consequences on both the individual and community cohesion (Dean *et al.*, 2019; Iyer and Pryce, 2024). Iyer and Pryce (2024) have theorised that living near a social frontier during childhood may have negative consequences for outcomes in adulthood, limiting educational and employment opportunities.

Objectives: This study aims to investigate whether living in close proximity to a country of birth social frontier during childhood has an impact on income and education levels in adulthood.

Methodology: The chapter implements nearest neighbour matching with replacement. Treated individuals - those who lived near a social frontier between 11 and 15 - and control individuals - those who did not - are matched based on their propensity score which is calculated based on a large number of covariates including parental, household and individual variables, before applying a regression model to identify the treatment effect. Social frontiers are identified using the PAM clustering method to identify spatial discontinuities in relation to the proportion of non-Western residents.

Results: A small negative association was identified between living near social frontiers during childhood and both yearly income and years spent in education for all ages between 20 and 25. On average, individuals earned approximately €640 less per year and spent 0.06 fewer years in education if they lived near a social frontier during childhood, a relatively minor impact in reality. Sensitivity analysis suggested that we cannot interpret these results causally, however, an association is identified.

Conclusion: This chapter shows the importance and potential long-term impacts of steep neighbourhood boundaries on outcomes in adulthood, with close proximity to these boundaries in childhood negatively associated with life course outcomes.

Further research is needed to determine if this relationship is causal and consistent across different cities or countries.

Introduction

As noted throughout the thesis, the impact of steep neighbourhood boundaries - social frontiers - on social and life course outcomes is relatively unknown, despite Legewie (2018) suggesting that the effects of neighbourhood boundaries are sizeable and potentially greater than that of neighbourhood characteristics. To understand why these boundaries have such an impact, we look to both the hypothesised and empirically proven consequences of social frontiers: increased feelings of relative deprivation, limited positive contact, and exacerbated territorial attitudes and behavior (Iyer and Pryce, 2024), along with higher crime rates close to social frontiers or contested boundaries (Legewie and Schaeffer, 2016; Dean *et al.*, 2019). I focus particularly on these consequences' ability to have a strong impact on individuals who experience them, explaining why close proximity to social frontiers is likely to be of great consequence for the individual. Despite their potential impact, these boundaries, and specifically their effects, are almost entirely unresearched. To demonstrate the importance of these steep boundary transitions, this chapter examines their impact on socioeconomic outcomes, especially focusing on their long-term effects on children.

We know that neighbourhoods have an influence on shaping people's lives and outcomes (Galster *et al.*, 2008; van Ham and Manley, 2010; Brattbakk and Wessel, 2012; Musterd *et al.*, 2012). This is particularly true for young people due to the extended period of time they spend socialising, living, and being educated within their local area (Sykes and Kuyper, 2009). Through their neighbourhoods, individuals are able to form networks which can be incredibly valuable for them, whether this be through increased access to opportunities that improve their economic outcomes, or the increased social network improving individuals' mental health by providing additional support (Putnam, 2007, pp. 137–138). However, Putnam (2007) proposed that residents in ethnically diverse neighbourhoods are likely to 'hunker down', meaning that they are unlikely to participate in neighbourhood activities, due to diminished trust in both their own communities and those that differ from them, along with a reduction in social capital. These mechanisms are also likely to be in play near social frontiers. As despite communities living separately from each other, they often have higher exposure levels due to their close proximity meaning there is an increased opportunity for contact (Dean *et al.*, 2019; Iyer and Pryce, 2024), which could bring about similar effects to those identified in the neighbourhood effects literature. In addition, social frontiers, especially when forming or when zones are being contested, come with an increased risk of territorial conflict along with a lack of social control (Legewie and Schaeffer, 2016; Iyer and Pryce, 2024). These factors are explored in relation to their impact upon the individual within my literature section, discussing the relationship with social frontiers.

There are very social frontier studies which have explored their effect and there is especially an absence of literature which considers how they impact individuals who

live in close proximity to them during childhood. This chapter specifically focuses on the impact of living in close proximity to social frontiers on income and educational attainment in adulthood. It does this by matching individuals who grew up near social frontiers with their counterparts in non-frontier areas to identify the causal effect of social frontiers on individual socioeconomic outcomes. The model aims to distinguish whether individuals are adversely affected by living close to the frontier or if any difference identified is due to a combination of multiple individual or neighbourhood factors (Bucheli, 2022). It does this by attempting to simulate a randomised control trial, with the vastness of the data meaning I should be able to create a similar treatment and control group (Thoemmes and Kim, 2011).

Before I apply the matching methods, I discuss existing neighbourhood and social frontier literature. Within the literature section, I briefly discuss existing neighbourhood effects literature, particularly focusing on the studies that explore the effect on children. In the following two parts of the literature discussion, I firstly link the theorised and established social frontier impacts with children's socioeconomic outcomes. Before, as a result of the established relationship between economic and country of birth segregation, I discuss economic neighbourhood effects literature and how these effects might present themselves in relation to social frontiers. The subsequent section introduces the data used within the study, particularly focusing on the range of covariates and their importance. The methodology section discusses model choice, including minimising differences between the treatment and control groups. I then present my results which include sensitivity testing and an analysis of the variation in the frontier effect depending on exposure length. Finally, I conclude with a summary of the findings, limitations of the study, and suggestions for future research.

Literature

Exploring childhood exposure effects

As discussed in the introduction, neighbourhood effects are often magnified on children with them having a substantial impact on development (Sykes and Kuyper, 2009; Rowe, Zimmer-Gembeck and Hood, 2016; Wang, Choi and Shin, 2020). The 'geography of opportunity' means that the environment that individuals experience can have a significant impact on their development, where their "values, aspirations and preferences are all seen as influenced by the neighbourhood" (Musterd, Ostendorf and De Vos, 2003, p. 880). There is a large amount of social science research exploring the effect of children's neighbourhoods on outcomes in adulthood, including on education (Garner and Raudenbush, 1991; Sykes and Kuyper, 2009; Troost, van Ham and Manley, 2023), income (Musterd *et al.*, 2012), health (Vartanian and Houser, 2010; Halonen *et al.*, 2014), and family life course (Buyukkececi, 2023). These have often included a discussion about the mechanisms within neighbourhoods that result in them having an effect, for example, access to

resources, strength of networks, and creating zones of mutual trust that increase cohesive attitudes in neighbourhoods and reduce stigmatisation (Sykes and Kuyper, 2009; Musterd *et al.*, 2012).

All of these effects could plausibly be recognised as impacts of social frontiers, which would suggest a strong similarity between the impacts of the neighbourhood and social frontiers. Indeed, in the presence of social frontiers, these effects could be magnified as, not only are individuals experiencing neighbourhood effects, but they are also experiencing spillover effects from their neighbouring communities. This effect is almost completely unexplored in the literature, in part due to the aspatial nature of traditional segregation indexes meaning that segregation is measured individually for each neighbourhood, with the spatial arrangement of the neighbourhoods ignored (Tselios *et al.*, 2015). Hence, studies often fail to consider how neighbourhoods interact with each other and whether the characteristics of neighbouring neighbourhoods have an impact (Dekker, Rijnks and Mierau, 2021).

This chapter focuses on the impact of country of birth social frontiers. Ethnic enclaves have been found to be beneficial for residents, but typically only if individuals have a high level of human capital (Cutler, Glaeser and Vigdor, 2008). Musterd *et al.* (2012, p. 98) found that living in neighbourhoods which had a low share of non-Westerners was associated with an increase in earnings for non-Westerners, whilst there was minor impact upon those born in the Netherlands. Koopmans and Schaeffer (2016, p. 863) found that a higher level of ethnic diversity was linked with lower levels of trust between individuals, however, contact between non-Western individuals and those born in the Netherlands was at its lowest when neighbourhoods were mostly non-Western (Gijssberts and Dagevos, 2007, p. 827). This literature suggests mixed effects for those living in more ethnically homogeneous neighbourhoods, and it is important to note the potential difference in the intersectional effect. Ethnic and income segregation is clearly intertwined, potentially meaning that the country of birth frontiers could have a similar impact to economic segregation. However, individuals divided on ethnicity are likely to also deal with additional historical patterns of discrimination and obstacles, especially in the housing and labour market (Witteveen and Alba, 2018; Fang and van Liempt, 2021). In the remainder of the literature section, I will link the existing neighbourhood effects literature to the established and theorised effects of social frontiers and also consider economic effects due to the close relationship between economic and ethnic segregation.

Linking the established and theorised effects of social frontiers to socioeconomic outcomes

The existing literature has tended to focus on the effect of social frontiers on crime and conflict (Legewie and Schaeffer, 2016; Legewie, 2018; Dean *et al.*, 2019; Křížková *et al.*, 2021), mental health (Maguire, French and O'Reilly, 2016), household mobility (Olnér *et al.*, 2024), and neighbourhood relations (Piekut *et al.*,

2024). Theoretical literature has hypothesised that frontiers could result in higher territorial effects, as well as increased feelings of relative deprivation and isolation within communities (Dean *et al.*, 2019; Iyer and Pryce, 2024). These factors can be linked to differences in the socioeconomic outcomes of individuals and potentially explain any effect of living close to social frontiers that we identify. Growing up near contested and acrimonious areas has been found to have negative consequences on both educational achievement and behaviour, with young people potentially feeling unsafe both in their neighbourhood and at school (Bowen and Bowen, 1999; Damm and Dustmann, 2014). Additionally, children with mental health problems often have lower levels of educational attainment and are more likely to be unemployed in adulthood, likely linked to their reduced attendance at school (Veldman *et al.*, 2014; Veldman *et al.*, 2015).

Household mobility can be linked to Iyer and Pryce's (2024, p. 790) proposed neighbourhood allegiance effect, with social frontiers potentially reducing geographical mobility. This is due to residents having a higher "sense of allegiance to remain in the neighbourhood to defend the frontier and preserve ingroup cultural identity" (Iyer and Pryce, 2024, p. 790). Oliner *et al.* (2024, p. 635) found that individuals aged between 28 and 37 who lived on the high non-Western side of the frontiers were less likely to move away from the social frontier, whereas the reverse was true for those living on the high Dutch born side. If communities feel their territory or culture is being eroded, they may feel a greater obligation to stay within their neighbourhood to defend their territory, potentially reducing both their education and employment opportunities (Iyer and Pryce, 2024). The frontier's territorial nature and the potential limited positive contact often seem to create a vicious circle. Residents feel an increased need to defend their territory, likely resulting in any contact between groups being negative, and this perpetuates as they cannot break the system, with the negative contact meaning that individuals are unable to reduce prejudice or fear by building relationships. I discussed in Chapter 2 that it is often expected that non-Westerners will remove themselves from their own culture in order to become more assimilated with those born in the Netherlands (Ghorashi, 2014, p. 102). This pressure, along with a greater feeling of needing to defend their neighbourhood, can result in individuals feeling increasingly isolated and segregated when living between contrasting communities, especially as social frontiers can act as a barrier between them.

These barriers come with consequences: they can prevent networks forming, limit the visibility of role models, and reduce access to facilities, all of which have consequences on socioeconomic outcomes in adulthood (Musterd *et al.*, 2012). Social capital can be transmitted through the formation of cultural networks, enabling others to gain access to information and opportunities (van der Laan Bouma-Doff, 2007, p. 1003). This transmission has the potential to promote ambitious social attitudes, leading to individuals being more likely to attend university along with an increased work ethic (Sykes and Kuyper, 2009, p. 2431; Troost, van Ham and

Manley, 2023). van der Laan Bouma-Doff (2007) refers to the concept of the isolation thesis where non-Westerners having limited relations with those born in the Netherlands means that they retain their own culture, thus resulting in them having fewer opportunities both within the labour market and in higher education. Individuals with many 'weak ties' are likely to have multiple bridges to different aspects of communities, meaning that they have greater reach with more opportunities to engage in upwards social mobility (Granovetter, 1973).

In some cases, living in concentrated ethnic areas or amongst those who have similar income levels is advantageous for residents: closer ethnic ties may make it easier for networks to form, thus improving social capital (Putnam, 2007). Bezzo and Vanhuysse (2024) found that those who grew up in social housing in a deprived neighbourhood had more frequent interactions with their peers compared to those in a less deprived neighbourhood. This contact could highlight that relationships can be formed over the similarities, such as similar social and human capital, or ethnicity, which is ultimately beneficial for social mixing. When residents form these networks, they are likely to feel more supported within their communities and potentially gain access to job opportunities, both of which improve their life satisfaction (Arat and Bilgili, 2021). This is especially the case for immigrants newly arriving to the country (Bolt and van Kempen, 2010). If these communities are well-organised, such as the Turkish community in the Netherlands, then the increased level of community participation may result in increased social capital (van Tubergen and Volker, 2015, p. 535). Additionally, an individual living in a neighbourhood with a high proportion of non-Westerners surrounded by social frontiers may actually benefit from close proximity to their Dutch born neighbours. Tselios *et al.* (2015) found that immigrants who live in highly ethnically dense low-economic development areas show higher levels of social integration if they border more prosperous, majority Dutch born neighbourhoods. The close proximity that comes with the neighbourhood structure naturally means that there is a greater likelihood of interaction which has the potential to be beneficial for the residents.

In their qualitative study Piekut *et al.* (2024) found that social frontiers in Rotherham, UK, often overlapped with physical barriers, thus impacting the ability of individuals to cross boundaries. They also identified a scarcity of community spaces to aid interaction close to social frontiers (Piekut *et al.*, 2024). This was especially apparent for young people, with the lack of youth centres, such as freely accessible buildings for social gatherings, likely resulting in interethnic tension due to the reduced availability of neutral locations which allow the formation of relationships (Piekut *et al.*, 2024). We know that shared community spaces can help to bridge the gaps between contrasting groups (Putnam, 2007, p. 164). Without them, territorial attitudes are more likely to be present as is reduced contact, which could be of great consequence for the residents. In regards to physical barriers, Maguire, French and O'Reilly (2016) found that the likelihood of an individual taking antidepressants increased by 19% if they lived near a 'Peace Wall' in Northern Ireland, and that in

particular the likelihood of anxiety disorders was even higher (39%), indicating that the physical segregation barriers created a high stress environment. Amongst children, a negative association has been identified between social economic status and mental health especially in adolescents (Weinberg *et al.*, 2019).

Economic neighbourhood effects on children

We discussed in previous chapters how in the Netherlands, areas of high poverty are often those with a high ethnic minority population (Gijsberts and Dagevos, 2007; Bolt, van Kempen and van Weesep, 2009; Feijten and van Ham, 2009). Thus, the effects of living close to deprived areas are frequently replicated in areas with a high proportion of non-Westerners. Additionally, I suggested in Chapter 4 and identified in Chapter 5 that there exists an overlap between economic and country of birth frontiers in the Netherlands.

Nieuwenhuis and Hooimeijer (2016) reviewed neighbourhood effects literature finding that both living in a neighbourhood with high levels of poverty and a high proportion of migrants is negatively associated with education attainment. Additionally, living in a deprived neighbourhood has been found to be negatively associated with labour market performance (based on an individual getting and retaining a job) (van Ham and Manley, 2010). Children who moved before the age of 13 from a high-poverty to a less-deprived neighbourhood had a greater likelihood of attending college and higher incomes in adulthood (Chetty, Hendren and Katz, 2016). Each extra year that a child is exposed to a “one standard deviation (std. dev.) better county increases income in adulthood by 0.5%” (Chetty and Hendren, 2018, p. 1163). When interpreting the mechanisms that may result in economically deprived areas having negative consequences on the socioeconomic outcomes of individuals, this could be a result of numerous additional issues that individuals may face in life, including higher levels of stress, exposure to crime, unstable household finances, challenging living conditions, and a reduction in facilities in deprived areas (Bezzo and Vanhuyse, 2024). Without shared community spaces, residents may struggle to form relationships across differing communities (Putnam, 2007) and lack ‘role models’ who can provide an aspirational figure in terms of education or future work ambitions (Jencks and Mayer, 1990, p. 113; Musterd, Ostendorff and De Vos, 2003, p. 880). In terms of how these effects may present themselves in social frontiers, several studies have explored the impact of close proximity or spillover between neighbourhoods. Troost, van Ham and Manley (2023) found that living amongst affluent neighbourhoods had a greater impact on educational attainment than the level of poverty in an area. The impact of poverty was almost completely non-existent when parental education was included in the study (Troost, van Ham and Manley, 2023). This impact of the greater exposure to affluence means that “well-off and more highly educated neighbours can transfer their social and cultural capital through shared social networks” (Troost, van Ham and Manley, 2023, p. 2). Cultural and social capital can be transmitted between neighbourhoods if a strong network exists (Toft and Ljunggren, 2016; Troost, van Ham and Manley, 2023).

Ferguson's (2006) review of social capital literature found that in the majority of studies a positive association was found between having a strong network and children's outcomes. This association was suggested to be a result of enhanced access to resources and individuals with information potentially about job or education opportunities (Ferguson, 2006, p. 5; Franzen and Hangartner, 2006, p. 355).

Iyer and Pryce's (2024) theoretical study suggested that social frontiers can reinforce feelings of being unable to engage in upwards social mobility if the boundaries feel impenetrable and thus have negative effects on life course outcomes. Especially for young people, an awareness of relative deprivation and increased social comparison plays a large role in shaping their life experiences, influencing their mental health, quality of life, and how they interact with others (Bezzo and Vanhuysse, 2024). These feelings of deprivation can have consequences for the neighbourhood as a feeling of unjustness and unfairness is likely to result in increased aggression (Greitemeyer and Sagioglou, 2016), disrupting community cohesion. However, as suggested above, the close proximity of more affluent neighbours can provide the opportunity to access better resources and role models. López Turley (2002, p. 683) found that relative deprivation, defined as having a lower-income than the comparison group, was positively associated with several factors, including problem-solving skills, mental health, and behaviour especially amongst children. The more extreme the difference between the groups, the more the lower-income individuals benefited (López Turley, 2002, p. 684). This would suggest that social frontiers, as they are located at the extremes, could actually be sites of opportunity for individuals, with the assumption that the frontiers are peaceful zones. If the transition zones are antagonistic, then the negative effect of this could override the positive implications of the exposure to affluence.

It is the combination of these mechanisms that explain the social frontier effects that are potentially consequential, especially for children who are exposed to them for an extended period of time whilst growing up. If social frontiers do restrict the formation of networks, potential upwards social mobility, and result in higher levels of conflict, then I would suspect that living near them during childhood has a negative impact on both income and educational attainment in adulthood. However, as discussed, there are likely advantages to living near the social frontiers, with the close proximity potentially increasing access to resources and role models, which could result in living near a frontier during adolescence being beneficial. I investigate and discuss the frontier effect in the forthcoming sections of this chapter.

Data

This study includes 305,276 individuals born between 1996 and 1998 who lived in the Netherlands between the ages of 10 and 15 and 20 and 25. For this study I identified social frontiers using the 500m by 500m grid cells resulting in frontiers

being identified across over 96,000 grid cells. The choice of 500m by 500m grid cells was in part due to the computational complexity – identifying the frontiers using 100m by 100m grid cells would have vastly increased the processing time of the study. The dataset would have been divided into over 2.4 million grid cells. Based on the processing time in Table 3.3 I estimate this would take over 100 hours for each year, as well as being incredibly computationally challenging for the remote access memory. I was limited in this study by the capability of the computer processing power within the remote access which would have made this level of study unachievable. Additionally, frontiers are more likely to have a long-lasting sustained impact if they not only exist for an extended period of time, but are also of substantial length which results in them increasing divisions and restricting movement between communities. In Chapter 3 I highlighted the importance of the length of the frontier in relation to its impact. I noted that shorter frontiers are less likely to have a territorial impact and could be less restrictive for movement across the frontiers for individuals living close to them (Iyer and Pryce, 2024). The heightened territorial attitudes that come with longer frontiers are indirectly linked to reductions in social mobility, increased crime and conflict, and feelings of isolation all of which are discussed in the literature section of this chapter (van der Laan Bouma-Doff, 2007; Musterd *et al.*, 2012; Iyer and Pryce, 2024). Thus, longer frontiers (using the 500m by 500m grid cells the boundaries will be a minimum of 500m long) are more likely to be impactful.

I define 'living close' to social frontiers as to be living in a grid cell where at least one of the boundaries is a social frontier, thus living within 500m (the grid cell size) of a social frontier. Although spatial by nature as the frontiers are identified in relation to the spatial organisation of the grid cells, as discussed in Chapter 1, the data is aggregated. There is no exact calculation of the resident's household distance to the frontier (additionally, as discussed in the conclusion the location of the frontiers is based on the choice of grid cell). This poses an issue as it is likely that the effect is impacted by the households distance from the centre of the community and to the boundary as suggested by Iyer and Pryce (2024). Future studies could consider whether the exact distance to the social frontier impacts the frontier effect. Iyer and Pryce (2024) suggested that close proximity heightens awareness of difference meaning the exposure has a greater impact (p. 795). The individual-level nature of the Dutch data means that this could be achieved as exact household location is available in the data, however, the distance calculation, as discussed earlier, will rely largely on the choice of grid cell meaning it may not create any more meaningful results for the study.

The treatment effect was measured using the income and education level of the individuals between the ages of 20 and 25. These individuals were required to have a complete household history in the Netherlands between the ages of 10 and 15 as well as having income and education data available between the ages of 20 and 25. I specifically chose those born between 1996 and 1998 as it enabled income and education to be measured up to the age of 25 (in the latest year 2023). Income at 25

is more likely to be representative of future adult outcomes rather than income at 18 as typically individuals have finished their education and began on their career scale. This chapter focuses on the impact of social frontiers on children as I am interested in how our experiences whilst growing up shape our future. Additionally, children spend the majority of their time in their neighbourhood meaning the social frontiers may be more consequential than on adults. Children who experience segregation during their childhood are likely to experience it on a more extreme level than adults due to the similarities between residential and school segregation meaning that the divisions are visible across multiple environments (McArdle and Acevedo-Garcia, 2018). In this chapter, the treated individuals are those who lived near social frontiers in their early adolescence (between 11 and 15) for all 5 years, whilst control individuals did not live near any social frontiers during any of the 5 years.

The strength and success of matching/stratification methods relies heavily on being able to confidently state that there are no unobservable characteristics which may have impacted upon the outcome. All variables that are known to impact upon the treatment assignment and outcome are required to be included in the study (Stuart, 2010). In the case of propensity scores, it comes at little cost to include variables that ultimately prove to have an almost non-existence impact or no impact at all upon the outcome, however, the reverse often results in bias (Stuart, 2010). It is also important that no variables are included that are measured after the treatment has occurred. Variables that are potentially impacted by the treatment may introduce bias resulting in a confounding estimate (Greenland, 2003; Stuart, 2010). The extensiveness of the data available from Statistics Netherlands means I am able to include a large number of variables which I believe are likely to be influential on income and education outcomes.

I identify these covariates based on previous studies and theories, dividing them into four categories: individual, parental, household, and neighbourhood. The individual covariates, gender and country of birth, are consistently included in both neighbourhood effect and segregation studies (Musterd *et al.*, 2012; Gorodzeisky and Semyonov, 2017; Bisschop, ter Weel and Zwetsloot, 2020; Andersson *et al.*, 2023). On average in the Netherlands, women's gross annual wage was 32% lower than men's, some of this difference can be explained by the variation in average working hours and the different types of work (CBS, 2025h). Additionally, within educational outcomes 54% of university enrolments were female in 2023 (CBS, 2023d). There are also discrepancies for country of birth, with non-Western individuals more than twice as likely to be unemployed (van den Bossche, no date), and ethnic minorities having a lower average income than those born in the Netherlands (van Elk, Jongen and Koot, 2019).

The household characteristics focus on the socioeconomic characteristics of the family, including household structure, income, and the characteristics of the home such as value, ownership status, and house usage. Household income is a key

factor in housing location choice, with resources not only impacting range of choice but also the ability to experience upwards socio-spatial mobility (Hedman, van Ham and Manley, 2011; Nieuwenhuis *et al.*, 2020). Additionally, studies have identified intergenerational income patterns where children's income is often similar to that of their parents (Blanden, 2019; Colagrossi, Geraci and Mazzarella, 2023). Household structure also plays a key role in household location choice, especially household composition as this determines the type, size, and price that is accessible or required for the household along with the mobility of the household (Nieuwenhuis *et al.*, 2020).

In addition to these characteristics, I also include parental characteristics including fundamental information about the parent's education, socio-economic status, age at the birth of their child, and country of birth. Having a parent with a high level of education can reduce the impact of poverty - children who have at least one parent with a high level of education are not affected by exposure to poorer neighbourhoods, with the neighbourhood proving to be less consequential than parental education on educational attainment (Troost, van Ham and Manley, 2023). Andersson *et al.* (2023, p. 895) found that socio-economic family characteristics variables explained 9.1% of the variance in education level in the Netherlands for those measured at 16. Moreover, Dutch born children's income distribution was on average 10-19 percentiles higher than children of migrants for the lowest parental income (van Elk, Jongen and Koot, 2019, p. 3). Additionally, we know that country of birth has an impact upon housing choice, with non-Western migrants likely to live in areas with high levels of social housing (Musterd and Fullaondo, 2008, p. 109) as well as areas where there is already a high proportion of non-Westerners (Boschman and van Ham, 2015, p. 1165). The socioeconomic status of parents is an important factor in determining income and education in adulthood, in part due to the resources that parents invest in their children (Tessema *et al.*, 2022).

As I am considering the impact of where people live, I also include variables that reflect the composition of the neighbourhood. We know from neighbourhood effects literature that neighbourhood characteristics and mechanisms play a large role in impacting opportunities available for individuals and thus their life outcomes in the Netherlands (Musterd, Ostendorf, De Vos, 2003; Musterd *et al.*, 2012). For these characteristics I include grid cell information: population, and the proportion of non-Westerners, unemployed, and work-age individuals.²⁵ Additionally, I include the average house value and the percentage of houses owned. In this study neighbourhoods are defined using 500m by 500m grid cells with an average grid cell population of 994 across the study period.

²⁵ Between the ages of 20 and 64.

Table 6.1: Descriptive statistics for imputed dataset.

	Social Frontier ²⁶	Non-Social Frontier
Categorical Covariates (%)		
Male	49.78%	49.82%
Non-Western Individual	1.87%	2.18%
Non-Western Mother	10.79%	13.03%
Non-Western Father	11.46%	13.68%
Parent's Education Level ²⁷		
Low	11.77%	11.99%
Secondary	21.60%	20.91%
High	17.27%	14.10%
Mother Employed	73.56%	69.18%
Father Employed	77.60%	75.98%
Mother SECM Status		
Employee	61.97%	58.63%
Director/major shareholder	0.96%	1.21%
Self-employed	8.14%	9.87%
Benefit Recipient	9.45%	9.08%
Pensioner	0.55%	0.57%
Without Income	18.37%	20.07%
Father SECM Status		
Employee	73.76%	69.96%
Director/major shareholder	4.11%	4.45%

²⁶ Individuals are grouped based on whether they lived near a social frontier for all five years between 11 and 15. The descriptive statistics for those who lived near a social frontier are listed in this table under the column name: 'Social Frontier', the statistics for those who did not live near a social frontier for any of the five years are listed under 'Non-Social Frontier'.

²⁷ Education level is aggregated using the SOI which is the standard classification of educational programs in the Netherlands, this is the first aggregation level, Table L.1 includes a detailed explanation of the levels. As I had a significant amount of missing data for parental education the variable was combined to describe the highest level of either parent's education.

Self-employed	9.70%	13.03%
Benefit Recipient	5.92%	5.75%
Pensioner	0.32%	0.34%
Without Income	2.56%	2.57%
Household Composition (Family Structure)		
Live with neither	0.64%	0.55%
Live with both	79.90%	82.31%
Live with Mum	17.58%	15.49%
Live with Dad	1.74%	1.51%
Proportion of houses owned	72.30%	73.22%
Continuous Variables	Mean (SD)	Mean (SD)
Mother Age at Birth	30.37 (4.49)	30.30 (4.45)
Father Age at Birth	33.04 (5.10)	33.07 (5.01)
House Value	€247,578 (€135,601)	€261,869 (€141,063)
Grid Cell House Value	€206,423 (€80,332)	€222,515 (€94,238)
Household Income	€54,481 (€49,095)	€51,127 (€42,054)
Median Grid Cell Household Income	€51,561 (€24,162)	€48,075 (€28,477)
Grid Cell Population Size	1267.52 (629.61)	1165.42 (1070.91)
Grid Cell Proportion of Non-Westerners	9.42% (9.02%)	10.76% (18.01%)
Grid Cell Proportion of Unemployed individuals	18.12% (4.75%)	18.66% (5.60%)
Grid Cell Proportion of Work-age population	60.82% (6.01%)	59.82% (6.88%)
These statistics are created using the non-public microdata from Statistics Netherlands.		

On initial inspection, the treated (those living in close proximity to a social frontier) and control individuals are relatively similar across the majority of variables, as visible in Table 6.1. There are a few covariates which the groups differ on: 1.87% of treated individuals were non-Western, however, a slightly higher percentage (2.18%) of control individuals were non-Western. This difference is consistent across the parental country of birth variable with control individuals having a higher proportion of

non-Western parents. There is also a difference in the parental variables visible in Table 6.1 in relation to education level and employment status. There are also differences in the neighbourhood covariates: both the average grid cell house prices and the house value were nearly €15,000 higher for the control individuals, however, they did have a higher standard deviation, indicating a larger range in house prices. The income variables were higher for the treated individuals by approximately €3,000 a year. The remaining neighbourhood characteristics were relatively homogeneous, with the grid cell level demographics similar for the control and treated individuals. Based on the standard deviations of the neighbourhood variables, I would suspect that there is more diversity in the neighbourhood characteristics of the control individuals - there is a bigger range for the population size, proportion of non-Westerners, unemployed individuals, and house prices. This is unsurprising as when identifying social frontiers we are looking for homogeneous neighbourhoods which contrast with each other.

Region of common support

Figure 6.1 shows the distribution of the propensity scores for the treated and control groups.²⁸ To accurately measure the treatment effect, it is required that there is an acceptable overlap between the treated and control groups propensity scores (Stuart, 2010; Garrido *et al.*, 2014). Although the control individuals are slightly more left-skewed with smaller propensity scores, there is still a large overlap particularly for the treated individuals suggesting that a suitable match within the control group is likely to be found, strengthening the findings.

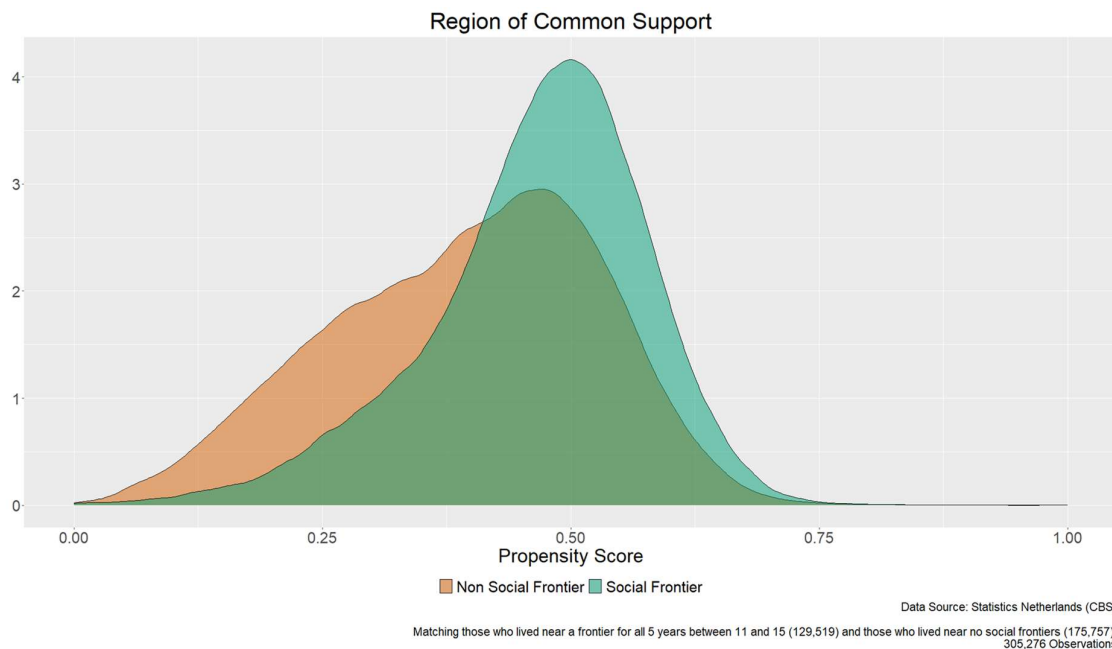


Figure 6.1: Region of common support

²⁸ The propensity score is the probability of an individual receiving a treatment. The score is explained in greater detail in the methodology section.

Outcome variables

The outcome variables in this study were income and education level between the ages of 20 and 25. For income, this is the total income per year from any type of work. The education level is the highest level of education achieved, converted into the typical number of years spent in education to obtain this level - this ranges from 4 (lower primary education) to 22 (doctorate degree), with each level described in Table L.1. Within Table 6.2, we can see a difference in the mean outcomes for both variables across all ages when comparing those who lived near social frontiers with those who did not. Individuals who did not live near social frontiers earned on average between €500 and €900 more each year across the ages compared to those who lived near social frontiers. The high standard deviation for the income of the control individuals indicates that there was greater dispersion across the income of these individuals, especially when measuring income at age 20 where there was a broad range of incomes.

For the education variable, individuals who did not live near social frontiers spent longer in education on average, albeit the difference was almost non-existent - at most 0.12 years. Additionally, as the standard deviation was nearly 2 years for most ages there was a large crossover between the educational outcomes. The differences in both outcomes within a larger perspective are relatively small and could be a result of alternative factors. However, the difference could be explained by social frontier proximity and thus to test this I implement a causal method as a further investigation.

Table 6.2: Descriptive statistics for outcome variables - Mean (SD)

	20	21	22	23	24	25
Income (€)						
SF	€6,877 (€7,785)	€10,289 (€9,506)	€14,165 (€11,860)	€17,827 (€12,602)	€22,020 (€18,247)	€26,860 (€16,601)
Non-SF	€7,431 (€15,882)	€11,026 (€12,399)	€15,017 (€13,603)	€18,764 (€15,751)	€22,921 (€17,759)	€27,696 (€30,109)
Years Spent in Education						
SF	13.57 (1.62)	14.12 (1.81)	14.63 (1.95)	15.04 (1.97)	15.33 (1.99)	15.52 (2)
Non-SF	13.66 (1.62)	14.24 (1.78)	14.73 (1.89)	15.11 (1.90)	15.38 (1.91)	15.56 (1.91)

These statistics are created using the non-public microdata from Statistics Netherlands.

Missing data

Although the data from Statistics Netherlands is incredibly detailed, it does have issues with missing data, the percentage of which is included in Table 6.3. In this study I include a missing data indicator for each of the covariates where there is

missing data. This works well for matching methods using propensity scores as they are not concerned specifically about the covariates within the model, but more the probability predictions (Rosenbaum and Rubin, 1984; Stuart, 2010). Using a missing data indicator, especially in the case when the data is not missing at random, results in smaller levels of bias (Choi, Dekkers and le Cessie, 2019, p. 32). In this study the data is unlikely to be missing randomly, there is a notable identified relationship between the missing data and the dataset.

The variable with the largest percentage of missing data is parent's education with over half of the information missing. I addressed this in some form by using a joint variable for the parent's education – including separate education variables for the mother and father resulted in a higher proportion of data being missing. The education dataset is created using a combination of various sources which should result in it having a high coverage rate. However, Troost, van Ham and Manley (2023) noted that there is an overrepresentation of migrants within the missing parent's education and there is less data available for the older population (Sivak and Stulp, 2024), potentially explaining why I identify such a high percentage of missing data. Due to the large percentage of data missing, I considered dropping the variable as it may add bias. However, I felt that as the data did not appear to be missing at random, removing the variable would be of greater cost than the potential bias introduced by the missing data. The parental educational variable is crucial for predicting socioeconomic outcomes in adulthood due to the importance that parents place upon education, access to resources, and inherited or learned traits (Buis, 2013; Erola, Jalonen and Lehti, 2016). The variable was highlighted as a crucial factor when measuring children's education – Troost, van Ham, and Manley (2023) found that having one higher educated parent could negate the impact of neighbourhood poverty with it acting as a protection from the effect of the neighbourhood. Additionally, the fact that the data is not missing at random is also informative for the study as it shows a pattern amongst the dataset which can be used to predict outcomes. To address the issue I created a missing data indicator, meaning that individuals are also matched on the pattern of missing data (Stuart, 2010). As the matching methods do not work with missing variables, in addition to the missing data indicator I also use a simple single imputation using the median value for continuous variables and the mode for the categorical variables. This is a simple and fast method of imputation which also maintains the balance for the covariates, especially for the extreme values, and keeps the missing data pattern consistent (Stuart, 2010, p. 17). In another study a more robust imputation method could be applied to strengthen the findings such as combining multiple imputation and missing data indicators although this can struggle when the data is not missing at random (Choi, Dekkers and le Cessie, 2019). Additionally, researchers would have to be careful using this approach as it can be computationally complex.

Table 6.3: Missing data

Variables	Missing data percentage
Parent's Education	51.45%
Household Composition	0.14%
Grid Cell Population	0.001%
Grid Cell Non-Western Population	0.001%
Grid Cell Unemployed Population	0.001%
Grid Cell Work Age Population	0.001%
Grid Cell Average Income	0.71%
Household Income	4.91%
Mother SECM Status	0.57%
Father SECM Status	3.79%
Mother age at birth	0.07%
Father age at birth	2.17%

Methodology

Social frontier identification

As this chapter identifies country of birth social frontiers on a country-wide scale, it implements the Partitioning Around Medoid (PAM) clustering method proposed in Chapter 3. The method choice is largely due to the size of the study – for each year the social frontiers were identified across a population of over 16.5 million (for 23 years). Based on the processing time in Table 3.3 this would be expected to take around 163 hours using the Lee and Mitchell (2013) method. The PAM clustering method identifies the frontiers significantly quicker than the established Lee and Mitchell (2013) method, taking approximately 2 and a half hours, whilst operating at a relatively high level of accuracy which means that we can be confident in the majority of frontiers identified. In a smaller study it would be preferable to use the Lee and Mitchell (2013) method as there is a higher level of certainty in the identified frontiers meaning we could be more confident that any effects identified are a direct consequence of the social frontiers. This is a limitation of the study and a suggestion for future research - researchers could choose to perform a smaller version of the study potentially focusing on a particular region to explore whether the effects are consistent when using the Lee and Mitchell (2013) method. However, this would come at the cost of generalisability and statistical power with the smaller study limiting the study's ability to claim that the results will be visible in different countries or that the findings are genuine and significant. An explanation of the method is

provided in Chapter 3, but the basic concept is that it groups grid cells based on the percentage of non-Westerners and if the adjacent grid cells are in different groups a social frontier exists between them. Social frontiers were identified for every year between 1999 and 2021 in the Netherlands.

Matching method choice

This study makes use of matching and stratification methods which aim to balance the observable characteristics between a treated and control group, making the distribution of the covariates as similar as possible. The methods aim to replicate a randomised control trial by creating two groups of people (control and treated) who are as identical as possible based on their observable characteristics (Stuart, 2010), varying only on their treatment variable. Their outcome is then compared to evaluate the treatment effect. Matching is specifically suitable for a situation where it is impossible for individuals to be randomly assigned to a treatment. Simply comparing the outcome of those who lived near social frontiers to those who did not may be misleading if there are systematic differences between the individuals. Which is likely to be the case as housing is not unbiased, people do not choose where they live randomly and there are many factors at play which influence household location: income, housing availability, ethnic homogeneity preference, and household type. It could be an amalgamation of these factors which result in individuals self-sorting into neighbourhoods prone to social frontiers or deprived neighbourhoods, which ultimately will have consequences for outcomes in adulthood. Understanding selection bias and reducing it, is crucial for neighbourhood effects studies to correctly identify the effect of the independent variable (Hedman and van Ham, 2012). The strength of this study comes from the extensive range of covariates available in the dataset, with matching methods only ever as strong as their covariates (Thoemmes and Kim, 2011). As I have access to a large amount of information about individuals, I am able to go beyond demographic data and thus can be more confident in obeying the strongly ignorable treatment assignment assumption.

Best practice in covariate balancing studies is to test multiple methods and use these findings to discuss the most suitable method, doing this by assessing the covariate balance and the ability of the methods to maintain the sample size (Thoemmes and Kim, 2011; Iacus, King and Porro, 2012). Within this study I applied seven different methods - five of which use propensity scores as their distance measure - to estimate the best matches for the control and treated individuals. Propensity scores are the probability that an individual was assigned to a treatment variable – in this case, this is the probability that an individual lived near a social frontier during their childhood (Rosenbaum and Rubin, 1984). In my study the propensity scores are calculated using the most common logistic regression model (Stuart, 2010), “in which treatment status is regressed on observed baseline characteristics” (Austin, 2011, p. 403). I applied seven matching/stratification methods: nearest neighbour; nearest neighbour with a caliper; nearest neighbour with replacement; nearest neighbour with replacement and a caliper; Coarsened Exact Matching (CEM); subclassification

and Generalised Full Matching (GFM). A description of these methods is included in Table M.1. To test the quality of the matches achieved by the methods I initially look at the Standardised Mean Difference (SMD) and variance ratio for each of the covariates. Rubin (2001, p. 174) suggests that for results to be reliable the absolute SMD scores should be less than 0.25 and the variance ratio should be close to 1. Austin (2009, p. 3090) suggests that the literature implies that this could be extended further, by requiring a standardised difference of less than 0.1 for good matching to have been achieved.

Table 6.4: Sample size for matching methods.

	Prior to Matching	Nearest Neighbour	Nearest Neighbour with Replacem ent	GFM	CEM	Subclassi fication
Matched						
Control (ESS)	175,757	129,519	74,877 (52,386)	175,757 (70,720)	1,820 (1,429)	175,757 (135,528)
Treated	129,519	129,519	129,519	129,519	1,678	129,519
Unmatched						
Control		46,238	100,880	0	173,937	0
Treated		0	0	0	127,841	0

The covariate balancing scores are shown in Figure 6.2 (Table N.1 also shows the scores) I have added lines to highlight the scores above 0.1 and 0.05 and below 0.1 and 0.05 which indicate weaker balance between covariates. I have not included the results from the caliper matching as including the caliper had very little impact upon the balancing scores compared to non-caliper. Based on a combination of the balancing scores and the sample sizes in Table 6.4, I identify the nearest neighbour matching with replacement, subclassification, and GFM as the best performing methods.

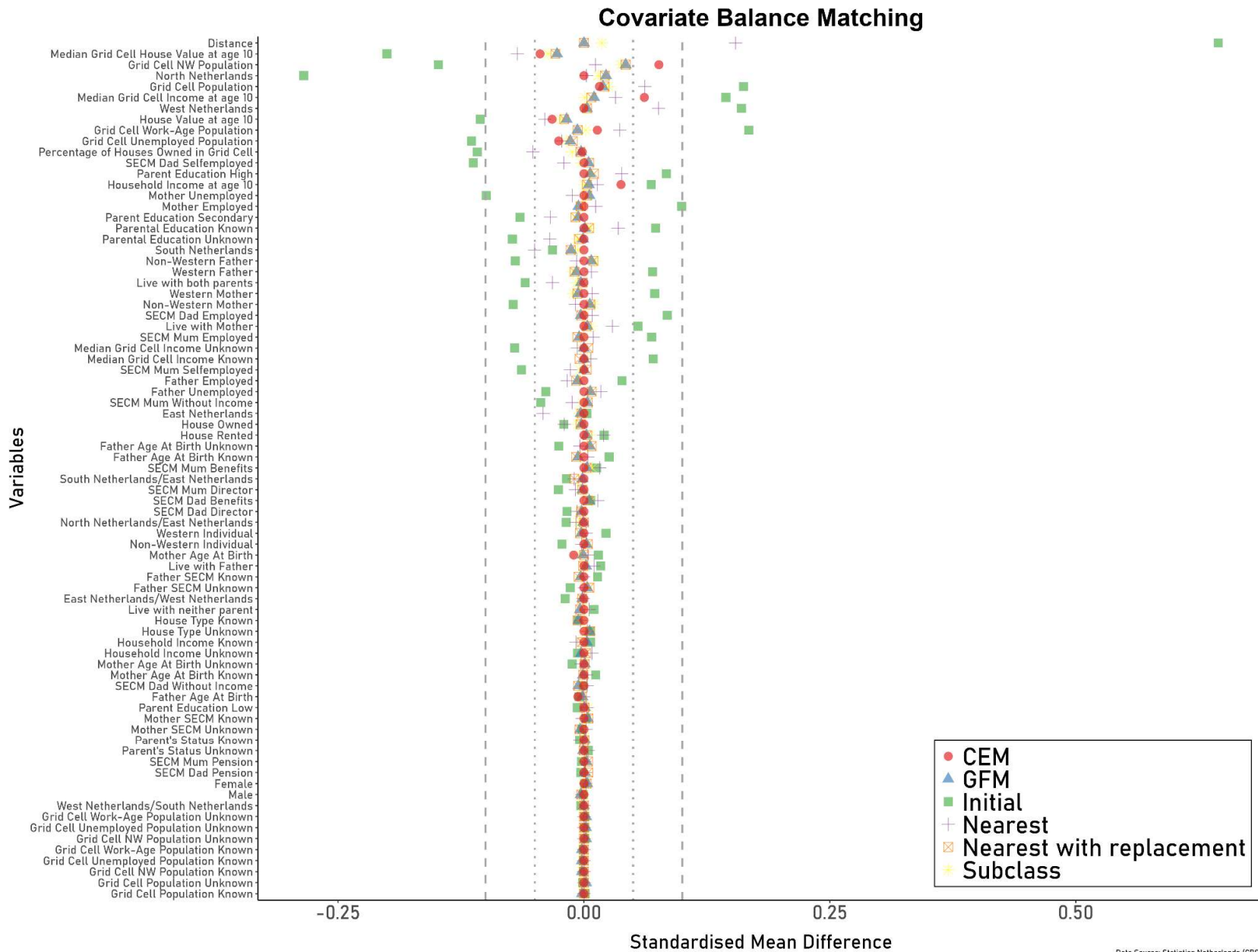


Figure 6.2: Covariate balancing love plot for all the matching methods.

To aid with my final choice, I include Figure 6.3 which is a plot showing the SMD for the three methods, along with the unmatched variables prior to matching. The aim of the figure is to identify if there are any covariates that were clearly unbalanced prior to matching and provide a closer inspection into the scores. Within Figure 6.3 we can identify the North Netherlands region (part of the NUTs zones covariates) and median grid cell house value as clearly misbalanced. Following matching, for all three methods the matching has been significantly improved, as has the matching for the majority of the covariates. There are a couple of covariates where the SMD has increased, but as these increases are minor this is unlikely to be problematic and is often the case when the covariates are similar prior to matching (Stuart, 2010, p. 15). From Figure 6.3 we can clearly see several patterns: firstly, the matching methods produce very similar SMD scores; secondly, all three methods show a clear improvement on average on the unmatched variables; and thirdly, none of the methods consistently have the smallest SMD across all the variables. In terms of the overall matching, the nearest neighbour with replacement matching has the smallest median SMD score, indicating that it is the most consistent lowest scoring matching method whilst also frequently either having the smallest SMD score or the second smallest, highlighting its stability and reliability. The subclassification method appears to be well-balanced, but it can have imbalance within the individual groups which could impact my results. Ultimately, I select the nearest neighbour matching with replacement as my method based on its ability to balance covariates and maintain a large sample.

Nearest neighbour matching with replacement

Nearest neighbour matching with replacement is one of the most commonly implemented matching methods. Each treated individual is matched to the control individual with the smallest difference between their propensity scores (Stuart, 2010). Allowing replacement means that control variables can be used multiple times as matches for the treated individuals, resulting in the matching order becoming inconsequential as treated individuals are always matched to their nearest control. The method can be unreliable if the results are based on a too small selection of control individuals. In this case, I am still using a large number of control individuals, there are around 1.5 times as many control as treated individuals, with control individuals typically matched twice. To calculate the ATT (average treatment effect on the treated) I implement a weighted linear regression model.

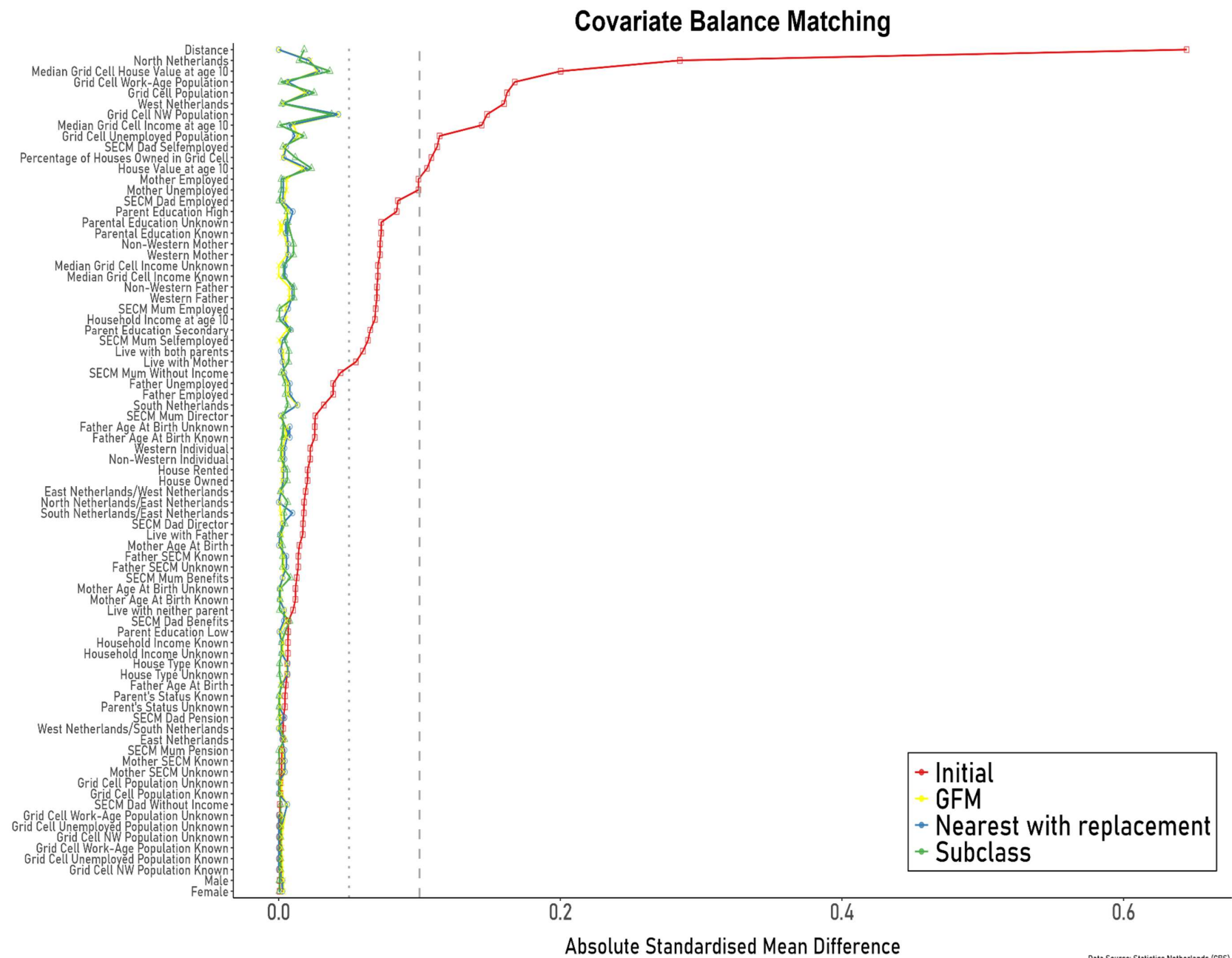


Figure 6.3: Covariate balancing plot for three chosen methods

Results

To estimate the treatment effect (ATT) on those who lived near social frontiers and the standard error I fitted 12 weighted linear regression models, using the income and years spent in education between the ages of 20 and 25 as the outcomes for the models.

Income

The results for the regression model estimating the ATT score for income are included in Table 6.5. Older individuals (those aged 23 or above) experienced a larger difference in income due to living near a social frontier compared to when younger. Living near a social frontier was associated with a reduction in income per year of €458 on average as a 20-year-old, whilst at age 25 it was €697 – this is as expected as younger individuals often earn less. This indicates that living near a social frontier during childhood has a clear and consistent, albeit relatively small, negative impact on income as an adult. The ATT does begin to reduce after 23, with a slightly smaller impact visible at 24 and 25 compared to 23. Although potentially not a large difference in earnings over the space of a year, my ATT findings do highlight the relationship between living near a social frontier and reduced income.

Table 6.5: Average Treatment Effect on the Treated (ATT) for income.

	Estimate	Std. Error	p-value	E-Value
Income at 20	-458	54.7	<0.001	1.215
Income at 21	-579	52.9	<0.001	1.272
Income at 22	-675	60	<0.001	1.275
Income at 23	-733	67.8	<0.001	1.269
Income at 24	-693	94.1	<0.001	1.228
Income at 25	-697	118	<0.001	1.187

Alternative exposure length

As an additional aspect of the study, I calculated the ATT for varying levels of exposure to social frontiers reporting the results in Table 6.6. First, I compared the outcomes of those who lived near social frontiers for all 5 years between 11 and 15 and those who could have lived near a social frontier at some point between 11 and 15 but not for all 5 years. Before estimating the effect of living near a social frontier for at least one year between 11 and 15 compared to those who were never exposed. In terms of the difference in the income results, unsurprisingly the largest effects were found, across the majority of the ages, when comparing those who lived near social frontiers for all 5 years with those who never lived near frontiers (Table 6.5). The exception being the ATT at the ages of 24 and 25 when any exposure to social frontiers resulted in a large impact - with individuals who lived near social

frontiers earning on average €745.22 less at 24. There was consistency across all the exposure lengths with a significant (based on p-value) negative association identified on income for all ages between 20 and 25 across all three varying levels of exposure.

Table 6.6: Average Treatment Effect on the Treated (ATT) for income for alternative exposure length.

	Exposed to SF for 5 years vs less than 5			Exposed to SF vs no exposure		
	Estimate	Std. Error	p-value	Estimate	Std. Error	p-value
Income at 20	-340	57.4	<0.001	-350	66.7	<0.001
Income at 21	-403	51.1	<0.001	-476	55.5	<0.001
Income at 22	-434	58.4	<0.001	-561	62.1	<0.001
Income at 23	-674	152	<0.001	-583	133.6	<0.001
Income at 24	-364	89.4	<0.001	-745	89.4	<0.001
Income at 25	-384	113	<0.001	-732	135.4	<0.001

Education

The education results for the regression model estimating the ATT score are included in Table 6.7. In contrast to the income variable, the pattern of the ATT was inconsistent throughout the years. The highest estimated treatment effect was identified when the individuals were 21 years old when those who lived near a social frontier averaged 0.0898 fewer years in education. By the time those who lived near social frontiers were 25, the average effect was 0.0313 fewer educational years. The difference in all the years was very minimal, as was the standard error - all of the average treatment effects were less than 0.1 - which in reality is an incredibly small difference. Both the inconsistency in the results and the minimal difference indicate that living near a social frontier during childhood has a relatively minimal, if not completely non-existent, impact on individuals' education.

Table 6.7: Average Treatment Effect on the Treated (ATT) for education years.

	Estimate	Std. Error	p-value	E-Value
Education Years at 20	-0.0648	0.0082	<0.001	1.233
Education Years at 21	-0.0898	0.0090	<0.001	1.267
Education Years at 22	-0.0815	0.0096	<0.001	1.242
Education Years at 23	-0.0541	0.0096	<0.001	1.189
Education Years at 24	-0.0409	0.0096	<0.001	1.160
Education Years at 25	-0.0313	0.0095	<0.001	1.137

Alternative exposure length

As with income, I also included the ATT scores calculated for different exposure lengths in Table 6.8. In contrast to the income results, the largest negative association was identified when comparing individuals who lived near a social frontier for at least one year with those who never did - with the maximum impact identified at the age of 21 where treated individuals spent on average 0.1 fewer years in education. As with the income outcomes, a consistent negative association was identified across the different exposure lengths and ages. The only point to note being the results at 24 and 25 which were not statistically significant with a p-value of 0.188, when comparing individuals who lived near social frontiers for all 5 years with those with less exposure time. Again, as with Table 6.7 the effects were minimal and unlikely to be consequential in reality.

Table 6.8: Average Treatment Effect on the Treated (ATT) for education years for alternative exposure length.

	Exposed for 5 years vs less than 5			Exposed to SF vs no exposure		
	Estimate	Std. Error	p-value	Estimate	Std. Error	p-value
Education Years at 20	-0.0466	0.0076	<0.001	-0.0831	0.0073	<0.001
Education Years at 21	-0.0591	0.0085	<0.001	-0.1040	0.0080	<0.001
Education Years at 22	-0.0481	0.0090	<0.001	-0.0991	0.0085	<0.001
Education Years at 23	-0.0364	0.0089	<0.001	-0.0787	0.0085	<0.001
Education Years at 24	-0.0183	0.0089	0.041	-0.0682	0.0085	<0.001
Education Years at 25	-0.0117	0.0089	0.188	-0.0518	0.0085	<0.001

Sensitivity analysis

Although the matching resulted in well-balanced covariates across the treated and control group, the results can only be accurate if we have considered the potential impact of unobserved covariates. To do this we need to measure whether there is hidden bias which is likely to impact the treatment effect. For this I implement E-value sensitivity testing, the E-value was introduced and defined by VanderWeele and Ding (2017, p. 268) as:

The minimum strength of association ... that an unmeasured confounder would need to have with both the treatment and the outcome to fully explain away a specific treatment-outcome association, conditional on the measured covariates.

The E-values for this study, visible in Tables 6.6 and 6.8, indicate that for unmeasured confounders to explain the treatment effect I have identified for income at 25, they would need to have an association equivalent to a risk ratio of 1.187 with both the treatment and outcome variables when the observed covariates are adjusted for. A weaker unobserved covariate would not be able to explain away this effect. The E-value scores are relatively consistent: the smallest associated risk ratio is for income at 25 whilst the highest is 1.75 for income at 22. For education, the results are quite similar, with an unmeasured confounder needing an association equivalent to a risk ratio of 1.267 and 1.137 for years spent in education at 21 and 25 respectively. These findings suggest that the treated effect identified between those who lived near a social frontier in childhood and outcomes in adulthood would be unlikely to be considered completely robust to unmeasured confounding, meaning that evidence of causality for this association is quite weak. However, this absence of causality does not imply that the treatment has no effect, nor that it has an insignificant impact upon those treated (VanderWeele and Ding, 2017, p. 273). Additionally, despite the E-value scores being relatively small, I have controlled for multiple covariates, indicating that the study should be more robust to unobserved confounder impact (VanderWeele and Ding, 2017, p. 271; Bi *et al.*, 2025, p. 10). Bi *et al.* (2025, p. 13) discuss how a rigorous matching method implementation will also provide robustness to a study. All of these factors do improve the likelihood of the results being causal, however, the sensitivity analysis results highlight that I must be cautious in reporting my findings, particularly with any claims of causality.

In addition to the E-value, I have included in Appendix O a range of additional sensitivity tests with the intent of finding evidence which supports the findings. I can conclude that living near a social frontier during childhood, specifically for all the years between 11 and 15 is negatively associated with income and education in adulthood between the ages of 20 and 25 in the Netherlands.

Conclusion

Summary of findings

This chapter makes multiple contributions to the social frontier literature including being: (i) the first robust test of whether there exists an association between living near social frontiers in childhood and outcomes in adulthood using individual-level longitudinal data, and (ii) the first identification and study of social frontiers in the Netherlands on a national scale. This depth and breadth of analysis is only possible due to the extensive data I am able to access, with the study strengthened due by its country-wide nature resulting in a sample size of over 300,000 individuals.

The chapter also includes multiple sensitivity analysis tests, which help gauge the robustness of the findings. In non-experimental studies these tests are particularly beneficial as we cannot test underlying assumptions. The E-value score means we

can quantify how much an unmeasured confounder would need to be associated with both the treatment and outcome variables to explain the average treatment effect. In this study, the results of the sensitivity tests prevent any causal claims being made, however, by that same logic they also reduce the chance of making false claims. This is the first use of nearest neighbour matching with replacement to identify social frontier effects and future studies could look to include more covariates to attempt to address the bias potentially introduced by the unobserved confounders.

Although the chapter is innovative in its methodology as a social frontier study, its main contribution to the literature is largely related to its study of the impact of living close to sharp transitions between communities. Through this, my study identifies the long-lasting impact of living near these social frontiers during childhood on income and educational outcomes in adulthood. When we interpret the mechanisms that make these frontiers consequential on society, we can look to the neighbourhood effects and social frontier literature. Putnam's (2007) theory of 'hunkering down' suggests that those in more diverse neighbourhoods tend to become more isolated, withdrawing themselves from the communities. Withdrawing from communities reduces the opportunity for residents to form weak ties, which potentially provide more opportunities than strong ties due to individuals being able to form networks and access the resources of multiple individuals (Granovetter, 1973). A lack of community cohesion - including the use of shared spaces - can reduce the relationships and networks formed, likely reducing individuals' ability to gain social capital, along with exposure and access to role models and resources (van der Laan Bouma-Doff, 2007; Putnam, 2007; Sykes and Kuyper, 2009; Musterd *et al.*, 2012; Iyer and Pryce, 2024). This, combined with frontiers becoming zones of conflict (Dean *et al.*, 2019), can help to explain the negative association I identify between socioeconomic outcomes and social frontier proximity in childhood.

Previously social frontier literature has focused on their immediate impact on an area bypassing any investigation into whether they have long-term consequences, especially on children who grow up close to them, this is along with not studying their impact on the individual. The negative association I identified, supports the theory of the 'geography of opportunities', where a person's, and particularly in this chapter's case a child's, neighbourhood is crucial for shaping their life outcomes (Musterd, Ostendorf and De Vos, 2003). The average treatment effect highlights that we cannot fully understand the importance of the neighbourhood unless we consider the multiple different aspects of the neighbourhood, including proximity to social frontiers as living near them clearly has an impact upon residents. This impact is particularly visible on income, with an average decrease of €697 in yearly earnings for treated individuals at age 25, whilst the difference in years spent in education was relatively minimal. Extended exposure to social frontiers during childhood, between the ages of 11 and 15, had a negative influence on the life course trajectories of those living in the Netherlands. Further research is needed to determine whether this is consistent

across different countries, social frontier dimensions, and whether the length of exposure changes the level of impact.

Limitations

There are several limitations to the study. Firstly, the matching methods are never going to achieve the same accuracy as a randomised control trial as they rely on the assumption that there exist no unobserved confounding factors. I control for this assumption as much as I can by matching on a large number of covariates, however, the sensitivity analysis does highlight that this is still a potential issue. The limitation means I have to be cautious with any causal claims. Future studies could potentially apply different methods, whether that be alternative matching methods or a synthetic control model, or include different variables to attempt to be able to state a causal claim.

Secondly, I faced a substantial issue with the missing data in the parental education variable. The absence of key data potentially results in increased bias within a study. In this study, this level of missing data was unavoidable based on the data that was available from Statistics Netherlands. I felt that to remove the variable would have been more detrimental to the study than using the variable with the missing data. Thus, I made the decision to impute the missing variable and use a missing data indicator variable. Additionally, I used a combined parents' education variable rather than individual mother and father variables which reduced the level of missing data. I was limited in this chapter based on data availability, however, future studies may be able to access data with fewer missing cases, potentially from different countries or regions, to create a stronger model. Furthermore, I could have employed a different technique to cope with the missing variables which may have strengthened the findings.

Suggestions for future research

This study is the first to empirically explore the impact of social frontier exposure on income and education outcomes. I hope that my study will lead to further discussions surrounding the effect of social frontiers on individuals. There are also opportunities to evaluate whether the effect varies based on exposure length, frontier strength and duration, frontier dimension, and the potential difference in the impact of intersectional frontiers. Within Chapter 5 and Iyer and Pryce's (2024) theoretical paper, it was suggested that if frontiers are co-occurring across multiple domains, then they have the potential to have a heightened impact. Additionally, tension levels can be at their highest when frontiers are forming, with greater territoriality issues (Legewie and Schaeffer, 2016; Iyer and Pryce, 2024), which could result in them being more consequential on individuals. This could be tested by exploring whether those who are exposed to young and not yet fully established frontiers experience a different effect to those exposed to long-lasting and well-established frontiers. Furthermore, future research could explore whether economic frontiers, such as those identified in Chapter 4, also have an influence on life course outcomes and

how this compares to the impact of country of birth frontiers. The role that social frontiers play in impacting individuals, especially those who experience them in childhood, is an area of great interest and yields a range of important insights into understanding the impact of childhood neighbourhoods on achievements and outcomes in adulthood.

Chapter 7 Conclusion

Overview

In this thesis I have investigated the measurement, prevalence, dynamics, and impacts of social frontiers. This has led to a number of important contributions to the literature, all of which further our understanding of residential segregation in the Netherlands and enhance the scope of future studies to explore whether these results are consistent beyond the Netherlands. These contributions include:

- Identifying the Partitioning Around Medoid (PAM) clustering as an alternative identification method to the established Lee and Mitchell (2013) method. The key recommendation that follows from this is that there are scenarios which suit either method. The Lee and Mitchell (2013) method is deliberately more cautious and thus is likely to only detect the strongest of frontiers. This makes it particularly suitable for studying the trends in social frontiers, both dynamics and prevalence as the researcher can be more confident in the frontiers identified. Hence its application in Chapters 4 and 5. The PAM clustering method plays a role in the larger studies, where it becomes computationally impossible or unfeasible to use the Lee and Mitchell (2013) method due to the processing time. Larger studies are often more generalisable with higher statistical power. Additionally, it can be used when making policies as due to the high TPR the method is very unlikely to miss any individuals and thus targeted interventions would likely be more successful. Finally, researchers could incorporate both methods with the PAM clustering method acting as a robustness test for the Lee and Mitchell (2013) method. In the case of Chapter 6, it would be useful to repeat the process on a smaller area using the Lee and Mitchell (2013) method to strengthen my findings. The PAM clustering method means that the results can be more generalisable with the larger dataset enhancing statistical power.
- Comparing the social frontier prevalence in multiple regions and dimensions using the aggregate prevalence descriptors: percentage, length, strength, and duration.
- Estimating empirical lifecycle models for social frontiers, exploring the variation in their strength depending on their lifecycle phase, along with including the rate of re-emergence.
- Providing a theoretical discussion about the definition of intersectional social frontiers – posing two different interpretations relating to sociology and existing social frontier literature.
- Presenting the first identification and discussion of economic frontiers, comparing their prevalence with the country of birth frontiers, and identifying how they intersect with each other to produce coterminous intersectional social frontiers.
- Analysing the effect of living close to social frontiers during childhood on income and educational outcomes in adulthood.

Each of the quantitative chapters in the thesis (Chapters 3, 4, 5, and 6) made distinct methodological, empirical, and/or conceptual contributions, summarised in Figure 7.1. Chapter 3 acts as an introduction to social frontiers in the Netherlands and is only the second time frontiers have been identified in the Netherlands. It does this by identifying country of birth social frontiers in the city of Amsterdam using both the Lee and Mitchell (2013) and PAM clustering methods, the first application of PAM clustering as a social frontier identification method. I was able to justify the use of PAM clustering as a social frontier identification method through the results of the simulation study which assessed the accuracy of multiple identification methods under a range of conditions.

In Chapter 4, I included multiple frontier dimensions identifying unemployment, and high- and low-income social frontiers. This chapter builds upon Chapter 3 by including the two prevalence measures from Chapter 3, along with several other measures to explore their prevalence in Amsterdam and the Hague regions. By identifying these frontiers over a 20-year period, and including duration and strength measurements, I was able to study frontier dynamics and produce social frontier lifecycle plots.

Chapter 5 added to the study of intersectional social frontiers, by making both conceptual and empirical contributions. Building upon intersectionality and social frontier literature, I proposed two definitions for intersectional social frontiers, describing them as coterminous and multidimensional. Identifying intersectional frontiers in the Netherlands enabled a comparison of their prevalence levels and that of single dimension frontiers. There were several occasions when the low-income non-Western and unemployment non-Western frontiers were more common than the single dimension low-income and unemployment frontiers respectively. In this scenario, individuals who fall into multiple groups are experiencing a greater level of division. This is important as it suggests that certain individuals are becoming increasingly segregated if they are part of multiple groups.

Finally, Chapter 6 focused on the effect of social frontiers on individuals, particularly how living close to these steep boundaries during childhood can be consequential for adult outcomes. Very few studies have empirically explored the effect of social frontiers and the majority of them have focused on their impact on crime, conflict, and housing characteristics (Dean *et al.*, 2019; Olnér *et al.*, 2024; Myatt, Lindgren and Osland, 2026).

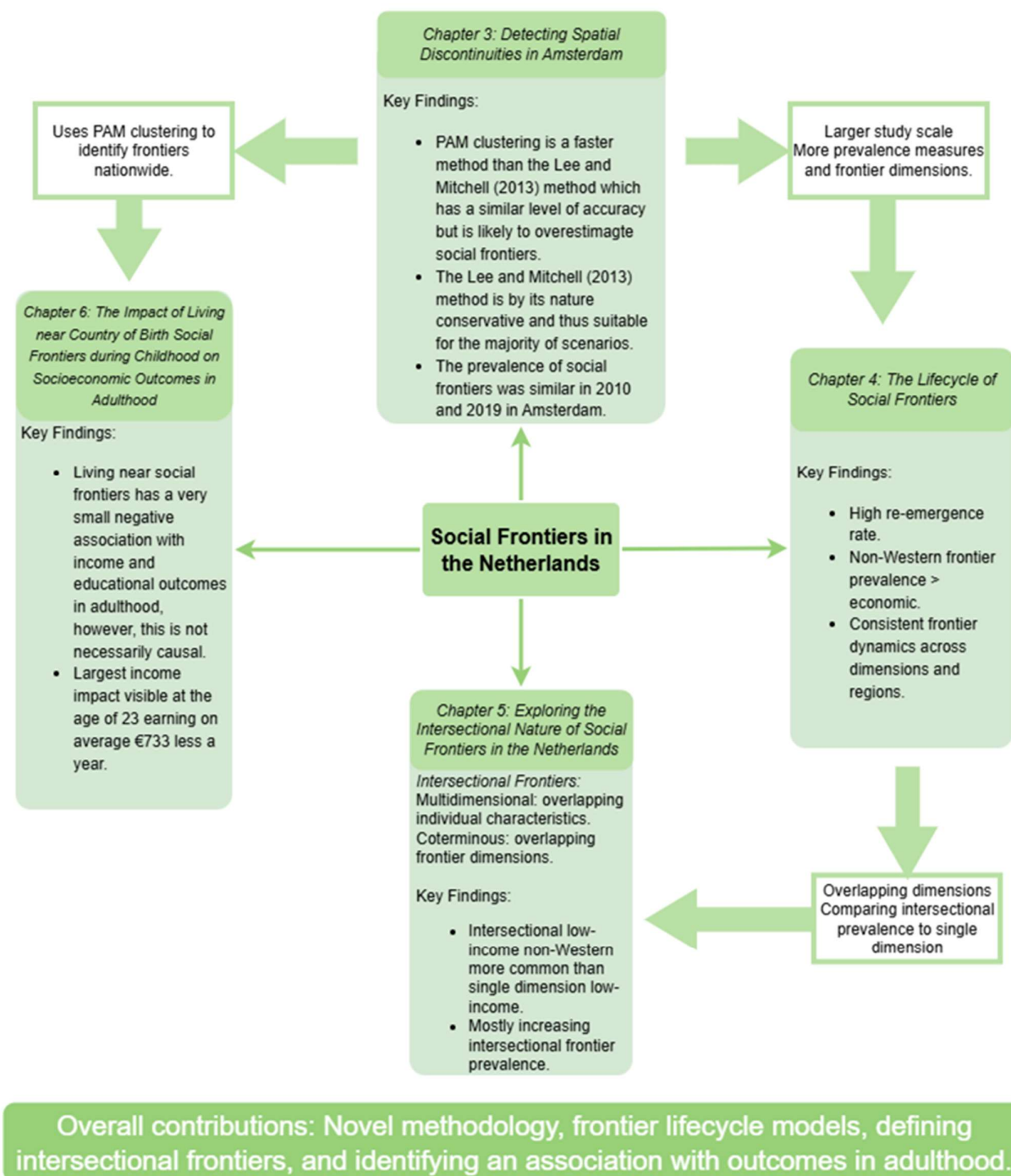


Figure 7.1: Key thesis findings

Key Findings and Conceptual Contributions

For each of the chapters I present a key finding or conceptual contribution (see Figure 7.1) along with potential methodology and policy recommendations based upon these.

Finding 1: PAM Clustering can act as an alternative, faster identification method to the Lee and Mitchell (2013) method.

The main aim of Chapter 3 was to discover whether there existed a faster, but similarly as accurate identification method as the Lee and Mitchell (2013) method, which also preferably displays some level of consistency with the frontiers identified by the Lee and Mitchell (2013) method. The Lee and Mitchell (2013) method is well-established in social frontier literature (Dean *et al.*, 2019; Křížková *et al.*, 2021; Olnér *et al.*, 2024) and is deliberately cautious by its nature – only the strongest social frontiers are identified. Thus, if we are studying the pattern of frontiers and considering how they evolve it is preferable to use the Lee and Mitchell (2013) method, as researchers can be more confident in their findings. However, it is computationally complex which has meant that frontier identification can be an incredibly time-consuming process when applied to large datasets, potentially inhibiting large scale studies. This could especially pose a problem when identifying frontiers in multiple regions, using a finer spatial scale, or if calculating the excess embrittlement level for multiple regions when cities would need to be shuffled and frontiers identified over 1,000 times for each city/year. Thus, I tested the accuracy of the Lee and Mitchell (2013) method in a simulation study alongside three alternatives: two non-spatial clustering methods and an absolute contiguous difference method.

After analysing the accuracy scores and computational complexity, I identified the PAM clustering method as the most successful alternative method for this study. I particularly noted its high true positive rate (its ability to successfully identify social frontiers), a contrast to the Lee and Mitchell (2013) method which by the nature of it being more conservative has a stricter criterion for a social frontier. Hence the Lee and Mitchell (2013) method is more likely to miss a frontier, making it useful when identifying areas for policies as areas are unlikely to be missed reducing the frontier impact. The PAM clustering also displays a notably shorter processing time than the Lee and Mitchell (2013) method meaning studies can be larger increasing generalisability.

The use of the PAM clustering method does come with compromises: it is more likely to identify false positives thus is likely to overestimate the level of divisions in a region. Secondly, it will always identify a frontier even if the spatial arrangement of the region is completely smooth due to the nature of the method. There are circumstances which suit each of the Lee and Mitchell (2013) and PAM clustering method as shown within this thesis. PAM clustering makes it feasible to identify

frontiers on a national scale, enhancing our understanding of how frontiers behave in different societies and over time. It made it feasible for me to identify social frontiers in the whole of the Netherlands over a 20-year period in Chapter 6 which had I used the Lee and Mitchell (2013) method would have required a large amount of processing power and memory for the method to execute. The conservative nature of the Lee and Mitchell (2013) method makes it particularly suited for when we study the pattern of frontiers, exploring how they vary in their strength and prevalence. It is the preferred method for prevalence studies as it only identifies the strongest frontiers making the study more substantial and reliable. In the case of identifying the frontier impact in Chapter 6, my study likely would have been advanced by using the Lee and Mitchell (2013) method as it would have measured the effect of the steepest frontiers and would ensure that 'treated' individuals really did live near social frontiers. However, as I was studying their effect on a large scale over an extended period of time it would have been incredibly time consuming and almost completely unfeasible to apply the Lee and Mitchell (2013) method in the remote access and produce the research within an appropriate time frame.

Recommendation:

The key methodological recommendation for future studies based on this chapter was that the PAM clustering method can be used as an alternative method if the size of the study makes it unfeasible or it is deemed suitable for the study. The methods could work alongside each other as the high true negative rate – low level of false positives - of the Lee and Mitchell (2013) method works well with the high true positive rate – low level of false negatives - of the PAM clustering method. However, if researchers only implement one method then the Lee and Mitchell (2013) method is suitable for the majority of cases as the more conservative nature of the method strengthens prevalence and dynamics findings as only the strongest frontiers are identified. Whether this is feasible is likely depends on the size and scope of the study. In Chapter 6, when I identify frontiers nationwide over a 20 year period, I applied only the PAM clustering to reduce the complexity of the study.

In terms of policy recommendations, I suggested that the location of the frontiers in Amsterdam was linked to gentrification in the city centre, highlighting the potential role that social frontier identification can play in local evidence-based policies as they can describe social divisions in an area. Addressing gentrification is a complicated issue. The selling off of social housing naturally leads to changes in neighbourhoods which, can in turn, lead to gentrification (Teernstra, 2014). Identifying the steep boundaries between neighbourhoods provides policymakers with a more concrete understanding of the specific areas they may need to target, if the frontiers are consequential for society (Spierenburg, van Cranenburgh and Cats, 2024). Additionally, advancements in social frontier identification methods can make frontier estimations more robust, with their identification potentially providing policymakers with a guideline of social divisions and inequalities.

Finding 2: Social frontier dynamics are consistent across regions and dimensions with a high rate of re-emergence. However, sharp country of birth transitions were more prevalent than economic frontiers in Amsterdam and the Hague.

Chapter 4 focused on exploring the evolution of social frontiers over their lifecycle, making it the first study to graph frontier strength. I found that, for frontiers that dissipated, their strength increased prior to and during their formation, before stabilising as the frontier became established, and finally before their demise. This was consistent across regions, frontier durations, and dimensions, suggesting that most frontiers have a very similar cycle. Particularly of note was the high level of re-emergence: frontiers that disappeared reappeared in over 50% of cases. The re-emergence combined with the frontiers often increasing in strength following their classification as non-social frontiers, leads to the conclusion that frontiers are likely to have long-lasting effects as they often never fully disappear with communities unlikely to completely merge.

An additional point to note is the higher occurrence rate and longer average length of country of birth frontiers compared to economic frontiers in Amsterdam and the Hague. This corresponds with the higher index of dissimilarity for the country of birth dimension for the majority of the time and the literature surrounding patterns of segregation in the Netherlands (Zorlu and Mulder, 2008; Boterman, Musterd and Manting, 2021).

Recommendation:

The main methodological recommendation of the chapter relates to the dynamics of social frontiers. Future studies should investigate how the frontier impact varies depending on at what stage of their lifecycle individuals are exposed to them at. Based on existing literature and this chapter, I would expect frontiers to be at their most impactful when they are forming, with both the territory and boundary unstable and clearly the communities experiencing the largest change as visible in Chapter 4 (Legewie and Schaeffer, 2016; Dean *et al.*, 2019; Iyer and Pryce, 2024). By including a duration measure in the chapter, I hope that future social frontier literature can be extended to explore whether the frontier effect differs depending on exposure length time and the duration of the frontier.

In terms of policy implications for this chapter, these mostly arise from the differences in the prevalence of the economic and country of birth frontiers. It is likely that the loss of social housing has contributed to the growing prevalence of low-income and unemployment frontiers with individuals potentially being forced out of the city centre by involuntary gentrification. If there is a need to reduce social frontiers, then this could likely be achieved in part by diversifying the housing stock to attract new households to the neighbourhood. However, this diversification is not to be at the cost of social housing which is still in high demand. Alternative policies could include the construction of more community spaces, with the aim of facilitating

positive contact between residents and encouraging them to build relationships (van Kempen and Bolt, 2019).

Conceptual Contribution: Intersectional social frontier definitions

Chapter 5 makes the largest conceptual contribution of the thesis to social frontier literature through the development of the intersectional social frontier definitions. Previous social frontier literature has touched upon intersectional frontiers, noting that we should consider boundaries which are “co-occurring across multiple socio-demographic domains” (Iyer and Pryce, 2024, p. 784) and that these frontiers may vary in their impact depending on the combination of the dimensions (Dean *et al.*, 2019, p. 286). However, this is the extent to which they have been discussed in the literature, leaving a large space for the interpretation of the term. In Chapter 5 I introduced two types of intersectional frontiers: coterminous and multidimensional. I discussed these multiple interpretations in relation to how they overlap, the role they may play in society, and their prevalence in the regions of Amsterdam and the Hague. This is the first comprehensive discussion of intersectional social frontiers and poses an important dilemma: how should policies be adapted when intersectional frontiers are increasing in prevalence whilst the opposite is true for the single dimension frontiers.

Recommendation:

This chapter can act as a framework for future research studying this topic. Researchers can choose which definition they believe is most suited to their area of research, depending on whether they are interested in how social frontiers overlap (coterminous) or how individuals who are part of multiple groups experience divisions in communities (multidimensional). It is particularly of interest to evaluate the difference in the effect of the intersectional frontiers compared to the single frontiers, along with the difference in the effect of the coterminous and multidimensional intersectional frontiers.

In relation to policies this is likely influenced by the difference in the effect of single dimension and intersectional frontiers, particularly the multidimensional ones. If there exists a difference then specific targeted policies would need to be introduced to address the intersectional frontiers. Crenshaw’s (1989) intersectionality literature indicates that current policies may be bypassing those who experience multiple levels of discrimination or divisions, with the nature of the divisions naturally more complex than single dimension frontiers. This is an important insight for urban policies aimed at reducing segregation or spatial exclusion. Strategies such as the Startblok programme (<https://startblokkriekerveen.nl/en/>) in Amsterdam aim to create a community of young people with a mixture of those born in the Netherlands and refugees - imposing a maximum yearly income with the aim of providing affordable housing. Initiatives such as these have been deemed as a softer entry into Dutch society, with an opportunity to have greater contact with those born in the Netherlands and a chance to learn Dutch (Kim and Smets, 2020). Studies have

identified that relationships and networks have been formed as a result of Startblok, with individuals from various ethnic backgrounds forming bonds with each other (Czischke and Huisman, 2018). This programme is providing support to those who are both low-income and non-Western, with the idea of creating communities across the different characteristics.

Finding 3: Living near social frontiers during childhood has a very small negative association with income and educational outcomes in adulthood.

As the final empirical chapter, Chapter 6 is the only study in the thesis to focus on the effect of social frontiers. It is one of the very few pieces of social frontier literature which considers the impact of the frontiers, particularly their impact on children and young people. This is despite neighbourhood effects literature highlighting the importance of the neighbourhood on children due to the extended period of time they spend there (Sykes and Kuyper, 2009). The chapter finds that living near social frontiers continuously between the ages of 11 and 15 is associated with a small reduction in earnings of between €458 and €733 a year between the ages of 20 and 25, and a very small reduction in years spent in education. This finding suggests that living near social frontiers can be detrimental for life outcomes in adulthood.

However, it is important to note that the effects I identify are very small and the sensitivity testing means that they cannot be claimed as causal. Thus, I must be cautious in making large claims in relation to the significance of social frontiers on individuals. The existence of an effect highlights the importance of future social frontier research – which could potentially identify which factors result in social frontiers having a greater or lesser impact - along with creating policies aimed at reducing divisions between communities.

Recommendation:

The negative association between country of birth social frontiers, and education and income outcomes in adulthood brings awareness to the need for policies aimed at reducing these frontiers or addressing the effects that these frontiers result in. In terms of reducing social frontiers this could potentially be altered by decisions related to neighbourhood planning and housing allocation. Bolt, Hooimeijer and van Kempen (2002) noted in the Netherlands that clusters of ethnic and low-income individuals often arise in areas where the housing is of poorer quality and thus to reduce the impact or presence of social frontiers improvements must be made to existing housing along with residents being given the opportunity to leave their neighbourhood if they desire. Based on the Chetty, Hendren and Katz (2016) paper the ability to move to a less deprived neighbourhood can be hugely beneficial for individuals. Smets and den Uyl (2008) suggested that to encourage social mixing, it may be best to not target neighbourhood mixing, instead focusing on education and workplace mixing. I discussed in Chapter 3 how the chapter, and in fact the whole thesis, focuses on place-based segregation, through applying a more person-based approach future researchers may be able to identify whether these effects are magnified if individuals are also divided in their workplace or social network.

Of potentially greater importance than reducing social frontiers, is addressing the mechanisms which explain the frontier effect on society: increasing crime rates (Dean *et al.*, 2019), feelings of relative deprivation (Iyer and Pryce, 2024), and becoming isolated from communities with limited access to role models and resources (Galster, 2012; Troost, van Ham and Manley, 2023). All of these factors could have significant consequences for the life trajectories of children growing up near social frontiers. Nieuwenhuis *et al.* (2017, p. 1902) said that policies should not aim to increase social mixing as this is unlikely to fix all the problems, instead it is about providing equal opportunities for those whose neighbourhood may be detrimental for them. Chetty, Hendren and Katz (2016) found that moving to a lower poverty area at a young age resulted in an increase in earnings of around \$302,000 across a lifetime, thus suggesting that it could be the factor of living in deprived areas which has a greater impact upon outcomes. In terms of addressing the issues of individuals feeling excluded or isolated these could be achieved through the introduction of more community spaces - specifically those crossing the social frontiers - aimed at encouraging positive communication and building relationships between contrasting communities (Putnam, 2007; Piekut *et al.*, 2024). Völker, Flap and Lindenberg (2007) found in that in the Netherlands communities were more likely to form when shared spaces and interests existed, as they provided individuals with an opportunity to meet and form networks. Shared community spaces encourage positive intergroup contact - by blurring the boundaries - which has a role in reducing conflict as well as reducing the formation of prejudiced beliefs, helping to build beneficial relationships for individuals in both communities (Paolini *et al.*, 2004; Al Ramiah and Hewstone, 2013).

Limitations and Future Research

Although this thesis makes a large contribution to the literature, it is not without its limitations and these bring with them key suggestions for future social frontier literature.

Part of what we discuss within the introduction (Chapter 1) is the conceptualisation of social frontiers, whether they exist in reality and what element of them is created through the manipulation of the data by the researcher. The nature of the study of social frontiers means that there is a requirement for the researcher to define what is classified as a 'neighbourhood' which in this study is described based on grid cells created by Statistics Netherlands. Thus, there is an element of social frontiers which have been manufactured to exist (be identified) along these grid cell boundaries and depend on the methods implemented.

Dean *et al.* (2019, p. 279) stated that social frontiers are perhaps better described as boundaries where there is an overlap between administrative boundaries and steep step changes in the socioeconomic characteristics of the neighbourhoods. As a

quantitative study I cannot guarantee that the frontiers I have identified truly exist for those who live within the communities, however, Zhang et al (2024) conducted a qualitative study using social frontiers identified by areal wombling and found that “participants were more likely to recognise boundaries with higher boundary values as local community borders” (p. 1). Highlighting that despite them being conceptualised and partially manufactured by the researcher within the space, social frontiers are likely to correspond with the neighbourhood boundaries identified by residents and as such have an impact upon them.

In addition to the impact of grid cell choice, there is a lack of consideration in this study for how the spatial arrangement of residents within the grid cells impacts the existence and effect of the frontiers. This is a clear limitation of the study; it is partially an example of the modified areal unit problem with the aggregated nature of the data meaning that part of the spatial element of where people live is ignored. The extent of the data available from Statistics Netherlands means that future studies could look to include more point-level data potentially including an exact distance calculation, like in Maguire, French and O'Reilly (2016), and considering the patterns within the grid cells. Additionally, when studying the frontier effect studies could explore whether the impact is greater in certain areas of the region by studying specific individuals or through the use a buffer zone similar to that in Harris (2014) where an area is drawn around the frontier to explore whether the effect varies based on the distance from the frontier.

Another aspect future researchers could consider is the relationship between social frontiers and physical boundaries, such as the ‘Peace Walls’ in Belfast, which are likely to divide communities but may not sit on administrative boundaries. Nor does it consider the existence of established communities within the city/region and how this is likely to impact the level of frontiers. Spierenburg *et al.* (2025) found no alignment between social frontiers and physical boundaries such as railways or roads, however, barriers such as the ‘Peace Walls’ in Belfast would be clear examples of physical barriers which have clearly divided communities and as such likely coincide with social frontiers. Prior research has identified the impact that these walls have on mental health (Maguire, French and O'Reilly, 2016) and contact between contrasting groups (Dixon *et al.*, 2025). Exploring how social frontiers coincide with physical barriers, and whether their impact differs depending on this, could be explored through mapping the social frontiers on physical maps which show nature and man-made barriers. This is related to my discussion about the excess embrittlement level in Chapter 3, about how the measure fails to consider the broader spatial arrangement and context beyond the city/region to consider how these impact the level of frontiers. As a whole, social frontier studies could benefit from discussing aspects beyond the area of study to understand why and how people choose to settle in an area and what impact this has on the frontier pattern and prevalence.

There is always a risk that the identified social frontiers are manufactured based on the methodological choices made within a study. Especially when using the PAM clustering method, as there is a risk that the method will artificially create social frontiers due to the nature of clustering always splitting the grid cells into two groups. Thus, a frontier will always be identified even if the actual arrangement of the grid cells results in a spatially smooth distribution. Thus, conclusions and findings must be stated cautiously with an element of trepidation, we cannot, without a qualitative study, be sure that the frontiers which have been identified are all equally experienced or even felt within a neighbourhood. However, the inclusion of the strength measurement does address this in part and it would be beneficial for this to be included within the causal study.

Additionally, throughout the thesis the social frontiers are all identified using a binary frontier definition. This is addressed in part in Chapter 4 through the inclusion of a strength measurement to explore the variation in the steepness of the frontiers. However, further studies could look to describe frontiers depending on their strength or include it within effects studies. Kramer (2017) used a staggered scale to describe boundaries which also involved a description of their consistency over time. With the introduction of the strength and duration measures within this thesis, future studies could look to describe the frontiers based on these properties and explore whether the strength and duration alters the impact of the frontiers.

Concluding Remarks

This thesis is by no means a complete exploration of social frontiers, nor is it without its limitations. I have highlighted these limitations and how future studies could address the gaps in the literature within this conclusion. However, the thesis does address several gaps in the literature, providing the first identification of social frontiers in multiple dimensions, time periods, and regions; investigating the effect of childhood exposure to social frontiers on outcomes in adulthood; plotting the lifecycle of these frontiers; and identifying and defining intersectional frontiers. I hope that throughout the thesis the research recommendations provide future researchers with exciting opportunities for creating a better understanding of social frontiers as well as yielding insights for social policies.

References

- Aalbers, M. B. and Deurloo, R. (2003) 'Concentrated and condemned? Residential patterns of immigrants from industrial and non-industrial countries in Amsterdam', *Housing, Theory and Society*, 20(4), pp. 197–208. doi: <https://doi.org/10.1080/14036090310018932>
- Aalbers, M. B. (2007) 'Place-Based and Race-Based Exclusion from Mortgage Loans: Evidence from Three Cities in the Netherlands', *Journal of Urban Affairs*, 29(1), pp. 1–29. doi: <https://doi.org/10.1111/j.1467-9906.2007.00320.x>
- Albada, K., Hansen, N. and Otten, S. (2021) 'Polarization in attitudes towards refugees and migrants in the Netherlands', *European Journal of Social Psychology*, 51(3), pp. 627–643. doi: <https://doi.org/10.1002/ejsp.2766>
- Allender, S. *et al.* (2012) 'Relative deprivation between neighbouring wards is predictive of coronary heart disease mortality after adjustment for absolute deprivation of wards', *Journal of Epidemiology and Community Health*, 66(9), pp. 803–808. doi: <https://doi.org/10.1136/jech.2010.116723>
- Allport, G. W. (1954) *The nature of prejudice*. Cambridge, Mass: Addison-Wesley.
- Al Ramiah, A. and Hewstone, M. (2013) 'Intergroup contact as a tool for reducing, resolving, and preventing intergroup conflict: Evidence, limitations, and potential', *American Psychologist*, 68(7), pp. 527–542. doi: <https://doi.org/10.1037/a0032603>
- Andersson, E. K. *et al.* (2018) 'A Comparative Study of Segregation Patterns in Belgium, Denmark, the Netherlands and Sweden: Neighbourhood Concentration and Representation of Non-European Migrants', *European Journal of Population*, 34(2), pp. 251–275. doi: <https://doi.org/10.1007/s10680-018-9481-5>
- Andersson, E. *et al.* (2023) 'Contextual poverty and obtained educational level and income in Sweden and the Netherlands: A multi-scale and longitudinal study', *Urban Studies*, 60(5), pp. 885–903. doi: <https://doi.org/10.1177/00420980221120492>
- Arat, E. and Biligili, O. (2021) 'Transnational and Local Co-ethnic Social Ties as Coping Mechanisms Against Perceived Discrimination - A Study on the Life Satisfaction of Turkish and Moroccan Minorities in the Netherlands', *Frontiers in Sociology*, 7. doi: <https://doi.org/10.3389/fsoc.2021.671897>
- Austin, P. C. (2009) 'Balance diagnostics for comparing the distribution of baseline covariates between treatment groups in propensity-score matched samples', *Statistics in Medicine*, 28(25), pp. 3083–3107. doi: <https://doi.org/10.1002/sim.3697>

- Austin, P. C. (2011) 'An Introduction to Propensity Score Methods for Reducing the Effects of Confounding in Observational Studies', *Multivariate Behavioral Research*, 46(3), pp. 399–424. doi: <https://doi.org/10.1080/00273171.2011.568786>
- Bailey, N. (2012) 'How Spatial Segregation Changes over Time: Sorting Out the Sorting Processes', *Environment and Planning A: Economy and Space*, 44(3), pp. 705–722. doi: <https://doi.org/10.1068/a44330>
- Bailey, N., van Gent, W. P. C. and Musterd, S. (2017) 'Remaking Urban Segregation: Processes of Income Sorting and Neighbourhood Change', *Population, Space and Place*, 23(3). doi: <https://doi.org/10.1002/psp.2013>
- Bannister, J., Kintrea, K. and Pickering, J. (2013) 'Young people and violent territorial conflict: exclusion, culture and the search for identity', *Journal of Youth Studies*, 16(4), pp. 474–490. doi: <https://doi.org/10.1080/13676261.2012.725835>
- Bauer, G. *et al.* (2021) 'Intersectionality in quantitative research: A systematic review of its emergence and applications of theory and methods', *SSM - Population Health*, 14. doi: <https://doi.org/10.1016/j.ssmph.2021.100798>
- Beckers, P. and Boschman, S. (2019) 'Residential choices of foreign highly skilled workers in the Netherlands and the role of neighbourhood and urban regional characteristics', *Urban Studies*, 56(4), pp. 760–777. doi: <https://doi.org/10.1177/0042098017741262>
- Belastingdienst (no date) *Box 1: uitleg en tarieven*. Available at: https://www.belastingdienst.nl/wps/wcm/connect/bldcontentnl/belastingdienst/prive/inkomstenbelasting/heffingskortingen_boxen_tarieven/boxen_en_tarieven/box_1/ (Accessed: 24 November 2025).
- Belfi, B. *et al.* (2022) 'Early career trajectories of first- and second-generation migrant graduates of professional university', *Journal of Ethnic and Migration Studies*, 48(10), pp. 2415–2435. doi: <https://doi.org/10.1080/1369183X.2021.1935666>
- Berdahl, J. L. and Moore, C. (2006) 'Workplace harassment: Double jeopardy for minority women', *Journal of Applied Psychology*, 91(2), pp. 426–436. doi: <https://doi.org/10.1037/0021-9010.91.2.426>
- Bezzo, F.B. and Vanhuyse, P. (2024) 'Better to Grow Up Poor in a Richer Place? Social Housing, Neighbourhood Comparisons, and English Teenagers' Well-Being', *Journal of Happiness Studies*, 25(3), pp. 1–20. doi: <https://doi.org/10.1007/s10902-024-00740-z>
- Bi, K. *et al.* (2025) 'Navigating Unmeasured Confounding in Nonexperimental Psychological Research: A Practical Guide to Computing and Interpreting E-Value',

Advances in Methods and Practices in Psychological Science, 8(2), pp. 1–21. doi: <https://doi.org/10.1177/25152459251326571>

Bisschop, P., ter Weel, B. and Zwetsloot, J. (2020) 'Ethnic Employment Gaps of Graduates in the Netherlands', *De Economist*, 168, pp. 577–598. doi: <https://doi.org/10.1007/s10645-020-09375-w>

Blackwell, M. *et al.* (2009) 'Cem: Coarsened Exact Matching in Stata', *The Stata Journal: Promoting Communications on Statistics and Stata*, 9(4), pp. 524–546. doi: <https://doi.org/10.1177/1536867X0900900402>

Blanden, J. (2019) 'Intergenerational income persistence', *IZA World of Labor*. doi: <https://doi.org/10.15185/izawol.176.v2>

Blankvoort, N. *et al.* (2021) 'Decolonising civic integration: a critical analysis of texts used in Dutch civic integration programmes', *Journal of Ethnic and Migration Studies*, 47(15), pp. 3511–3530. doi: <https://doi.org/10.1080/1369183X.2021.1893668>

Boal, F. W. and Livingstone, D. N. (1984) 'The Frontier in the City: Ethnonationalism in Belfast', *International Political Science Review / Revue Internationale de Science Politique*, 5(2), pp. 161–179. doi: <https://doi.org/10.1177/019251218400500207>

Bolt, G., Hooimeijer, P. and van Kempen, R. (2002) 'Ethnic Segregation in the Netherlands: New Patterns, New Policies?', *Tijdschrift voor Economische en Sociale Geografie*, 93(2), pp. 214–220. doi: <https://doi.org/10.1111/1467-9663.00196>

Bolt, G., van Kempen, R. and van Ham, M. (2008) 'Minority Ethnic Groups in the Dutch Housing Market: Spatial Segregation, Relocation Dynamics and Housing Policy', *Urban Studies*, 45(7), pp. 1359–1384. doi: <https://doi.org/10.1177/0042098008090678>

Bolt, G., van Kempen, R. and van Weesep, J. (2009) 'After Urban Restructuring: Relocations and Segregation in Dutch Cities', *Tijdschrift voor Economische en Sociale Geografie*, 100(4), pp. 502–518. doi: <https://doi.org/10.1111/j.1467-9663.2009.00555.x>

Bolt, G. and van Kempen, R. (2010) 'Ethnic Segregation and Residential Mobility: Relocations of Minority Ethnic Groups in the Netherlands', *Journal of Ethnic and Migration Studies*, 36(2), pp. 333–354. doi: <https://doi.org/10.1080/13691830903387451>

Bolt, G. and van Kempen, R. (2011) 'Successful Mixing? Effects of Urban Restructuring Policies in Dutch Neighbourhoods', *Tijdschrift voor Economische en Sociale Geografie*, 102(3), pp. 361–368. doi: <https://doi.org/10.1111/j.1467-9663.2011.00668.x>

Bonjour, S. and Duyvendak, J. W. (2018) 'The "migrant with poor prospects": racialized intersections of class and culture in Dutch civic integration debates', *Ethnic and Racial Studies*, 41(5), pp. 882–900. doi: <https://doi.org/10.1080/01419870.2017.1339897>

Boschman, S. and van Ham, M. (2015) 'Neighborhood selection of non-Western ethnic minorities: testing the own-group effects hypothesis using a conditional logit model', *Environment and Planning A: Economy and Space*, 47(5), pp. 1155–1174. doi: <https://doi.org/10.1177/0308518X15592300>

Bosma, U. (2012a) 'Introduction: Post-colonial immigrants and identity formations in the Netherlands', in Bosma, U. (ed.) *Post-colonial Immigrants and Identity Formations in the Netherlands*. Amsterdam: Amsterdam University Press, pp. 7–26.

Bosma, U. (2012b) 'Why is there no post-colonial debate in the Netherlands?', in Bosma, U. (ed.) *Post-colonial Immigrants and Identity Formations in the Netherlands*. Amsterdam: Amsterdam University Press, pp. 193–212.

Bosswick, W. and Heckmann, F. (2006) *Integration of Migrants: The Role of Local and Regional Authorities*. Available at: <https://coilink.org/20.500.12592/rk1xv2> (Accessed: 21 October 2025).

Boterman, W. R., Musterd, S. and Manting, D. (2021) 'Multiple dimensions of residential segregation. The case of the metropolitan area of Amsterdam', *Urban Geography*, 42(4), pp. 481–506. doi: <https://doi.org/10.1080/02723638.2020.1724439>

Boterman, W. and van Gent, W. (2023) *Making the middle-class city: the politics of gentrifying Amsterdam*. New York, NY: Palgrave Macmillan.

Bowen, N. K. and Bowen, G. L. (1999) 'Effects of Crime and Violence in Neighborhoods and Schools on the School Behavior and Performance of Adolescents', *Journal of Adolescent Research*, 14(3), pp. 319–342. doi: <https://doi.org/10.1177/0743558499143003>

Brattbakk, I. and Wessel, T. (2012) 'Long-term Neighbourhood Effects on Education, Income and Employment among Adolescents in Oslo', *Urban Studies*, 50(2), pp. 391–406. doi: <https://doi.org/10.1177/0042098012448548>

Bucheli, J. F. (2022) 'The influence of the 'Fragmented City' on wellbeing: Do mixed neighbourhoods affect young adults' life trajectories in Bogota?', *Journal of Urban Management*, 11(4), pp. 450–466. doi: <https://doi.org/10.1016/j.jum.2022.08.003>

Buis, M. L. (2013) 'The Composition of Family Background: The Influence of the Economic and Cultural Resources of both Parents on the Offspring's Educational

Attainment in the Netherlands between 1939 and 1991', *European Sociological Review*, 29(3), pp. 593–602. doi: <https://doi.org/10.1093/esr/jcs009>

Burgers, J. and Musterd, S. (2002) 'Understanding Urban Inequality: A Model Based on Existing Theories and an Empirical Illustration', *International Journal of Urban and Regional Research*, 26(2), pp. 403–413. doi: <https://doi.org/10.1111/1468-2427.00387>

Buyukkececi, Z. (2023) 'Neighbourhood effects on early adulthood family life courses: A trajectory-based approach', *Population, Space and Place*, 29(2), pp. 1–14. doi: <https://doi.org/10.1002/psp.2609>

CBS (2022a) 'Population; sex, age, generation and migration background, 1 Jan; 1996-2022'. Available at: <https://opendata.cbs.nl/#/CBS/en/dataset/37325eng/table?ts=1755862763498> (Accessed: 22 August 2025).

CBS (2022b) *How many people immigrate to the Netherlands?* Available at: <https://www.cbs.nl/en-gb/dossier/asylum-migration-and-integration/how-many-people-immigrate-to-the-netherlands-> (Accessed: 22 August 2025).

CBS (2022c) *New classification of population by origin*. Available at: <https://www.cbs.nl/en-gb/longread/statistische-trends/2022/new-classification-of-population-by-origin?onpage=true#c-2--The-new-classification-by-origin-in-brief> (Accessed: 27 August 2025).

CBS (2023a) *Over 40 percent more asylum applications in 2022*. Available at: <https://www.cbs.nl/en-gb/news/2023/05/over-40-percent-more-asylum-applications-in-2022> (Accessed: 22 August 2025).

CBS (2023b) 'External migration; sex, age, marital status, country of birth; 1995-2022'. Available at: <https://opendata.cbs.nl/#/CBS/en/dataset/03742eng/table?ts=1766405804377> (Accessed: 22 December 2025).

CBS (2023c) *Maps regional divisions 2023*. Available at: https://www.cbs.nl/-/media/_pdf/2016/42/2016ov69-coropgebieden-2016-web.pdf (Accessed: 20 December 2025).

CBS (2023d) *More women than men in higher education for 23 consecutive years*. Available at: <https://www.cbs.nl/en-gb/news/2023/10/more-women-than-men-in-higher-education-for-23-consecutive-years> (Accessed: 23 July 2025).

CBS (2024a) *New method of measuring poverty: 540 thousand people in poverty in 2023*. Available at: <https://www.cbs.nl/en-gb/news/2024/42/new-method-of->

measuring-poverty-540-thousand-people-in-poverty-in-2023 (Accessed: 24 November 2025).

CBS (2024b) *Most family migrants stay in the Netherlands*. Available at: <https://www.cbs.nl/en-gb/news/2024/44/most-family-migrants-stay-in-the-netherlands> (Accessed: 18 August 2025).

CBS (2024c) *Integratie en samenleven 2024*. Available at: <https://www.cbs.nl/en-gb/publication/2024/47/report-on-integration-and-society-2024> (Accessed: 18 August 2025).

CBS (2024d) *Slight increase in number of unemployed persons*. Available at: <https://www.cbs.nl/en-gb/news/2024/42/slight-increase-in-number-of-unemployed-persons> (Accessed: 15 November 2024).

CBS (2024e) *User's Guide Educational Information*. Available at: <https://www.cbs.nl/-/media/cbs-op-maat/zelf-onderzoek-doen/handleiding-ssb-onderwijs.pdf> (Accessed: 14 October 2025).

CBS (2025a) *Person with a non-western migration background*. Available at: <https://www.cbs.nl/en-gb/our-services/methods/definitions/person-with-a-non-western-migration-background> (Accessed: 27 August 2025).

CBS (2025b) *Person with a western migration background*. Available at: <https://www.cbs.nl/en-gb/our-services/methods/definitions/person-with-a-western-migration-background> (Accessed: 20 November 2025).

CBS (2025c) 'Population; sex, age, country of origin, country of birth'. Available at: <https://opendata.cbs.nl/#/CBS/en/dataset/85384ENG/table?ts=1755862973851> (Accessed: 22 August 2025).

CBS (2025d) 'External migration; sex, age, marital status and country of birth'. Available at: <https://opendata.cbs.nl/#/CBS/en/dataset/85469ENG/table?ts=1755863698501> (Accessed: 22 August 2025).

CBS (2025e) 'Asylum requests and family reunification; nationality, sex and age'. Available at: <https://opendata.cbs.nl/#/CBS/en/dataset/83102ENG/table> (Accessed: 3 September 2025).

CBS (2025f) 'Housing stock; characteristics of dwellings and occupancy on December 31, region'. Available at: <https://opendata.cbs.nl/#/CBS/nl/dataset/86079NED/table?searchKeywords=woning%20voorraad%20naar%20eigendom> (Accessed: 7 November 2025).

CBS (2025g) 'Key figures for districts and neighborhoods 2020.' Available at: <https://opendata.cbs.nl/#/CBS/nl/dataset/84799NED/table?dl=41062> (Accessed: 16 December 2025).

CBS (2025h) *Gender pay gap is narrowing*. Available at: <https://www.cbs.nl/en-gb/news/2025/18/gender-pay-gap-is-narrowing> (Accessed: 23 July 2025).

CBS (2026) 'Existing own homes; average purchase prices, region'. Available at: <https://opendata.cbs.nl/#/CBS/en/dataset/83625ENG/table> (Accessed: 28 May 2026).

Cheshire, P. (2009) 'Policies for Mixed Communities: Faith-Based Displacement Activity? Faith-Based Displacement Activity?', *International Regional Science Review*, 32(3), pp. 343–375. doi: <https://doi.org/10.1177/0160017609336080>

Chesnaye, N. C. *et al.* (2022) 'An introduction to inverse probability of treatment weighting in observational research', *Clinical Kidney Journal*, 15(1), pp. 14–20. doi: <https://doi.org/10.1093/ckj/sfab158>

Chetty, R., Hendren, N. and Katz, L. F. (2016) 'The Effects of Exposure to Better Neighborhoods on Children: New Evidence from the Moving to Opportunity Experiment.', *American Economic Review*, 106(4), pp. 855–902. doi: <https://doi.org/10.1257/aer.20150572>

Chetty, R. and Hendren, N. (2018) 'The Impacts of Neighborhoods on Intergenerational Mobility II: County-Level Estimates', *The Quarterly Journal of Economics*, 133(3), pp. 1163–1228. doi: <https://doi.org/10.1093/qje/qjy006>

Choi, J., Dekkers, O. M. and le Cessie, S. (2019) 'A comparison of different methods to handle missing data in the context of propensity score analysis', *European journal of epidemiology*, 34(1), pp. 23–36. doi: <https://doi.org/10.1007/s10654-018-0447-z>

Christafore, D. and Leguizamon, S. (2019) 'Neighbourhood inequality spillover effects of gentrification', *Papers in Regional Science*, 98(3), pp. 1469–1485. doi: <https://doi.org/10.1111/pirs.12405>

Chyn, E. and Katz, L. F. (2021) 'Neighborhoods Matter: Assessing the Evidence for Place Effects', *Journal of Economic Perspectives*, 35(4), pp. 197–222. doi: <https://doi.org/10.1257/jep.35.4.197>

Coenders, Y. (2024) 'Colonial recursion: state categories of race and the emergence of the "Non-Western Allochthone"', *American Journal of Cultural Sociology*, 12(4), pp. 698–729. doi: <https://doi.org/10.1057/s41290-024-00214-y>

Colagrossi, M., Geraci, A. and Mazzarella, G. (2023) 'Intergenerational mobility in the Netherlands: models, outcomes and trends', *The Journal of Economic Inequality*, 21, pp. 775–788. doi: <https://doi.org/10.1007/s10888-023-09569-7>

'Coterminous (adj.), sense 1' (no date) in *Merriam-Webster*. Available at: <https://www.merriam-webster.com/dictionary/coterminous> (Accessed: 9 December 2025).

Crenshaw, K. (1989) 'Demarginalizing the Intersection of Race and Sex: A Black Feminist Critique of Antidiscrimination Doctrine, Feminist Theory and Antiracist Politics', *University of Chicago Legal Forum*, 1989(1), pp. 139–167. Available at: <http://chicagounbound.uchicago.edu/uclf/vol1989/iss1/8> (Accessed: 14 July 2025).

Crenshaw, K. (1991) 'Mapping the Margins: Intersectionality, Identity Politics, and Violence against Women of Color', *Stanford Law Review*, 43(6), pp. 1241–1299. doi: <https://doi.org/10.2307/1229039>

Cutler, D. M., Glaeser, E. L. and Vigdor, J. L. (2008) 'When are ghettos bad? Lessons from immigrant segregation in the United States', *Journal of Urban Economics*, 63(3), pp. 759–774. doi: <https://doi.org/10.1016/j.jue.2007.08.003>

Czischke, D. and Huisman, C. (2018) 'Integration through Collaborative Housing? Dutch Starters and Refugees Forming Self-Managing Communities in Amsterdam', *Urban Planning*, 3(4), pp. 156–165. doi: <https://doi.org/10.17645/up.v3i4.1727>

Damm, A. P. and Dustmann, C. (2014), 'Does Growing Up in a High Crime Neighborhood Affect Youth Criminal Behavior?', *American Economic Review*, 104(6), pp. 1806–1832. doi: <https://doi.org/10.1257/aer.104.6.1806>

De Backer, M. (2022) 'Between place and territory: Young people's emotional geographies of security and insecurity in Brussels' deprived areas', *Emotion, Space and Society*, 45. doi: <https://doi.org/10.1016/j.emospa.2022.100911>

de Vuijst, E. and van Ham, M. (2019) 'Parents and Peers: Parental Neighbourhood- and School-Level Variation in Individual Neighbourhood Outcomes Later in Life', *European Sociological Review*, 35(1), pp. 15–28. doi: <https://doi.org/10.1093/esr/jcy041>

Dean, N. *et al.* (2019) 'Frontiers in residential segregation: understanding neighbourhood boundaries and their impacts', *Tijdschrift Voor Economische en Sociale Geografie*, 110(3), pp. 271–288. doi: <https://doi.org/10.1111/tesg.12316>

Dekker, L. H., Rijnks, R. H. and Mierau, J. O. (2021) 'The health potential of neighborhoods: A population-wide study in the Netherlands', *SSM - Population Health*, 15. doi: <https://doi.org/10.1016/j.ssmph.2021.100867>

Dixon, J. *et al.* (2020) 'Parallel lives: Intergroup contact, threat, and the segregation of everyday activity spaces', *Journal of Personality and Social Psychology*, 118(3), pp. 457–480. doi: <https://doi.org/10.1037/pspi0000191>

Dixon, J. *et al.* (2022) 'Navigating the divided city: Place identity and the time-geography of segregation', *Journal of Environmental Psychology*, 84. doi: <https://doi.org/10.1016/j.jenvp.2022.101908>

Dixon, J. *et al.* (2025) "Crossing the line": A boundary transgression model of resistance to desegregation', *European Review of Social Psychology*, 36, pp. 1–42. doi: <https://doi.org/10.1080/10463283.2025.2550110>

Duncan, O. D. and Duncan, B. (1955) 'A Methodological Analysis of Segregation Indexes', *American Sociological Review*, 20(2), pp. 210–217. doi: <https://doi.org/10.2307/2088328>

Durlauf, S. N. (2004) 'Chapter 50 - Neighborhood Effects', in Henderson, J. V. and Thisse, J-F. (eds.) *Handbook of Regional and Urban Economics*. Edition 1. Elsevier, pp. 2173–2242.

Dutch News (2024) 'The rich, white Dutch live the most segregated lives: CBS', *Dutch News*, 23 February. Available at: <https://www.dutchnews.nl/2024/02/the-rich-white-dutch-live-the-most-segregated-lives-cbs/> (Accessed: 30 July 2025).

Entzinger, H. (2019) 'A Tale of Two Cities: Rotterdam, Amsterdam and Their Immigrants', in: Scholten, P., Crul, M. and van de Laar, P. (eds.) *Coming to Terms with Superdiversity*. Cham, Switzerland: Springer, pp. 173–189.

Erola, J., Jalonen, S. and Lehti, H. (2016) 'Parental education, class and income over early life course and children's achievement', *Research in Social Stratification and Mobility*, 44, pp. 33–43. doi: <https://doi.org/10.1016/j.rssm.2016.01.003>

European Commission (2024) *2024 Country Report - Netherlands*. Available at: https://economy-finance.ec.europa.eu/publications/2024-european-semester-country-reports_en (Accessed: 14 October 2024).

Fainstein, S. S. (1997) 'The Egalitarian city: The restructuring of Amsterdam', *International Planning Studies*, 2(3), pp. 295–314. doi: <https://doi.org/10.1080/13563479708721686>

Fang, C. and van Liempt, I. (2021) "We prefer our Dutch": International students' housing experiences in the Netherlands', *Housing Studies*, 36(6), pp. 822–842. doi: <https://doi.org/10.1080/02673037.2020.1720618>

- Feijten, P. and van Ham, M. (2009) 'Neighbourhood Change... Reason to Leave?', *Urban Studies*, 46(10), pp. 2103–2122. doi: <https://doi.org/10.1177/0042098009339430>
- Feitosa, F. F. *et al.* (2007) 'Global and local spatial indices of urban segregation', *International Journal of Geographical Information Science*, 21(3), pp. 299–323. doi: <https://doi.org/10.1080/13658810600911903>
- Ferguson, K. M. (2006) 'Social capital and children's wellbeing: a critical synthesis of the international social capital literature', *International Journal of Social Welfare*, 15(1), pp. 2–18. doi: <https://doi.org/10.1111/j.1468-2397.2006.00575.x>
- Fowler, C. S. (2015) 'Segregation as a multiscalar phenomenon and its implications for neighborhood-scale research: the case of South Seattle 1990–2010', *Urban Geography*, 37(1), pp. 1–25. doi: <https://doi.org/10.1080/02723638.2015.1043775>
- Franzen, A. and Hangartner, D. (2006) 'Social Networks and Labour Market Outcomes: The Non-Monetary Benefits of Social Capital', *European Sociological Review*, 22(4), pp. 353–368. doi: <https://doi.org/10.1093/esr/jcl001>
- Galster, G. C. *et al.* (2008) 'Does neighborhood income mix affect earnings of adults? New evidence from Sweden', *Journal of Urban Economics*, 63(3), pp. 858–870. doi: <https://doi.org/10.1016/j.jue.2007.07.002>
- Galster, G. C. (2012) 'The Mechanism(s) of Neighbourhood Effects: Theory, Evidence, and Policy Implications', in: van Ham, M. *et al.* (eds.) *Neighbourhood Effects Research: New Perspectives*. Dordrecht: Springer, pp 23–56.
- Galster, G. C. and Sharkey, P. (2017) 'Spatial Foundations of Inequality: A Conceptual Model and Empirical Overview', *RSF: The Russell Sage Foundation Journal of the Social Sciences*, 3(2), pp. 1–33. doi: <https://doi.org/10.7758/rsf.2017.3.2.01>
- Garner, C. L. and Raudenbush, S. W. (1991) 'Neighborhood effects on educational attainment: A multilevel analysis', *Sociology of Education*, 64(4), pp. 251–262. doi: <https://doi.org/10.2307/2112706>
- Garrido, M. M. *et al.* (2014) 'Methods for constructing and assessing propensity scores', *Health Services Research*, 49(5), pp. 1701–1720. doi: <https://doi.org/10.1111/1475-6773.12182>
- Ghorashi, H. (2014) 'Racism and “the Ungrateful Other” in the Netherlands', in: Essed, P. and Hoving, I. (eds.) *Dutch Racism*. Amsterdam, Netherlands: Rodopi, pp. 101–116.

- Gijsberts, M. and Dagevos, J. (2007) 'The Socio-cultural Integration of Ethnic Minorities in the Netherlands: Identifying Neighbourhood Effects on Multiple Integration Outcomes', *Housing Studies*, 22(5), pp. 805–831. doi: <https://doi.org/10.1080/02673030701474768>
- Glaeser, E. (2011) *Triumph of the City: How Our Greatest Invention Makes Us Richer, Smarter, Greener, Healthier and Happier*. United Kingdom: Pan Macmillan.
- Gold, J. R. (1982) 'Territoriality and human spatial behaviour', *Progress in Human Geography*, 6(1), pp. 44–67. doi: <https://doi.org/10.1177/030913258200600102>
- Goodman, E. *et al.* (2001) 'Adolescents' perceptions of social status: development and evaluation of a new indicator', *Pediatrics*, 108(2). doi: <https://doi.org/10.1542/peds.108.2.e31>
- Gorodzeisky, A. and Semyonov, M. (2017) 'Labor force participation, unemployment and occupational attainment among immigrants in West European countries', *PloS One*, 12(5). doi: <https://doi.org/10.1371/journal.pone.0176856>
- Government of the Netherlands (no date-a) *Civic integration (inburgering) in the Netherlands*. Available at: <https://www.government.nl/topics/integration-in-the-netherlands/civic-integration-in-the-netherlands> (Accessed: 29 August 2025).
- Government of the Netherlands (no date-b) *Rented Housing*. Available at: <https://www.government.nl/topics/housing/rented-housing> (Accessed: 6 April 2023).
- Granovetter, M. S. (1973) 'The Strength of Weak Ties', *American Journal of Sociology*, 78(6), pp. 1360–1380. Available at: <http://www.jstor.org/stable/2776392> (Accessed: 1 December 2025).
- Greenland, S. (2003) 'Quantifying Biases in Causal Models: Classical Confounding vs Collider-Stratification Bias', *Epidemiology*, 14(3), p 300–306. doi: <https://doi.org/10.1097/01.EDE.0000042804.12056.6C>
- Greifer, N. and Stuart, E. A. (2021) 'Matching Methods for Confounder Adjustment: An Addition to the Epidemiologist's Toolbox', *Epidemiologic Reviews*, 43(1), pp. 118–129. doi: <https://doi.org/10.1093/epirev/mxab003>
- Greifer, N. (2025) *Estimating Effects After Matching*. Available at: <https://cran.r-project.org/web/packages/MatchIt/vignettes/estimating-effects.html> (Accessed: 24 December 2025)
- Greitemeyer, T. and Sagioglou, C. (2016) 'Subjective socioeconomic status causes aggression: A test of the theory of social deprivation', *Journal of Personality and Social Psychology*, 111(2), pp. 178–194. doi: <https://doi.org/10.1037/pspi0000058>

- Haandrikman, K. *et al.* (2023) 'Socio-economic segregation in European cities. A comparative study of Brussels, Copenhagen, Amsterdam, Oslo and Stockholm', *Urban Geography*, 44(1), pp. 1–36. doi: <https://doi.org/10.1080/02723638.2021.1959778>
- Hainmueller, J. (2012) 'Entropy Balancing for Causal Effects: A Multivariate Reweighting Method to Produce Balanced Samples in Observational Studies', *Political Analysis*, 20(1), pp. 25–46. doi: <https://doi.org/10.1093/pan/mpr025>
- Halonen, J. I. *et al.* (2014) 'Adverse experiences in childhood, adulthood neighbourhood disadvantage and health behaviours', *Journal of Epidemiology and Community Health*, 68(8), pp. 741–746. doi: <https://doi.org/10.1136/jech-2013-203441>
- Harris, R. (2014) 'Measuring Changing Ethnic Separations in England: A Spatial Discontinuity Approach', *Environment and Planning A: Economy and Space*, 46(9), pp. 2243–2261. doi: <https://doi.org/10.1068/a130021p>
- Hartog, J. and Zorlu, A. (2009) 'Ethnic segregation in The Netherlands: An analysis at neighbourhood level', *International Journal of Manpower*, 3(1–2), pp. 15–25. doi: <https://doi.org/10.1108/01437720910948366>
- Heath, A. F., Rethon, C. and Kilpi, E. (2008) 'The second generation in Western Europe: Education, unemployment, and occupational attainment', *Annual Review of Sociology*, 34, pp. 211–235. doi: <https://doi.org/10.1146/annurev.soc.34.040507.134728>
- Hedman, L., van Ham, M. and Manley, D. (2011) 'Neighbourhood Choice and Neighbourhood Reproduction', *Environment and Planning A: Economy and Space*, 43(6), pp. 1381–1399. doi: <https://doi.org/10.1068/a43453>
- Hedman, L. and van Ham, M. (2012) 'Understanding Neighbourhood Effects: Selection Bias and Residential Mobility', in: van Ham, M. *et al.* (eds.) *Neighbourhood Effects Research: New Perspectives*. Dordrecht: Springer, pp. 79–99.
- Hekwolter of Hekhuis, M., Nijskens, R. and Heeringa, W. (2017) *The housing market in major Dutch cities*, DNB Occasional Studies Volume 15 – 1. Available at: https://www.dnb.nl/media/ykmhc2el/201705_nr_1_-2017-_the_housing_market_in_major_dutch_cities.pdf (Accessed: 31 May 2026).
- Henning, C. and Lieberg, M. (1996) 'Strong ties or weak ties? Neighbourhood networks in a new perspective', *Scandinavian Housing and Planning Research*, 13(1), pp. 3–26. doi: <https://doi.org/10.1080/02815739608730394>
- Hill Collins, P. and Bilge, S. (2020) *Intersectionality*. 2nd edition. Cambridge: Polity.

- Ho, D. *et al.* (2011) 'MatchIt: Nonparametric Preprocessing for Parametric Causal Inference', *Journal of Statistical Software*, 42(8), pp. 1–28. doi: <https://doi.org/10.18637/jss.v042.i08>
- Hochstenbach, C. (2017) *Inequality in the gentrifying European city*. PhD Thesis, University of Amsterdam. Available at: <https://dare.uva.nl/search?identifier=c6943254-c4aa-4c31-b164-57382ee698aa> (Accessed: 15 September 2025).
- Hochstenbach, C. and Musterd, S. (2017) 'Gentrification and the suburbanization of poverty: changing urban geographies through boom and bust periods', *Urban Geography*, 39(1), pp. 26–53. doi: <https://doi.org/10.1080/02723638.2016.1276718>
- Hochstenbach, C. and Musterd, S. (2021) 'A regional geography of gentrification, displacement, and the suburbanisation of poverty: Towards an extended research agenda', *Area*, 53(3), pp. 481–491. doi: <https://doi.org/10.1111/area.12708>
- Hochstenbach, C. (2024) 'Framing the housing crisis: politicization and depoliticization of the Dutch housing debate', *Housing Studies*, 40(5), pp. 1226–1251. doi: <https://doi.org/10.1080/02673037.2024.2344844>.
- Hoekstra, M. (2015) 'Diverse cities and good citizenship: how local governments in the Netherlands recast national integration discourse', *Ethnic and Racial Studies*, 38(10), pp. 1798–1814. doi: <https://doi.org/10.1080/01419870.2015.1015585>
- Hoekstra, M. (2018) 'Governing difference in the city: urban imaginaries and the policy practice of migrant incorporation', *Territory, Politics, Governance*, 6(3), pp. 362–380. doi: <https://doi.org/10.1080/21622671.2017.1306456>
- Hofstede, G. (2001) *Culture's Consequences: Comparing Values, Behaviors, Institutions, and Organizations Across Nations*. 2nd edn. Thousand Oaks: Sage.
- Howard, A., Hochstenbach, C. and Ronald, R. (2023) 'Rental sector liberalization and the housing outcomes for young urban adults', *Urban Geography*, 45(7), pp. 1173–1194. doi: <https://doi.org/10.1080/02723638.2023.2279873>
- Iacus, S. M., King, G. and Porro, G. (2012) 'Causal Inference without Balance Checking: Coarsened Exact Matching', *Political Analysis*, 20(1), pp. 1–24. doi: <https://doi.org/10.1093/pan/mpr013>
- 'Intersectionality (n.), sense 2' (2023) in *OED Online*. Available at: <https://doi.org/10.1093/OED/8904687553>.
- Iyer, A. and Pryce, G. (2024) 'Theorising the causal impacts of social frontiers: The social and psychological implications of discontinuities in the geography of

residential mix', *Urban Studies*, 61(5), pp. 782–798. doi: <https://doi.org/10.1177/00420980231194834>

Jakubs, J. F. (1981) 'A distance-based segregation index', *Socio-Economic Planning Sciences*, 15(3), pp. 129–136. doi: [https://doi.org/10.1016/0038-0121\(81\)90028-8](https://doi.org/10.1016/0038-0121(81)90028-8)

James, G. *et al.* (2023) *An Introduction to Statistical Learning*. 2nd edn. New York: Springer.

Jencks, C. and Mayer, S. E. (1990) 'The social consequences of growing up in a poor neighborhood', in Lynn, L. E. and McGeary, M. G. H. (eds.) *Inner-city Poverty in the United States*. Washington, D.C: National Academy Press, pp. 111–186.

Jennissen, R. P. W. *et al.* (2023) *Migration diversity and social cohesion: reassessing the Dutch policy agenda*. Cham: Springer Nature.

Kalaiselvi, T., Sriramakrishnan, P. and Vasanthi, R. (2015) 'Brain Tumor Boundary Detection from MRI Brain Scans using Edge Indication Map', *National Conferences on New Horizons in Computational Intelligence and Information Systems (NHCIIS)*. Vol. 1, pp. 154–155. Available at: https://www.researchgate.net/publication/293401839_Brain_Tumour_Boundary_Detection_from_MRI_Brain_Scans_Using_Edge_Indication_Map (Accessed: 3 August 2023).

Karsten, S. *et al.* (2006) 'Choosing Segregation or Integration?: The Extent and Effects of Ethnic Segregation in Dutch Cities: The Extent and Effects of Ethnic Segregation in Dutch Cities', *Education and Urban Society*, 38(2), pp. 228–247. doi: <https://doi.org/10.1177/0013124505282606>

Kauppinen, T. M. and van Ham, M. (2019) 'Unravelling the demographic dynamics of ethnic residential segregation', *Population, Space and Place*, 25(2). doi: <https://doi.org/10.1002/psp.2193>

Kaufman, L. and Rousseeuw, P. J. (1990) *Finding groups in data: an introduction to cluster analysis*. New York: Wiley.

Kim, K. and Smets, P. (2020) 'Home experiences and homemaking practices of single Syrian refugees in an innovative housing project in Amsterdam', *Current Sociology*, 68(5), pp. 607–627. doi: <https://doi.org/10.1177/0011392120927744>

King, D. K. (1988) 'Multiple Jeopardy, Multiple Consciousness: The Context of a Black Feminist Ideology', *Signs: Journal of Women in Culture and Society*, 14(1), pp. 42–72. doi: <https://doi.org/10.1086/494491>

- Kloosterman, R. and Priemus, H. (2001) 'The Hague — A Dual City? Causes and Policy Responses: An Introduction', *Built Environment* (1978-), 27(3), pp. 167–175. Available at: <http://www.jstor.org/stable/23288412> (Accessed: 15 December 2025).
- Koopmans, R. and Schaeffer, M. (2016) 'Statistical and Perceived Diversity and Their Impacts on Neighborhood Social Cohesion in Germany, France and the Netherlands', *Social Indicators Research*, 125(3), pp. 853–883. doi: <https://doi.org/10.1007/s11205-015-0863-3>
- Kramer, R. (2017) 'Defensible Spaces in Philadelphia: Exploring Neighborhood Boundaries Through Spatial Analysis', *RSF: The Russell Sage Foundation Journal of the Social Sciences*, 3(2). doi: <https://doi.org/10.7758/rsf.2017.3.2.04>
- Kristof, L. K. D. (1959) 'The nature of frontiers and boundaries', *Annals of the Association of American Geographers*, 49(3), pp. 269–282. doi: <https://doi.org/10.1111/j.1467-8306.1959.tb01613.x>
- Křížková, I. *et al.* (2021) 'Social Frontiers: Estimating the Spatial Boundaries Between Residential Groups and Their Impacts on Crime', in Pryce, G. *et al.* (eds.) *Urban Inequality and Segregation in Europe and China*. Cham: Springer, pp. 285–304.
- Krysan, M. and Crowder, K. (2017) *Cycle of Segregation: Social Processes and Residential Stratification*. Russell Sage Foundation.
- Lamont, M. and Molnár, V. (2002) 'The study of boundaries in the social sciences', *Annual Review of Sociology*, 28, pp. 167–195. doi: <https://doi.org/10.1146/annurev.soc.28.110601.141107>
- Lau, D. C. and Murnighan, J. K. (1998) 'Demographic Diversity and Faultlines: The Compositional Dynamics of Organizational Groups', *The Academy of Management Review*, 23(2), pp. 325–340. doi: <https://doi.org/10.2307/259377>
- Leclerc, C., Vink, M. and Schmeets, H. (2023) 'Does Residential Context Matter? Neighborhood Migrant Concentration and Citizenship Acquisition in the Netherlands', *International Migration Review*, 57(4), pp. 1456–1485. doi: <https://doi.org/10.1177/01979183221115165>
- Lee, D. and Mitchell, R. (2013) 'Locally Adaptive Spatial Smoothing Using Conditional Auto-Regressive Models', *Journal of the Royal Statistical Society Series C (Applied Statistics)*, 62(4), pp. 593–608. doi: <https://doi.org/10.1111/rssc.12009>
- Lee, D., Pryce, G. and Ramos, M. (2025) 'Do Social Frontiers Matter for Mental Health?', *Sheffield Economic Research Paper Series*, 2025012. Available at: <https://sheffield.ac.uk/media/110411/download?attachment> (Accessed: 17 December 2025).

- Legewie, J. and Schaeffer, M. (2016) 'Contested Boundaries: Explaining Where Ethnoracial Diversity Provokes', *American Journal of Sociology*, 122(1), pp. 125–161. doi: <https://doi.org/10.1086/686942>
- Legewie, J. (2018) 'Living on the Edge: Neighborhood Boundaries and the Spatial Dynamics of Violent Crime', *Demography*, 55(5), pp. 1957–1977. doi: <https://doi.org/10.1007/s13524-018-0708-1>
- Li, Q-Q. *et al.* (2022) 'Towards a new paradigm for segregation measurement in an age of big data', *Urban Informatics*, 1(5). doi: <https://doi.org/10.1007/s44212-022-00003-3>
- López Turley, R. N. (2002) 'Is relative deprivation beneficial? The effects of richer and poorer neighbors on children's outcomes', *Journal of Community Psychology*, 30(6), pp. 671–686. doi: <https://doi.org/10.1002/jcop.10033>
- Lu, H. and Carlin, B. P. (2005) 'Bayesian Areal Wombling for Geographical Boundary Analysis', *Geographical Analysis*, 37(3), pp. 265–285. doi: <https://doi.org/10.1111/j.1538-4632.2005.00624.x>
- Madhulatha, T. (2012) 'An Overview on Clustering Methods', *IOSR Journal of Engineering*, 2(4), pp. 719–725. doi: <https://doi.org/10.9790/3021-0204719725>
- Maguire, A., French, D. and O'Reilly, D. (2016) 'Residential segregation, dividing walls and mental health: a population-based record linkage study', *Journal of Epidemiology and Community Health*, 70(9), pp. 845–854. doi: <https://doi.org/10.1136/jech-2015-206888>
- Manley, D. *et al.* (2023) 'Embrittlement as the Missing Dimension of Segregation Research', [PowerPoint presentation]. Social Frontiers Symposium. Available at: https://docs.google.com/presentation/d/1KXVrVu9dg_pWQhCOUwiLPo1kIDBgjZL6/edit?usp=drive_link&ouid=100348419262356483753&rtpof=true&sd=true (Accessed: 17 December 2025).
- Marcuse, P. M. (1985) 'Gentrification, Abandonment, and Displacement: Connections, Causes, and Policy Responses in New York City', *Washington University Journal of Urban and Contemporary Law*, 28(1), pp. 195–240. Available at: https://openscholarship.wustl.edu/cgi/viewcontent.cgi?article=1396&context=law_urbanlaw (Accessed: 7 November 2025).
- Marcuse, P. and van Kempen, R. (2000) 'Introduction', in Marcuse, P. and van Kempen, R. (eds.) *Globalizing cities: a new spatial order?* Oxford: Blackwell, pp. 14–24.

- Martinovic, B., van Tubergen, F. and Maas, I. (2009) 'Changes in immigrants' social integration during the stay in the host country: The case of non-western immigrants in the Netherlands', *Social Science Research*, 38(4), pp. 870–882. doi: <https://doi.org/10.1016/j.ssresearch.2009.06.001>
- Massey, D. S. and Denton, N. A. (1985) 'Spatial Assimilation as a Socioeconomic Outcome', *American Sociological Review*, 50(1), pp. 94–106. doi: <https://doi.org/10.2307/2095343>
- Massey, D. S. and Denton, N. A. (1988) 'The Dimensions of Residential Segregation', *Social Forces*, 67(2), pp. 281–315. doi: <https://doi.org/10.1093/sf/67.2.281>
- Mauritti, R. *et al.* (2016) 'The social structure of European inequality: A multidimensional perspective', *Sociologia, Problemas e Práticas*, 81, pp. 75–93. doi: <https://doi.org/10.7458/SPP2016818798>
- McArdle, N. and Acevedo-Garcia, D. (2018) 'Consequences of Segregation for Children's Opportunity and Wellbeing', in Herbert, C. *et al.* (eds.) *A Shared Future: Fostering Communities of Inclusion in an Era of Inequality*. Massachusetts: Joint Center for Housing Studies of Harvard University, pp. 80–95.
- McPherson, M., Smith-Lovin, L. and Cook, J. M. (2001) 'Birds of a Feather: Homophily in Social Networks', *Annual Review of Sociology*, 27, pp. 415–444. doi: <https://doi.org/10.1146/annurev.soc.27.1.415>
- Meijers, E. *et al.* (2014) 'City profile: The Hague', *Cities*, 41(A), pp. 92–100. doi: <https://doi.org/10.1016/j.cities.2014.05.012>
- Mingle, K. (2018) *Bijlmer (City of the Future, Part 1)*. Available at: <https://99percentinvisible.org/episode/bijlmer-city-future-part-1/> (Accessed: 20 June 2023).
- Mitchell, A. and Kelly, L. (2011) 'Peaceful Spaces? "Walking" through the New Liminal Spaces of Peacebuilding and Development in North Belfast', *Alternatives: Global, Local, Political*, 36(4), pp. 307–325. doi: <https://doi.org/10.1177/0304375411431761>
- Modai-Snir, T. and van Ham, M. (2020) 'Reordering, inequality and divergent growth: processes of neighbourhood change in Dutch cities', *Regional Studies*, 54(12), pp. 1668–1679. doi: <https://doi.org/10.1080/00343404.2020.1747607>.
- Municipality of Amsterdam - Research, Information and Statistics (2022) *LAYOUT_DISTRICT* [Shapefile]. Available at: https://maps.amsterdam.nl/open_geodata/?k=202 (Accessed: 9 August 2023).

- Musterd, S. *et al.* (2012) 'Neighbourhood composition and economic prospects: A longitudinal study in the Netherlands', *Tijdschrift voor Economische en Sociale Geografie*, 103(1), pp. 85–100. doi: <https://doi.org/10.1111/j.1467-9663.2011.00678.x>
- Musterd, S. *et al.* (2016) 'Adaptive behaviour in urban space: Residential mobility in response to social distance', *Urban Studies*, 53(2), pp. 227–246. doi: <https://doi.org/10.1177/0042098014562344>
- Musterd, S. and Fullaondo, A. (2008) 'Ethnic Segregation and the Housing Market in Two Cities in Northern and Southern Europe: The Cases of Amsterdam and Barcelona', *Architecture*, 8(8), pp. 94–115. doi: <https://doi.org/10.5821/ace.v3i8.2459>
- Musterd, S., Hochstenbach, C. and Boterman, W. (2020) 'Ripples of structural economic transformation: The changing social geographies of Dutch metropolitan regions', *Applied Geography*, 116. doi: <https://doi.org/10.1016/j.apgeog.2020.102151>
- Musterd, S., Ostendorf, W. and De Vos, S. (2003) 'Neighbourhood effects and social mobility: a longitudinal analysis', *Housing Studies*, 18(6), pp. 877–892. doi: <https://doi.org/10.1080/0267303032000135483>
- Musterd, S. and Ostendorf, W. (2009) 'Residential Segregation and Integration in the Netherlands', *Journal of Ethnic and Migration Studies*, 35(9), pp. 1515–1532. doi: <https://doi.org/10.1080/13691830903125950>
- Musterd, S. and van Kempen, R. (2009) 'Segregation and Housing of Minority Ethnic Groups in Western European Cities', *Tijdschrift voor Economische en Sociale Geografie*, 100(4), pp. 559–566. doi: <https://doi.org/10.1111/j.1467-9663.2009.00558.x>
- Myatt, A. J. W., Lindgren, U. and Osland, L. (2026) 'Housing Prices and Segregation: The Case of Social Frontiers.', *Journal of Regional Science*, 66(2), pp. 516–541. doi: <https://doi.org/10.1111/jors.70044>
- The Netherlands Institute for Social Research (2024) *The social environments of the rich and the poor*. Available at: <https://www.scp.nl/site/binaries/site-content/collections/documents/2024/04/05/de-leefwerelden-van-arm-en-rijk/Summary+The+social+environments+of+the+rich+and+the+poor+april+2024.pdf> (Accessed: 1 November 2024).
- Newman, D. and Paasi, A. (1998) 'Fences and neighbours in the postmodern world: boundary narratives in political geography', *Progress in Human Geography*, 22(2), pp. 186–207. doi: <https://doi.org/10.1191/030913298666039113>

- Newman, D. (2003) 'Boundaries', in Agnew, J., Mitchell, K. and Toal, G. (eds.) *A Companion to Political Geography*. Massachusetts: Blackwell Publishers, pp. 123–131.
- Newman, D. (2006) 'The lines that continue to separate us: borders in our 'borderless' world', *Progress in Human Geography*, 30(2), pp. 143–161. doi: <https://doi.org/10.1191/0309132506ph599xx>
- Nieuwenhuis, J. *et al.* (2017) 'Being Poorer Than the Rest of the Neighborhood: Relative Deprivation and Problem Behavior of Youth', *Journal of Youth and Adolescence*, 46(9), pp. 1891–1904. doi: <https://doi.org/10.1007/s10964-017-0668-6>
- Nieuwenhuis, J. *et al.* (2020) 'Does segregation reduce socio-spatial mobility? Evidence from four European countries with different inequality and segregation contexts', *Urban Studies*, 57(1), pp. 176–197. doi: <https://doi.org/10.1177/0042098018807628>
- Nieuwenhuis, J. and Hooimeijer, P. (2016) 'The association between neighbourhoods and educational achievement, a systematic review and meta-analysis', *Journal of Housing and the Built Environment: HBE*, 31(2), pp. 321–347. doi: <https://doi.org/10.1007/s10901-015-9460-7>
- NL Times (2021) '400% gap between home prices in Amsterdam, The Hague's cheapest & most expensive neighborhoods', *NL Times*, 30 August. Available at: <https://nltimes.nl/2021/08/30/400-gap-home-prices-amsterdam-hagues-cheapest-expensive-neighborhoods> (Accessed: 28 May 2026).
- OECD (2018) *Working Together for Local Integration of Migrants and Refugees in Amsterdam*. Paris: OECD Publishing. doi: <https://doi.org/10.1787/9789264299726-en>
- OECD (2023) *Policy Options for Labour Market Challenges in Amsterdam and Other Dutch Cities*. doi: <https://doi.org/10.1787/181c0fff-en>
- Olnier, D. *et al.* (2024) 'The conflicting geographies of social frontiers: Exploring the asymmetric impacts of social frontiers on household mobility in Rotterdam', *Environment and Planning B: Urban Analytics and City Science*, 51(3), pp. 625–640. doi: <https://doi.org/10.1177/23998083231173696>
- OpenStreetMap (2023) OSM Standard. Available at: <https://www.openstreetmap.org/#map=11/52.3673/4.9129> (Accessed: 10 June 2023). Licensed under the Open Data Commons Open Database License (ODbL): <http://opendatacommons.org/licenses/dbcl/1.0/>.

O'Sullivan, D. and Wong, D. W. S. (2007) 'A Surface-Based Approach to Measuring Spatial Segregation', *Geographical Analysis*, 39(2), pp. 147–168. doi: <https://doi.org/10.1111/j.1538-4632.2007.00699.x>

Owen, G. *et al.* (2021) 'Multi-scale Inequality and Segregation: Theory and Estimation', in: Pryce, G. *et al.* (eds.) *Urban Inequality and Segregation in Europe and China*. Cham: Springer, pp. 207–231.

Owen, G. *et al.* (2023) 'Residential segregation of migrants: Disentangling the intersectional and multiscale segregation of migrants in Shijiazhuang, China', *Urban Studies*, 60(1), pp. 166–182. doi: <https://doi.org/10.1177/00420980221076802>

Paolini, S. *et al.* (2004) 'Effects of direct and indirect cross-group friendships on judgments of Catholics and Protestants in Northern Ireland: the mediating role of an anxiety-reduction mechanism', *Personality & social psychology bulletin*, 30(6), pp. 770–786. doi: <https://doi.org/10.1177/0146167203262848>

Pavlačka, D. *et al.* (2023) 'Changes in spatial discontinuity in settlement patterns in the Czech-Polish border area: A case study of Těšín Silesia', *Papers in Regional Science*, 102(3), pp. 565–588. doi: <https://doi.org/10.1111/pirs.12732>

Peters, K. and de Haan, H. (2011) 'Everyday spaces of inter-ethnic interaction: the meaning of urban public spaces in the Netherlands', *Leisure/Loisir*, 35(2), pp. 169–190. doi: <https://doi.org/10.1080/14927713.2011.567065>

Pettigrew, T. F. and Tropp, L. R. (2005) 'Allport's Intergroup Contact Hypothesis: Its History and Influence', in Dovidio, J. F., Glick, P. and Rudman, L. A. (eds.) *On the nature of prejudice: Fifty years after Allport*. Blackwell Publishing, pp. 262–277.

Piantadosi, S., Byar, D. P. and Green, S. B. (1988) 'The Ecological Fallacy', *American Journal of Epidemiology*, 127(5), pp. 893–904. doi: <https://doi.org/10.1093/oxfordjournals.aje.a114892>

Piekut, A. *et al.* (2024) *Social frontiers and community life in Rotherham West*. Available at: https://orda.shef.ac.uk/articles/report/Social_Frontiers_and_Community_Life_in_Rotherham_West_Life_at_the_Frontier_Research_Briefing_1/25027763?file=44157623 (Accessed: 9 September 2025).

Piekut, A., Pryce, G. and van Gent, W. (2019) 'Segregation in the Twenty First Century: Processes, Complexities and Future Directions', *Tijdschrift Voor Economische en Sociale Geografie*, 110(3), pp. 225–234. doi: <https://doi.org/10.1111/tesg.12355>

Piekut, A. (2021) 'Re-Theorising Spatial Segregation: A European Perspective', in: Pryce, G. *et al.* (eds.) *Urban Inequality and Segregation in Europe and China*. Cham: Springer, pp. 13–38.

Putnam, R. D. (2007) '*E Pluribus Unum*: Diversity and Community in the Twenty-first Century The 2006 Johan Skytte Prize Lecture', *Scandinavian Political Studies*, 30(2), pp. 137–174. doi: <https://doi.org/10.1111/j.1467-9477.2007.00176.x>

Rajkumar, K. *et al.* (2022) 'A causal test of the strength of weak ties', *Science*, 377(6612), pp. 1304–1310. doi: <https://doi.org/10.1126/science.abl4476>

Reardon, S. F. and O'Sullivan, D. (2004) 'Measures of Spatial Segregation', *Sociological Methodology*, 34(1), pp. 121–162. doi: <https://doi.org/10.1111/j.0081-1750.2004.00150.x>

Rich, M. A. (2009) "'It Depends on How You Define Integrated": Neighborhood Boundaries and Racial Integration in a Baltimore Neighborhood', *Sociological Forum*, 24(4), 828–853. doi: <https://doi.org/10.1111/j.1573-7861.2009.01138.x>

Rosenbaum, P. R. and Rubin, D. B. (1984) 'Reducing Bias in Observational Studies Using Subclassification on the Propensity Score', *Journal of the American Statistical Association*, 79(387), pp. 516–524. doi: <https://doi.org/10.2307/2288398>

Rowe, S. L., Zimmer-Gembeck, M. J. and Hood, M. (2016) 'From the child to the neighbourhood: Longitudinal ecological correlates of young adolescents' emotional, social, conduct, and academic difficulties', *Journal of Adolescence*, 49(1), pp. 218–231. doi: <https://doi.org/10.1016/j.adolescence.2016.04.001>

Rubin, D.B. (2001) 'Using Propensity Scores to Help Design Observational Studies: Application to the Tobacco Litigation', *Health Services and Outcomes Research Methodology*, 2(3–4), pp. 169–188. doi: <https://doi.org/10.1023/A:1020363010465>

Sampson, R. J. and Sharkey, P. (2008) 'Neighborhood selection and the social reproduction of concentrated racial inequality', *Demography*, 45(1), pp. 1–29. doi: <https://doi.org/10.1353/dem.2008.0012>

San Millán, J., Cottineau-Mugadza, C. and Van Ham, M. (2025) 'The Economic Urban Divide: A Detailed Study of Income Inequality and Segregation in Dutch Urban Areas (2011–2022)', *Tijdschrift voor Economische en Sociale Geografie*, 116(4), pp. 508–530. doi: <https://doi.org/10.1111/tesg.70011>

Sävje, F., Higgins, M. J. and Sekhon, J.S. (2021) 'Generalized Full Matching', *Political Analysis*, 29(4), pp. 423–447. doi: <https://doi.org/10.1017/pan.2020.32>

Schelling, T. C. (1971) 'Dynamic Models of Segregation', *Journal of Mathematical Sociology*, 1(2), pp. 143–186. doi: <https://doi.org/10.1080/0022250X.1971.9989794>

- Schnittker, J. (2004) 'Social Distance in the Clinical Encounter: Interactional and Sociodemographic Foundations for Mistrust in Physicians', *Social Psychology Quarterly*, 67(3), pp. 217–235. doi: <https://doi.org/10.1177/019027250406700301>
- Schubert, E. and Rousseeuw, P.J. (2019) 'Faster *k*-Medoids Clustering: Improving the PAM, CLARA, and CLARANS Algorithms', in: Amato, G. *et al.* (eds.) *Similarity Search and Applications 12th International Conference, SISAP 2019*. Cham: Springer, pp. 171–187.
- 'Segregation (n.), sense 1.a' (2024) in *OED Online*. Available at: <https://doi.org/10.1093/OED/7461759097>
- Sivak, L. and Stulp, G. (2024) Intro to the CBS data. Available at: <https://stulp.gmw.rug.nl/prefer/posts/posts/2024-07-16-intro-csb-datasets.html> (Accessed: 9 October 2025).
- Sleutjes, B. and Boterman, W. R. (2016) 'Stated residential preferences of highly skilled international migrants', in: Musterd, S., Bontje, M. and Rouwendal, J. (eds.) *Skills and Cities*. London: Routledge, pp. 145–169.
- Sleutjes, B., de Valk, H. A. G. and Ooijevaar, J. (2018) 'The Measurement of Ethnic Segregation in the Netherlands: Differences Between Administrative and Individualized Neighbourhoods', *European Journal of Population*, 34(2), pp. 195–224. doi: <https://doi.org/10.1007/s10680-018-9479-z>
- Sleutjes, B., Ooijevaar, J. and de Valk, H. A. G. (2019) 'Residential Segregation in the Amsterdam Metropolitan Region: A Longitudinal Analysis Using Scalable Individualised Neighbourhoods', *Tijdschrift voor Economische en Sociale Geografie*, 110(3), pp. 359–377. doi: <https://doi.org/10.1111/tesg.12356>
- Smets, P. and den Uyl, M. (2008) 'The Complex Role of Ethnicity in Urban Mixing: A Study of Two Deprived Neighbourhoods in Amsterdam', *Urban Studies*, 45(7), pp. 1439–1460. doi: <https://doi.org/10.1177/0042098008090683>
- Smith, H. J. and Huo, Y. J. (2014) 'Relative Deprivation: How Subjective Experiences of Inequality Influence Social Behavior and Health', *Policy Insights from the Behavioral and Brain Sciences*, 1(1), pp. 231–238. doi: <https://doi.org/10.1177/2372732214550165>
- Spielman, S. E. and Logan, J. R. (2013) 'Using High-Resolution Population Data to Identify Neighborhoods and Establish Their Boundaries', *Annals of the Association of American Geographers*, 103(1), pp. 67–84. doi: <https://doi.org/10.1080/00045608.2012.685049>
- Spierenburg, L. *et al.* (2025) 'Does Residential Segregation Align with Urban Barriers?'. [Preprint]. doi: <https://doi.org/10.21203/rs.3.rs-7085658/v1>

- Spierenburg, L., van Cranenburgh, S. and Cats, O. (2024) 'Evolution of residential segregation patterns in the Netherlands between 2015 and 2020', *Cities*, 150. doi: <https://doi.org/10.1016/j.cities.2024.105038>
- Stolle, D., Soroka, S. and Johnston, R. (2008) 'When Does Diversity Erode Trust? Neighborhood Diversity, Interpersonal Trust and the Mediating Effect of Social Interactions', *Political Studies*, 56(1), pp. 57–75. doi: <https://doi.org/10.1111/j.1467-9248.2007.00717.x>
- Stuart, E. A. and Green, K. M. (2008) 'Using full matching to estimate causal effects in nonexperimental studies: examining the relationship between adolescent marijuana use and adult outcomes', *Developmental Psychology*, 44(2), pp. 395–406. doi: <https://doi.org/10.1037/0012-1649.44.2.395>
- Stuart, E. A. (2010) 'Matching methods for causal inference: A review and a look forward', *Statistical Science: A Review Journal of the Institute of Mathematical Statistics*, 25(1), pp. 1–21. doi: <https://doi.org/10.1214/09-STS313>
- Sykes, B. and Kuyper, H. (2009) 'Neighbourhood Effects on Youth Educational Achievement in the Netherlands: Can Effects Be Identified and Do They Vary by Student Background Characteristics?', *Environment and Planning A: Economy and Space*, 41(10), pp. 2417–2436. doi: <https://doi.org/10.1068/a41255>
- Tammaru, T. et al. (2016) 'A multi-factor approach to understanding socio-economic segregation in European capital cities', in Tammaru, T. et al. (eds.) *Socio-Economic Segregation in European Capital Cities: East Meets West*. London: Routledge, pp. 1–29.
- Tammaru, T. et al. (2019) 'Relationship between income inequality and residential segregation of socioeconomic groups', *Regional Studies*, 54(4), pp. 450–461. doi: <https://doi.org/10.1080/00343404.2018.1540035>
- Teernstra, A. B. (2014) *Disentangling processes of neighbourhood change: Towards a better understanding of upgrading and downgrading of neighbourhoods in the highly-regulated context of the Netherlands*. PhD Thesis. University of Amsterdam. Available at: <https://dare.uva.nl/search?identifier=4196e5d2-5b6d-497c-a389-44b7fe7e3634> (Accessed: 26 November 2025).
- Tessema, M. et al. (2022) 'The Link between Parental Socioeconomic Status (SES) and Children's Future Employment Prospects: A Conceptual Framework', *Academy of Strategic Management Journal*, 21(5), pp. 1–16. Available at: https://www.researchgate.net/publication/363761748_THE_LINK_BETWEEN_PARENTAL_SOCIOECONOMIC_STATUS_SES_AND_CHILDREN'S_FUTURE_EMPLOYMENT_PROSPECTS_A_CONCEPTUAL_FRAMEWORK (Accessed: 16 October 2025).

Thijssen, L., Coenders, M. and Lancee, B. (2021) 'Ethnic Discrimination in the Dutch Labor Market: Differences Between Ethnic Minority Groups and the Role of Personal Information About Job Applicants—Evidence from a Field Experiment', *Journal of International Migration and Integration*, 22(3), pp. 1125–1150. doi: <https://doi.org/10.1007/s12134-020-00795-w>

Thoemmes, F. J and Kim, E. S. (2011) 'A Systematic Review of Propensity Score Methods in the Social Sciences', *Multivariate Behavioral Research*, 46(1), pp. 90–118. doi: <https://doi.org/10.1080/00273171.2011.540475>

Thomas, R. (2023) *The Utility of Crime Harm Measurement in the Spatial Analysis of Police Reported Victim Based Crime*. PhD Thesis. University of Sheffield. Available at: <https://etheses.whiterose.ac.uk/id/eprint/35554/> (Accessed: 8 June 2026).

Tilly, C. (2004) 'Social Boundary Mechanisms', *Philosophy of the Social Sciences*, 34(2), pp. 211–236. doi: <https://doi.org/10.1177/0048393103262551>

Tobler, W. R. (1970) 'A Computer Movie Simulating Urban Growth in the Detroit Region', *Economic Geography*, 46, pp. 234–240. doi: <https://doi.org/10.2307/143141>

Toft, M. and Ljunggren, J. (2016) 'Geographies of class advantage: The influence of adolescent neighbourhoods in Oslo', *Urban Studies*, 53(14), pp. 2939–2955. doi: <https://doi.org/10.1177/0042098015601770>

Troost, A. A., van Ham, M. and Manley, D. J. (2023) 'Neighbourhood effects on educational attainment. What matters more: Exposure to poverty or exposure to affluence?', *Plos One*, 18(3). doi: <https://doi.org/10.1371/journal.pone.0281928>

Tselios, V. *et al.* (2015) 'Integration of Immigrants, Bridging Social Capital, Ethnicity, and Locality', *Journal of Regional Science*, 55(3), pp. 416–441. doi: <https://doi.org/10.1111/jors.12160>

Urria, J. *et al.* (2025) 'The spatio-temporal evolution of social inequalities in cities: a multidimensional, multiscalar and longitudinal approach for neighbourhood classification', *Cities*, 165. doi: <https://doi.org/10.1016/j.cities.2025.106089>

van Amersfoort, H. and van Niekerk, M. (2006) 'Immigration as a Colonial Inheritance: Post-Colonial Immigrants in the Netherlands, 1945–2002', *Journal of Ethnic and Migration Studies*, 32(3), pp. 323–346. doi: <https://doi.org/10.1080/13691830600555210>

van den Bossche, C. (no date) *Inequalities in the Netherlands*. Available at: <https://www.sdgwatcheurope.org/wp-content/uploads/2019/06/10.3.a-report-NL.pdf> (Accessed: 23 July 2025).

- van der Laan Bouma-Doff, W. (2007) 'Confined Contact: Residential Segregation and Ethnic Bridges in the Netherlands', *Urban Studies*, 44(5–6), pp. 997–1017. doi: <https://doi.org/10.1080/00420980701255965>
- van der Laan, M., Pollard, K. and Bryan, J. (2003) 'A new partitioning around medoids algorithm', *Journal of Statistical Computation and Simulation*, 73(8), pp. 575–584. doi: <https://doi.org/10.1080/0094965031000136012>
- VanderWeele, T. J. and Ding, P. (2017) 'Sensitivity Analysis in Observational Research: Introducing the E-Value', *Annals of Internal Medicine*, 167(4), pp. 268–274. doi: <https://doi.org/10.7326/M16-2607>
- van Deursen, H. (2023) *The People's Housing: Woningcorporaties and the Dutch Social Housing System Part 2: The Mechanics*. Available at: <https://www.jchs.harvard.edu/research-areas/working-papers/peoples-housing-woningcorporaties-and-dutch-social-housing-system-0> (Accessed: 10 November 2025).
- van de Water, M. (2024) 'CBS: Mensen met Nederlandse ouders en hoog inkomen leven het meest in hun eigen bubbel', *de Volkskrant*, 23 February. Available at: <https://www.volkskrant.nl/binnenland/cbs-mensen-met-nederlandse-ouders-en-hoog-inkomen-leven-het-meest-in-hun-eigen-bubbel~b340d901/> (Accessed: 30 July 2025).
- van Duijne, R.J. and Ronald, R. (2018) 'The unraveling of Amsterdam's unitary rental system', *Journal of Housing and the Built Environment*, 33(4), pp. 633–651. doi: <https://doi.org/10.1007/s10901-018-9601-x>
- van Elk, R., Jongen, E. and Koot, P. (2019) *Income differences across migrant groups in the Netherlands: an intergenerational perspective*. Available at: <https://www.cpb.nl/sites/default/files/omnidownload/cpb-achtergronddocument-income-differences-across-migrant-groups-in-the-netherlands.pdf> (Accessed: 3 December 2024).
- van Gaalen, T. (2025) 'Fighting for Political Influence', *Left History*, 26(2), pp. 7–32. doi: <https://doi.org/10.25071/1913-9632.39681>
- van Gent, W. P. C. (2013) 'Institutional neoliberalization and gentrification in Amsterdam', *International Journal of Urban and Regional Research*, 37(2), pp. 503–522. doi: <https://doi.org/10.1111/j.1468-2427.2012.01155.x>
- van Gent, W. and Boterman, W. (2019) 'Gentrification Of The Changing State', *Tijdschrift voor Economische en Sociale Geografie*, 110(1), pp. 35–46. doi: <https://doi.org/10.1111/tesg.12331>

- van Gent, W. and Hochstenbach, C. (2020) 'The impact of gentrification on social and ethnic segregation', in Musterd, S. (ed.) *Handbook of Urban Segregation*. Cheltenham, UK: Edward Elgar Publishing, pp. 306–324.
- van Ham, M. and Manley, D. (2010) 'The effect of neighbourhood housing tenure mix on labour market outcomes: a longitudinal investigation of neighbourhood effects', *Journal of Economic Geography*, 10(2), pp. 257–282. doi: <https://doi.org/10.1093/jeg/lbp017>
- van Ham, M. and Tammaru, T. (2016) 'New perspectives on ethnic segregation over time and space. A domains approach', *Urban Geography*, 37(7), pp. 953–962. doi: <https://doi.org/10.1080/02723638.2016.1142152>
- van Kempen, R. and Bolt, G. (2009) Social cohesion, social mix, and urban policies in the Netherlands, *Journal of Housing and the Built Environment*, 24(4), pp. 457–475. doi: <https://doi.org/10.1007/s10901-009-9161-1>
- van Schie, G. (2018) 'Origins: A History of Race-ethnic Categorisation in the Dutch Governmental Data Ontology (1899–2018)', *TMG Journal for Media History*, 21(2), pp. 67–88. doi: <https://doi.org/10.18146/2213-7653.2018.367>
- van Tubergen, F. and Volker, B. (2015) 'Inequality in Access to Social Capital in the Netherlands', *Sociology*, 49(3), pp. 521–538. doi: <https://doi.org/10.1177/0038038514543294>
- Vartanian, T. P. and Houser, L. (2010) 'The Effects of Childhood Neighborhood Conditions on Self-reports of Adult Health', *Journal of Health and Social Behavior*, 51(3), pp. 291–306. doi: <https://doi.org/10.1177/0022146510378241>
- Veldman, K. *et al.* (2014) 'Mental Health Problems and Educational Attainment in Adolescence: 9-Year Follow-Up of the TRAILS Study', *Plos one*, 9(7). doi: <https://doi.org/10.1371/journal.pone.0101751>
- Veldman, K. *et al.* (2015) 'Mental health trajectories from childhood to young adulthood affect the educational and employment status of young adults: results from the TRAILS study', *Journal of Epidemiology and Community Health*, 69(6), pp. 588–593. doi: <https://doi.org/10.1136/jech-2014-204421>
- Vervoort, M., Flap, H. and Dagevos, J. (2011) 'The Ethnic Composition of the Neighbourhood and Ethnic Minorities' Social Contacts: Three Unresolved Issues', *European Sociological Review*, 27(5), pp. 586–605. doi: <https://doi.org/10.1093/esr/jcq029>
- Vervoort, M. (2012) 'Ethnic Concentration in the Neighbourhood and Ethnic Minorities' Social Integration: Weak and Strong Social Ties Examined', *Urban Studies*, 49(4), pp. 897–915. doi: <https://doi.org/10.1177/0042098011408141>

- Völker, B., Flap, H. and Lindenberg, S. (2007) 'When Are Neighbourhoods Communities? Community in Dutch Neighbourhoods', *European Sociological Review*, 23(1), pp. 99–114. doi: <https://doi.org/10.1093/esr/jcl022>
- Wang, D., Choi, J-K. and Shin, J. (2020) 'Long-term Neighborhood Effects on Adolescent Outcomes: Mediated through Adverse Childhood Experiences and Parenting Stress', *Journal of Youth Adolescence*, 49(10), pp. 2160–2173. doi: <https://doi.org/10.1007/s10964-020-01305-y>
- Weinberg, D. *et al.* (2019) 'Adolescent Socioeconomic Status and Mental Health Inequalities in the Netherlands, 2001–2017', *International Journal of Environmental Research and Public Health*, 16(19). doi: <https://doi.org/10.3390/ijerph16193605>
- Weldon, S.L. (2008) 'Intersectionality', in G. Goertz and A.G. Mazur (eds.) *Politics, Gender, and Concepts: Theory and Methodology*. Cambridge: Cambridge University Press, pp. 193–218.
- Wickes, R. *et al.* (2025) 'Contested Boundaries and Perceptions of Inter-ethnic Disorder', *The British Journal of Criminology*. doi: <https://doi.org/10.1093/bjc/azaf094>
- Wijburg, G. (2021) 'The governance of affordable housing in post-crisis Amsterdam and Miami', *Geoforum*, 119, pp. 30–42. doi: <https://doi.org/10.1016/j.geoforum.2020.12.013>
- Witteveen, D. and Alba, R. (2018) 'Labour Market Disadvantages of Second-Generation Turks and Moroccans in the Netherlands: Before and during the Great Recession', *International Migration*, 56(2), pp. 97–116. doi: <https://doi.org/10.1111/imig.12411>
- Wong, D. W. S. (1993) 'Spatial Indices of Segregation', *Urban Studies*, 30(3), pp. 559–572. doi: <https://doi.org/10.1080/00420989320080551>
- Wong, D. W. S. (2004) 'Comparing Traditional and Spatial Segregation Measures: A Spatial Scale Perspective', *Urban Geography*, 25(1), pp. 66–82. doi: <https://doi.org/10.2747/0272-3638.25.1.66>
- Xu, D. and Tian, Y. (2015) 'A Comprehensive Survey of Clustering Algorithms', *Annals of Data Science*, 2(2), pp. 165–193. doi: <https://doi.org/10.1007/s40745-015-0040-1>
- Xu, Y. *et al.* (2019) 'Quantifying segregation in an integrated urban physical-social space', *Journal of the Royal Society Interface*, 16(160). doi: <https://doi.org/10.1098/rsif.2019.0536>
- Yanow, D. and van der Haar, M. (2013) 'People out of place: allochthony and autochthony in the Netherlands' identity discourse — metaphors and categories in

action', *Journal of International Relations and Development*, 16(2), pp. 227–261. doi: <https://doi.org/10.1057/jird.2012.13>

Yao, J. *et al.* (2019) 'Spatial Segregation Measures: A Methodological Review', *Tijdschrift Voor Economische en Sociale Geografie*, 110(3), pp. 235–250. doi: <https://doi.org/10.1111/tesg.12305>

Zahirović, S. and Sterk, B. (no date) '*The Bijlmer: a Dutch Approach to Multiculturalism*', Humanity in Action. Available at: https://humanityinaction.org/knowledge_detail/the-bijlmer-a-dutch-approach-to-multiculturalism/ (Accessed: 10 Jan 2023).

Zhang, J. (2011) 'Tipping and residential segregation: A unified Schelling model', *Journal of Regional Science*, 51(1), pp. 161–193. doi: <https://doi.org/10.1111/j.1467-9787.2010.00671.x>

Zhang, M. L. *et al.* (2024) 'Using residents and experts to evaluate the validity of areal wombling for detecting social boundaries: A small-scale feasibility study', *Plos one*, 19(8). doi: <https://doi.org/10.1371/journal.pone.0305774>

Zhao, Q. Y. *et al.* (2021) 'Propensity score matching with R: conventional methods and new features', *Annals of Translational Medicine*, 9(9). doi: <https://doi.org/10.21037/atm-20-3998>

Ziller, C. and Spörlein, C. (2020) 'Residential Segregation and Social Trust of Immigrants and Natives: Evidence From the Netherlands', *Frontiers in Society*, 5(45). doi: <https://doi.org/10.3389/fsoc.2020.00045>

Zorlu, A. and Mulder, C. H. (2008) 'Initial and Subsequent Location Choices of Immigrants to the Netherlands', *Regional Studies*, 42(2), pp. 245–264. doi: <https://doi.org/10.1080/00343400601145210>

Zwiers, M., van Ham, M. and Manley, D. (2018) 'Trajectories of ethnic neighbourhood change: Spatial patterns of increasing ethnic diversity', *Population, Space and Place*, 24(2). doi: <https://doi.org/10.1002/psp.2094>

R Packages and Programs

Results are based on analysis using non-public microdata from Statistics Netherlands (**University of Sheffield - project number 9197**).

All analysis was performed using R Statistical Software (4.2.3). This study uses functions from the socialFrontiers package available at: <https://github.com/life-at-the-frontier/socialFrontiers>. The data used within this thesis is not public access but for reproducibility the code is available at: <https://github.com/eleanorbale/Thesis-Code.git>.

Butts CT (2023). `_sna: Tools for Social Network Analysis_`. R package version 2.7-1, <https://CRAN.R-project.org/package=sna>

Csardi G, Nepusz T (2006). "The igraph software package for complex network research." `_InterJournal_`, *Complex Systems*, 1695. <https://igraph.org>

Dowle M, Srinivasan A (2023). `_data.table: Extension of data.frame`_`. R package version 1.14.8, <https://CRAN.R-project.org/package=data.table>

Dragulescu A, Arendt C (2020). `_xlsx: Read, Write, Format Excel 2007 and Excel 97/2000/XP/2003 Files_`. R package version 0.6.5, <https://CRAN.R-project.org/package=xlsx>

H. Rue, A. Riebler, S. H. Sorbye, J. B. Illian, D. P. Simpson, and F. K. Lindgren. Bayesian computing with INLA: A review. *Annual Reviews of Statistics and Its Applications*, 4(March):395-421, 2017. URL: <http://arxiv.org/abs/1604.00860>

Harrell Jr F (2024). `_Hmisc: Harrell Miscellaneous_`. R package version 5.1-2, <https://CRAN.R-project.org/package=Hmisc>

Ho D, Imai K, King G, Stuart E (2011). "MatchIt: Nonparametric Preprocessing for Parametric Causal Inference." `_Journal of Statistical Software_`, *42*(8), 1-28. <https://doi.org/10.18637/jss.v042.i08>

Maechler, M., Rousseeuw, P., Struyf, A., Hubert, M., Hornik, K. (2023). `cluster: Cluster Analysis Basics and Extensions`. R package version 2.1.6.

Novomestky F (2022). `_matrixcalc: Collection of Functions for Matrix Calculations_`. R package version 1.0-6, <https://CRAN.R-project.org/package=matrixcalc>

Pebesma E (2018). "Simple Features for R: Standardized Support for Spatial Vector Data." *The R Journal*, **10**(1), 439–446. doi: <https://doi.org/10.32614/RJ-2018-009>

Pebesma E, Bivand R (2023). `_Spatial Data Science With Applications in R_`. Chapman & Hall. <https://r-spatial.org/book/>

- Pebesma E, Bivand R (2023). *Spatial Data Science: With applications in R*. Chapman and Hall/CRC. doi: <https://doi.org/10.1201/9780429459016>
- Pebesma, E.J., R.S. Bivand, 2005. Classes and methods for spatial data in R. R News 5 (2), <https://cran.r-project.org/doc/Rnews/>
- Roger S. Bivand, Edzer Pebesma, Virgilio Gomez-Rubio, 2013. Applied spatial data analysis with R, Second edition. Springer, NY. <https://asdar-book.org/>
- Signorell A (2024). `_DescTools: Tools for Descriptive Statistics_`. R package version 0.99.54, <https://CRAN.R-project.org/package=DescTools>
- Tennekes M (2018). "tmap: Thematic Maps in R." `_Journal of Statistical Software_`, 84(6), 1-39. doi:10.18637/jss.v084.i06 <https://doi.org/10.18637/jss.v084.i06>
- Tillé Y, Matei A (2023). `_sampling: Survey Sampling_`. R package version 2.10, <https://CRAN.R-project.org/package=sampling>
- Wickham H. *ggplot2: Elegant Graphics for Data Analysis*. Springer-Verlag New York, 2016.
- Wickham H, *et al.* (2019). "Welcome to the tidyverse." `_Journal of Open Source Software_`, 4(43), 1686. doi: <https://doi.org/10.21105/joss.01686>
- Wickham H, Bryan J (2023). `_readxl: Read Excel Files_`. R package version 1.4.2, <https://CRAN.R-project.org/package=readxl>
- Wickham H, François R, Henry L, Müller K, Vaughan D (2023). `_dplyr: A Grammar of Data Manipulation_`. R package version 1.1.4, <https://CRAN.R-project.org/package=dplyr>
- Wickham H, Miller E, Smith D (2023). `_haven: Import and Export 'SPSS', 'Stata' and 'SAS' Files_`. R package version 2.5.2, <https://CRAN.R-project.org/package=haven>
- JGraph. (2021) *draw.io* (15.5.2). [Software]. Accessed on: 1 November 2025.

Appendices

Appendix A: The Boroughs of Amsterdam

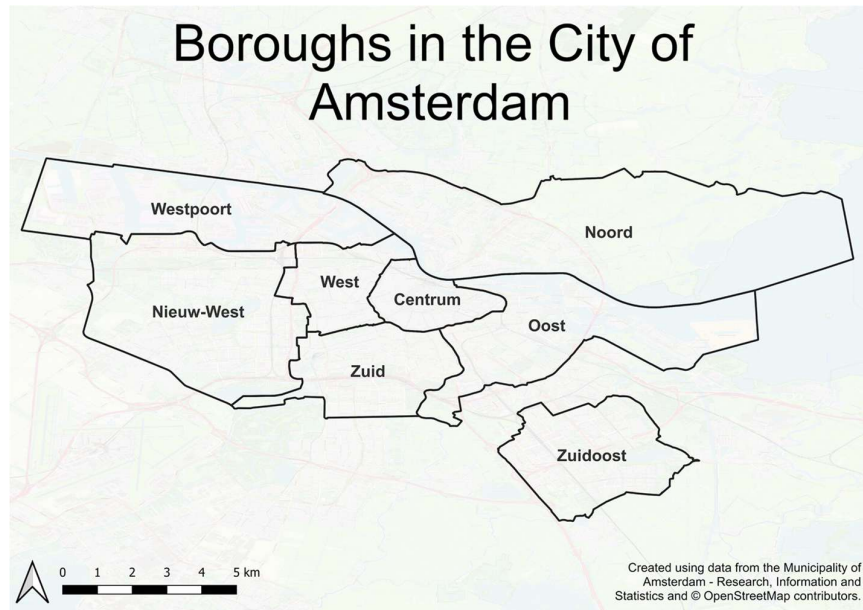


Figure A.1: City of Amsterdam with the boroughs labelled.
Created using shapefiles from the Municipality of Amsterdam -
Research, Information and Statistics (2022) and
OpenStreetMap (2023).

Appendix B: Absolute Contiguous Difference Threshold Choice

A crucial element of the absolute difference method is the choice of threshold when identifying a social frontier. Chapter 3 tested nine different thresholds varying from the 99th percentile - which classifies the top 1% of absolute difference scores as social frontiers - to the 5th percentile - which identifies the top 95% of the absolute difference scores as social frontiers. These are examples of extreme classifications, with the 99th percentile likely to classify almost all boundaries as non-social frontiers and the reverse true for the 5th percentile. To test which threshold is the most accurate based on the simulation study, nine thresholds were applied to the simulated data. The accuracy scores are reported in Tables B.1 and B.2. The 75th percentile returned the highest accuracy across the two years on average and thus was selected for this chapter.

Table B.1: 2010 Threshold choices for absolute difference social frontier identifier simulation accuracy results.

	0.2/0.8		0.3/0.7		0.4/0.6		Median Scores		
Thres hold	TPR	TNR	TPR	TNR	TPR	TNR	TPR	TNR	Average
99th	0.025	0.993	0.012	0.990	0	0.989	0.012	0.990	0.501
95th	0.239	0.984	0.123	0.966	0.043	0.951	0.123	0.966	0.544
90th	0.555	0.984	0.405	0.956	0.104	0.901	0.405	0.956	0.680
80th	0.982	0.945	0.896	0.931	0.509	0.857	0.896	0.931	0.913
75th	0.994	0.887	0.963	0.883	0.749	0.842	0.963	0.889	0.923
50th	1	0.593	1	0.593	0.982	0.589	1	0.593	0.796
25th	1	0.297	1	0.297	0.994	0.296	1	0.297	0.648
10th	1	0.120	1	0.120	0.994	0.118	1	0.120	0.560
5th	1	0.060	1	0.060	1	0.060	1	0.060	0.530

Table B.2: 2019 Threshold choices for absolute difference social frontier identifier simulation accuracy results.

	0.2/0.8		0.3/0.7		0.4/0.6		Median scores		
Thres hold	TPR	TNR	TPR	TNR	TPR	TNR	TPR	TNR	Average
99th	0.037	0.993	0.015	0.990	0.007	0.989	0.015	0.990	0.502
95th	0.324	0.992	0.195	0.972	0.029	0.948	0.195	0.972	0.584
90th	0.699	0.992	0.544	0.968	0.180	0.912	0.544	0.968	0.756
80th	1	0.924	0.978	0.920	0.765	0.887	0.978	0.920	0.949
75th	1	0.866	0.993	0.865	0.927	0.855	0.993	0.865	0.929
50th	1	0.578	1	0.578	0.993	0.577	1	0.578	0.789
25th	1	0.289	1	0.289	0.993	0.288	1	0.289	0.644
10th	1	0.116	1	0.116	0.993	0.115	1	0.116	0.558
5th	1	0.058	1	0.058	0.993	0.057	1	0.058	0.529

Appendix C: Simulation Study Results Table for 2019

Table C.1: Accuracy scores (TPR and TNR) for 2019 simulation study.									
Level of Difference between grid cells	Population Size	Detection Method							
		Lee and Mitchell (2013)		k-means		PAM Clustering		Abs Difference (75th Percentile)	
		TPR	TNR	TPR	TNR	TPR	TNR	TPR	TNR
Small Difference	Small Pop	0.701	<u>0.997</u>	<u>0.900</u>	0.890	<u>0.900</u>	0.890	0.615	0.873
	Real Pop	0.755	<u>0.996</u>	<u>0.923</u>	0.919	<u>0.923</u>	0.919	0.724	0.910
	Large Pop	0.870	<u>0.997</u>	<u>0.956</u>	0.967	<u>0.956</u>	0.969	0.927	0.978
Medium Difference	Small Pop	0.808	<u>0.997</u>	<u>0.943</u>	0.890	<u>0.943</u>	0.890	0.854	0.873
	Real Pop	0.839	<u>0.997</u>	<u>0.958</u>	0.961	<u>0.958</u>	0.963	0.897	0.963
	Large Pop	0.874	<u>0.997</u>	0.962	0.966	<u>0.969</u>	0.975	0.935	0.981
Large Difference	Small Pop	0.816	<u>0.999</u>	0.958	0.953	<u>0.985</u>	0.955	0.906	0.972
	Real Pop	0.879	<u>0.997</u>	<u>0.977</u>	0.981	<u>0.977</u>	0.982	0.954	0.987
	Large Pop	0.969	<u>0.999</u>	0.992	0.992	<u>0.996</u>	0.997	0.985	0.998
Median Score		0.816	<u>0.997</u>	<u>0.958</u>	0.961	<u>0.958</u>	0.963	0.906	0.972

Appendix D: Simulation Study Scores for 2010

True Boundaries in the City of Amsterdam in 2010

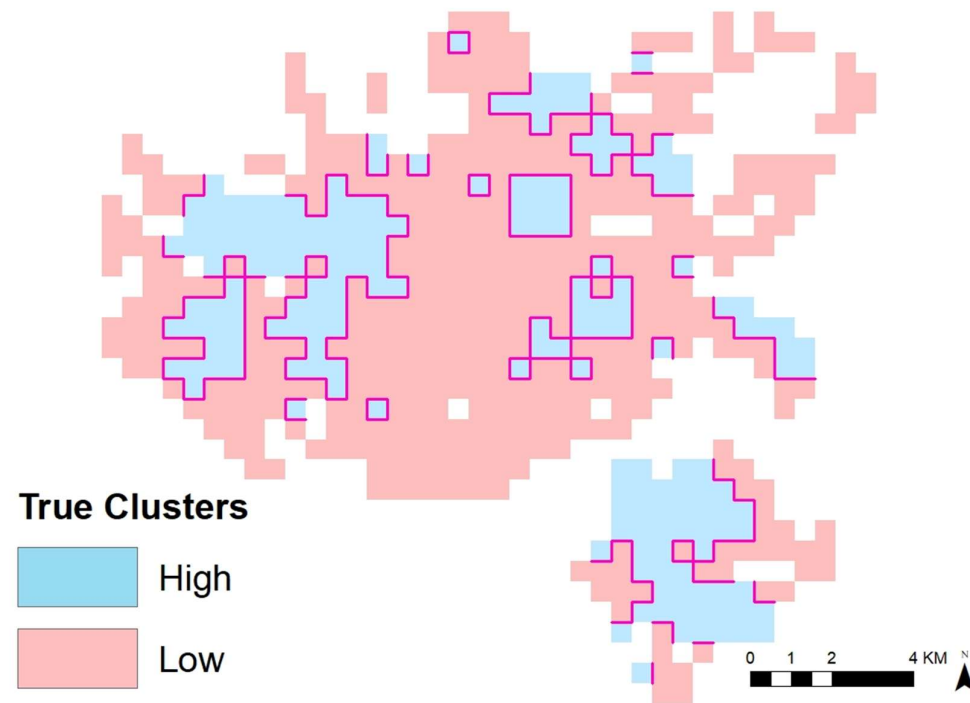
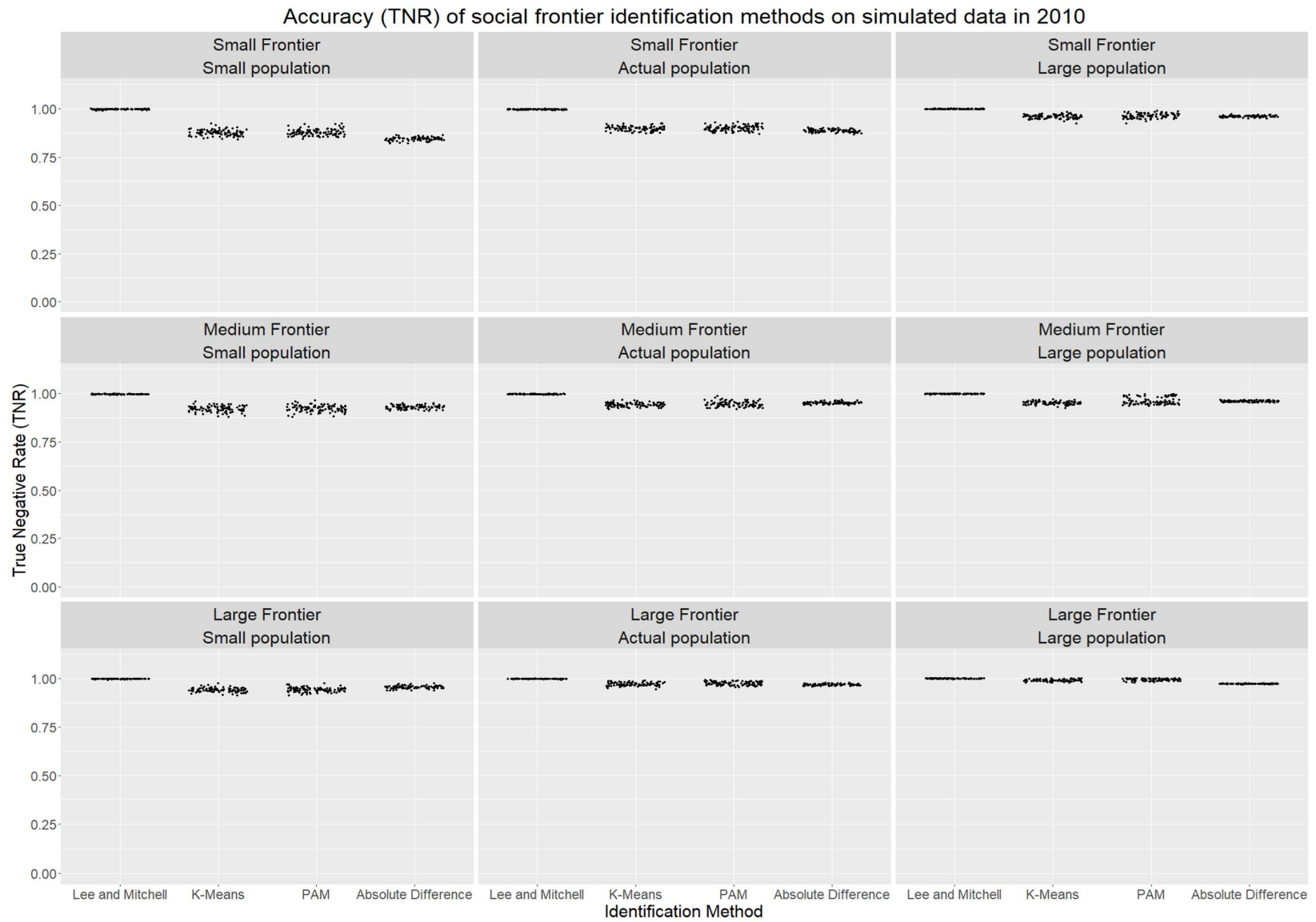


Figure D.1: The true social frontiers for 2010 data



(a)

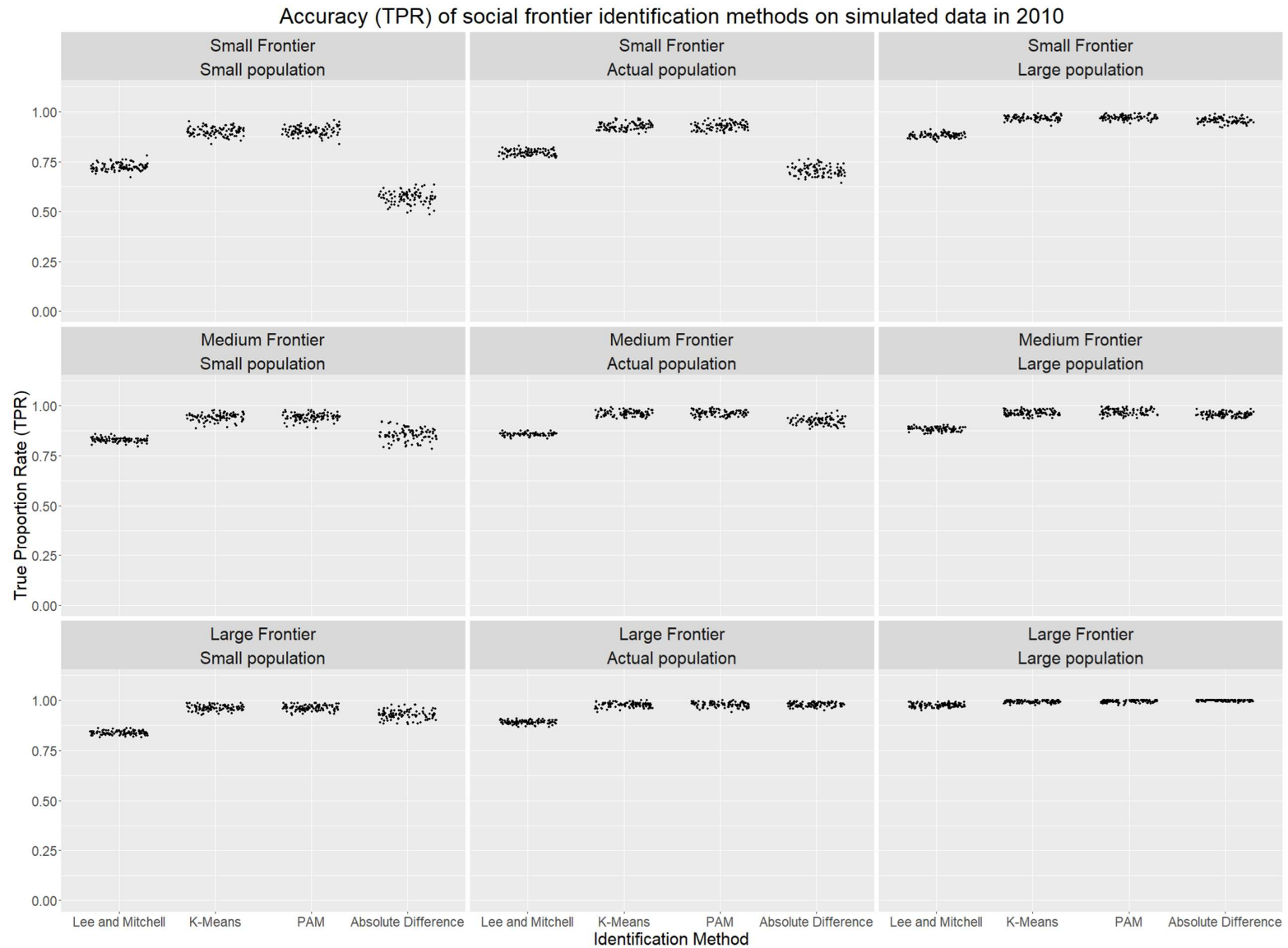


Figure D.2: Jitter plots showing the accuracy scores (TNR - a) and (TPR - b) for the 2010 simulation study for the four identification methods.

Table D.1: Accuracy scores (TPR and TNR) for 2010 simulation study.

Level of Difference between grid cells	Population Size	Detection Method							
		Lee and Mitchell (2013)		k-means		PAM Clustering		Abs Difference (75th Percentile)	
		TPR	TNR	TPR	TNR	TPR	TNR	TPR	TNR
Small Difference	Small Pop	0.719	<u>0.996</u>	0.901	0.874	<u>0.904</u>	0.875	0.575	0.844
	Real Pop	0.793	<u>0.995</u>	0.922	0.895	<u>0.927</u>	0.898	0.703	0.884
	Large Pop	0.883	<u>0.997</u>	0.966	0.958	<u>0.967</u>	0.962	0.927	0.959
Medium Difference	Small Pop	0.828	<u>0.996</u>	0.946	0.920	<u>0.948</u>	0.922	0.853	0.930
	Real Pop	0.862	<u>0.996</u>	0.961	0.943	<u>0.966</u>	0.944	0.927	0.951
	Large Pop	0.884	<u>0.997</u>	0.963	0.966	<u>0.968</u>	0.975	0.957	0.981
Large Difference	Small Pop	0.836	<u>0.996</u>	<u>0.961</u>	0.940	<u>0.961</u>	0.941	0.927	0.954
	Real Pop	0.888	<u>0.997</u>	0.978	0.969	0.978	0.974	<u>0.981</u>	0.967
	Large Pop	0.974	<u>1</u>	0.991	0.990	0.996	0.995	<u>1</u>	0.972
Median Score		0.862	<u>0.996</u>	0.961	0.943	<u>0.966</u>	0.962	0.927	0.954

Appendix E: Municipalities in COROP Region Definitions

Table E.1: Municipalities included within COROP region definitions			
Groot Amsterdam (Greater Amsterdam)		Agglomeratie 's-Gravenhage (Hague Metropolitan Area)	
0358	Aalsmeer	0637	Zoetermeer
0362	Amstelveen	0629	Wassenaar
0363	Amsterdam	0603	Rijswijk
0384	Diemen	1926	Pijnacker-Nootdorp
0385	Edam-Volendam	1916	Leidschendam-Voorburg
0394	Haarlemmermeer	0518	The Hague
0415	Landsmeer		
0431	Oostzaan		
0437	Ouder-Amstel		
0439	Purmerend		
0451	Uithoorn		
0852	Waterland		

Appendix F: Income Cut-offs

Table F.1: Cut-offs for low- and high-income population in the Netherlands between 1999 and 2021.

Year	Low-Income (€)	High-Income (€)
1999	11,076	40,608
2000	11,532	42,468
2001	12,096	44,928
2002	12,516	46,692
2003	12,768	47,703
2004	12,816	48,636
2005	12,864	50,004
2006	13,152	53,712
2007	13,704	56,052
2008	14,328	58,656
2009	14,652	60,000
2010	14,892	60,891
2011	15,156	62,004
2012	15,240	63,516
2013	14,436	61,104
2014	14,568	61,872
2015	14,688	63,192
2016	14,880	64,896
2017	15,132	66,828
2018	15,408	68,508
2019	16,056	70,632
2020	16,452	72,348
2021	17,184	75,468

Appendix G: Lee and Mitchell (2013) Frontier Prevalence for Chapter 4

Table G.1: Social frontier prevalence table for Lee and Mitchell (2013) method							
		Percentage		Average Length		Average Difference	
		Amsterdam	Hague	Amsterdam	Hague	Amsterdam	Hague
Country of Birth (Non-Western)	1999	9.36%	8.09%	248.94m	240.09m	27.3%	22.96%
	2010	9%	10.85%	237.8m	260.46m	29.7%	27.45%
	2021	10.55%	11.53%	245.56m	281.32m	29.35%	28.34%
High-Income	1999	3.37%	3.49%	183.05m	172.06m	21.76%	21.48%
	2010	3.93%	3.2%	191.55m	164.18m	21.55%	22.82%
	2021	4.08%	3.42%	200.44m	168.3m	20.88%	23.08%
Low-Income	1999	0.36%	0.38%	160.47m	164.29m	21.08%	22.45%
	2010	0.64%	0.73%	160m	170.97m	22.73%	24.75%
	2021	4.27%	5.39%	203.66m	207.4m	25.69%	28.96%
Unemployment	1999	2.25%	2.36%	194.17m	173.49m	29.76%	27.54%
	2010	3.75%	4.11%	199.25n	199m	29.13%	29.28%
	2021	4.57%	5.37%	219.74m	206.91m	29.33%	29.38%

Appendix H: PAM Clustering Social Frontiers for Chapter 4

As a robustness check for the Lee and Mitchell (2013) identification method, I also identified the social frontiers in Chapter 4 using the PAM clustering method proposed in Chapter 3. The methodology for the PAM clustering method can be found in Chapter 3.

Methodological differences

Although Chapter 4 does not focus on the social frontier identification methods, the PAM clustering method was also included as a secondary method to act as a robustness check for the Lee and Mitchell (2013) method. There is variation in the results of the two methods that can partly be explained by the findings of the simulation study in Chapter 3. The strength measurement shows that the PAM clustering method has a lower average difference, indicating that it is more likely to identify a social frontier. For the Lee and Mitchell (2013) method, the average frontier difference varied between 20-30%, whereas the PAM method has a larger range, varying from as low as a 9% difference up to a 22% difference. This difference is particularly clear for the non-Western, unemployment, and high-income frontiers where the percentage prevalence trends are similar but there is a large difference in the social frontier percentages of the two methods.

The two main findings in the simulation study in Chapter 3 were that, firstly, the PAM clustering method struggled to correctly identify non-frontiers when the frontier was weaker and especially when the population was small. Secondly, the Lee and Mitchell (2013) method has the opposite issue in that when the population and frontier strength was small, the true positive score for the method reduced considerably, highlighting the conservative nature of the method. In contrast, the PAM method is more likely to overestimate the frontiers but is unlikely to miss what could be a crucial boundary which the Lee and Mitchell (2013) method may overlook. This is shown in the results for all four dimensions, with the PAM method identifying considerably more social frontiers.

These concepts can help to explain the difference in the results of the high-income social frontiers. The high-income frontiers were the least common when identified using the Lee and Mitchell (2013) method but the most frequently identified by PAM clustering. The high-income variable was the smallest population of the four dimensions, with the differences between the cells minimised. In the simulation study, the Lee and Mitchell (2013) method averaged a 79.2% true positive rate, indicating that it was likely to underestimate the number of social frontiers, whilst the PAM method had a 90% true negative rate, meaning that 10% of the boundaries were incorrectly identified as social frontiers when the frontier strength and population were small. These findings can help to explain why the difference in the

social frontier percentages were extreme for the high-income dimension. It is likely that the true prevalence falls somewhere between the two methods.

For Chapter 4, I have chosen to base the main findings on the Lee and Mitchell (2013) method as this is the more conservative estimate. Therefore, the results were more likely to underestimate any trends, meaning that the frontiers identified are strong and more likely to have an influence on a community.

PAM social frontier results

In contrast to the Lee and Mitchell (2013) method, the non-Western dimension was the least frequent for PAM social frontiers. However, the pattern is similar, increasing between 1999 and 2021 in both Amsterdam and the Hague. The percentage of non-Western social frontiers in Amsterdam in 2021 is 1.5 times that of 1999, and 1.49 times for the Hague. For 15 of the years in the Hague and 16 in Amsterdam the percentage increased year on year with it never once decreasing in consecutive years. The decreases mainly occurred between 2000 and 2010 in Amsterdam, and 2010 and 2021 in the Hague. Despite this, the year-on-year changes were minimal with the maximum decrease of 0.81 and increase of 1.38. It has a very similar pattern as the Lee and Mitchell (2013) method, indicating that the frontiers are increasing. However, it does suggest that the Lee and Mitchell (2013) method slightly underestimates the percentage of frontiers, although likely not as significantly as the other dimensions.

Similarly, the high-income frontiers showed a very similar pattern to the Lee and Mitchell (2013) method with little change over the time-period; this was also true for the length of the frontiers. There is a large difference in the percentage of social frontiers identified, approximately 20 percentage points, but the pattern is consistent. This difference in percentage is likely due to the difference in the simulation study results discussed previously in Chapter 3.

Likewise, the percentage of unemployment PAM social frontiers is consistent throughout the years and similar to the pattern I identified for the high-income frontiers. The social frontier percentage in 2021 was not largely different to that of 1999 for both regions - a slight decrease for the Hague. However, in comparison to the high-income frontiers, the unemployment frontiers experience a large number of peaks and troughs, especially between 2010 and 2021. Despite these changes, by 2021 they were the most commonly identified social frontiers in both regions when using the PAM method. This is not largely dissimilar to the Lee and Mitchell (2013) method where they were the 2nd most common, indicating that it is likely to be a strong dividing factor between communities.

Our final frontier dimension is the one that has the greatest range in results when comparing the methods and also when considering the changes of the frontier percentages year on year: low-income frontiers. Interestingly, the PAM method

showed almost the opposite pattern to the Lee and Mitchell (2013), with them beginning as the most common frontiers: almost 40% for Amsterdam and 35% for the Hague, before reducing over the 20 years by around 8 percentage points for each region. The pattern is inconsistent, with large periods of decreases especially around 2016 and the percentage tending to be more likely to decrease in years than increase, the opposite pattern to what we see in the Lee and Mitchell (2013) method. It also has some of the largest decreases of any of the variables with a 4.8 percentage point decrease between 2016 and 2017 in Amsterdam and similar in the Hague. These inconsistencies between the results of the Lee and Mitchell (2013) and PAM clustering methods results potentially suggest inaccuracies. We saw an increase in both the low-income percentage and low-income index of dissimilarity, neither of which were replicated in the PAM social frontier percentage.

Table H.1: Social frontier prevalence table for PAM clustering.

		Percentage		Average Length		Average Difference	
		Amsterdam	Hague	Amsterdam	Hague	Amsterdam	Hague
Country of Birth (Non-Western)	1999	14.14%	13.3%	564.33m	497.75m	22.23%	19.33%
	2010	17.56%	16.32%	584.28m	587.19m	21.16%	22.03%
	2021	22.3%	19.88%	665.69m	675.63m	19.86%	21.03%
High-Income	1999	25.7%	21.65%	585.88m	610.42m	13%	14.07%
	2010	26.45%	21.9%	607.32m	593.82m	13.33%	14.61%
	2021	26.56%	21.36%	526.43m	500.97m	13.82%	15.32%
Low-Income	1999	39.42%	35.6%	673.93m	700.48m	9.7%	9.49%
	2010	38.89%	35.05%	701.78m	749.78m	10.49%	10.51%
	2021	31.09%	24.75%	540.8m	618.96m	14.36%	16.07%
Unemployment	1999	33.26%	31.32%	605.30m	674.03m	14.89%	15.47%
	2010	32.48%	31.67%	645.16m	683.46m	15.5%	16.35%
	2021	32.17%	28.33%	581.24m	651.62m	14.4%	17.18%

Appendix I: Frontier Lifecycle Plots

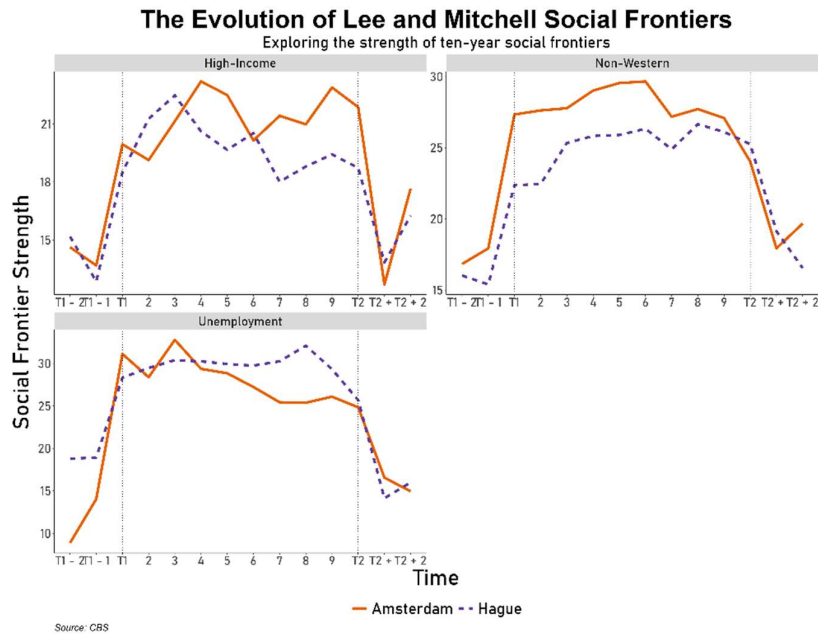


Figure I.1: Frontier lifecycle plot showing the evolution of the strength (median difference) of the 10-year long social frontiers identified using the Lee and Mitchell (2013) method in Amsterdam and the Hague between 1999 and 2021. There were no low-income social frontiers which consistently lasted for 10 or 23 years and thus there is no low-income dimension in the plot.

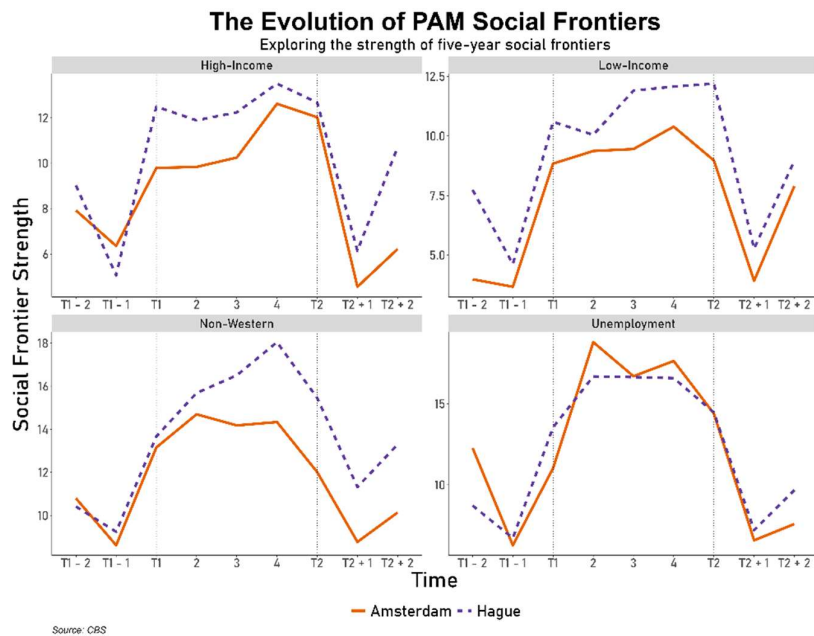


Figure I.2: Frontier lifecycle plot showing the evolution of the strength (median difference) of the 5-year long social frontiers identified using the PAM clustering method in Amsterdam and the Hague between 1999 and 2021.

Appendix J: Descriptive Statistics for Intersectional Frontiers in Amsterdam and the Hague

Table J.1: Amsterdam descriptive statistics							
Year	Population	Non-Western Low-Income Population	Non-Western Unemployed Population	Non-Western High-Income Population	Non-Western Low-Income %	Non-Western Unemployed %	Non-Western High-Income %
1999	1,103,145	35,418	58,452	3,423	3.21%	5.30%	0.31%
2000	1,111,611	36,223	58,090	4,116	3.26%	5.23%	0.37%
2001	1,119,196	35,620	58,803	4,378	3.18%	5.25%	0.39%
2002	1,127,106	36,251	61,076	4,524	3.22%	5.42%	0.40%
2003	1,136,468	36,103	64,986	4,451	3.18%	5.72%	0.39%
2004	1,146,135	38,211	69,810	4,459	3.33%	6.09%	0.39%
2005	1,150,360	38,786	69,696	4,826	3.37%	6.06%	0.42%
2006	1,154,787	39,684	66,432	5,141	3.44%	5.75%	0.45%
2007	1,164,816	44,421	62,703	5,496	3.81%	5.38%	0.47%
2008	1,176,969	41,088	61,555	5,902	3.49%	5.23%	0.50%
2009	1,193,311	42,445	64,837	6,145	3.56%	5.43%	0.51%
2010	1,206,965	44,062	67,094	6,498	3.65%	5.56%	0.54%
2011	1,218,855	45,882	69,808	6,778	3.76%	5.73%	0.56%
2012	1,234,173	47,732	73,220	7,028	3.87%	5.93%	0.57%
2013	1,245,879	49,139	77,186	7,394	3.94%	6.20%	0.59%
2014	1,258,820	50,507	79,229	8,131	4.01%	6.29%	0.65%
2015	1,268,148	53,635	78,928	8,713	4.23%	6.22%	0.69%
2016	1,279,994	55,239	78,893	9,362	4.32%	6.16%	0.73%
2017	1,292,499	65,328	77,916	10,454	5.05%	6.03%	0.81%
2018	1,302,673	65,972	76,625	11,690	5.06%	5.88%	0.90%
2019	1,318,742	68,597	76,163	13,220	5.20%	5.78%	1.00%
2020	1,316,946	71,913	78,905	14,175	5.46%	5.99%	1.08%
2021	1,322,412	73,505	79,782	16,596	5.56%	6.03%	1.25%

These statistics are created using the non-public microdata from Statistics Netherlands.

Table J.2: The Hague descriptive statistics.

Year	Population	Non-Western Low-Income Population	Non-Western Unemployed Population	Non-Western High-Income Population	Non-Western Low-Income %	Non-Western Unemployed %	Non-Western High-Income %
1999	715,557	19,401	31,591	2,143	2.71%	4.41%	0.30%
2000	723,882	20,417	32,580	2,431	2.82%	4.50%	0.34%
2001	730,196	21,162	33,664	2,678	2.90%	4.61%	0.37%
2002	738,257	21,529	35,785	2,831	2.92%	4.85%	0.38%
2003	747,061	22,137	39,435	2,851	2.96%	5.28%	0.38%
2004	753,257	23,696	42,748	2,819	3.15%	5.68%	0.37%
2005	759,109	24,750	44,145	2,881	3.26%	5.82%	0.38%
2006	760,098	24,967	41,962	3,035	3.28%	5.52%	0.40%
2007	764,505	24,806	39,524	3,292	3.24%	5.17%	0.43%
2008	775,020	25,851	39,657	3,460	3.34%	5.12%	0.45%
2009	782,662	26,515	41,872	3,646	3.39%	5.35%	0.47%
2010	791,528	27,594	44,486	3,752	3.49%	5.62%	0.47%
2011	799,654	29,143	46,436	4,182	3.64%	5.81%	0.52%
2012	804,889	30,495	49,294	4,342	3.79%	6.12%	0.54%
2013	809,601	31,016	51,438	4,311	3.83%	6.35%	0.53%
2014	817,629	32,261	53,353	4,772	3.95%	6.53%	0.58%
2015	821,159	34,514	54,440	4,985	4.20%	6.63%	0.61%
2016	828,758	35,205	54,712	5,270	4.25%	6.60%	0.64%
2017	838,441	41,331	53,540	5,624	4.93%	6.39%	0.67%
2018	846,372	41,106	52,014	6,146	4.86%	6.15%	0.73%
2019	856,449	42,750	51,106	6,696	4.99%	5.97%	0.78%
2020	859,893	44,143	52,643	7,099	5.13%	6.12%	0.83%
2021	863,025	45,110	52,448	7,825	5.23%	6.08%	0.91%

These statistics are created using the non-public microdata from Statistics Netherlands.

Appendix K: Intersectional Social Frontier Prevalence

Table K.1: Intersectional social frontier percentage and median length (% , m) in the Hague

Intersectional Frontiers							Individual Frontiers			
Year	Non-Western Unemployment		Non-Western Low-Income		Non-Western High-Income		Non-Western	High-Income	Low-Income	Unemployment
	Coterminals	Multidimensional	Coterminals	Multidimensional	Coterminals	Multidimensional				
1999	0.632% (100m)	2.729% (200m)	0.522% (100m)	1.197% (100m)	0.372% (100m)	0.000%	8.08%	3.49%	0.38%	2.36%
2000	1.805% (100m)	2.401% (200m)	1.805% (100m)	1.515% (200m)	0.425% (100m)	0.000%	9.30%	3.28%	0.46%	2.52%
2001	0.338% (100m)	3.037% (200m)	0.338% (100m)	1.400% (200m)	0.440% (100m)	0.000%	8.98%	3.21%	0.39%	2.71%
2002	0.362% (100m)	2.804% (200m)	0.362% (100m)	1.209% (100m)	0.177% (100m)	0.008% (100m)	8.98%	3.32%	0.34%	3.1%
2003	0.450% (100m)	3.113% (200m)	0.450% (100m)	1.268% (100m)	0.848% (100m)	0.008% (100m)	8.90%	3.20%	0.43%	3.23%
2004	1.871% (100m)	3.036% (200m)	1.871% (100m)	1.284% (100m)	1.069% (100m)	0.000%	9.22%	3.42%	0.43%	3.46%
2005	0.636% (100m)	3.533% (200m)	0.636% (100m)	1.265% (100m)	0.863% (100m)	0.000%	8.85%	3.25%	0.58%	3.7%
2006	0.398% (100m)	2.973% (100m)	0.398% (100m)	1.519% (100m)	1.186% (100m)	0.000%	9.63%	3.18%	0.65%	3.62%
2007	0.395% (100m)	2.815% (100m)	0.395% (100m)	1.530% (200m)	1.049% (100m)	0.022% (150m)	10.14%	3.19%	0.50%	3.39%
2008	1.818% (100m)	2.936% (100m)	1.818% (100m)	1.429% (100m)	1.408% (100m)	0.014% (200m)	10.89%	3.00%	0.62%	3.7%
2009	1.775% (100m)	3.089% (200m)	1.775% (100m)	1.524% (100m)	1.048% (100m)	0.000%	10.14%	3.02%	0.63%	3.66%
2010	0.833%	3.243%	0.833%	1.542%	1.583%	0.000%	10.85%	3.20%	0.73%	4.11%

	(100m)	(100m)	(100m)	(100m)	(100m)					
2011	1.737%	3.378%	1.737%	1.696%	1.340%	0.000%	10.49%	3.08%	0.86%	4.6%
	(100m)	(100m)	(100m)	(100m)	(100m)					
2012	1.418%	3.562%	1.418%	1.877%	1.158%	0.014%	10.68%	3.28%	1.09%	4.81%
	(100m)	(100m)	(100m)	(100m)	(100m)	(200m)				
2013	1.852%	3.670%	1.818%	1.832%	1.347%	0.041%	10.39%	3.50%	1.05%	5.14%
	(100m)	(100m)	(100m)	(100m)	(100m)	(150m)				
2014	2.444%	3.771%	2.444%	1.926%	1.298%	0.048%	11.11%	3.62%	1.13%	5%
	(100m)	(100m)	(100m)	(100m)	(100m)	(150m)				
2015	1.096%	4.092%	1.089%	2.239%	1.300%	0.047%	11.19%	3.61%	1.64%	5.38%
	(100m)	(100m)	(100m)	(100m)	(100m)	(200m)				
2016	2.181%	4.375%	2.174%	2.694%	1.354%	0.054%	10.79%	3.90%	1.77%	5.43%
	(100m)	(100m)	(100m)	(100m)	(100m)	(100m)				
2017	0.909%	4.559%	0.909%	4.290%	1.480%	0.074%	11.20%	3.71%	4.11%	5.28%
	(100m)	(100m)	(100m)	(200m)	(100m)	(150m)				
2018	0.929%	4.418%	0.929%	4.317%	1.406%	0.034%	11.48%	3.64%	4.21%	5.4%
	(100m)	(100m)	(100m)	(100m)	(100m)	(100m)				
2019	0.838%	4.412%	0.838%	4.472%	1.347%	0.067%	11.48%	3.42%	5.31%	5.25%
	(100m)	(150m)	(100m)	(200m)	(100m)	(200m)				
2020	2.774%	4.411%	2.774%	4.358%	1.455%	0.141%	11.47%	3.62%	5.10%	5.2%
	(100m)	(100m)	(100m)	(200m)	(100m)	(100m)				
2021	2.181%	4.461%	2.181%	4.748%	1.453%	0.121%	11.53%	3.42%	5.39%	5.37%
	(100m)	(200m)	(100m)	(200m)	(100m)	(100m)				

Table K.2: Intersectional social frontier percentage and median length (% , m) in Amsterdam.

	Non-Western Unemployment		Non-Western Low-Income		Non-Western High-Income		Non-Western	High-Income	Low-Income	Unemployment
Year	Cotermi nus	Multidime nsional	Cotermi nus	Multidime nsional	Cotermi nus	Multidime nsional				
1999	0.922%	3.100%	1.364%	1.429%	1.248%	0.000%	9.36%	3.37%	0.36%	2.25%
	(100m)	(100m)	(100m)	(100m)	(100m)					

2000	0.912% (100m)	3.011% (200m)	1.646% (100m)	1.543% (100m)	1.182% (100m)	0.000%	8.70%	3.22%	0.32%	2.41%
2001	0.867% (100m)	2.979% (200m)	1.733% (100m)	1.230% (100m)	1.229% (100m)	0.000%	8.89%	3.57%	0.22%	2.17%
2002	0.925% (100m)	2.918% (100m)	0.999% (100m)	1.207% (100m)	1.120% (100m)	0.026% (250m)	8.87%	3.48%	0.30%	2.43%
2003	0.938% (100m)	3.008% (100m)	1.362% (100m)	1.231% (100m)	1.183% (100m)	0.010% (100m)	9.00%	3.44%	0.28%	2.31%
2004	1.004% (100m)	3.016% (100m)	1.305% (100m)	1.241% (100m)	1.155% (100m)	0.024% (100m)	8.70%	3.41%	0.23%	2.63%
2005	1.150% (100m)	3.390% (200m)	0.929% (100m)	1.274% (100m)	1.303% (100m)	0.041% (100m)	9.24%	3.56%	0.26%	2.97%
2006	1.107% (100m)	3.140% (100m)	1.529% (100m)	1.346% (100m)	1.152% (100m)	0.000%	9.03%	3.46%	0.25%	2.86%
2007	1.063% (100m)	3.195% (100m)	1.236% (100m)	1.761% (100m)	1.382% (100m)	0.010% (100m)	8.31%	3.53%	0.98%	3.06%
2008	1.138% (100m)	3.108% (100m)	1.671% (100m)	1.642% (100m)	1.287% (100m)	0.005% (100m)	8.84%	3.51%	0.61%	3.19%
2009	1.131% (100m)	3.251% (100m)	1.415% (100m)	1.599% (100m)	1.223% (100m)	0.014% (150m)	8.37%	3.77%	0.64%	3.39%
2010	1.355% (100m)	3.708% (100m)	1.214% (100m)	1.721% (100m)	1.286% (100m)	0.032% (100m)	9.00%	3.93%	0.64%	3.75%
2011	1.586% (100m)	3.827% (100m)	1.605% (100m)	1.726% (100m)	1.376% (100m)	0.019% (100m)	9.04%	3.89%	0.78%	4.10%
2012	1.739% (100m)	4.004% (100m)	1.428% (100m)	1.986% (100m)	1.470% (100m)	0.009% (200m)	9.50%	3.79%	0.87%	4.41%
2013	1.830% (100m)	4.251% (100m)	1.728% (100m)	2.116% (100m)	1.396% (100m)	0.018% (100m)	9.90%	3.84%	0.95%	4.41%
2014	1.788% (100m)	4.497% (200m)	1.368% (100m)	2.152% (100m)	1.329% (100m)	0.041% (100m)	9.80%	3.88%	0.94%	4.40%
2015	1.834% (100m)	4.358% (200m)	1.669% (100m)	2.395% (100m)	1.490% (100m)	0.051% (200m)	9.96%	4.01%	1.11%	4.62%

2016	1.872% (100m)	4.619% (100m)	1.351% (100m)	2.639% (100m)	1.573% (100m)	0.041% (150m)	9.96%	4.22%	1.34%	4.79%
2017	1.949% (100m)	4.796% (200m)	1.507% (100m)	3.906% (200m)	1.519% (100m)	0.014% (100m)	10.30%	4.11%	2.77%	4.83%
2018	1.810% (100m)	4.732% (200m)	0.845% (100m)	4.024% (100m)	1.713% (100m)	0.109% (100m)	10.48%	4.00%	2.94%	4.59%
2019	1.982% (100m)	4.818% (200m)	1.391% (100m)	4.590% (200m)	1.485% (100m)	0.142% (100m)	10.23%	4.23%	4.11%	4.76%
2020	2.086% (100m)	4.758% (200m)	1.225% (100m)	4.668% (100m)	1.258% (100m)	0.109% (100m)	10.38%	4.18%	4.20%	4.81%
2021	1.938% (100m)	4.606% (200m)	1.284% (100m)	4.717% (100m)	0.967% (100m)	0.192% (100m)	10.55%	4.08%	4.27%	4.57%

Appendix L: Aggregated Education Levels

Table L.1: Aggregated education levels - the information is from CBS (2024e).

Aggregate Education Level	Education Levels	Education Years
Low	Lower Primary Education	4
	Upper Primary Education	8
	Practical education/VMBO lower /MBO1 ²⁹	12
	VMBO higher/HAVO ³⁰ /VWO lower ³¹	13
	HAVO upper school	13
Secondary	MBO2/VWO upper school	14
	MBO3	15
	MBO4	16
	HBO associate degree ³²	15
High	HBO bachelor's degree/WO bachelor's degree ³³	17
	HBO Masters/WO Masters	18
	Doctorate	22

²⁹ VMBO is a Vocational Education Stream completed at 16. MBO is vocational training.

³⁰ HAVO is a preparatory stream for higher professional education, completed at 17.

³¹ VWO is the university preparatory stream aimed for students aiming at research universities, completed at 18.

³² HBO is the University of Applied Sciences that focuses on professional skills.

³³ WO is the University which focuses on academic research.

Appendix M: Matching Methods Descriptions

Table M.1: Description of matching methods

Matching Method	Explanation
1-1 Nearest Neighbour	Treated individuals are matched to control individuals with the smallest difference in their propensity score. Those with the highest propensity scores are matched first. There are several different options within the method: - With replacement: Control individuals can either be matched multiple times or only used once. Allowing control variables to be reused often decreases bias as a better match for the treated individual is identified (Stuart, 2010). - With caliper: Setting a caliper implements a maximum distance allowed between the propensity scores of the treated and control individuals. The inclusion of a caliper reduces “98% of the bias in a normally distributed covariate” (Stuart, 2010, p. 8).
Coarsened Exact Matching (CEM)	CEM works by grouping individuals into bins based on categories that have been created by temporarily categorising continuous or ordinal variables into broader categories. The individuals are then matched using exact matching on the coarsened version of the covariates before retaining only the original values for the variables (Iacus, King and Porro, 2012). The method often works faster and is easier to understand compared to traditional matching methods (Blackwell <i>et al.</i>, 2009).
Subclassification	A stratification method which groups individuals into classes based on quintiles of their propensity score; it is able to identify both the ATT and the ATE (Stuart, 2010; Zhao <i>et al.</i>, 2021).
Generalised Full Matching	Generalised full matching (GFM) is a slightly adapted version of full matching introduced by Sävje, Higgins and Sekhon (2021). Full matching groups individuals into subclasses which contain at least 1 control and treated individual, with similar individuals grouped together (Stuart and Green, 2008). GFM uses less memory and operates using a faster algorithm in order to run far quicker than full matching, especially on large datasets (Sävje, Higgins and Sekhon, 2021).

Appendix N: Covariate Balancing Scores

Table N.1: Covariate balance for matching at the age of 10 - values over 0.1 are highlighted.

Variables	Initial Balance	Nearest Neighbour ³⁴	Nearest Neighbour with Replacement ³	GFM	CEM	Subclassification
	SMD (Var. Ratio)	SMD (Var. Ratio)	SMD (Var. Ratio)	SMD (Var. Ratio)	SMD (Var. Ratio)	SMD (Var. Ratio)
distance	0.645	0.154	0.000	0.000		0.018
Gender						
Male	-0.001	-0.002	-0.003	-0.003	0.000	-0.000
Female	0.001	0.002	0.003	0.003	0.000	0.000
Mother Age at Birth	0.015	0.004	0.000	-0.001	-0.010	-0.003
Father Age at Birth	-0.005	-0.000	-0.002	-0.002	-0.006	-0.002
Individual Country of Birth						
Western	0.022	0.002	-0.004	-0.003	0.000	-0.002
Non-Western	-0.022	-0.002	0.004	0.003	0.000	0.002
Mother Country of Birth						
Western	0.072	0.009	-0.007	-0.006	0.000	-0.011
Non-Western	-0.072	-0.009	0.007	0.006	0.000	0.011
Father Country of Birth						
Western	0.070	0.008	-0.010	-0.008	0.000	-0.011
Non-Western	-0.070	-0.008	0.010	0.008	0.000	0.011

³⁴ The results for including a caliper into the matching were very similar and thus are not included within this table.

Parent's Education Level						
Low	-0.007	0.003	0.001	0.001	0.000	0.005
Secondary	-0.065	-0.034	-0.008	-0.006	0.000	-0.008
High	0.084	0.038	0.010	0.006	0.000	0.006
NUTS Zone						
North Netherlands	-0.285	0.002	0.021	0.023	0.000	0.015
North Netherlands/East Netherlands	-0.018	-0.008	0.000	-0.001	0.000	-0.006
East Netherlands	0.003	-0.042	-0.002	-0.003	0.000	-0.004
East Netherlands/West Netherlands	-0.019	-0.001	-0.002	-0.002	0.000	-0.001
West Netherlands	0.160	0.076	0.003	0.003	0.000	0.002
West Netherlands/South Netherlands	-0.003	0.000	0.000	-0.000	0.000	-0.001
South Netherlands	-0.032	-0.050	-0.013	-0.013	0.000	-0.007
South Netherlands/East Netherlands	-0.018	-0.010	-0.010	-0.002	0.000	-0.004
Mother Employment						
Unemployed	-0.099	-0.012	0.003	0.006	0.000	0.002
Employed	0.099	0.012	-0.003	-0.006	0.000	-0.002
Father Employment						
Unemployed	-0.039	0.017	0.008	0.007	0.000	0.005
Employed	0.039	-0.017	-0.008	-0.007	0.000	-0.005
Household Composition						
Live with neither parent	0.010	0.005	-0.004	-0.004	0.000	0.001

Live with both parents	-0.060	-0.032	-0.002	-0.003	0.000	-0.007
Live with Mum	0.055	0.029	0.003	0.003	0.000	0.007
Live with Dad	0.017	0.010	-0.001	0.002	0.000	0.001
Mother SECM Status						
Employee	0.069	0.009	-0.007	-0.005	0.000	-0.001
Director	-0.026	-0.009	-0.002	-0.002	0.000	-0.003
Self-employed	-0.064	-0.014	0.003	0.000	0.000	-0.004
Benefits Recipient	0.013	0.016	0.003	0.003	0.000	0.008
Pension Recipient	-0.002	0.001	0.004	0.003	0.000	0.000
Without income	-0.044	-0.012	0.004	0.004	0.000	-0.002
Father SECM Status						
Employee	0.085	0.009	-0.003	-0.003	0.000	0.000
Director	-0.017	-0.007	-0.003	-0.002	0.000	-0.005
Self-employed	-0.113	-0.020	0.005	0.005	0.000	-0.003
Benefits Recipient	0.007	0.014	0.004	0.005	0.000	0.008
Pension Recipient	-0.003	0.002	0.004	0.001	0.000	0.000
Without income	-0.001	0.003	-0.006	-0.006	0.000	-0.001
Household Ownership Status						
Owned	-0.021	-0.020	-0.003	-0.003	0.000	-0.006
Rented	0.021	0.020	0.003	0.003	0.000	0.006
Percentage of Houses Owned in Grid Cell	-0.108	-0.052	-0.004	-0.003	-0.002	-0.019
Grid Cell Population	0.162	0.062	0.021	0.019	0.016	0.025

Grid Cell Non-Western Population	-0.148	0.012	0.042	0.042	0.076	0.038
Grid Cell Work-Age Population	0.168	0.036	-0.006	-0.007	0.014	0.002
Grid Cell Unemployed Population	-0.114	-0.023	-0.011	-0.014	-0.026	-0.018
House Value	-0.105	-0.040	-0.020	-0.018	-0.032	-0.023
Grid Cell Median House Value	-0.200	-0.068	-0.029	-0.027	-0.045	-0.036
Household Income	0.068	0.014	0.003	0.005	0.038	0.000
Grid Cell Median Household Income	0.144	0.032	0.008	0.010	0.061	0.001
Missing Variables Indicators						
Mother Age at Birth Variable						
Known	0.012	0.003	-0.001	-0.001	0.000	0.001
Unknown	-0.012	-0.003	0.001	0.001	0.000	-0.001
Father Age at Birth Variable						
Known	0.026	0.004	-0.008	-0.006	0.000	-0.003
Unknown	-0.026	-0.004	0.008	0.006	0.000	0.003
Parental Education Variable						
Known	0.073	0.035	0.005	0.001	0.000	0.007
Unknown	-0.073	-0.035	-0.005	-0.001	0.000	-0.007
Household Composition Variable						
Known	-0.005	-0.004	0.000	0.001	0.000	-0.000
Unknown	0.005	0.004	-0.000	-0.001	0.000	0.000
Mother SECM Variable						

Known	0.002	-0.002	0.004	0.004	0.000	-0.000
Unknown	-0.002	0.002	-0.004	-0.004	0.000	0.000
Father SECM Variable						
Known	0.014	-0.001	-0.006	-0.003	0.000	-0.003
Unknown	-0.014	0.001	0.006	0.003	0.000	0.003
Household Ownership Status Variable						
Known	-0.006	-0.004	-0.006	-0.006	0.000	-0.001
Unknown	0.006	0.004	0.006	0.006	0.000	0.001
Grid Cell Population Variable						
Known	0.001	0.000	0.000	-0.002	0.000	-0.001
Unknown	-0.001	0.000	0.000	0.002	0.000	0.001
Grid Cell Non-Western Population Variable						
Known	-0.001	0.000	0.000	-0.002	0.000	-0.001
Unknown	0.001	0.000	0.000	0.002	0.000	0.001
Grid Cell Work-Age Population Variable						
Known	-0.001	0.000	0.000	-0.002	0.000	-0.001
Unknown	0.001	0.000	0.000	0.002	0.000	0.001
Grid Cell Unemployed Population Variable						
Known	-0.001	0.000	0.000	-0.002	0.000	-0.001
Unknown	0.001	0.000	0.000	0.002	0.000	0.001
Median Grid Cell Income Variable						
Known	0.070	0.007	-0.004	0.000	0.000	0.003

Unknown	-0.070	-0.007	0.004	0.000	0.000	-0.003
Household Income Variable						
Known	0.007	-0.008	-0.002	0.003	0.000	-0.002
Unknown	-0.007	0.008	0.002	-0.003	0.000	0.002

Appendix O: Sensitivity Analysis Results

Entropy balancing

The first analysis I apply is entropy balancing. This is a technique that aims to achieve covariate balance where individuals are reweighted to fulfil a set of balance constraints (Hainmueller, 2012). I can then use these weights to find the treatment effect of the weighted sample by fitting a linear model (Greifer, 2025), the results of which are reported in Tables O.1 and O.2. The results of the entropy balancing are very similar to the nearest neighbour matching with replacement regression included in Tables 6.6 and 6.8. With even the pattern of peaking at 23 reflected for the ATT on income and 21 for the education ATT. Thus, the results from the entropy balancing seem to support the findings of the matching results.

Table O.1: Average Treatment Effect on the Treated (ATT) for income using entropy balancing

	Estimate	Std. Error	p-value
Income at 20	-463.90	40.30	<0.001
Income at 21	-613.33	39.76	<0.001
Income at 22	-701.48	46.27	<0.001
Income at 23	-761.21	53.12	<0.001
Income at 24	-701.81	65.49	<0.001
Income at 25	-631.29	79.04	<0.001

Table O.2: Average Treatment Effect on the Treated (ATT) for education years using entropy balancing

	Estimate	Std. Error	p-value
Education Years at 20	-0.0733	0.00620	<0.001
Education Years at 21	-0.0904	0.00688	<0.001
Education Years at 22	-0.0791	0.00734	<0.001
Education Years at 23	-0.0524	0.00740	<0.001
Education Years at 24	-0.0354	0.00747	<0.001
Education Years at 25	-0.0246	0.00747	0.00101

Inverse Probability Weighting (IPW)

Inverse Probability Weighting (IPW) weights individuals by the inverse probability of receiving the treatment. The weights are created for the treated individuals using $1/Propensity\ Score$ and $1/(1 - Propensity\ Score)$ for the control group (Chesnaye, 2022, p. 16). For example, an individual who was unlikely to receive the treatment, e.g. a propensity score of 0.25 who received the treatment would have the weight of

4, whilst if a control individual had the same score their weight would be 1.33. Therefore, the treated individual would have a larger influence on the effect (Chesnaye, 2022, p. 16). To improve the study and the model, I trim the extreme propensity scores and weights which limits the impact of large weights. As with the entropy balancing, the IPW weights are used within a linear regression model. The results, visible in Tables O.3 and O.4, are also similar to the matching and entropy balancing models, again providing support for the findings.

Table O.3: Average Treatment Effect on the Treated (ATT) for income using IPW

	Estimate	Std. Error	p-value
Income at 20	-450.40	43.83	<0.001
Income at 21	-588.37	39.35	<0.001
Income at 22	-671.12	45.48	<0.001
Income at 23	-742.61	51.89	<0.001
Income at 24	-685.90	63.91	<0.001
Income at 25	-655.11	85.08	<0.001

Table O.4: Average Treatment Effect on the Treated (ATT) for education years using IPW

	Estimate	Std. Error	p-value
Education Years at 20	-0.0703	0.00584	<0.001
Education Years at 21	-0.0895	0.00650	<0.001
Education Years at 22	-0.0801	0.00696	<0.001
Education Years at 23	-0.0565	0.00704	<0.001
Education Years at 24	-0.0421	0.00712	<0.001
Education Years at 25	-0.0319	0.00714	<0.001

Matching with common support

My final sensitivity analysis technique is matching with common support. The common support is the overlap of the propensity scores for the treated and control individuals. When restricting the common support, I remove control individuals whose propensity score did not fall within the common support region. This can in some situations reduce potential unmeasured confounding (Greifer and Stuart, 2021, p. 120). The results for matching with common support, visible in Tables O.5 and O.6, are incredibly similar to the results of the matching, with very few control individuals unsuitable for matching.

Table O.5: Average Treatment Effect on the Treated (ATT) for income for common support

	Estimate	Std. Error	p-value
Income at 20	-457	54.6	<0.001
Income at 21	-578	52.9	<0.001
Income at 22	-673	59.9	<0.001
Income at 23	-731	67.9	<0.001
Income at 24	-692	94.1	<0.001
Income at 25	-697	118	<0.001

Table O.6: Average Treatment Effect on the Treated (ATT) for education years for common support

	Estimate	Std. Error	p-value
Education Years at 20	-0.0647	0.00817	<0.001
Education Years at 21	-0.0897	0.00904	<0.001
Education Years at 22	-0.0814	0.00957	<0.001
Education Years at 23	-0.054	0.00956	<0.001
Education Years at 24	-0.0409	0.00956	<0.001
Education Years at 25	-0.0314	0.00951	<0.001