MOBI-TWIN

TWIN TRANSITION AND CHANGING PATTERNS OF SPATIAL MOBILITY: A REGIONAL APPROACH

MOBI-TWIN D3.1 METHODOLOGICAL REPORT DESCRIBING THE MOBI-TWIN MODEL

Work package	WP3: Assessing the impact of spatial mobility on EU regions. Updates and next steps
Task	Task 3.1: Design and development of the MOBI-TWIN model.
Authors	**Dimitris Ballas-RUG (main author), Thanasis Ziogas-RUG (main author), Jianhua Zhang, Wander Jager, Junyao He** Contributors:
Lead Beneficiary	RUG
Reviewer(s)	AUTH - Anastasia Panori POLIMI – Manos Matsaganis
Status	Draft
Version	1.3
Due Date	31/08/2025
Issue Date	31/08/2025
Dissemination Level	PU

Funded by the European Union

MOBI-TWIN has received funding from the European Union's Horizon Europe Research and Innovation Programme under Grant Agreement no. 101094402. Content reflects only the authors' view and European Commission is not responsible for any use that may be made of the information it contains.

Abstract	This document contains a methodological report presenting the MOBITWIN model. The report includes an introduction of the key methods that are being used in the context of the state of the art and previous relevant work. It also considers the relevance of the model to the MOBITWIN project objectives as well as other deliverables. The report also includes a discussion of how the methods were operationalised (with extracts of illustrative code in R) and used. It also includes model outputs and validation in relation to the analysis conducted so far. The model outputs presented include maps of estimated population subgroups, which should be seen as a demonstrator and an example of the potential of the model for the next steps in terms of scenario analysis. The model outputs presented here will be the basis for further dynamic analysis and agent-based modelling, building on the frameworks developed so far.
Disclaimer	

Funded by the European Union

MOBI-TWIN has received funding from the European Union's Horizon Europe Research and Innovation Programme under Grant Agreement no. 101094402. Content reflects only the authors' view and the European Commission is not responsible for any use that may be made of the information it contains.

Version	Date	Partner	Description
0.1	11/07/2025	RUG	Draft outline
0.2	31/07/2025	RUG	First draft (sent to reviewers - POLIMI and AUTH)
1.0	20/08/2025	AUTH	Review
1.1	29/08/2025	POLIMI	Review
2.0	29/08/2025	RUG	Final version

EXECUTIVE SUMMARY

This deliverable presents the methodological foundations, implementation, and initial outputs of the MOBI-TWIN spatial microsimulation model. The work reported here falls under Work Package 3 (WP3), work task 3.1 which focuses on the development of tools and methods to estimate residential mobility and to support future scenario-based analyses developed under work task 3.2 . This deliverable is closely aligned with the broader objectives of the MOBI-TWIN project, particularly the goal of generating synthetic microdata that can be used to inform policy and strategic planning in the context of the Twin Transition (Green and Digital). The deliverable is also directly connected to other project deliverables both in terms of useful information and model inputs as well as next tasks that will build on these models for dynamic and agent-based simulations in the next project stages.

A key contribution of this deliverable is the development and operationalization of spatial microsimulation models for five pilot regions: Central Macedonia (Greece), Castile-La Mancha (Spain), Lombardy (Italy), Groningen (Netherlands), and North and East Finland (Finland). For each pilot, the model combines country-specific subsets of the MOBI-TWIN survey with census or administrative data obtained from national statistical authorities (ELSTAT, INE, ISTAT, CBS, Tilastokeskus). Constraint variables, such as age, biological sex, employment status, education level, income, marital status, and ethnicity (where available), were selected based on their relevance to mobility behavior, availability in both survey and census data, and ability to represent socio-demographic diversity. The resulting synthetic populations assign realistic characteristics and geolocation to individuals, enabling the estimation of probabilities of residential mobility over different time horizons (6 months, 1 year, 5 years, 10 years).

The main results presented in this deliverable include:

- Model outputs in the form of maps of simulated population groups and sub-groups which can be seen as illustrative examples for the next stages.
- Internal validation graphs for all pilot regions, demonstrating high accuracy in reproducing the observed distributions of the constraint variables at small-area levels
- External validation results, where tax return data at the prefecture level confirmed that the simulated counts of low-income individuals closely match real-world data

- Socio-demographic profiles and spatial patterns of mobility across the pilot regions, highlighting areas with higher or lower mobility potential and

providing insights into possible demographic imbalances (e.g., aging populations or areas with low relocation propensity)

These outputs demonstrate the robustness and policy relevance of the MOBI-TWIN microsimulation approach and serve as a foundation for future analyses. In subsequent phases, the models developed in this deliverable will be used to generate dynamic simulations and support agent-based modelling, allowing the project to explore what-if scenarios under the Twin Transition and to assess the implications of changing demographic and socio-economic conditions on residential mobility.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

List of terms and abbreviations

ABM	Agent-Based Model
EU-SILC	European Union Statistics on Income and Living Conditions
EU	European Union
IPF	Iterative Proportional Fitting
NUTS2	Nomenclature of Units for Territorial Statistics. Level 2.
NUTS3	Nomenclature of Units for Territorial Statistics. Level 3.
RRI	Responsible Research and Innovation
TT	Twin Transition

1 INTRODUCTION

This report presents deliverable 3.1 (Methodological report describing the MOBI-TWIN model) which is the output of task 3.1, contributing to MOBI-TWIN objective 3. Work under this task involved the development of a spatial microsimulation model aimed at creating small area microdata that can be used to evaluate the impact of mobility relating to the Twin Transition (TT). It also includes a conceptual agent-based model for the simulation of mobilities pertaining to the TT. Further, it includes the use of Geographical Information Systems (GIS) mapping and econometric analysis of the model outputs. This report presents the data, methods and outputs and is accompanied by code in R and the datasets used.

As indicated in the MOBI-TWIN application and research proposal, work under task 3.1 involves the use of spatial microsimulation. As discussed in some of the work of the co-authors of this report (Ballas *et al*., 2005a and 2005b and Panori, Ballas and Psycharis 2017, upon which we draw on to explain the method in this introductory section) microsimulation models aim to build large-scale data sets on the attributes of individuals or households (and/or on the attributes of individual firms or organisations) and to analyse policy impacts on these micro-units. These micro-units are often described as 'synthetic' individuals or households because they have the most likely attributes of an individual unit, but do not represent a *real* individual.[1] The analysis at the level of the synthetic individual or household enables the assessment of variations in the distributional effects of different policies. Microsimulation is a technique that has been broadly developed and used by economists and other social scientists. The results of national (aspatial) microsimulation models are typically quoted in the media when covering the possible impact of government budget changes upon different types of households. Microsimulation models become spatial when geographical information about the simulated entities is available (or estimated). Spatial microsimulation involves the creation of large-scale population microdata sets and the analysis of the impacts of any policy changes which change the attributes contained in these micro databases in some way (Ballas *et al*., 2005b). Adding spatial detail to traditional microsimulation involves creating geographically-referenced microdata that refers to a particular locality. Since there are very few sources of geographically detailed microdata, there is a need to create these data using spatial microsimulation techniques by merging small area data (typically census of population data) and social survey data to simulate a population of individuals within households (for different geographical units) whose characteristics are as close to the real population as it is possible to estimate.

[1] In contrast, in other types of microsimulation (i.e. EUROMOD) the attributes refer to actual individuals (see Sutherland and Figari (2013)).

Spatial microsimulation is also increasingly seen as a powerful method for small area estimation which can be used for the analysis of migration and other forms of mobility and there have already been some efforts to that direction (including the work of some of the co-authors of this report; see Ballas et al. , 2005a and 2005b).[2]

The work presented in this report builds on previous research by the one of the co-authors in the development and application of spatial microsimulation models that are capable of generating small area microdata based on estimation methods ranging from Iterative Proportional Fitting-based deterministic reweighting (Panori, Ballas and Psycharis, 2017; Lovelace and Ballas, 2013; Ballas, 2004) and synthetic reconstruction (Ballas and Clarke, 2000) to combinatorial optimisation heuristic techniques such as simulated annealing (Kavroudakis, Ballas and Birkin, 2011). The simulation outputs (i.e. the attributes of the synthetic individuals) include a wide range of policy relevant variables such as earned income, tenure status, household type, socio-economic group, consumption patterns, car ownership and so forth. Workpackage 3 also benefits from and builds upon the work of colleagues from POLIMI on the Tax-Benefit Microsimulation work for the European Union EUROMOD[3]. In particular, MOBI-TWIN task 3.1 involves a combination of social survey data (including the MOBITWIN survey but also the EU-SILC dataset and its EUROMOD version) with small area data to build spatial microsimulation models for the pilot areas under investigation, using variables deemed relevant to migration and other spatial mobility patterns as well as the TT (e.g. socio-economic group, occupation, educational qualifications, age, gender). To that end data from the MOBI-TWIN survey (prepared and collected under D1.2) were combined with spatial data from suitable sources to build a simulated intra-regional area microdata sets that can be used to identify, map and analyse synthetic population groups and their attributes that may be most affected by TT related mobility and migration.

1.1 Document structure

The remainder of this report is structured as follows: Section 2 provides an overview of Work Package 3 together with an introduction to the spatial microsimulation methodology and its application for the analysis of spatial mobility as well as a description of the data that have been used for each of the pilot areas. Section 3 focuses on the five pilot cases, with subsections focusing on the validation of the spatial microsimulation model, population description, and regression analysis for each of the pilot cases. Section 4 discusses the incorporation of Responsible Research and Innovation (RRI) pillars while section 5

[2] https://www.youtube.com/watch?v=g0I87SuRSWg
[3] https://euromod-web.jrc.ec.europa.eu

offers some concluding comments. The report also included accompanying annexes.

2 SOCIAL SIMULATIONS: SPATIAL MICROSIMULATION AND AGENT-BASED MODELLING

2.1 OVERVIEW

Spatial microsimulation models build on a long tradition of modelling for the analysis of urban and regional policy analysis, economic impact assessment and spatial analysis of social policies by economists and regional scientists such as geographers and spatial planners. In particular, regional economists and demographers have long been involved in the development of methodological frameworks such as population projection methods, migration modelling as well as urban and regional policy analysis and regional multiplier and input-output models to estimate the economic impact of major economic events (such as major investments and disinvestments) and external shocks upon regions and industrial sectors of the economy (McCann, 2013; Ballas and Clarke, 2009; Clarke, 1996). Most of these modelling efforts typically aim at estimating the macro and regional-level impacts of major events, but they do not disaggregate the analysis at the individual or household level. On the other hand, economists have long been involved in the development of microsimulation models that are capable of analysing the impact of policies and economics for different types of households. The first microsimulation effort (without any attempt to consider geographical aspects) dates back to the efforts of Guy Orcutt (1957). Since then there has been significant progress and the results of national (*aspatial*) microsimulation models are widely quoted in the media when covering the possible impact of government budget changes upon different types of households. Microsimulation models aim to build large-scale data sets on the attributes of individuals or households (including demographic attributes such as age, gender, marital status and household type as well as socio-economic information including job status and occupation, educational qualifications and income) and to analyse policy impacts on these micro-units; by permitting analyses at the level of the individual, family or household they provide the means of assessing variations in the distributional effects of different policies (e.g. see Mitton *et al*., 2000; Redmond *et al*., 1998). Amongst the most successful and impactful microsimulation efforts is the EUROMOD model and our project benefits from the participation of a key expert that has long been involved in its development and applications, Prof Manos Matsaganis from partner POLIMI.

In addition, the MOBITWIN project benefits from previous efforts of the RUG and AUTH team members in adding geography to microsimulation models, which involves adding spatial information about the simulated entities is available (or estimated). The conceptual framework and blueprint for a geographical dimension to be added to aspatial microsimulation was developed over 50 years ago (Wilson and Pownall, 1976). Since then there has been a growing number of relevant spatial microsimulation studies and applications (for comprehensive reviews see Ballas and Clarke, 2009; Ballas et al., 2019).

The deliverable presented here is the output of a task aimed at adding geography to microsimulation modelling. The deliverable includes implementing code in the programming language R that can be used to create small area population microdata for the MOBI-TWIN pilot regions (with the possibility of adapting the model for any other region), which can then be used to identify populations and population sub-groups that are affected (and in turn are likely to have an impact on regions) by TT related mobility. In particular, the models presented in this report involve the use of a suitable social survey microdata set that has been developed for the purposes of the MOBI-TWIN project under D1.2 (as well as the EU-SILC dataset and in principle any other equivalent microdata) for the pilot study regions, including information on relevant demographic and socio-economic characteristics (including age, occupation, social grouping and income) of individuals and households. This survey micro-dataset is combined with small area information (typically available from sources such as the census of population or administrative data) that paint a picture of the social and economic geography of the MOBI-TWIN pilot areas and the scenarios that are explored in the context of the project.

As also noted in the introduction, spatial microsimulation involves the creation of large-scale population microdata sets and the analysis of policy impacts at the micro-level. *Population microdata* contains information on individuals rather than aggregate data. *Population microdata* can be distinguished between *individual microdata* that contain *information on individuals*, *household microdata* which may contain *household information* only and household microdata which may contain *individual* and *household information*. In the context of the proposed project, we have at our disposal the MOBI-TWIN survey microdata set (individual level) and the EU-SILC. These are representative survey microdata on the social situation of individuals across Europe. These social surveys microdata may be presented in the following format (see Table 1) of a list of individuals within households (this is an illustration drawing on earlier blueprint work of the method by Ballas *et al.*, 2005b using the British Household Panel Survey):

PERSON	AHID	PID	AAGE12	SEX	AJBSTAT	…	AHLLT	AQFVOC	ATENURE	AJLSEG	…
1	1000209	10002251	91	2	4	…	1	1	6	9	…
2	100038	100044	28	1	3	…	2	0	7	-8	…

	1	91									
3	1000381	10004521	26	1	3	...	2	0	7	-8	...
4	1000667	10007857	58	2	2	...	2	1	7	-8	...
5	1001221	10014578	54	2	1	...	2	0	2	-8	...
6	1001221	10014608	57	1	2	...	2	1	2	-8	...
7	1001418	10016813	36	1	1	...	2	1	3	-8	...
8	1001418	10016848	32	2	-7	...	2	-7	3	-7	...
9	1001418	10016872	10	1	-8	...	-8	-8	3	-8	...
10	1001507	10017933	49	2	1	...	2	0	2	-8	...
11	1001507	10017968	46	1	2	...	2	0	2	-8	...
12	1001507	10017992	12	2	-8	...	-8	-8	2	-8	...

TABLE 1. SOCIAL SURVEY MICRODATA EXAMPLE (EXTRACT FROM THE BRITISH HOUSEHOLD PANEL SURVEY); AFTER BALLAS ET AL., 2005B)

where:

Person	person number
AHID	household identifier (number of household to which the listed individual belongs)
PID	person identifier (a unique number to identify the individual)
AAGE12	Age
SEX	Sex
AJBSTAT	Current labour force status (e.g. self-employed, in paid employment, unemployed, family care etc.) in 1991
AHLLT	Health status
AQFVOC	Vocational qualifications
ATENURE	Tenure status
AJLSEG	Socio-economic group: last job

The remainder of this section draws on previous work (such as Ballas et al., 2005b) to give a brief historical overview and introduction of the methodological approach. Spatial microsimulation techniques involve the merging of survey microdata with census and other geographical area data to simulate a population of individuals within households (for different geographical units), whose characteristics are as close to the real population as it is possible to estimate. In other words, geographical microsimulation models simulate *virtual populations* in given geographical areas, so that the characteristics of these populations are as close as possible to their “real world” counterparts. One of the major advantages of microsimulation is that it can be a substitute for conducting detailed surveys to produce survey data such as the example described above at the small area level.

The procedure described above is typically referred to as *static spatial microsimulation*. This involves the *reweighting of an existing microdata sample* (which is only available at coarse levels of geography), so that it would fit small area population statistics tables for one point in time. For instance, an existing microdata set such as the MOBI-TWIN survey as well as EU-SILC microdata and any other suitable social survey microdataset can be reweighted to "populate" small areas. The MOBI-TWIN survey and EU-SILC provides detailed records for a sample of individuals. Reweighting methods aim to find the set of individual records that best matches the population described in the relevant pilot regions in Greece, Spain, Italy, the Netherlands and Finland using IPF algorithm**.** First, a series of small area tables (e.g. from the Census or other sources) that describe the small area of interest must be selected. For example, a reweighting method would sample from MOBI-TWIN or any other survey (i.e. EU-SILC microdata) to find a suitable combination of individuals that would fit the following statistical tables (which can be considered as small area constraints in a combinatorial optimisation approach aimed at finding the best sect of records from a social survey microdata sets that fits the small area descriptions; see Ballas et al., 2005a and 2005b) in two hypothetical areas or neighbourhoods (see Table 2) within Greece, Spain, Italy, the Netherlands and Finland:

Small area table 1 (household type)	Small area table 2 (economic activity of household head)	Small area table 3 (tenure status)
Area 1	Area 1	Area 1
60 married couple households	80 employed/self-employed	60 owner occupier
20 Single-person households	10 unemployed	20 Local Authority or Housing association
20 Other	20 other	20 Rented privately
Area 2	Area 2	Area 2
40 married couple households	60 employed/self-employed	60 owner occupier
20 Single-person households	20 unemployed	20 Local Authority or Housing association
40 Other	30 other	20 Rented privately

TABLE 2. EXAMPLES OF SMALL AREA CONSTRAINTS

The task would be to select the records of the social survey microdata that best match these tables (which can also be defined as 'small area constraints' in a combinatorial optimisation effort) using statistical matching or geographical microsimulation reweighting techniques (Williamson *et al*., 1998; Ballas *et al.*, 2005a and 2005b). However, there are a vast number of possible sets of households that can be drawn from the social survey microdata sample. There is a wide range of techniques that can be employed to find a set that fits the target tables well.

The spatial microsimulation model that we developed and used, builds on previous relevant work and models developed by us and other scholars. The first example (and the original idea) of combining a national survey data with small area census data to estimate small area microdata is presented by Williamson et al (1998) in a study that combined the United Kingdom Sample of Anonymised Records with small area census data with the use of three alternative algorithms (programmed in Fortran): hill-climbing, simulated annealing and genetic algorithms (all based on random number generating). Following this the SimBritain model *SimBritain* model, programmed in Java (Ballas *et al*., 2005a and 2005b) offered an alternative approach that adopted a deterministic (without using any random number generators) approach to reweighting survey microdata so that they fit given small area statistics tables. In particular, this methodology was used to estimate a wide range of non-Census variables (including household income) at the small area level. This model has been used to assess the socio-economic as well as geographical impact of a wide range of national social policy changes in the UK (Ballas *et al*., 2007). A more recent example of relevant work is the *SimAthens* model (Panori, Ballas and Psycharis, 2017), which used small-area demographic and socioeconomic information available from the Greek census of population with data from the European Union Statistics on Income and Living Conditions (EU-SILC). In particular, the *SimAthens* model built on the SimBritain model and adopted the deterministic reweighting Iterative Proportional Fitting (IPF) algorithm to reweight EU-SILC social microdata records to fit in small-area (municipal level) descriptions for Athens based on 2001 and 2011 censuses. This was achieved by using demographic and socioeconomic characteristics as small area constraint variables.

The MOBITWIN model adopts the approach described above to generate suitable regional and sub-regional level microdata for all pilot areas. In addition, it includes dynamic microsimulation aspects by utilising relevant information on mobility and migration probabilities from the MOBITWIN survey. Dynamic microsimulation incorporates behavioural responses under different policy scenarios. In addition, microsimulation models can become *geographical* when spatial information about the simulated entities is available (or estimated). For example, dynamic microsimulation involves forecasting past changes forward to produce as best an estimate as possible of an individual's circumstances in the future - were current trends to continue, or were they to change under different policy scenarios.

Another social simulation approach that is conceptually similar to microsimulation and especially dynamic microsimulation is Agent Based Modelling (ABM). A key feature of ABM (and advantage over microsimulation) is that it considers and allows for interaction between the simulated units (agents). In particular, ABM is a method where people (or other agents) are being represented as computational agents that can interact with each other through a network. This ABM allows for simulating how a group of people with different properties can influence each other.

The origins of ABM go back to 1971, when Thomas Schelling published his widely cited paper on segregation, demonstrating with a simple agent-based model (ABM) that ethnic groups could completely segregate, whilst each individual would be happy in a mixed neighbourhood, only not preferring a minority position. This influential paper demonstrated the intriguing capacity of ABMs to dynamically connect the individual micro level with the macro societal level (see Figure 1).

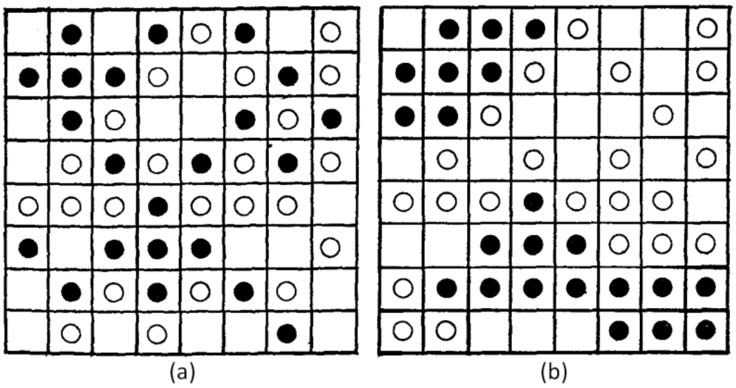


FIGURE 1: THE INITIAL CONDITION (A) AND A STABLE SEGREGATED PATTERN OBTAINED AFTER SEVERAL ITERATIONS (B) (SCHELLING, 1974). TO EXPERIMENT WITH THIS MODEL ONLINE VISIT: HTTP://WWW.NETLOGOWEB.ORG/LAUNCH#HTTP://WWW.NETLOGOWEB.ORG/ASSETS/MODELSLIB/SAMPLE%20MODELS/SOCIAL%20SCIENCE/SEGREGATION.NLOGO

The finding that simple rules for interaction generated fascinating and sometimes surprising group phenomena, triggered a lot of scholars to explore the possibilities of ABMs to study dynamical phenomena, such as opinion dynamics and conflicts, the diffusion of ideas, products and social practices such as fashions, and the activities of networks, both physical as well as online. Because computation traditionally cannot be found at the top of most social scientists' methodological toolboxes, many ABMs were developed by computational scientists instead, using assumptions on micro behaviour that were often poorly grounded in behavioural theory and data, if at all.

Despite the behavioural limitations of the first ABM models, the potential of the methodology was recognized by several researchers because climate related problems are often rooted in our social behaviour. The capacity of ABMs to show how social practices such as dietary preferences (e.g., plant based) or modality choice (e.g., cycling) can "grow" from minority behaviour towards a common practice, and explore the dynamics of such social driven processes bears relevance for policy testing, as the forecasting of societal dynamics in relation to different policies becomes a possibility. This makes the method also very relevant in studying emergent patterns in migration. To be useful however, human behaviour should be represented in a sufficiently realistic manner in ABMs. For that, we developed the HUMAT framework (Jager et al, 2025) in an earlier Horizon 2020 project focusing on local social innovation, called SMARTEES[4].

For the HUMAT integrated model presented here, we decided to connect the key drivers and processes in a relatively simple dynamical framework. HUMAT has been developed as an integrated framework to support the development of social simulation models that address a number of key processes on both the micro and macro level. It supports the implementation of social simulation models for different cases, and provides a perspective for different types of data that can be used to parameterise the model. Thus, HUMAT can be considered as a more generic framework, and the basic implementation in code has supported the development of computational models in several different cases, such as a referendum on making a park car-free, an island transitioning to energy independence (Bouman et al. 2021; Jager et al. 2024), opinion dynamics and polarisation on Covid-19 vaccinations (Li & Jager 2023) and more (see Figure 2).

HUMAT serves as an integrated framework for modelling different drivers and processes driving opinion dynamics in communities (Antosz et al., 2020; Jager et al., 2025). Having an Agent Based Modelling approach at its core, the use of HUMAT allows for a multi-method approach, where qualitative and quantitative data can be used for parameterising a model, and experimentation with policies in a modelled case can be done in a co-creative process with stakeholders.

The HUMAT socio-cognitive architecture can be applied on synthetic populations composed of agents having dynamic beliefs about how satisfying behavioural alternatives are for their needs and values, and employ social networks to communicate with one another about these beliefs. The HUMAT framework represents social influence in the context of the (dis)satisfaction of different needs and values as motives for action. Multiple needs can be grouped into three basic categories of (1) experiential needs related to short-term outcomes, (2) social needs related to fitting in the group, and (3) values. Internal conflicts can emerge between experiential needs, social needs and value needs, which translate into experienced cognitive dissonance. A dissonant state will motivate the agent to reduce the

[4] https://local-social-innovation.eu

inconsistency: try to persuade other agents in its ego-network to change their minds, change their own opinion, or end certain social connections and form new connections. Communication between the agents is based on network theory. Agents that are more similar regarding their spatial location and their socio-demographic properties (age, education, income) are more likely to interact. Also, agents are more likely to listen to and be convinced by agents having a higher reputation. Hence the HUMAT framework serves as an integrated modelling framework allowing for the simulation of complex patterns of normative and informative influences within a community.

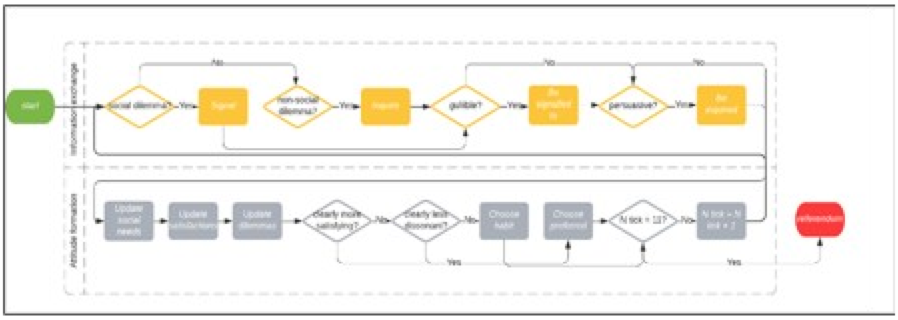


FIGURE 2. A FLOW DIAGRAM OF THE HUMAT FRAMEWORK FOR A REFERENDUM CASE

The HUMAT framework is being applied now to several cases, and much experience has been built up in implementing and experimenting with various models. In Jager et al (2025)[5] an extensive description can be found including exemplary cases, formulas and a description following the widely used ODD protocol. This is also a framework that is part of the MOBITWIN model and it is envisaged that it will be used to simulate mobility decisions and their implications.

2.2 RELEVANCE TO MOBITWIN AIMS AND OBJECTIVES

This deliverable has direct relevance to ***objective 3*** of the MOBITWIN project, which is as follows:

Analyse the effects of spatial mobility on EU demographics, society, welfare systems and labour market using microsimulation and agent-based modelling.

In particular, the MOBITWIN model is capable of generating regional and sub-regional level population microdata that can be used to identify particular demographic and socio-economic groups and individuals that are likely to be

[5] https://www.jasss.org/28/1/4.html#ref-li2023

affected by spatial mobility and to also estimate the demographic and fiscal impacts of such mobilities.

The envisaged outputs of the MOBITWIN model (to be produced and analysed in the forthcoming tasks 3.3 and 3.4) can also be used to inform the open policy discussion with relevant stakeholders and society on the effects of the TT on spatial mobility and therefore the work presented here is also relevant to MOBITWIN objectives 4 and 5:

- Initiate an open policy discussion and engage stakeholders and society on the effects of Twin Transition on spatial mobility for addressing regional inequalities across EU regions.
- Maximise MOBI-TWIN's impact by disseminating project's outcomes to a wide audience and engage multiple types of stakeholders.

2.3 Link to other deliverables

The MOBITWIN model makes heavy use of the MOBITWIN survey dataset (deliverable D1.2) as input for the microsimulation and agent-based modelling frameworks. The following figure (Figure 3) depicts the links between all WP3 tasks and links to other deliverables and outputs.

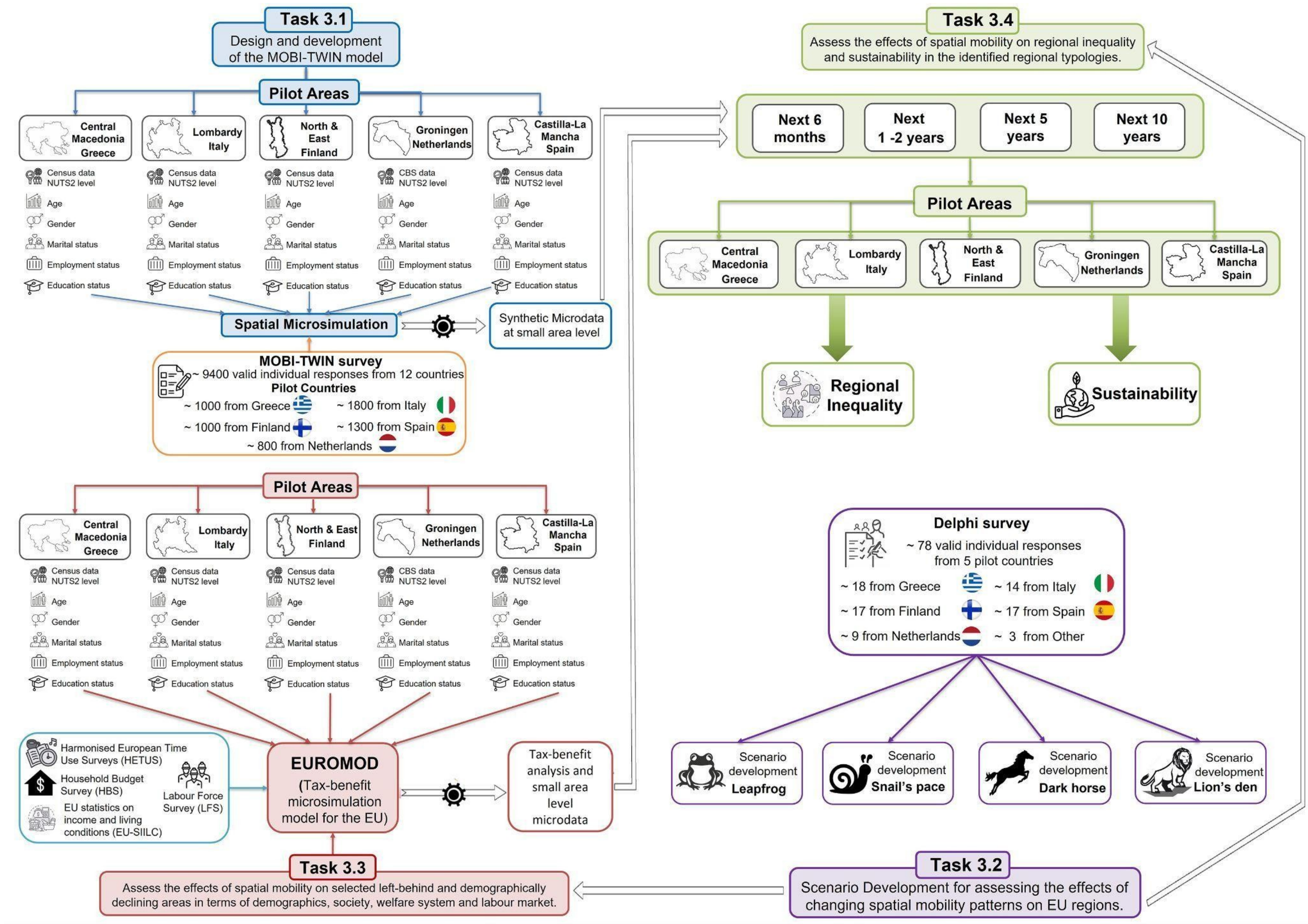


FIGURE 3. OVERVIEW OF WORK PACKAGE 3 AND ITS TASKS

2.4 SPATIAL MICROSIMULATION MODELLING

In this section we present examples of the code that implements the methods described in the previous section. The code that is part of this deliverable can be adapted for different areas and what-if scenario analysis. To that end it adapts previous work with the use of a suitable computer programming language such as R. The software provided is open-source to enable the application of the methods in different pilot areas and types of climate-change induced events and more widely for the transfer of knowledge to support further research in this field. The developed software is designed to be as flexible as possible so that it can be adapted to work for any cities and regions in Europe and elsewhere by changing the appropriate spatial and attribute data to match the appropriate locations.

In particular, the following code examples implement a spatial microsimulation procedure using the iterative proportional fitting (IPF) method[6]. The aim is to generate synthetic populations at the small-area level that are statistically consistent with known aggregate data (marginal totals) for key demographic characteristics. This is achieved by adjusting the weights of individual-level survey data so that, when aggregated, the synthetic populations match external control totals for variables such as income, gender, employment status, education, and age. The process is executed iteratively, progressively refining the weights in each step until the synthetic dataset closely aligns with the area-level constraints. What follows is a step-by-step explanation of the main components of the script, including initialization, application of demographic constraints, and the iterative reweighting process.

This section of the script initializes the spatial microsimulation procedure by setting up the number of iterations and establishing the base structures needed for the IPF process. The variable "n_iter" defines how many times the reweighting process will be repeated, allowing the model to progressively adjust weights until the synthetic microdata closely matches the known marginal totals for each geographic unit. The matrix "weights0" is initialized as a two-dimensional array filled with ones, with dimensions corresponding to the number of survey respondents (rows) and the number of small geographic areas (columns). This initial weight matrix assumes, at the outset, that all individuals are equally representative across all areas. This uniform starting point is essential for iterative refinement. Next, the script initializes "survey_agg", a matrix that stores the aggregated synthetic population characteristics for each area, using the initial weights. The for-loop populates this matrix by computing, for each area, the column-wise sum of the categorical variables ("survey.cat_pilot_area") weighted by the corresponding entries in "weights0". This step serves as the initial benchmark against which future iterations will adjust the weights in order to bring the synthetic data into alignment with the actual population distributions.

[6] The code below refers to a generic case and hence the constraint variables that have been used do not necessarily reflect a specific pilot area. By adjusting the code to the constraint variables of each pilot area, we manage to obtain the simulated weights,

```r
# Setting number of iterations
n_iter <- 10

# Initialise weights0 ONCE (not inside the loop)
weights0 <- array(1, dim = c(nrow(survey_pilot_area), nrow(all.msim_pilot_area)))

# Initialise survey_agg correctly
survey_agg <- array(dim = c(nrow(all.msim_pilot_area), ncol(all.msim_pilot_area)))
colnames(survey_agg) <- colnames(all.msim_pilot_area)
for (i in 1:nrow(all.msim_pilot_area)) {
  survey_agg[i, ] <- colSums(survey.cat_pilot_area * weights0[, i])
}
```

This section initiates the first iteration of the spatial microsimulation loop and applies the first constraint: aligning the synthetic population with the known income distribution in each small area. The variable "weights1" is initialized as a matrix of ones, while having the dimensions of the survey-area combination. For each small area j, the code loops over the three income categories k, calculating adjustment factors that rescale the weights of individuals in the survey who belong to that income group. Specifically, it divides the known marginal total for income group k in area j (from "all.msim_pilot_area") by the current estimated total from the synthetic survey ("survey_agg"). After reweighting, the script calculates "survey_agg1", the updated synthetic marginal totals for each area, by summing the weighted categorical survey data across individuals, using the product of the initial weights ("weights0") and the new income-specific weights ("weights1"). This updated aggregate will be used as the input for the next constraint, thereby progressively improving the alignment between the synthetic and known data distributions.

```r
for (iter in 1:n_iter) {
  # Constraint 1: Income level
  weights1 <- array(1, dim = c(nrow(survey_pilot_area), nrow(all.msim_pilot_area)))
  for (j in 1:nrow(all.msim_pilot_area)) {
    for (k in 1:3) {
      denom <- ifelse(is.na(survey_agg[j, k]) | survey_agg[j, k] == 0, 1e-6, survey_agg[j, k])
      weights1[which(survey_pilot_area$income_stat == k), j] <- all.msim_pilot_area[j, k] / denom
    }
  }
  survey_agg1 <- array(dim = c(nrow(all.msim_pilot_area), ncol(all.msim_pilot_area)))
  for (i in 1:nrow(all.msim_pilot_area)) {
    survey_agg1[i, ] <- colSums(survey.cat_pilot_area * weights0[, i] * weights1[, i])
  }
```

Following the income constraint, the script proceeds with the second constraint, which adjusts the synthetic population to match the known distribution of biological sex across small areas. A new matrix, "weights2", is initialized to store the sex-specific adjustment factors. For each area j, the loop iterates over the two categories of gender (typically male and female), adjusting the weights of individuals whose gender corresponds to category k. The relevant column index ("colidx") is computed as 3 + k, based on the structure of the "all.msim_pilot_area" data, where gender categories follow the income columns. The adjusted weights are then applied to all individuals in the corresponding gender group. After the adjustment, the updated synthetic aggregates are stored in "survey_agg2", which is computed using the cumulative product of all weights applied up to this point ("weights0 * weights1 * weights2"). This updated aggregate serves as the input for the next constraint.

```r
# Constraint 2: Biological sex
weights2 <- array(1, dim = c(nrow(survey_pilot_area), nrow(all.msim_pilot_area)))
for (j in 1:nrow(all.msim_pilot_area)) {
  for (k in 1:2) {
    colidx <- 3 + k
    denom <- ifelse(is.na(survey_agg1[j, colidx]) | survey_agg1[j, colidx] == 0, 1e-6, survey_agg1[j, colidx])
    weights2[which(survey_pilot_area$gender_stat == k), j] <- all.msim_pilot_area[j, colidx] / denom
  }
}
survey_agg2 <- array(dim = c(nrow(all.msim_pilot_area), ncol(all.msim_pilot_area)))
for (i in 1:nrow(all.msim_pilot_area)) {
  survey_agg2[i, ] <- colSums(survey.cat_pilot_area * weights0[, i] * weights1[, i] * weights2[, i])
}
```

In this section, the script applies the third constraint, which ensures that the employment status distribution in the synthetic dataset matches known area-level totals. A new weight matrix "weights3" is initialized. For each area j, the code loops over the five employment categories k, adjusting the weights of individuals in the survey whose employment status matches k. The appropriate column index in "all.msim_pilot_area" is computed as 5 + k, reflecting that employment-related columns follow those for income and gender (as established earlier). These adjustments are stored in "weights3". The updated synthetic totals after applying the employment constraint are calculated and stored in "survey_agg3", using the combined product of all previously applied weights ("weights0", "weights1", and "weights2") along with "weights3". These updated weights form the new basis for applying the next constraint.

```r
# Constraint 3: Employment status
weights3 <- array(1, dim = c(nrow(survey_pilot_area), nrow(all.msim_pilot_area)))
for (j in 1:nrow(all.msim_pilot_area)) {
  for (k in 1:5) {
    colidx <- 5 + k
    denom <- ifelse(is.na(survey_agg2[j, colidx]) | survey_agg2[j, colidx] == 0, 1e-6, survey_agg2[j, colidx])
    weights3[which(survey_pilot_area$employment_stat == k), j] <- all.msim_pilot_area[j, colidx] / denom
  }
}
survey_agg3 <- array(dim = c(nrow(all.msim_pilot_area), ncol(all.msim_pilot_area)))
for (i in 1:nrow(all.msim_pilot_area)) {
  survey_agg3[i, ] <- colSums(survey.cat_pilot_area * weights0[, i] * weights1[, i] * weights2[, i] * weights3[, i])
}

```

This part of the script introduces the fourth constraint, adjusting for educational attainment across the synthetic population. A new weight matrix "weights4" is created, and for each small area j, the loop iterates over the three education categories k. The column index "colidx" is calculated as 10 + k, following the assumed ordering of variables in "all.msim_pilot_area", where education-related columns come after those for employment status. Individuals in the survey whose education level corresponds to category k receive updated weights for each area. After applying these adjustments, the script calculates "survey_agg4", which represents the new synthetic marginal totals after incorporating the education constraint. This is done by summing the product of all weights applied so far ("weights0" through "weights4") across the categorical variables. The result sets the stage for the final constraint in the loop.

```
# Constraint 4: Education level
weights4 <- array(1, dim = c(nrow(survey_pilot_area), nrow(all.msim_pilot_area)))
for (j in 1:nrow(all.msim_pilot_area)) {
  for (k in 1:3) {
    colidx <- 10 + k
    denom <- ifelse(is.na(survey_agg3[j, colidx]) | survey_agg3[j, colidx] == 0, 1e-6, survey_agg3[j, colidx])
    weights4[which(survey_pilot_area$education_stat == k), j] <- all.msim_pilot_area[j, colidx] / denom
  }
}
survey_agg4 <- array(dim = c(nrow(all.msim_pilot_area), ncol(all.msim_pilot_area)))
for (i in 1:nrow(all.msim_pilot_area)) {
  survey_agg4[i, ] <- colSums(survey.cat_pilot_area * weights0[, i] * weights1[, i] * weights2[, i] * weights3[, i] * weights4[, i])
}
```

This final constraint introduces adjustments for age group distributions within each small area. The code begins by defining "age_bins", a list of 8 age intervals that correspond to age brackets used in the marginal totals (e.g., 20–29, 30–39, etc.). The weight matrix "weights5" is then initialized. For each area j, the loop iterates through all 8 age groups k, and for each, it identifies the subset of individuals in the survey (idx) whose age falls within the defined bin. The adjustment factor is calculated using the known area-level age totals from "all.msim_pilot_area" (columns 15–22, as suggested by "colidx <- 13 + k)" divided by the corresponding synthetic total in "survey_agg4". These age-specific adjustments are applied only to individuals within each age bracket. After all adjustments are made, the updated aggregate totals are calculated and stored in survey_agg5, using the cumulative product of all constraints applied so far ("weights0" through "weights5"). This completes the reweighting phase for a single iteration, with the synthetic population now aligned to all five demographic constraints.

```
# Constraint 5: Age group
age_bins <- list(
  c(0, 19),  # 20-29
  c(20, 29),  # 20-29
  c(30, 39),  # 30-39
  c(40, 49),  # 40-49
  c(50, 59),  # 50-59
  c(60, 69),  # 60-69
  c(70, 79),  # 70-79
  c(80, Inf)  # 80+
)

weights5 <- array(1, dim = c(nrow(survey_pilot_area), nrow(all.msim_pilot_area)))
for (j in 1:nrow(all.msim_pilot_area)) {
  for (k in 1:8) {
    colidx <- 13 + k  # columns 15-22 in all.msim_pilot_area
    idx <- which(survey_pilot_area$age >= age_bins[[k]][1] & survey_pilot_area$age <= age_bins[[k]][2])
    denom <- ifelse(is.na(survey_agg4[j, colidx]) | survey_agg4[j, colidx] == 0, 1e-6, survey_agg4[j, colidx])
    weights5[idx, j] <- all.msim_pilot_area[j, colidx] / denom
  }
}
survey_agg5 <- array(dim = c(nrow(all.msim_pilot_area), ncol(all.msim_pilot_area)))
for (i in 1:nrow(all.msim_pilot_area)) {
  survey_agg5[i, ] <- colSums(survey.cat_pilot_area * weights0[, i] * weights1[, i] * weights2[, i] * weights3[, i] * weights4[, i] * weights5[, i])
}
```

This concluding section of the loop finalizes each iteration of the spatial microsimulation process. First, the script computes "weights6" as the combined product of all previously calculated weight matrices ("weights0" through "weights5"). This cumulative weight matrix represents the full adjustment for all five demographic constraints applied so far. Using these final weights, the script calculates "survey_agg6", the synthetic aggregate characteristics for each area after all constraints have been applied in the current iteration. This is done by multiplying the categorical variables in the survey data with the final weights and summing across individuals, as in earlier aggregation steps.

Next, "weights0" and "survey_agg" are updated to reflect the results of this iteration, ensuring that the next loop will begin with the latest weight estimates. This step is crucial for the convergence of the IPF procedure, as each round of adjustments builds on the last.

Finally, the script exports the updated weight matrix for the current iteration as an Excel file using "write_xlsx", which allows for external review or later analysis of the iteration-by-iteration reweighting results.

```
  # Update weights6 as product of all weights
  weights6 <- weights0 * weights1 * weights2 * weights3 * weights4 * weights5

  # Final aggregate for this iteration
  survey_agg6 <- array(dim = c(nrow(all.msim_pilot_area), ncol(all.msim_pilot_area)))
  for (i in 1:nrow(all.msim_pilot_area)) {
    survey_agg6[i, ] <- colSums(survey.cat_pilot_area * weights6[, i])
  }

  # Prepare for next iteration
  weights0 <- weights6
  survey_agg <- survey_agg6

  # Save weights to Excel file
  dfweights0 <- as.data.frame(weights0)
  write_xlsx(dfweights0, paste0("/Path/mobitwin_weights0_pilot_area_iteration_", iter, ".xlsx"))

  cat("Completed iteration", iter, "\n")
}
```

2.4.1 Convergence of weights to "true" values

Once the spatial microsimulation models have been estimated, we proceed with a visualisation of the convergence of the weights. The convergence graphs presented in the next section provide a visual summary of the iterative adjustment process used in spatial microsimulation. Each graph illustrates how the weights assigned to individual survey respondents evolve across successive iterations of the algorithm, specifically for a selected set of regions and individuals.

In essence, these plots allow us to assess the stability and performance of the fitting procedure. A key objective in iterative spatial microsimulation (e.g. using methods such as IRF or similar routines) is to ensure that individual weights converge to stable values that satisfy the known marginal totals for each spatial unit. Convergence here implies that, after a certain number of iterations, the weight values do not fluctuate substantially and, as a result, are suitable for generating the synthetic microdata.

Anomalies in convergence graphs such as highly volatile weights for an individual in a certain spatial unit or divergence may indicate problematic constraints, mismatches between sample and population distributions, or the need for further iterations. However, as we present in the next section, the vast majority of weights either stabilized quickly or showed minimal fluctuation in the later stages, reinforcing confidence in the microsimulation outputs.

In each plot, the horizontal axis represents selected spatial units (municipalities or neighborhoods), and the vertical axis shows the weights assigned to one individual across multiple iterations. Different colors correspond to different iterations, enabling easy tracking of the convergence path for each region for the same individual. If the points stabilize towards the final iterations, this is strong evidence that convergence has been achieved.

Moreover, the decision to visualise a small number of individuals across a subset of spatial units serves both technical and communicative purposes. On the technical side, it allows for detailed inspection of weight behavior. On the communicative side, it facilitates the validation and transparency of the spatial microsimulation process, supporting the overall reliability of the synthetic datasets.

Finally, these graphs also offer a helpful way to identify errors in the spatial microsimulation code and in general they provide model transparency, particularly in interdisciplinary projects such as MOBI-TWIN, where visual communication of technical processes is important. The following code is designed to process and prepare the output of the spatial microsimulation model for visualisation and diagnostic purposes.

```r
# Step 1: Read all 10 Excel files for the pilot area under investigation
weight_list_pilot_area <- lapply(1:10, function(i) {
  df <- read_excel(paste0("/path/mobitwin_weights0_pilot_area_iteration_", i, ".xlsx"))
  df$Individual <- 1:nrow(df)
  df$Iteration <- i
  df
})

# Step 2: Combine all iterations into one long dataframe
weights_long_pilot_area <- bind_rows(weight_list_pilot_area)

# Step 3: Keep only individuals 1, 2, 3, 4 and 5
selected_inds_pilot_area <- c(1, 2, 3, 4, 5)
weights_subset_pilot_area <- weights_long_pilot_area %>%
  filter(Individual %in% selected_inds_pilot_area)

# Step 4: Convert wide to long format (Region, Weight)
weights_tidy_pilot_area <- weights_subset_pilot_area %>%
  pivot_longer(cols = -c(Individual, Iteration), names_to = "Region", values_to = "Weight")

# Rename Region values with meaningful names
region_names <- c("subregion_name_1", "subregion_name_2", "subregion_name_3", ... , "subregion_name_N")

# Replace "V1", "V2", etc., with real region names
weights_tidy_pilot_area <- weights_tidy_pilot_area %>%
  mutate(Region = factor(Region, levels = paste0("V", 1:12), labels = region_names))
```

Specifically, the code above handles the weight files produced during the iterative fitting procedure across ten iterations for a given pilot area. The goal is to track how the weights assigned to selected individuals change across iterations and spatial units, enabling assessment of convergence and model behaviour.

In Step 1, the code reads ten Excel files, each representing the output from one iteration of the spatial microsimulation algorithm. The file naming convention is assumed to follow the format "mobitwin_weights0_pilot_area_iteration_i.xlsx", where "*i*" ranges from 1 to 10. For each file, a new variable "Individual" is added to uniquely identify each row (corresponding to a person in the survey), and another variable "Iteration" is added to mark the iteration number. These individual data frames are then stored as a list.

In Step 2, all ten data frames are combined into a single long-format data frame ("weights_long_pilot_area") using "bind_rows()". This newly created dataset stacks the rows vertically while preserving the iteration and individual identifiers. The resulting structure enables longitudinal tracking of weights across iterations.

In Step 3, the dataset is filtered to retain only a small, manageable subset of individuals (in this case, individuals 1 to 5). This step is useful for illustrative purposes, allowing the analyst to explore how weights evolve for specific individuals across different spatial units and iterations.

Step 4 transforms the data from wide format (where each column represents a spatial unit) to long format using the "pivot_longer()" function. This step creates three columns: "Individual, Iteration", and "Region" (which holds the original column names like "V1", "V2", etc.), as well as a fourth column, “Weight”, which holds the weight values. This tidy format is essential for plotting and analyzing the data effectively by using the “ggplot2” function.

Finally, the code replaces the generic region identifiers (e.g., "V1", "V2", ..., "Vn") with meaningful names corresponding to actual subregions (e.g., municipalities or neighborhoods). This is done by converting the “Region” variable into a factor with specified levels and labels using the “mutate()” function. The new names, stored in the “region_names” object, enhance the interpretability of subsequent plots by showing familiar or policy-relevant geographic labels rather than abstract codes.

Together, this code sets the stage for generating clear, interpretable convergence plots that show how individual weights are distributed across geographic units and how they evolve during the iterative fitting process.

```r
# Step 5: Create and save individual plots
for (ind in selected_inds_pilot_area) {
  plot_data <- weights_tidy_pilot_area %>%
    filter(Individual == ind) %>%
    mutate(Iteration = as.factor(Iteration))  # ensure discrete coloring

  p <- ggplot(plot_data, aes(x = Region, y = Weight, color = Iteration, group = Iteration)) +
    geom_line(size = 1.2) +
    geom_point(size = 2) +
    scale_color_manual(
      values = c(
        RColorBrewer::brewer.pal(9, "Set1"),  # the 9 Set1 colors
        "black"                               # manually add black as the 10th color
      )
    ) +
    geom_hline(yintercept = 0, linetype = "dashed", color = "gray40", linewidth = 0.8) +
    labs(
      title = paste("Individual", ind, "- Weight Convergence (Pilot area)"),
      x = "Region",
      y = "Weight",
      color = "Iteration"
    ) +
    theme_bw() +
    theme(
      plot.title = element_text(size = 14, face = "bold"),
      axis.text.x = element_text(angle = 45, hjust = 1),
      legend.position = "bottom"
    )

  ggsave(
    filename = paste0("individual_", ind, "_weights_pilot_area.png"),
    plot = p,
    width = 8, height = 6, dpi = 300
  )
}
```

The final section of the code (Step 5) is responsible for generating and saving individual-level convergence plots that visually represent how the spatial weights of selected individuals evolve across the ten iterations of the spatial microsimulation models. These plots serve as a tool in order to assess the stability of the models and convergence behaviour across geographical subregions.

For each selected individual (in this example, Individuals 1 through 5), the code filters the previously created tidy data frame ("weights_tidy_pilot_area") to extract only the weight values associated with that individual. It then ensures that the "Iteration" variable is treated as a categorical (factor) variable, which is necessary for consistent color mapping across the ten iterations in the plots.

The plotting is performed using the "ggplot2" package. Each plot shows the weights assigned to that individual across subregions (x-axis) and their corresponding values (y-axis) for each iteration, with lines and points grouped and colored by iteration. The use of "geom_line()" and "geom_point()" ensures that the trajectory of weight updates is clearly visible for each region-iteration combination.

To enhance visual clarity, the code applies a custom color palette from the "RColorBrewer" package (Set1) for the first nine iterations and manually adds black for the 10th iteration. This ensures consistent and distinguishable coloring. A horizontal dashed line is added at "y = 0" to serve as a reference baseline, helping the viewer interpret the relative magnitude and direction of weight changes.

The plot is customized with a clear title indicating the individual ID and pilot area, along with axis labels and a discrete color legend. The "theme_bw()" style is applied for a clean appearance, and x-axis labels are rotated for better readability when region names are long. Finally, each plot is saved as a high-resolution PNG file (300 dpi) with dimensions suitable for inclusion in reports or presentations. The file naming convention ("individual_[ID]_weights_pilot_area.png") ensures clear identification and organization of output files.

Overall, the convergence graphs shown in the following section provide an intuitive and transparent way to monitor the evolution of synthetic population weights across spatial subregions and iterations. These visualisations not only confirm that the IPF algorithm is functioning as expected, but also reveal distinct convergence patterns that shed light on how the model adjusts individual weights to satisfy regional constraints. For example, rapidly decreasing weights suggest that certain individuals were initially overrepresented in specific regions and their influence had to be quickly reduced to better align with the known marginal distributions. In contrast, gradually decreasing weights indicate a more subtle and incremental adjustment, where the individual moderately matched the regional profile but required fine-tuning over several iterations. On the other hand, rapidly increasing weights reflect cases where an individual's attributes were underrepresented in a region's synthetic population, prompting the algorithm to significantly boost their weight early in the process. Finally, gradually increasing weights point to profiles that become more relevant over time, as the algorithm incrementally recognizes their fit to the regional constraints. These distinct trajectories provide insight into the stability, sensitivity, and convergence dynamics of the microsimulation model and can help identify potential mismatches or overfitting issues. In this way, the convergence plots serve as both a diagnostic tool and a conceptual window into how spatial microsimulation reconciles microdata with aggregate targets.

2.5 Data and Methods per study case

This section outlines the data sources and methodological choices applied in the spatial microsimulation models across the five pilot regions of the MOBI-TWIN project. For each pilot, we combine individual-level data from the country specific subsample of the MOBI-TWIN survey with aggregated demographic and socio-economic data obtained from national statistical agencies. This integration allows us to generate synthetic microdata representative of small-area populations, which serve as the basis for estimating future mobility probabilities at various temporal horizons (e.g., 6 months, 1 year, 5 years, 10 years).

Given the heterogeneity in data availability and administrative structures across countries, the geographical resolution of the models varies accordingly. Specifically, the spatial units of analysis range from NUTS-3 regions in Spain and Italy to municipalities in Greece, neighbourhoods in the Netherlands, and postal code areas in Finland. Each of the following sub-sections details the data inputs, constraint variables, and methodological adaptations specific to each country, providing transparency about how the spatial microsimulation models were operationalized across diverse regional settings.

2.5.1 Greece: Region of Central Macedonia

Although the Greek pilot of the MOBI-TWIN project is nominally focused on the region of Central Macedonia, the spatial microsimulation work has been conducted using data at the municipal level across the entire country of Greece. This broader spatial granularity allows for a more detailed modeling approach of internal mobility patterns. The MOBI-TWIN survey data, filtered to include only Greek respondents (N = 1281), is combined with municipal-level demographic and socio-economic data from ELSTAT (Hellenic Statistical Authority). The constraint variables selected for Greece include:

Age:

- 0–9, 10–19, 20–29, 30–39, 40–49, 50–59, 60–69, 70–79, 80+

Biological Sex:

- Male, Female

Marital Status:

- Single, Married, Widowed, Separated

Employment Status:

- Employed, Unemployed, Students, Retired, Other

Education Level:

- Primary, Secondary, Tertiary

These variables were chosen based on their relevance to mobility behavior, their availability in both the survey and census data, and their capacity to capture key household and demographic diversity. The ultimate goal is to use the resulting synthetic microdata to estimate the probability of individuals moving within defined time horizons (e.g., 6 months, 1-2 years, 5 years and 10 years), thereby enabling scenario-based forecasting such as estimating how many people are likely to relocate from a specific municipality within the next six months.

The convergence graphs presented below in Figure 4 correspond specifically to the Greek pilot, using a sample of 5 individuals from the synthetic population. As discussed earlier, these visualizations help track the evolution of individual weights across iterations of the fitting algorithm. In the context of the Greek pilot, they offer insights into how the model adjusts weights to align individuals with regional demographic and socio-economic constraints at the municipal level. The observed convergence patterns serve as a quality check for the simulation results and provide assurance that the model produces stable and interpretable synthetic microdata for forecasting internal mobility under various scenarios.

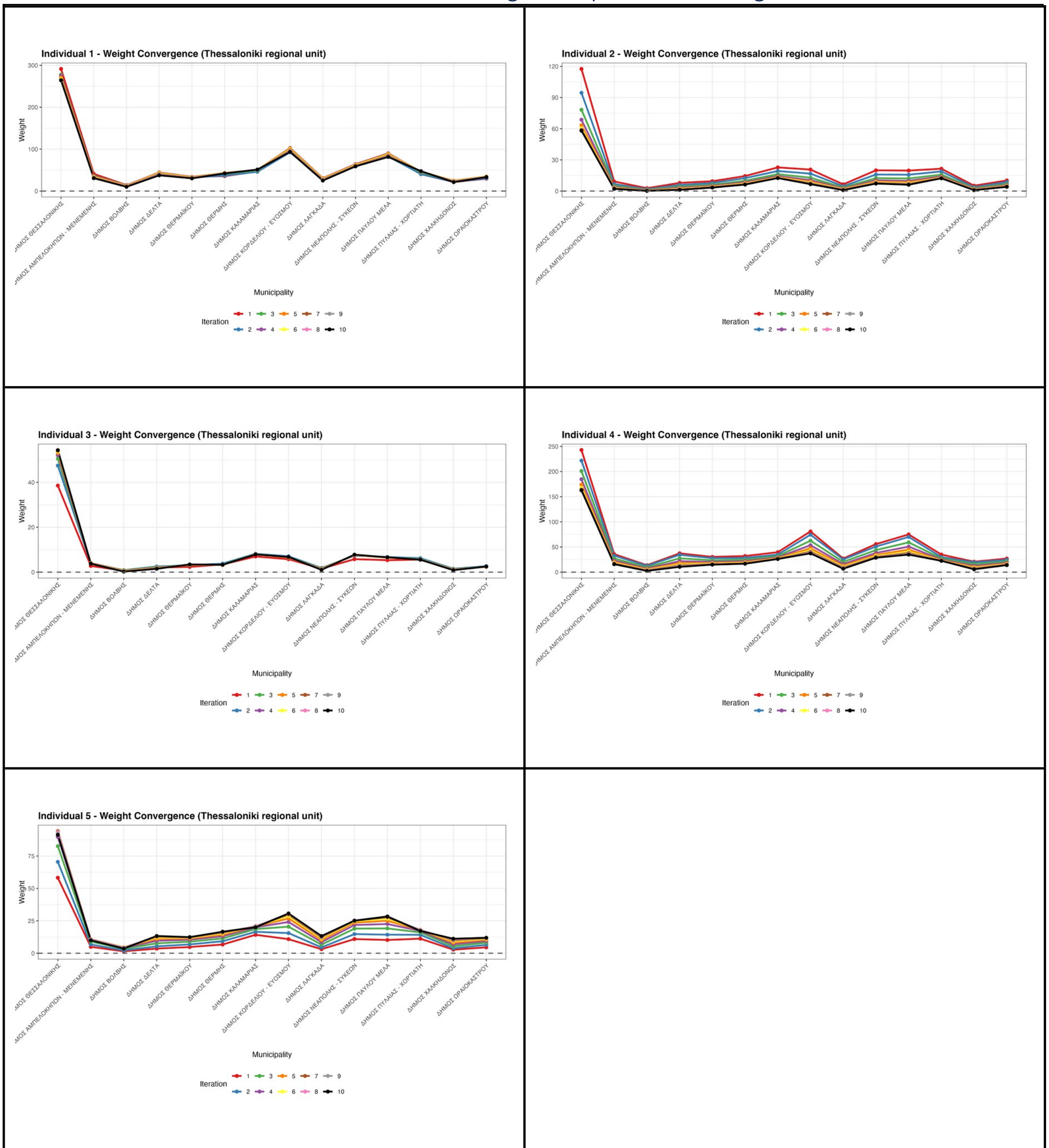


FIGURE 4. CONVERGENCE OF INDIVIDUALS' WEIGHTS ACROSS ITERATIONS FOR PILOT 1 (GREECE)

2.5.2 SPAIN: CASTILE LA-MANCHA

For the Spanish pilot, spatial microsimulations are performed at the NUTS-3 level within the region of Castile-La Mancha. Specifically, simulations cover the five provinces: Albacete, Ciudad Real, Cuenca, Guadalajara, and Toledo. Only Spanish respondents from the MOBI-TWIN survey are used (N = 1551), ensuring national consistency. These are integrated with demographic and socio-economic data provided by the Spanish National Statistics Institute (INE). The selected constraint variables mirror those used in Greece and include:

Age:

- 0–19, 20–29, 30–39, 40–49, 50–59, 60–69, 70–79, 80+

Biological Sex:

- Male, Female

Marital Status:

- Single, Married, Widowed, Separated

Employment Status:

- Employed, Unemployed, Students, Retired, Other

Education Level:

- Primary, Secondary, Tertiary

These variables were selected because they align with the survey data as well as with the known population margins of the region. The resulting synthetic microdata enables the estimation of individualized probabilities of mobility over time.

The convergence graphs in Figure 5 for the Spanish pilot reveal a different pattern compared to Greece. While the iterative fitting process does lead to convergence, the adjustment of weights continues beyond the first two or three iterations for several individuals, especially in less populated regions. This suggests a less immediate alignment between the MOBI-TWIN survey data and the regional constraint variables, potentially due to more uneven distributions or fewer representative cases in the sample. Nonetheless, the eventual stabilization of weights confirms the algorithm's ability to adapt and fine-tune the synthetic population over time.

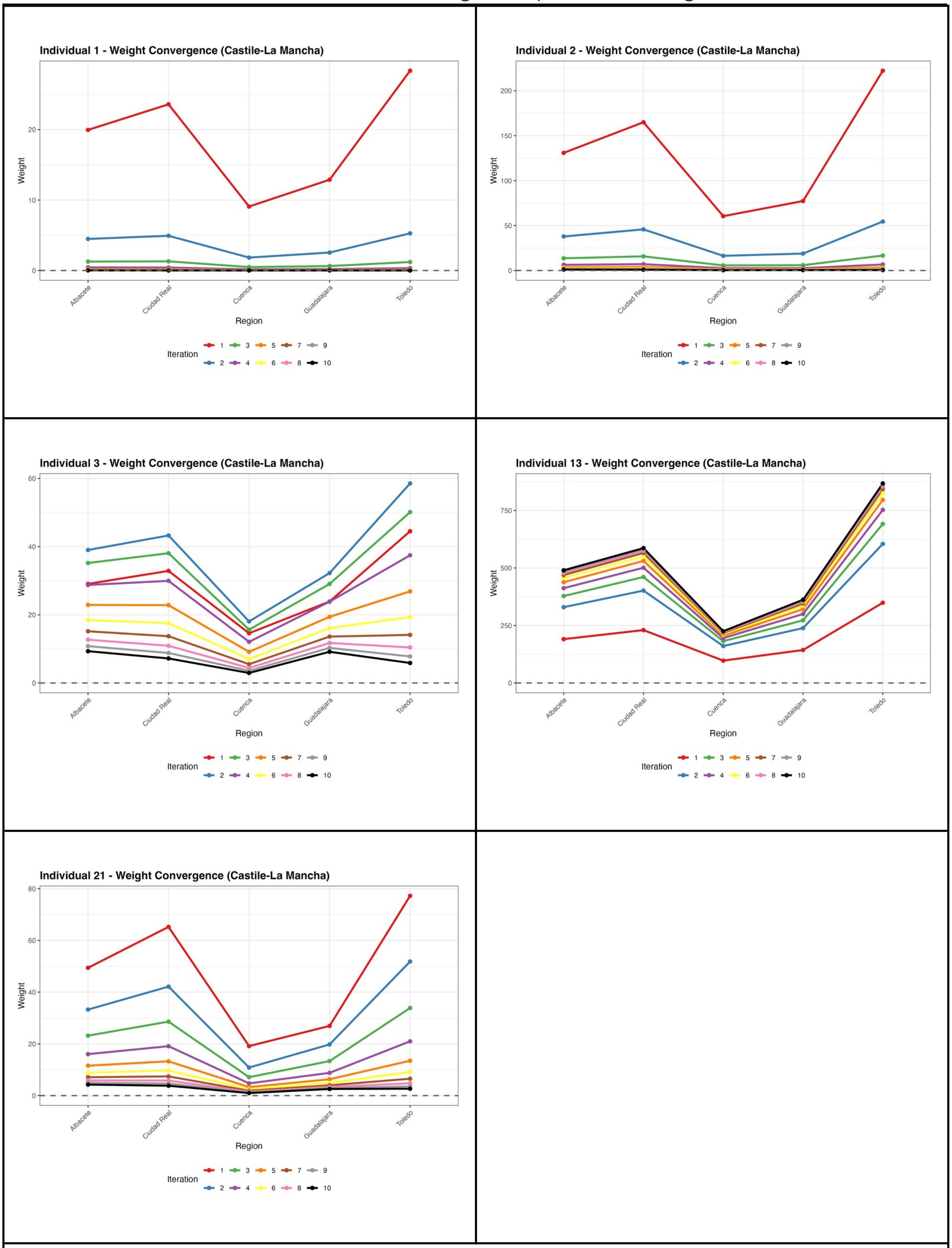


FIGURE 5. CONVERGENCE OF INDIVIDUALS' WEIGHTS ACROSS ITERATIONS FOR PILOT 2 (SPAIN)

2.5.3 Italy: Lombardy

In the case of Italy, the pilot focuses on the economically significant region of Lombardy. The spatial microsimulations are conducted at the NUTS-3 level for all twelve provinces within the region: Bergamo, Brescia, Como, Cremona, Lecco, Lodi, Mantova, Milano, Monza and Brianza, Pavia, Sondrio, and Varese. The MOBI-TWIN survey data, filtered for Italian respondents (N = 1986), is combined with official census data from ISTAT (Italian National Institute of Statistics). The selected constraint variables are once again the following:

Age:

- 0–19, 20–29, 30–39, 40–49, 50–59, 60–69, 70–79, 80+

Biological Sex:

- Male, Female

Marital Status:

- Single, Married, Widowed, Separated

Employment Status:

- Employed, Unemployed, Students, Retired, Other

Education Level:

- Primary, Secondary, Tertiary

The convergence graphs for the Italian pilot in Figure 6 closely resemble those observed in the Greek case. Most individual weights stabilize quickly, often within the first two or three iterations, indicating that the initial alignment between the survey data and the regional constraint variables is strong. This rapid convergence reinforces confidence in the internal consistency of the synthetic population and supports its use in scenario-based mobility forecasting at the subregional level.

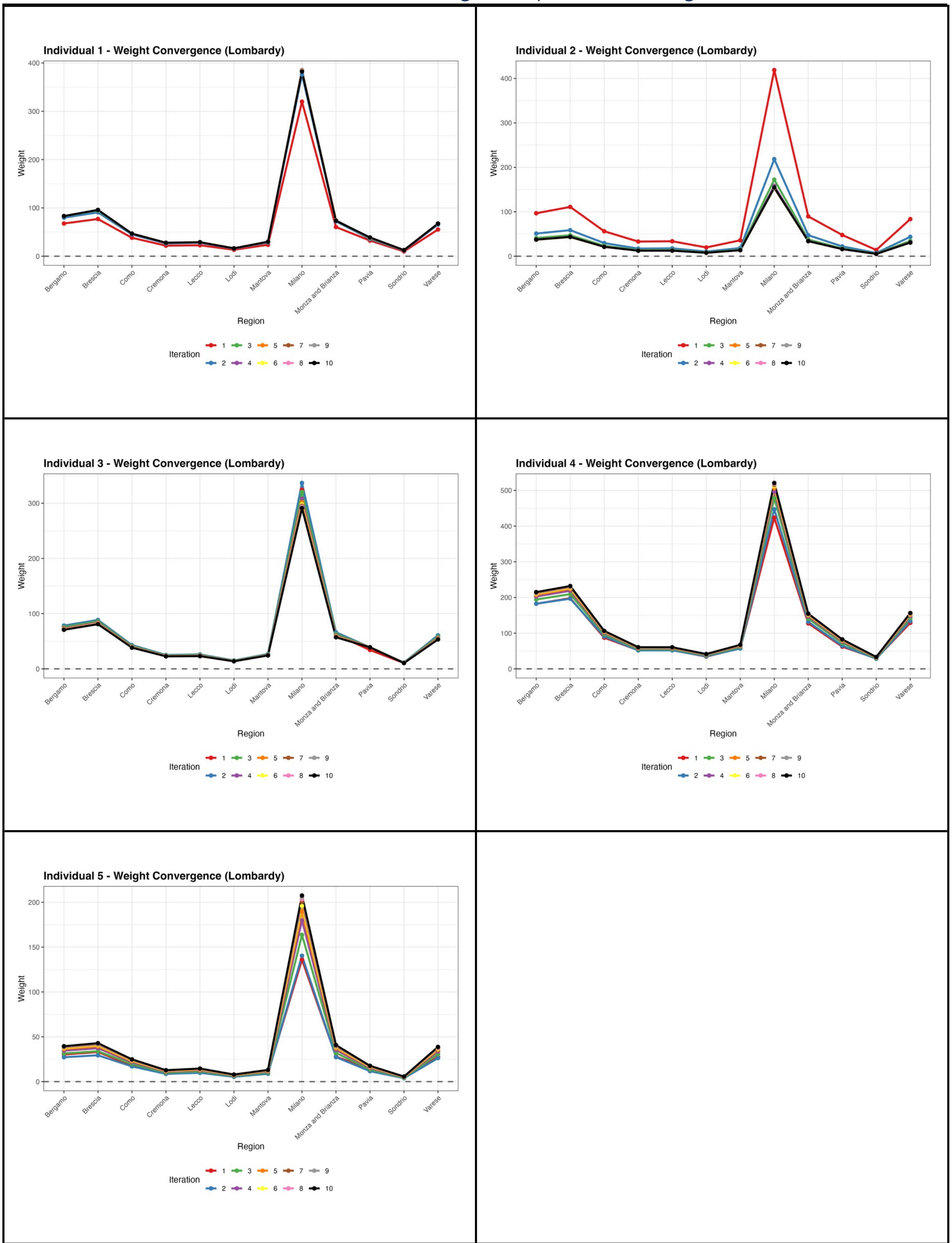


FIGURE 6. CONVERGENCE OF INDIVIDUALS' WEIGHTS ACROSS ITERATIONS FOR PILOT 3 (ITALY)

2.5.4 Netherlands: Groningen

In the Netherlands, the pilot area is the province of Groningen. Spatial microsimulations are carried out at the fine-grained neighbourhood level across the entire province. Dutch respondents from the MOBI-TWIN survey (N = 1254) are used in conjunction with administrative and demographic data sourced from CBS (Statistics Netherlands). The Dutch case includes a more diverse set of constraint variables, reflecting both the richness of the CBS data and the unique socio-demographic dimensions relevant to mobility behavior in the Netherlands. Constraints include:

Age:

- 0–14, 15–24, 25–44, 45–64, 65+

Biological Sex:

- Male, Female

Income Level:

- Lowest income, Middle income, Highest income

Education Level:

- Primary, Secondary, Tertiary

Marital Status:

- Single, Married, Widowed, Separated

Ethnicity:

- Dutch, Western migrant, Non-Western migrant

Kids in the household:

- No kids, With kids

This broader set of variables allows for a more detailed and multidimensional simulation of individual mobility.

The convergence graphs for the Netherlands in Figure 7 indicate that the model achieved stability very rapidly. The lines representing different iterations nearly overlap, demonstrating that minimal adjustments were necessary beyond the initial iteration. This suggests that the survey data and regional constraints were closely aligned from the beginning. Consequently, the synthetic population generated for the Dutch pilot appears to be accurate and reliable early in the iterative fitting process.

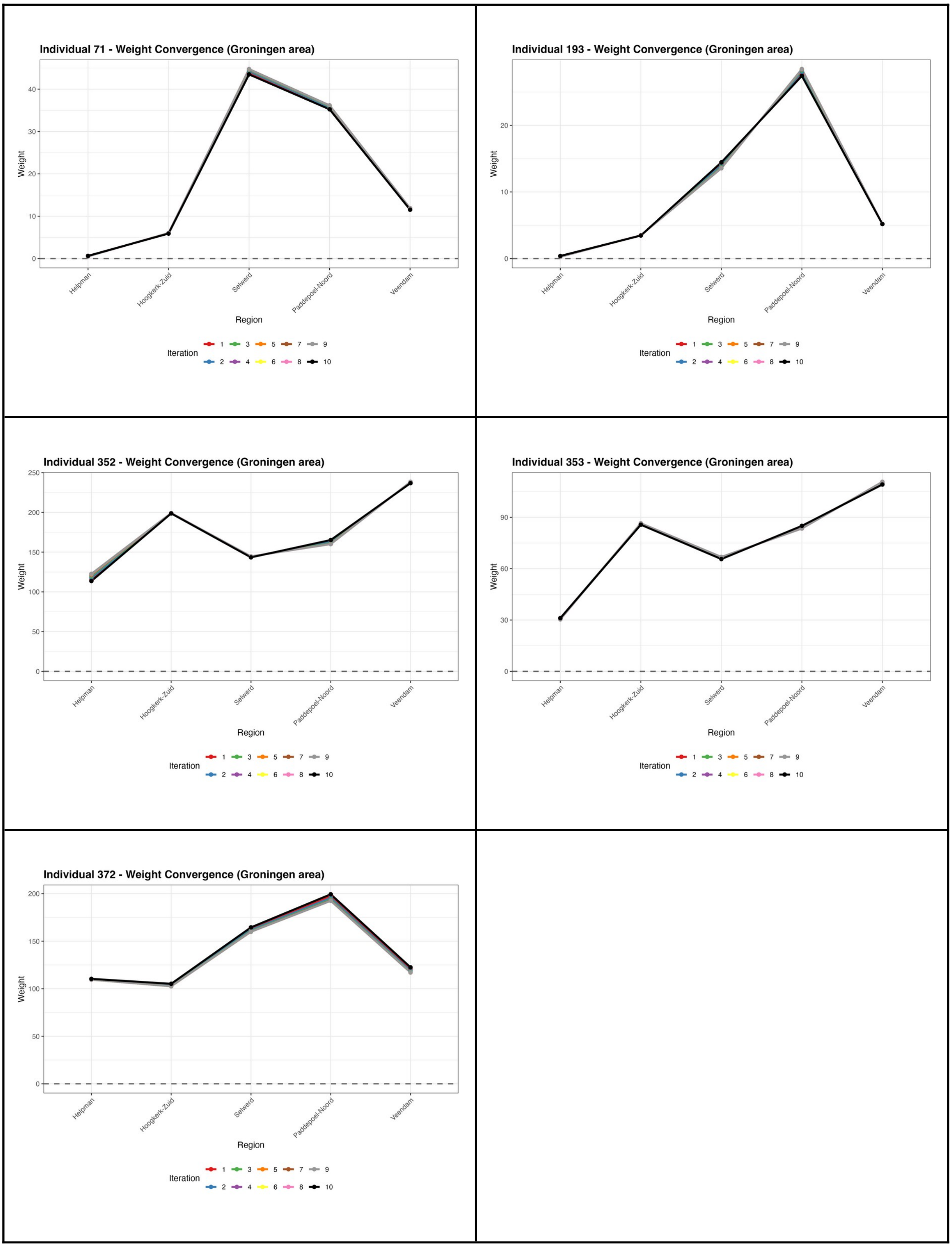


FIGURE 7. CONVERGENCE OF INDIVIDUALS' WEIGHTS ACROSS ITERATIONS FOR PILOT 4 (NETHERLANDS)

2.5.5 FINLAND: NORTH AND EAST FINLAND

Regarding the Finnish pilot, although it formally targets the sparsely populated and geographically extensive region of North and East Finland, the spatial microsimulation models are executed using postal code level data across the entire country. This nationwide coverage ensures a more comprehensive understanding of mobility in Finland, especially given the availability of detailed postal-code level data from Tilastokeskus (Statistics Finland). Finnish respondents from the MOBI-TWIN survey (N = 986) are matched to this administrative data using a consistent set of constraint variables:

Age:

- 0–19, 20–29, 30–39, 40–49, 50–59, 60–69, 70–79, 80+

Biological Sex:

- Male, Female

Income Level:

- Lowest income, Middle income, Highest income

Employment Status:

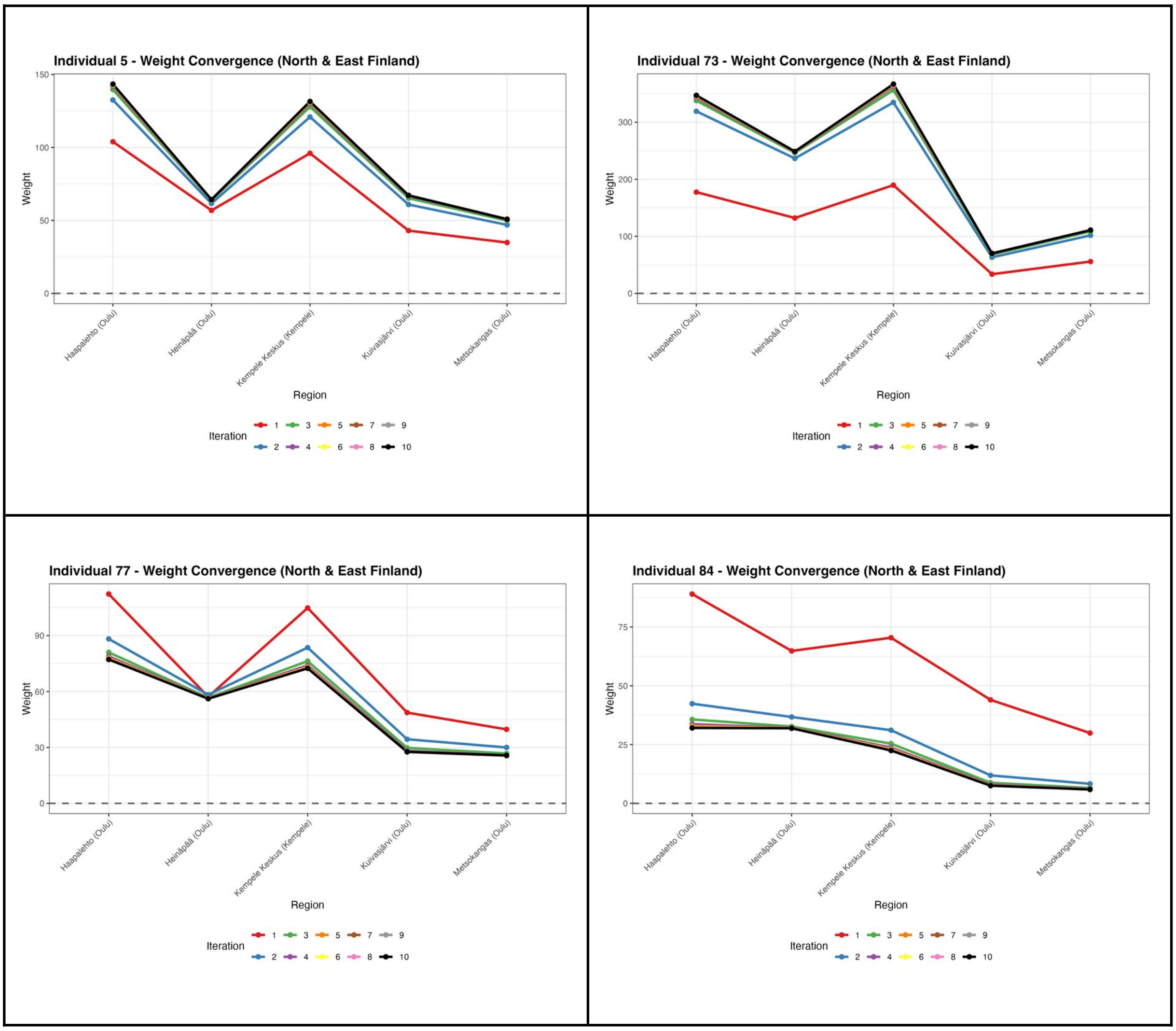
- Employed, Unemployed, Students, Retired, Other

Education Level:

- Primary, Secondary, Tertiary

This selection allows for high-resolution modeling of population movement, particularly valuable in the Finnish context. The synthetic data outputs can be used to forecast movements within or between postal code areas, for example estimating the probability that unemployed individuals aged 20–29 in Oulu region will move within the next six months.

In Figure 8, we present the convergence graphs for Finland reveal a distinct pattern where the first iteration stands apart from the subsequent ones, while iterations two through ten cluster closely together. This indicates that although the initial fit required some adjustment, the algorithm quickly stabilized after the first iteration. The close alignment of later iterations suggests a consistent and robust convergence, providing confidence in the reliability of the synthetic microdata for this pilot area.



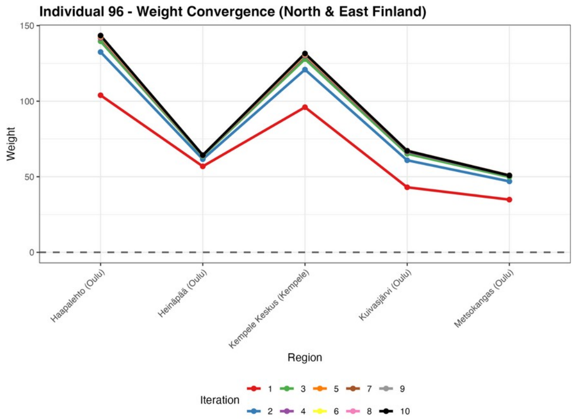


FIGURE 8. CONVERGENCE OF INDIVIDUALS' WEIGHTS ACROSS ITERATIONS FOR PILOT 5 (FINLAND)

2.6 AGENT-BASED MODELLING OF MIGRATION TOWARDS LEFT-BEHIND AREAS: A CONCEPTUAL PERSPECTIVE

This section presents a prototype ABM approach to modelling mobility and migration with a particular focus on migration dynamics towards left-behind regions in Europe in the context of digital and green transition. Many factors will determine if a person or a family decides to migrate to a left-behind region, and oftentimes difficult choices will have to be made, where the social context will often play a prominent role. The success or failure of certain people migrating is likely to have an impact on the perception and attractiveness of certain regions for other potential migrants, leading to emerging patterns of migration.

The remainder of this section provides a conceptual and methodological framework to support the development of an ABM simulation model aimed at identifying the complex dynamics of such emerging migration patterns and exploring the impact of policies on these dynamics.

2.6.1 MOTIVATIONS AND DECISIONS TO MIGRATE OR NOT

The decision to migrate is oftentimes a difficult choice, because it results in a change in a variety of outcomes. For example, moving to a left-behind area may provide a young family with more spacious dwelling, a cleaner and safer environment and more nature, but the educational and economic facilities may be worse, and the distance to family and friends may also result in more social isolation. If, however, a community in a left-behind area is emerging with a more vivid social structure and activities, and remote working becomes easier, migration may become more attractive. This also implies that the context for the decision to migrate changes as a function of how many people migrated before.

Key to understanding the motives to migrate or not is the multiple need satisfaction that people experience in the current situation versus the anticipated multiple need satisfaction of migrating to a new place. Many theories on human needs are available, the pyramid of Maslow (1954) being one of the most widely known. A more elaborate perspective on human needs is provided by Max Neef (1992) who sets nine 'axiological' needs – subsistence, protection, affection, understanding, participation, identity, idleness, creation, and freedom – against four 'existential' categories: being, doing, having and interacting.

As long as people's needs are sufficiently satisfied, they will be happy and not very motivated to migrate. However, the more needs are getting depleted, the more people are motivated to look for alternative places to migrate to. When people start considering an alternative location to migrate to, a decision-making process starts where several behavioural and psychological processes play a role: the different needs that drive the desire for change (or not), the immediate rewarding (or punishing) experiences that people experience now (e.g., small house, noise, facilities available, income) versus their anticipation of such outcomes for different alternative locations, the attitudes a person has (e.g., preferences for urban or rural environments), their adherence to norms (what do their friends and family think and do), their handling of dissonant information (the difficult trade-offs to make in a migration decision), changing needs due to critical events (e.g., child-birth) and the formation of (on-line) networks of people that plan to or did migrate to certain places, and share experiences. All of these drivers and processes can be linked conceptually to the many behavioural and psychological theories that are available in the social sciences. To name a few key theories, Maslow (1954) and Max-Neef (1992) provide theories on human needs, Pavlov (1927) provided a theory on (classical) conditioning, Skinner (1953) on operant conditioning, Festinger (1957) proposed the cognitive dissonance theory on resolving internal conflicts, Bandura (1962) proposed a theory on social learning and imitation, Cialdini et al. (1991) developed a theory on normative conduct, Ajzen (1985) integrated attitudes, social norms and behavioural control in the Theory of Planned Behaviour, Petty & Cacioppo (1986) provide a theory on communication and persuasion.

All these processes and drivers play a role in the decision to migrate or not. Changing personal situations, and observed migration behaviour of other people may lead to the decision to migrate towards a certain place, which in turn will impact the decisional context of other people. This may lead to emerging patterns, where certain locations, such as left-behind areas, become increasingly attractive for e.g., online working young families. A left-behind place may become more attractive for a young family if more young families have migrated there. The emergence of such patterns of migration may be triggered by infrastructural and economical triggers, such as making online working practical by offering good internet connection (fiberglass and increasingly satellite connection) and working opportunities allowing for living in left-behind areas (e.g., combining online work with small-scale agriculture).

In studying social complex processes like emerging migration, the method of Agent Based Modelling allows for experimentation with artificial populations and policies to better understand how processes of change, in this case migration towards left-behind areas, operate, and can be supported. In particular in a case like migration, the heterogeneity of the people as well as social influences on the decision to migrate make ABM a suitable tool for exploring the associated complex dynamics.

2.6.2 Agent Based Modelling

To use HUMAT in simulating emergent processes of migration in the MOBI-TWIN context we have to represent the target groups and the key properties of the current location of the target group and the properties of the left-behind areas. This implies the construction of a specific micro simulation (also known as synthetic population) for the segment of people that is of interest for the migration dynamics in question. Increasingly, micro simulations are being constructed to run computational social simulations on more realistic artificial populations, e.g., Chapuis et al., (2022). For several regional cases experience has been built up in constructing empirically based artificial simulations using different data sources, see e.g., Panori et al., (2017) which will serve as inspiration. Also, in the SMARTEES project we already used micro simulations representing different socio-demographic factors.

For building a micro simulation in the context of migration we will use socio-demographic data for the target group as collected in the MOBI-TWIN project. In particular we will focus on gender, age, education and skills as critical factors. The family composition will also be a key factor, as being single, a couple, or a family with young (school going) children will be relevant in the decision to migrate. Also, we want to include values regarding life aspirations (including rural versus urban orientation) and social connectivity regarding the current location (local ties with family and friends).

The aforementioned social and economic microdata will be combined with location specific data of both the current area as well as the left-behind area. Here we will have to decide on the attributes that are relevant for the multiple satisfaction of the people, and represent these in the simulated places. Aspects of interest are online connectivity (remote working possibilities), physical connectivity (travel times to relevant places and public transportation) local labour opportunities (related to e.g., green deal supported projects), housing prices/quality, facilities (e.g., shops, schools) and social attractivity, related to the attractiveness of the community regarding population composition and communal activities.

An important aspect of the implementation is the specificity/generalist approach of the simulation model. Whereas it is possible to develop a very detailed and realistic model for a specific case, it is important to realise that the simulation experiments are aimed at getting insights in the complex dynamics of migrating to left-behind areas. The simulation experiments should shed light on the key factors and processes that play a role in migration towards left-behind areas in general, and specifically how improvements in remote working possibilities and investments in local green projects may affect these migration dynamics. Hence in the simulation experiments we aim at modelling relative stylized environments, aiming to (1) get generic insights in the complex dynamics of migration, and (2) explore how selected policy interventions may benefit the migration process towards left-behind areas.

Building on the conceptual and methodological framework described above, the following section presents the specific design and implementation plan for the Agent-Based Model in the MOBI-TWIN context. This model operationalizes the behavioural mechanisms, environmental factors, and policy scenarios into a computational framework to simulate migration dynamics towards left-behind areas.

2.6.3 Conceptual Design of Agent-based Model

The agent-based model consists of four hierarchical and inter-related layers of data: agents, environment, rules, and scenarios. Specifically, the conceptual model includes:

Agents:

- Agents represent individual human beings (or, to reduce computational costs, sample subsets of individuals). They are of different types and can evaluate the attributes of their place of residence. Based on their own characteristics and the attributes of their location, they can make migration and mobility decisions, moving between regions. The types and initial proportions of agents are determined from real-world social attributes, such as gender, age, and education level. The collective aggregation of agents' behaviors influences the multi-dimensional attributes of the locations in which they reside.

Environment:

- The environment consists of geographical units in space, each with multi-dimensional attributes. These attributes change over time due to natural socio-economic and cultural evolution or because of policy interventions. Examples include demographic characteristics, housing prices, employment opportunities, transport accessibility, and online connectivity. Their initial values are determined from real-world data and evolve dynamically under the influence of agents and external policies. In the NetLogo model, shapefiles containing spatial attributes are imported to represent the geographical units.

Rules:

- Rules define the interaction logic between agents and the environment, between agents themselves, and between different environmental attributes. Mathematically, these are expressed as utility functions, correlations and/or dependences, which define the relationship between changes in agents and changes in the environment. They are informed by empirical data from this project as well as broader existing research. Specifically, the utility functions linking agents' mobility intentions (preferences) with environmental attributes will be based on the MOBI-TWIN survey results and the findings presented in this report. For agent–agent interactions, it will refer to the HUMAT framework, with the relevant utility functions to be defined. The interactions between environmental attributes will be determined from other literature and subsequent empirical studies.

Scenarios:

- Scenarios are alternative future development pathways designed by the modeler to simulate regional development trajectories, especially for left-behind areas. This ABM will include four scenarios, Leap Frog, Dark Horse, Snail Pace, and Lion's Den, plus a baseline reference scenario. The likelihoods and impacts of each scenario will be informed by the Delphi survey results. Each scenario corresponds to a specific set of rules, implemented through scenario-specific utility multipliers and unique rule combinations. For example:

- Leap Frog: Increased accessibility and digitalization, reduced migration costs, and expanded job opportunities.
- Dark Horse: Rapid increase in average wages, but with advantages concentrated in central regions.
- Snail Pace: Increased rents, reduced accessibility and digitalization, and significantly slowed demographic change.
- Lion's Den: Restricted reduction in housing costs and increased migration costs.

In these examples, scenarios influence individual migration decisions and environmental feedback through multipliers applied to wages, rents, accessibility, digitalization level, job vacancy rate, and migration costs.

2.6.4 Initialisation, Experiment, and Simulation

The conceptual model will first be initialized and tested to examine its stability and convergence properties, determine the required computational resources, and provide insights for further optimization. The prototype stage will focus on building the core model based on the utility functions linking agents' migration intentions (preferences) to environmental attributes, and will then gradually incorporate additional modules (sub-models) for agent interaction and environmental attributes to enrich the ABM. The required inputs for model initialization include:

Geospatial files for each pilot region:

- They contain population demographic information, socio-economic characteristics, and other attributes for each geographical unit, derived from micro-simulation.

Region-specific utility functions/dependences/correlations:

- Since pilot regions differ in social, cultural, and economic contexts, agents' migration intentions and preferences may vary accordingly. These differences will be reflected in the utility functions linking agents' migration intentions to environmental attributes, which will be determined from the analysis of MOBI-TWIN survey samples in each region.

Scenario:

- Choosing one of the four scenarios according to simulation strategies.

Random seed:

- To control the reproducibility and variability of simulation experiments.

Number of time steps:

- The length of the simulation run.

Two main approaches will be used to conduct simulation experiments: scenario-based approach and mechanism-based approach. The current conceptual model design follows a scenario-based approach, which answers the question “What would happen if…?” rather than “What will happen?”. Once supported by empirical evidence, the conceptual model can be further refined into a mechanism-based predictive model. Specifically:

Scenario-based approach:

- Uses existing local data to establish the initial state and apply different scenarios to simulate dynamic changes. This method is relatively straightforward and can forecast relative trends and changes in spatial patterns under alternative scenarios. However, it is less suitable for producing precise numerical forecasts.

Mechanism-based approach:

- Starts from theoretical mechanisms (e.g., migration utility, household decision-making, regional attractiveness) to define parameter ranges, and then simulates possible outcomes from the complex interactions of multiple parameters. This approach requires a more rigorous theoretical model, along with supporting prior and posterior empirical data, but can produce more precise numerical forecasts.

Following the scenario-based pathway, we developed a prototype version of ABM in Netlogo (Figure 9). After being expanded and adjusted based on empirical insights, this model will be applied to simulations for the five experimental regions, with the aim of projecting parameters related to inequality and sustainability. The model will embed an indicator framework to assess inequality and sustainability at specified time intervals.

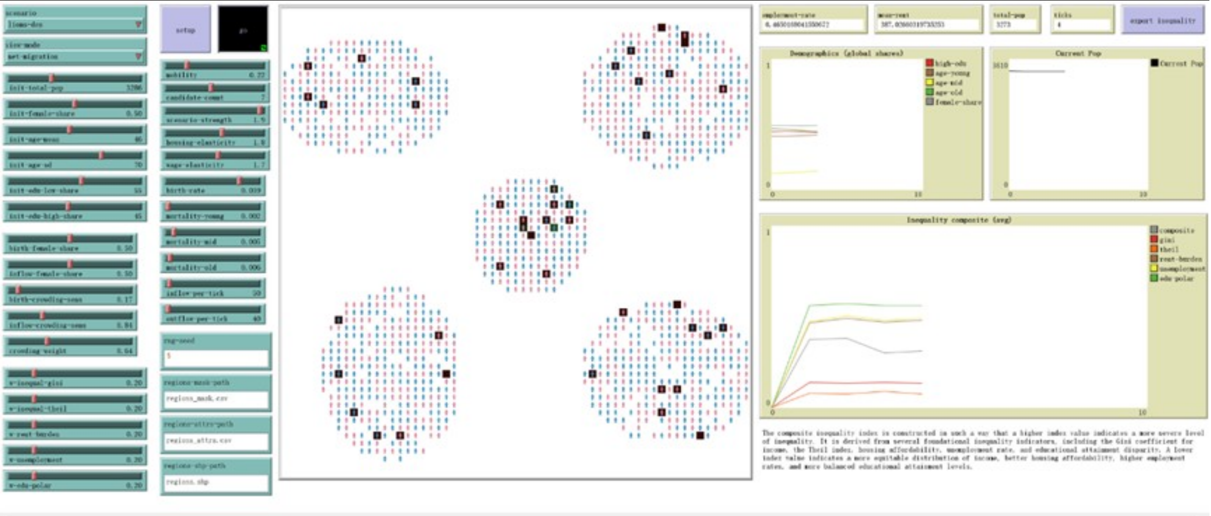


FIGURE 9. AN INTERFACE OF ABM PROTOTYPE

In the prototype version, inequality is measured through several dimensions. We used several common indicators and calculated an inequality index using a weighted method as an example. These indicators including:

Income Gini coefficient:

- Measures income distribution disparity.

Theil index:

- An entropy-based measure of inequality.

Rent burden ratio:

- Proportion of the population whose rent exceeds 40% of income.

Unemployment rate:

- Proportion of unemployed individuals in the population.

Educational polarization:

- Absolute difference between the proportion of highly educated and low-educated populations.

At the end of each simulation time-step (tick), the model will calculate inequality indicators at the cell level and aggregate them into a weighted composite inequality index (weights are adjustable and dimensions can be expanded). The following picture shows the interface of a prototype of this agent-based model.

2.6.5 SCALABILITY, ROBUSTNESS, AND CONVERGENCE

The prototype was designed as a scalable model. First, in terms of the selection of simulation paths, the logic of the simulation can be adjusted based on the degree of support from empirical data, i.e., choosing between following appropriate assumptions and following more evidence. Second, the content of the rules can be flexibly adjusted according to the simulation objectives and data availability, especially when considering the dynamic interactions between agents and between environmental variables.

Regarding the differences between regions/areas that may be more important than the attributes of individual regions/areas, this model also retains the possibility of extending the simulation environment to a network-based environment. This is because the intent of migration includes both push (from the source regions) and pull (from the target regions). In this case, different spatial units will be treated as nodes, and the flow of different agents between them will be abstracted as edges in the network. In other words, the environment will be expanded into a network space with multidimensional attributes, rather than being defined solely as a geographical space with multiple attributes. When data for left-behind regions and their related regions becomes available, we will be able to structure such differences between regions using a matrix structure and further incorporate these differences into the simulation process.

In addition, to ensure that the simulation results of the model are robust and convergent, we will adopt a diversified strategy in the actual simulation, specifically including:

Switching experimental subjects:

- Running the model for different pilot regions to check whether the model mechanisms are overly sensitive to specific regional characteristics. If key output trends remain consistent across regions, this suggests good cross-regional applicability.

Batch runs:

- Using NetLogo BehaviorSpace to conduct multiple independent simulations under the same parameter settings, then analyzing the output distributions. This identifies random biases in single runs and enables estimation of confidence intervals.

Parameter perturbation:

- Applying small random variations (e.g., ±5% or ±10%) to fixed parameters to assess the sensitivity of model outputs to minor changes in inputs. Robust models will maintain consistent trends and structural stability under reasonable perturbations.

Changing random seeds:

- Testing how stochastic variation affects the results. If output distributions remain similar across different seeds, this indicates results are driven by core mechanisms rather than chance.

Extended run length:

- Monitoring whether key output variables (e.g., demographic composition, inequality indices, migration flows) stabilize or fluctuate within a stable range over long runs. Stability indicates convergence, whereas persistent erratic drift suggests structural instability requiring model or parameter adjustments.

3 MAIN RESULTS

This section presents the main outcomes of the spatial microsimulation models developed for the five pilot regions. Drawing on the MOBI-TWIN survey and country-specific census or administrative data, we generated synthetic populations in which individuals are assigned weights that match the socio-demographic characteristics of the pilot areas. The constraint variables and target variables are linked at the appropriate small-area level for each country, allowing us to produce geographically disaggregated estimates of residential mobility. This way we can infer the probabilities of changing residence within defined time horizons (e.g., 6 months, 1-2 year, 5 years and 10 years) for each of the pilot areas.

The results are organized into three subsections. We begin by evaluating the performance of the spatial microsimulation models through validation techniques. While internal validation is systematically applied across all pilot regions, external validation was only possible for Greece, where suitable tax return data were available. We then describe the socio-demographic composition of the synthetic populations for each pilot area. Finally, we investigate the estimated probabilities of residential mobility using regression analysis (OLS, Logit, and Probit models). The analysis allows us to detect spatial patterns of residential mobility, highlight regional imbalances, such as ageing populations or concentrations of low-income individuals. It also allows us to compare mobility trends across different areas. These insights can contribute to more informed policy-making and strategic planning.

3.1 INTERNAL AND EXTERNAL VALIDATION OF THE SPATIAL MICROSIMULATION MODELS

Before moving to the analysis of the spatial microsimulation results, it is important to first assess how well our models perform through internal validation for each pilot area. Internal validation involves comparing the values of variables generated by our synthetic populations against the actual values provided by the official statistical authorities for the same geographical units. In practice, this means that we take variables that were already used as constraints in the spatial microsimulation process and check how closely the synthetic estimates reproduce the observed reality.

Specifically, for all pilots except Groningen, we compare the actual number of unemployed individuals in each area with the synthetic number of unemployed individuals generated using the microsimulation weights. For the Netherlands, due to data availability, the comparison is performed using the number of married individuals instead. These comparisons allow us to visually and quantitatively confirm that the spatial microsimulation procedure reproduces the distributions of the constraint variables with high accuracy. Because these variables are directly involved in the construction of the model, this process is formally known as internal validation.

The internal validation results are presented in the form of scatterplots, one for each pilot region, in Figures 10 to Figure 14. Each graph compares the actual number of individuals in a selected constraint category (e.g., unemployed or married) on the horizontal axis with the corresponding synthetic number of individuals generated through our spatial microsimulation models on the vertical axis. A 45-degree reference line (y = x) is included in each plot, which represents the scenario of perfect agreement between the simulated and observed values. Points lying exactly on this line indicate that the model has reproduced the true number of the variable of interest with high accuracy.

Across most pilot regions, the scatterplots show that the majority of points lie very close to or directly on the 45-degree line, demonstrating that the spatial microsimulation models perform with high internal accuracy. This indicates that the model is robust in replicating the distributions of the constraint variables at the small-area level. Minor deviations occur only in a small number of units, reflecting a common difficulty of perfectly matching all values in microsimulation.

An interesting exception is observed in the Greek pilot where a systematic pattern emerges. In some of the more populated municipalities, the simulated number of unemployed individuals tends to be slightly underestimated, which is reflected by the corresponding points falling below the 45-degree line (since the vertical axis represents the synthetic values and the horizontal axis represents the actual values). For a clearer view, we also present the internal validation graph for a smaller sample of municipalities (with less than 50000 inhabitants) where the same pattern of underestimation emerges. Nevertheless, the deviations remain modest, and the overall alignment of points with the reference line confirms a strong internal validity of the models across all pilots.

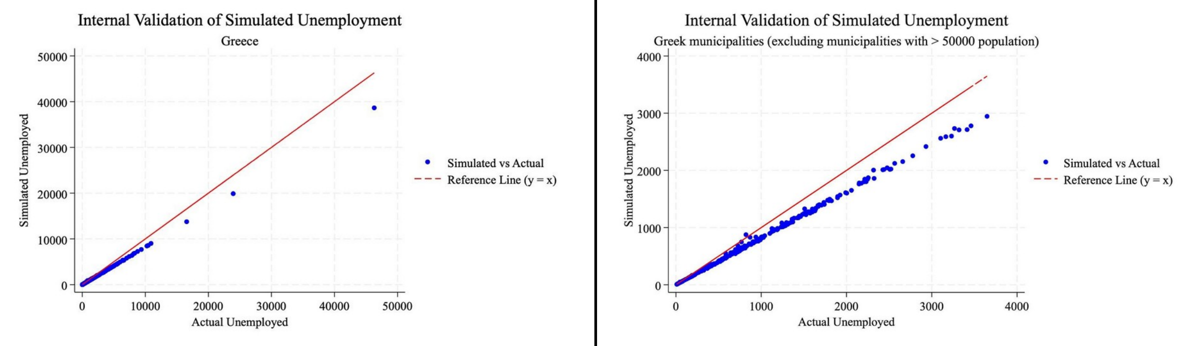


FIGURE 10. INTERNAL VALIDATION OF THE SPATIAL MICROSIMULATION MODELS FOR GREECE

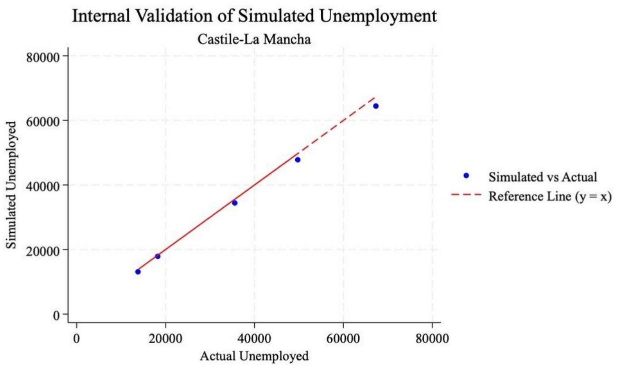


FIGURE 11. INTERNAL VALIDATION OF THE SPATIAL MICROSIMULATION MODELS FOR SPAIN

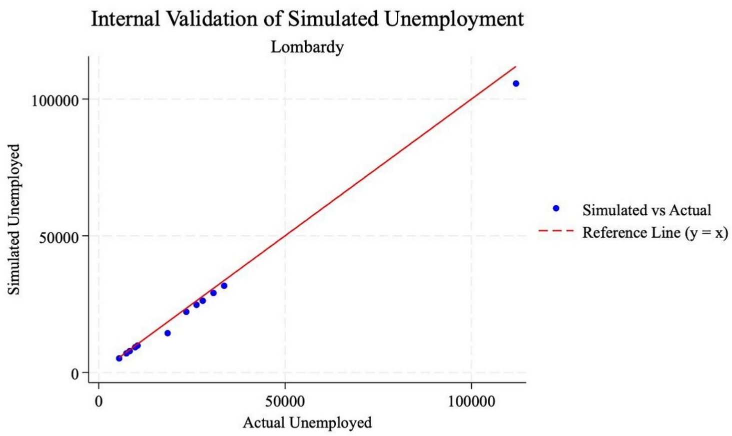


FIGURE 12. INTERNAL VALIDATION OF THE SPATIAL MICROSIMULATION MODELS FOR ITALY

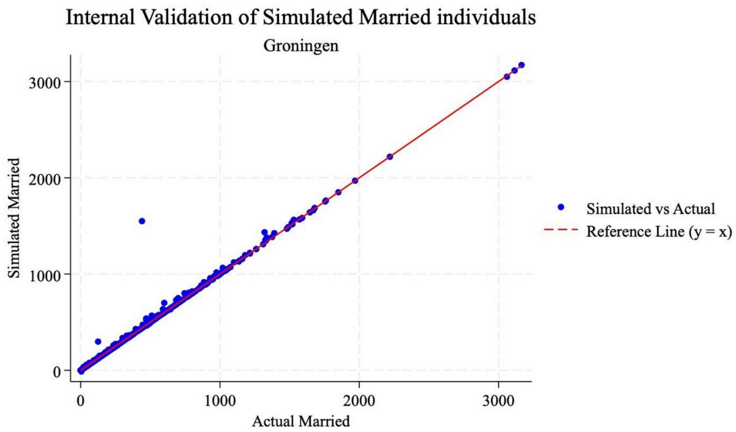


FIGURE 13. INTERNAL VALIDATION OF THE SPATIAL MICROSIMULATION MODELS FOR THE NETHERLANDS

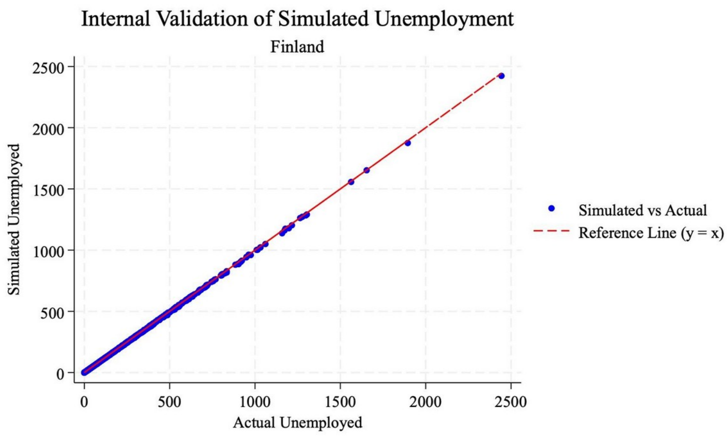


FIGURE 14. INTERNAL VALIDATION OF THE SPATIAL MICROSIMULATION MODELS FOR FINLAND

To complement the internal validation, we also conducted external validation of the spatial microsimulation model wherever suitable data were available. External validation assesses how well the synthetic population reproduces the statistical properties of variables that were not used as constraints in the model. This requires access to independent datasets that provide actual observations for a variable that can be generated from the synthetic microdata. In other words, we compare simulated estimates for an external variable with real-world data to assess the model's predictive validity beyond the constraint variables.

External validation was only feasible for the Greek pilot, as we were able to access tax return data at the prefecture level. Following the same principle as in the internal validation, we focused on the number of individuals with an annual income below €15,000. Using the synthetic microdata, we estimated these counts by combining the spatial microsimulation weights with the income variable from the MOBI-TWIN survey and then compared them with the actual prefecture-level data from tax returns.

Because Greek prefectures vary significantly in population size, we present four external validation graphs to better visualise the results across different subsets of prefectures:

1. All prefectures in Greece
2. All prefectures excluding Athens and Thessaloniki, to reduce the influence of the largest urban areas
3. Prefectures of Central Macedonia (pilot region)

4. Prefectures of Central Macedonia excluding Thessaloniki, allowing for a clearer view of prefectures with comparable population sizes

Across all four graphs in Figure 15, a consistent pattern emerges: the scatterplot points fall very close to, or overlap with, the 45° reference line (y = x), indicating a strong alignment between the simulated and actual number of individuals earning less than €15,000. This high degree of agreement provides robust evidence of the validity of our spatial microsimulation model for the Greek pilot and increases confidence in its ability to produce reliable small-area estimates for variables not explicitly used as constraints.

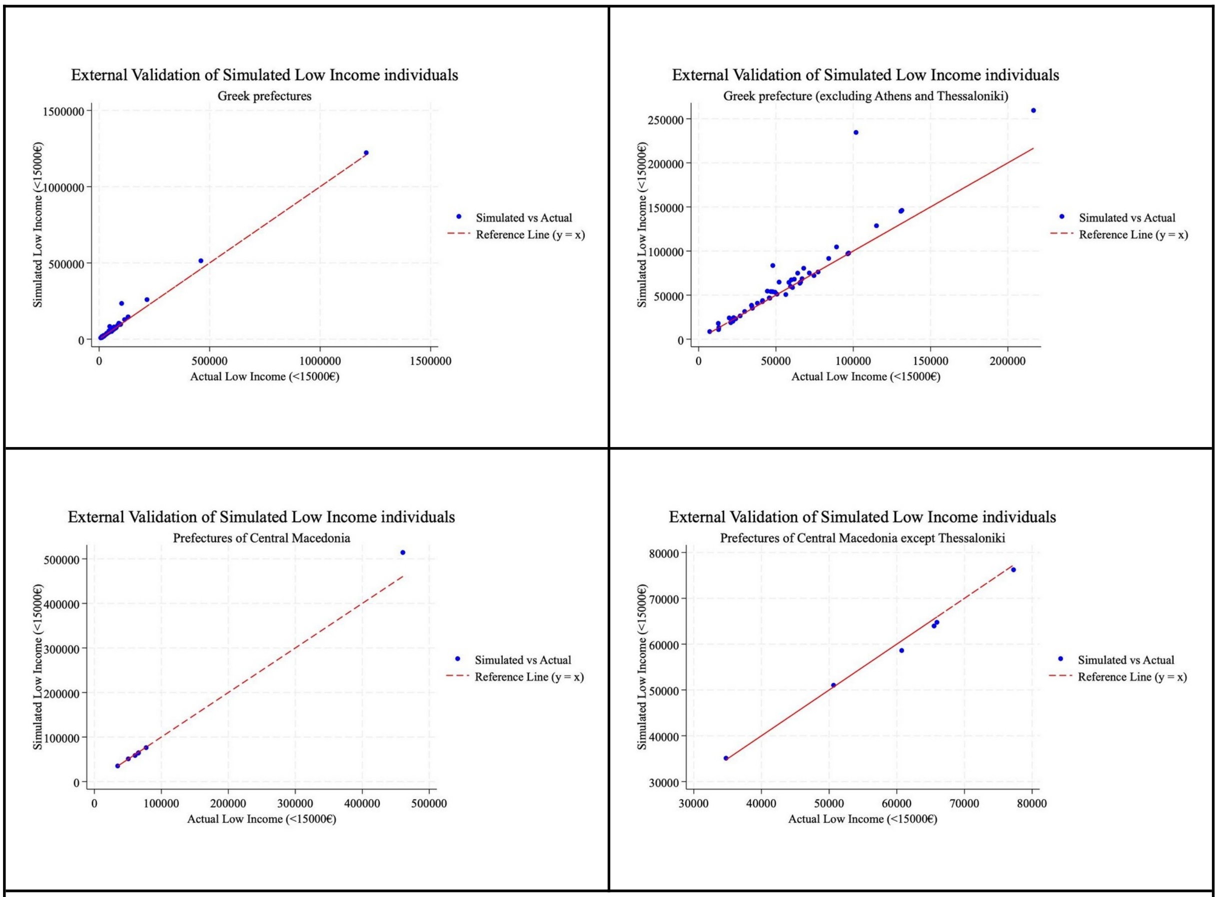


FIGURE 15. EXTERNAL VALIDATION OF THE SPATIAL MICROSIMULATION MODELS FOR GREECE

3.2 INTENTION TO CHANGE RESIDENCE PER PILOT REGION

After validating the spatial microsimulation models, we proceed to generate the target variables of interest, focusing on individuals' intentions to change their place of residence over different time horizons. The MOBI-TWIN survey provides direct measures of these intentions, covering four distinct periods: the next 6 months, the next 1–2 years, the next 5 years, and the next 10 years. Using the microsimulation weights, we transform these survey responses into synthetic variables that can be mapped across all sub-regions in the pilot areas. This enables us to visualise and compare how relocation intentions evolve spatially across the different time horizons, and to identify which areas are more likely to experience higher out-migration in the future.

For each pilot area, we present four thematic maps (Figures 16–20), one for each time horizon. In the cases of Greece, the Netherlands, and Finland, where the number of spatial units is large (municipalities, neighbourhoods, or postal codes), we have used the same number of bins and identical threshold values in the map legends across all four time horizons. This approach facilitates a direct comparison over time, allowing us to clearly observe the gradual higher mobility intentions as the time horizon expands from the short term (6 months) to the long term (10 years). In contrast, for Spain and Italy, the analysis is carried out at the NUTS-3 level, covering only 5 provinces in Castile-La Mancha and 12 provinces in Lombardy, respectively. Due to this smaller number of spatial units, it was not feasible to apply uniform thresholds across all four maps without producing legends that lacked overlap or comparability.

Despite these differences in spatial resolution, a consistent pattern emerges across all pilots, that is, mobility intentions increase with longer time horizons. For Greece, Spain, and Italy, the proportion of individuals intending to move within the next 6 months is very low (generally below 3%). However, this share rises substantially over time, reaching approximately 15–20% when the horizon extends to 10 years. In contrast, in the Netherlands and Finland, the baseline level of short-term mobility is already much higher, even within the next 6 months, as there are numerous areas where more than 10% of individuals express an intention to relocate. When extended to the 10-year horizon, these shares grow substantially, often reaching or even exceeding 30% in certain areas.

These findings not only confirm the expected increase in relocation intentions over longer time frames, but also highlight important cross-country differences in mobility patterns. While Southern European regions such as Greece, Spain, and Italy exhibit relatively conservative short-term mobility intentions that gradually increase in the long run, Northern regions such as the Netherlands and Finland display higher baseline mobility and more pronounced long-term projections. This divergence may reflect broader socio-economic and cultural differences in housing markets, labour mobility, and demographic structures, which will be further investigated in the subsequent stages of the project.

3.2.1 Greece: Region of Central Macedonia

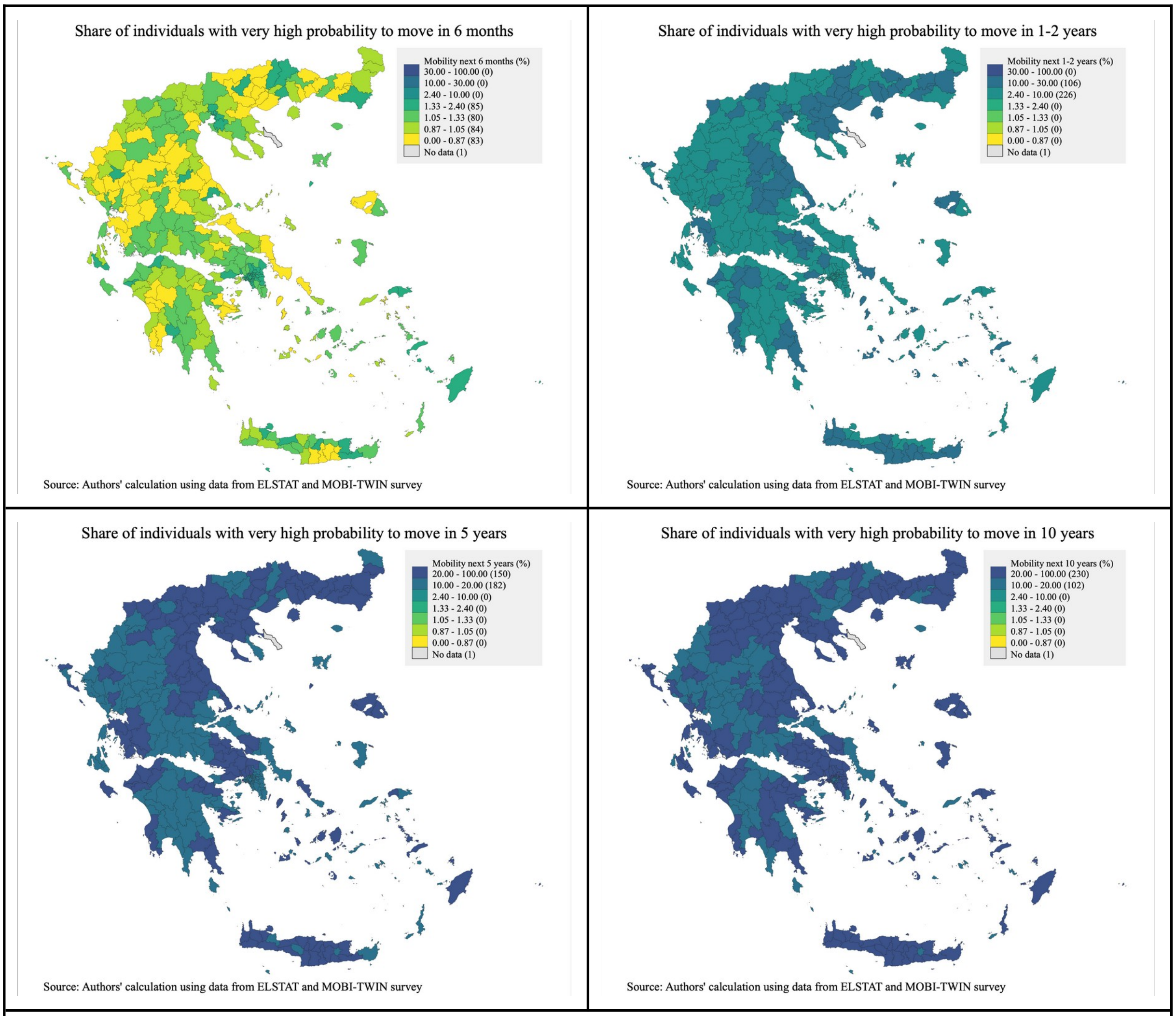


Figure 16. Spatial distribution of probabilities to change residence in different time Horizons in Greece

3.2.2 Spain: Castile La-Mancha

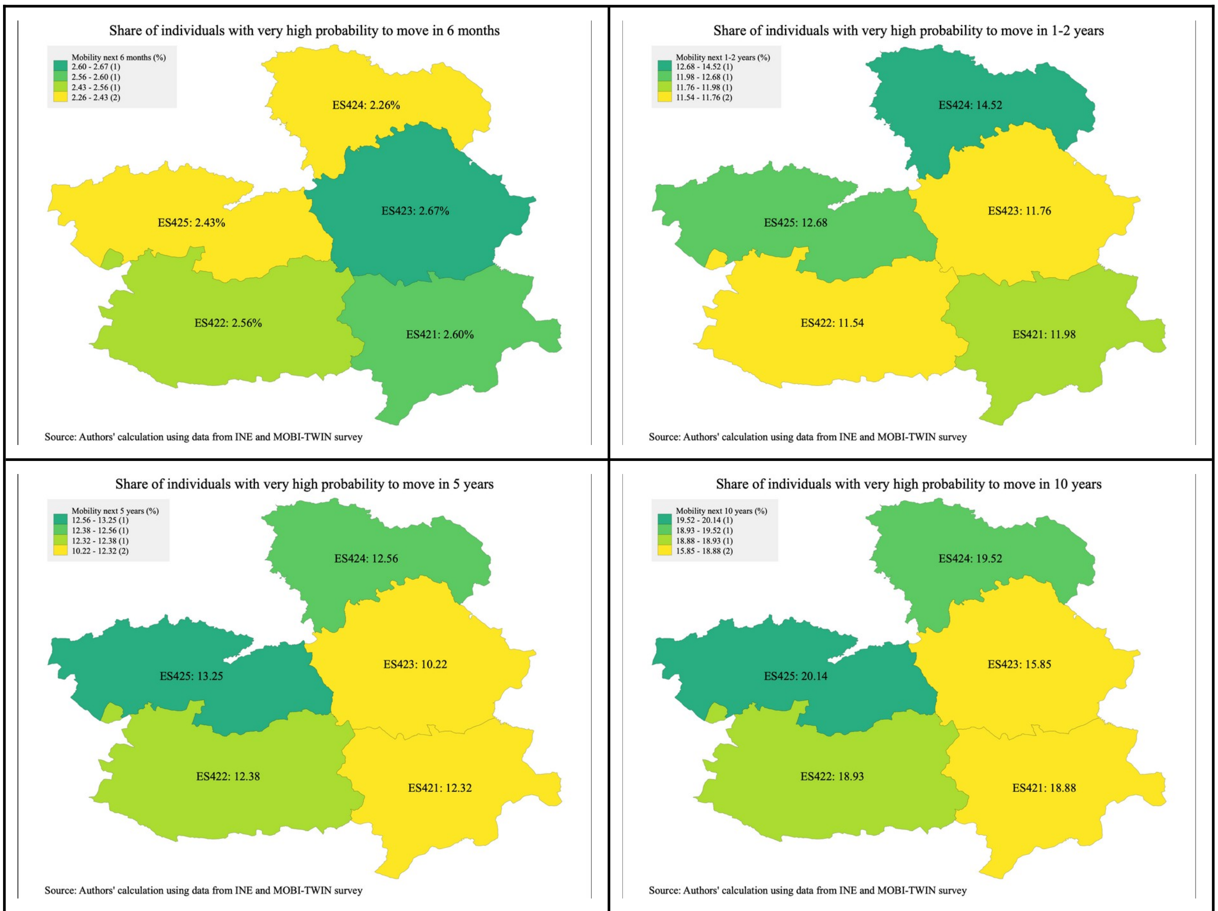


Figure 17. Spatial distribution of probabilities to change residence in different time Horizons in Castile-La Mancha

3.2.3 Italy: Lombardy

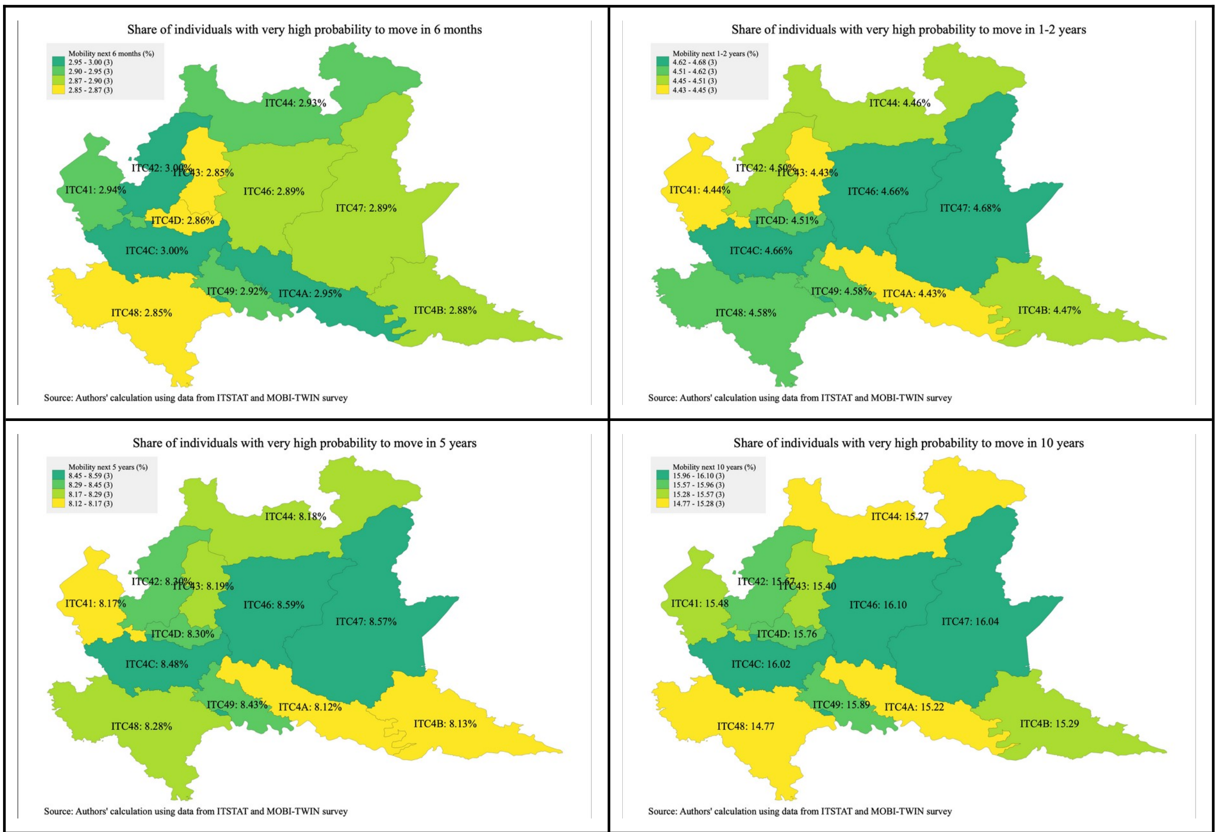


Figure 18. Spatial distribution of probabilities to change residence in different time Horizons in Lombardy

3.2.4 Netherlands: Groningen

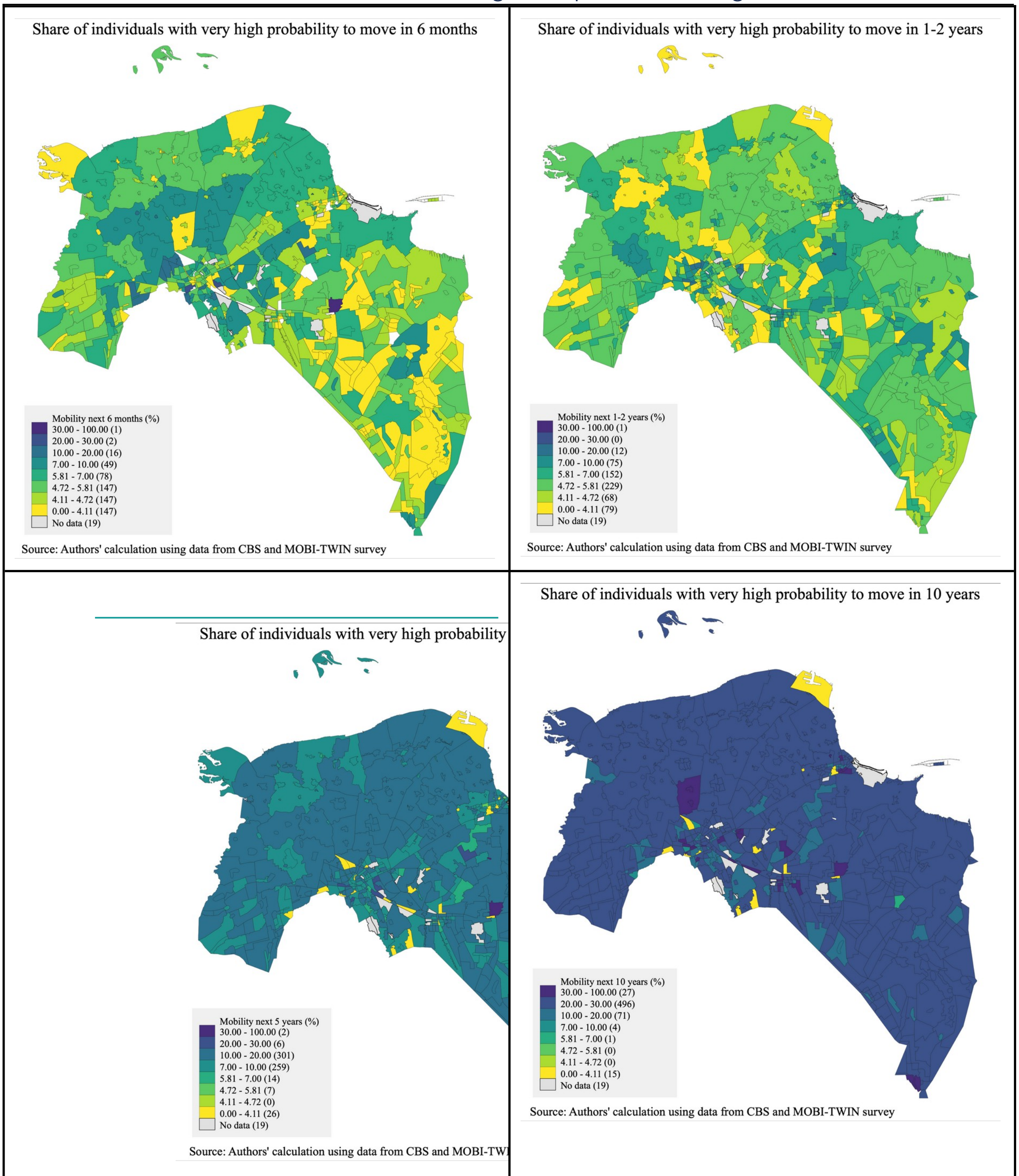
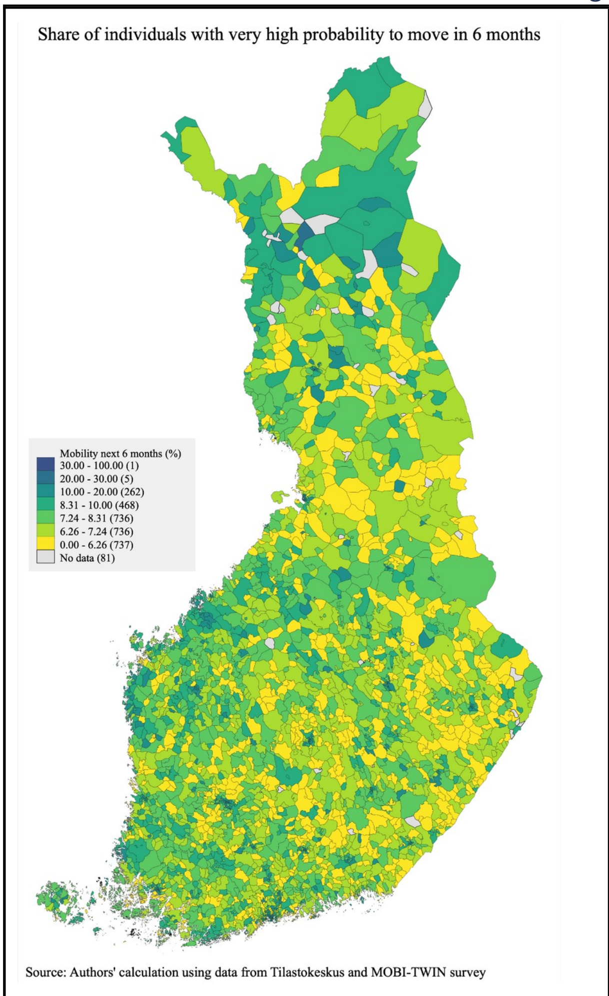
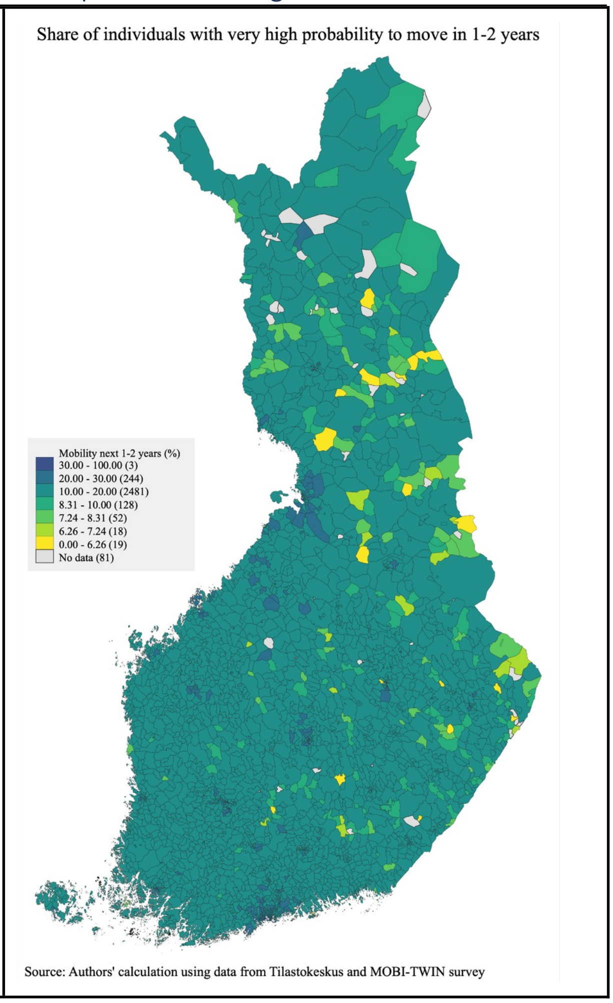


FIGURE 19. SPATIAL DISTRIBUTION OF PROBABILITIES TO CHANGE RESIDENCE IN DIFFERENT TIME HORIZONS IN GRONINGEN

3.2.5 FINLAND: NORTH AND EAST FINLAND


Share of individuals with very high probability to move in 6 months
Mobility next 6 months (%)
30.00 - 100.00 (1)
20.00 - 30.00 (5)
10.00 - 20.00 (262)
8.31 - 10.00 (468)
7.24 - 8.31 (736)
6.26 - 7.24 (736)
0.00 - 6.26 (737)
No data (81)
Source: Authors' calculation using data from Tilastokeskus and MOBI-TWIN survey


Share of individuals with very high probability to move in 1-2 years
Mobility next 1-2 years (%)
30.00 - 100.00 (3)
20.00 - 30.00 (244)
10.00 - 20.00 (2481)
8.31 - 10.00 (128)
7.24 - 8.31 (52)
6.26 - 7.24 (18)
0.00 - 6.26 (19)
No data (81)
Source: Authors' calculation using data from Tilastokeskus and MOBI-TWIN survey

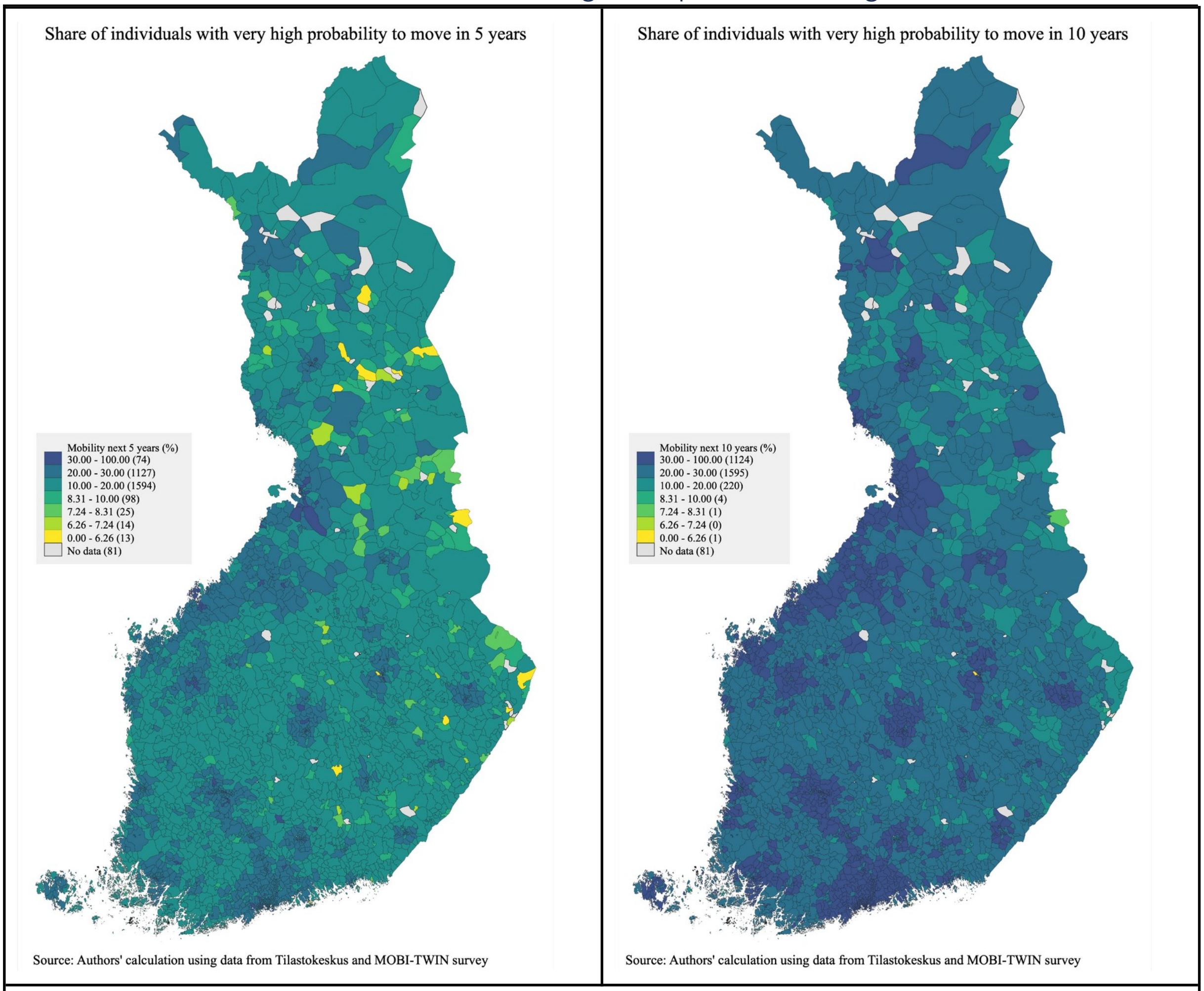


FIGURE 20. SPATIAL DISTRIBUTION OF PROBABILITIES TO CHANGE RESIDENCE IN DIFFERENT TIME HORIZONS IN GREECE

3.3 PROBABILITIES OF CHANGING RESIDENCE BASED ON REGRESSION ANALYSIS

In the final step of our analysis, we estimate the probabilities of changing residence for different socio-demographic groups using regression models. The analysis is based on the MOBI-TWIN survey, restricted to the national subsamples corresponding to each pilot area. We apply three types of regression models such as OLS, Logit, and Probit, however, for clarity and comparability, the aggregated results presented in Tables 3 to 7 are based on Logit models.[7]

Our modelling strategy proceeds in two stages. First, the five-point categorical variable from the MOBI-TWIN survey on mobility intentions (1 = "Very low

[7] The results, however, are very similar between the three approaches.

probability", 2 = "Low probability", 3 = "Moderate probability", 4 = "High probability", 5 = "Very high probability") is transformed into five binary indicators. Each dummy variable represents whether or not an individual reported belonging to one of the five categories for each time horizon (6 months, 1–2 years, 5 years, and 10 years). These binary indicators are then used as the dependent variables in the regression models. Second, after estimating the models, we calculate the predicted probabilities and then aggregate them by the socio-demographic categories used as independent variables.

The independent variables included in the models are age group, educational level, and gender. These three dimensions were selected because they are the only socio-demographic characteristics used consistently as constraint variables across all five spatial microsimulation models, making them the most robust basis for comparison. This allows us to answer questions such as: "*What is the average predicted probability of reporting a very low probability of moving in the next 6 months among women?*" or "*What is the probability that men report a high likelihood of moving in the next 10 years?*".

The results reveal some clear patterns. Across all five pilot areas, around 70% of the population reports a very low probability of moving within the next 6 months, while only about 5% report a very high probability. As the time horizon increases, these shares shift significantly. For the 1–2 year horizon, roughly 50% still report a very low probability, while 10% report a very high probability. At the 5-year horizon, the balance shifts further: only about 30% report a very low probability, while around 20% report a very high probability. By the 10-year horizon, these trends are most pronounced with only 25% of individuals reporting a very low probability, while nearly 30% report a very high probability of moving. These patterns are consistent across all five pilot regions, indicating a strong and systematic relationship between time horizon and mobility intentions.

When focusing on socio-demographic differences, some additional insights emerge. Younger age groups consistently show much higher probabilities of intending to move, particularly over the longer time horizons. Educational attainment also matters as individuals with tertiary education display higher probabilities of moving in the long term (10 years), suggesting that more highly educated groups are more open to relocation opportunities over time. By contrast, gender differences are less systematic. While no consistent patterns are observed across all pilot areas, some variation appears when examining specific countries, which suggests that a more careful, country-level analysis is required before drawing firm conclusions.

Overall, the regression analysis confirms that mobility intentions increase with longer time horizons and that socio-demographic characteristics, particularly age and education, play an important role in shaping these intentions. These findings reinforce the importance of integrating socio-demographic heterogeneity into future mobility scenarios and provide a strong basis for extending the analysis in the next phases of the project.

	The next six months					
Greece	**Very low probability**	**Low probability**	**Moderate probability**	**High probability**	**Very high probability**	**Total**
MALES	73.70%	10.33%	7.09%	6.77%	4.37%	100.00%
FEMALES	68.73%	10.64%	10.40%	6.45%	5.49%	100.00%
18-25	67.78%	8.16%	9.64%	9.55%	5.45%	100.00%
26-40	63.01%	14.42%	10.11%	7.26%	5.79%	100.00%
41-65	81.44%	6.10%	6.44%	3.71%	3.21%	100.00%
65+	94.43%				6.08%	100.00%
PRIMARY	84.96%		17.12%			100.00%
SECONDARY	73.05%	9.51%	8.42%	6.07%	4.52%	100.00%
TERTIARY	69.30%	10.98%	8.55%	6.87%	5.17%	100.00%

	The next 1-2 years					
Greece	**Very low probability**	**Low probability**	**Moderate probability**	**High probability**	**Very high probability**	**Total**
MALES	47.90%	17.44%	14.06%	11.11%	10.90%	100.00%
FEMALES	42.61%	14.50%	18.55%	13.13%	12.30%	100.00%
18-25	28.30%	15.77%	24.66%	16.12%	15.13%	100.00%
26-40	36.02%	17.97%	17.42%	14.19%	14.37%	100.00%
41-65	62.22%	14.74%	9.80%	7.72%	5.51%	100.00%
65+	89.17%	5.03%		5.63%		100.00%
PRIMARY	62.63%	6.70%	8.95%	15.23%	8.93%	100.00%
SECONDARY	50.12%	14.51%	16.38%	9.30%	11.40%	100.00%
TERTIARY	41.54%	17.25%	16.82%	13.43%	11.90%	100.00%

	The next 5 years					
Greece	**Very low probability**	**Low probability**	**Moderate probability**	**High probability**	**Very high probability**	**Total**
MALES	31.46%	12.08%	19.06%	17.46%	22.32%	100.00%
FEMALES	28.97%	12.19%	20.27%	20.55%	19.60%	100.00%
18-25	6.14%	6.16%	16.05%	33.44%	39.14%	100.00%
26-40	23.21%	12.20%	22.96%	20.20%	21.83%	100.00%
41-65	47.80%	15.98%	19.32%	9.92%	8.24%	100.00%
65+	78.92%	10.67%	5.80%	5.87%		100.00%
PRIMARY	41.77%	6.62%	27.41%		26.67%	100.00%
SECONDARY	36.90%	11.82%	17.86%	16.84%	17.87%	100.00%
TERTIARY	25.99%	12.65%	20.13%	20.20%	22.09%	100.00%

	The next 10 years					
Greece	**Very low probability**	**Low probability**	**Moderate probability**	**High probability**	**Very high probability**	**Total**
MALES	24.32%	12.31%	19.23%	15.64%	31.85%	100.00%
FEMALES	22.64%	12.40%	18.53%	17.77%	30.60%	100.00%
18-25	6.24%	2.77%	11.85%	22.86%	56.42%	100.00%
26-40	16.05%	12.92%	20.87%	18.94%	31.36%	100.00%
41-65	36.43%	17.70%	20.64%	11.43%	15.29%	100.00%
65+	78.96%		17.03%	5.56%		100.00%
PRIMARY	42.59%		27.61%	14.95%	18.12%	100.00%
SECONDARY	27.21%	13.27%	17.60%	14.77%	29.96%	100.00%
TERTIARY	20.31%	11.90%	18.98%	17.85%	32.55%	100.00%

TABLE 3. PROBABILITIES OF CHANGING RESIDENCE BASED ON REGRESSION ANALYSIS FOR PILOT 1 (GREECE)

	The next six months					
Spain	**Very low probability**	**Low probability**	**Moderate probability**	**High probability**	**Very high probability**	**Total**
MALES	77.14%	9.99%	5.54%	4.69%	3.40%	100.00%
FEMALES	74.45%	9.08%	7.17%	4.35%	5.54%	100.00%
18-25	67.77%	14.31%	6.17%	3.85%	8.03%	100.00%
26-40	70.62%	10.59%	8.45%	5.68%	4.65%	100.00%
41-65	85.23%	5.78%	3.56%	3.26%	2.13%	100.00%
65+	94.86%	4.52%				100.00%
PRIMARY	80.22%	2.67%	5.28%	2.63%	11.70%	100.00%
SECONDARY	79.73%	8.18%	5.15%	3.29%	4.48%	100.00%
TERTIARY	73.43%	10.65%	7.07%	5.23%	4.13%	100.00%

	The next 1-2 years					
Spain	**Very low probability**	**Low probability**	**Moderate probability**	**High probability**	**Very high probability**	**Total**
MALES	50.35%	16.47%	18.34%	8.38%	8.42%	100.00%
FEMALES	49.50%	14.12%	16.29%	10.04%	11.21%	100.00%
18-25	30.59%	17.42%	22.36%	15.34%	13.75%	100.00%
26-40	40.75%	15.80%	20.71%	10.32%	12.39%	100.00%
41-65	71.51%	11.76%	9.17%	3.90%	3.83%	100.00%
65+	73.83%	24.57%				100.00%
PRIMARY	58.92%	6.60%	19.89%	3.03%	14.41%	100.00%
SECONDARY	58.32%	13.35%	14.74%	8.04%	7.28%	100.00%
TERTIARY	45.01%	16.80%	18.39%	10.19%	10.91%	100.00%

	The next 5 years					
Spain	**Very low probability**	**Low probability**	**Moderate probability**	**High probability**	**Very high probability**	**Total**
MALES	32.41%	11.60%	20.84%	17.59%	18.50%	100.00%
FEMALES	34.95%	10.60%	19.01%	17.73%	18.27%	100.00%
18-25	7.96%	6.50%	21.03%	31.39%	31.46%	100.00%
26-40	27.27%	11.10%	21.05%	18.17%	22.41%	100.00%
41-65	54.92%	14.86%	15.93%	8.40%	6.45%	100.00%
65+	60.71%	4.82%	30.50%		5.11%	100.00%
PRIMARY	52.37%	5.11%	12.79%	21.17%	10.84%	100.00%
SECONDARY	43.80%	9.20%	16.92%	15.36%	15.23%	100.00%
TERTIARY	27.34%	12.44%	21.88%	18.61%	20.47%	100.00%

	The next 10 years					
Spain	**Very low probability**	**Low probability**	**Moderate probability**	**High probability**	**Very high probability**	**Total**
MALES	26.16%	9.56%	17.15%	16.26%	30.96%	100.00%
FEMALES	28.74%	9.12%	17.10%	14.32%	30.75%	100.00%
18-25	5.53%	5.27%	11.34%	20.69%	55.57%	100.00%
26-40	19.70%	10.42%	19.09%	15.80%	34.97%	100.00%
41-65	47.16%	10.96%	18.95%	10.22%	13.46%	100.00%
65+	61.85%	5.04%	10.10%	21.21%	5.01%	100.00%
PRIMARY	53.59%	5.13%	10.41%	15.83%	16.37%	100.00%
SECONDARY	37.07%	8.19%	14.39%	12.85%	27.12%	100.00%
TERTIARY	20.91%	10.19%	18.96%	16.48%	33.69%	100.00%

TABLE 4. PROBABILITIES OF CHANGING RESIDENCE BASED ON REGRESSION ANALYSIS FOR PILOT 2 (SPAIN)

	The next six months					
Italy	Very low probability	Low probability	Moderate probability	High probability	Very high probability	Total
MALES	74.95%	10.67%	6.81%	2.84%	4.71%	100.00%
FEMALES	73.60%	8.86%	7.27%	4.70%	5.62%	100.00%
18-25	70.85%	9.96%	7.50%	5.74%	6.04%	100.00%
26-40	65.65%	12.90%	10.26%	4.23%	6.98%	100.00%
41-65	86.05%	6.26%	2.82%	2.27%	2.57%	100.00%
65+	92.27%	1.67%	2.50%	1.80%	1.75%	100.00%
PRIMARY	81.03%	7.68%	7.01%	1.06%	3.36%	100.00%
SECONDARY	81.35%	7.73%	5.03%	2.62%	3.28%	100.00%
TERTIARY	70.18%	10.90%	8.09%	4.58%	6.27%	100.00%
	The next 1-2 years					
Italy	Very low probability	Low probability	Moderate probability	High probability	Very high probability	Total
MALES	51.32%	16.05%	14.16%	10.77%	8.22%	100.00%
FEMALES	48.23%	16.23%	15.26%	10.67%	9.94%	100.00%
18-25	32.66%	23.79%	19.26%	13.06%	11.42%	100.00%
26-40	35.15%	20.26%	18.32%	14.69%	11.70%	100.00%
41-65	74.42%	7.66%	8.82%	4.90%	3.95%	100.00%
65+	95.60%	1.69%	0.84%	1.84%		100.00%
PRIMARY	67.90%	10.67%	11.34%	3.54%	6.95%	100.00%
SECONDARY	57.18%	14.98%	14.06%	7.03%	7.27%	100.00%
TERTIARY	44.75%	17.09%	15.28%	13.08%	10.15%	100.00%
	The next 5 years					
Italy	Very low probability	Low probability	Moderate probability	High probability	Very high probability	Total
MALES	32.87%	14.37%	18.82%	17.77%	16.21%	100.00%
FEMALES	33.76%	10.93%	19.01%	18.58%	17.66%	100.00%
18-25	10.66%	8.29%	23.22%	29.81%	28.32%	100.00%
26-40	19.42%	15.09%	22.89%	21.64%	21.03%	100.00%
41-65	59.96%	12.58%	12.25%	8.60%	6.32%	100.00%
65+	82.83%	6.78%	7.48%	0.86%	1.72%	100.00%
PRIMARY	53.87%	10.98%	14.22%	8.39%	11.10%	100.00%
SECONDARY	38.06%	11.40%	19.18%	16.57%	14.81%	100.00%
TERTIARY	29.59%	13.33%	19.07%	19.63%	18.42%	100.00%
	The next 10 years					
Italy	Very low probability	Low probability	Moderate probability	High probability	Very high probability	Total
MALES	26.80%	10.42%	17.15%	15.48%	30.21%	100.00%
FEMALES	27.28%	9.84%	15.23%	16.67%	30.86%	100.00%
18-25	5.90%	4.57%	11.62%	22.98%	55.33%	100.00%
26-40	15.16%	11.91%	19.29%	19.31%	34.44%	100.00%
41-65	48.98%	11.51%	15.94%	9.22%	13.90%	100.00%
65+	82.19%	6.75%	6.50%	2.59%	1.64%	100.00%
PRIMARY	51.84%	7.47%	12.80%	5.60%	20.19%	100.00%
SECONDARY	32.08%	8.55%	16.39%	13.02%	29.96%	100.00%
TERTIARY	22.89%	11.10%	16.26%	18.34%	31.49%	100.00%

TABLE 5. PROBABILITIES OF CHANGING RESIDENCE BASED ON REGRESSION ANALYSIS FOR PILOT 3 (ITALY)

	The next six months					
Netherlands	**Very low probability**	**Low probability**	**Moderate probability**	**High probability**	**Very high probability**	**Total**
MALES	73.39%	11.51%	6.47%	4.47%	5.30%	100.00%
FEMALES	72.71%	11.05%	8.02%	3.54%	5.51%	100.00%
18-25	62.48%	17.48%	9.56%	6.36%	4.39%	100.00%
26-40	68.17%	12.15%	9.03%	5.21%	5.70%	100.00%
41-65	83.43%	6.14%	4.07%	0.60%	5.86%	100.00%
65+	94.15%		2.18%		3.82%	100.00%
PRIMARY	63.06%	16.08%	21.56%			100.00%
SECONDARY	78.98%	9.63%	6.80%	1.68%	3.61%	100.00%
TERTIARY	70.60%	11.82%	6.61%	5.13%	6.34%	100.00%

	The next 1-2 years					
Netherlands	**Very low probability**	**Low probability**	**Moderate probability**	**High probability**	**Very high probability**	**Total**
MALES	52.18%	17.84%	12.93%	9.71%	10.00%	100.00%
FEMALES	46.91%	13.27%	16.97%	11.70%	12.73%	100.00%
18-25	30.85%	15.82%	22.91%	15.78%	15.18%	100.00%
26-40	37.02%	19.54%	16.27%	12.63%	15.03%	100.00%
41-65	73.29%	9.35%	9.62%	4.70%	4.08%	100.00%
65+	87.06%		6.09%		7.68%	100.00%
PRIMARY	55.01%	15.44%	33.09%			100.00%
SECONDARY	54.56%	13.43%	17.52%	7.91%	8.01%	100.00%
TERTIARY	46.50%	16.47%	12.57%	12.17%	13.16%	100.00%

	The next 5 years					
Netherlands	**Very low probability**	**Low probability**	**Moderate probability**	**High probability**	**Very high probability**	**Total**
MALES	29.78%	14.18%	19.32%	18.23%	19.40%	100.00%
FEMALES	31.44%	7.45%	21.52%	17.62%	23.04%	100.00%
18-25	4.94%	7.53%	23.87%	29.23%	34.42%	100.00%
26-40	17.97%	11.47%	22.62%	21.96%	26.58%	100.00%
41-65	58.21%	12.08%	16.74%	7.21%	7.05%	100.00%
65+	74.40%	7.39%	10.26%	3.62%	5.74%	100.00%
PRIMARY	40.26%		45.85%	11.81%		100.00%
SECONDARY	39.85%	5.37%	20.56%	20.03%	14.11%	100.00%
TERTIARY	25.27%	13.45%	18.80%	17.21%	24.97%	100.00%

	The next 10 years					
Netherlands	**Very low probability**	**Low probability**	**Moderate probability**	**High probability**	**Very high probability**	**Total**
MALES	23.65%	8.68%	15.53%	15.55%	36.35%	100.00%
FEMALES	23.93%	5.12%	19.90%	14.23%	36.81%	100.00%
18-25	4.68%	2.19%	10.14%	19.07%	63.17%	100.00%
26-40	13.39%	7.92%	18.36%	15.57%	44.48%	100.00%
41-65	46.22%	8.18%	20.12%	12.71%	13.19%	100.00%
65+	53.19%	7.27%	28.10%	5.28%	6.95%	100.00%
PRIMARY	42.84%		18.11%	13.48%	24.09%	100.00%
SECONDARY	34.72%	3.66%	14.16%	12.55%	34.58%	100.00%
TERTIARY	16.97%	8.49%	19.62%	16.16%	38.41%	100.00%

TABLE 6. PROBABILITIES OF CHANGING RESIDENCE BASED ON REGRESSION ANALYSIS FOR PILOT 4 (NETHERLANDS)

	The next six months					
Finland	**Very low probability**	**Low probability**	**Moderate probability**	**High probability**	**Very high probability**	**Total**
MALES	73.76%	11.72%	4.82%	5.38%	4.37%	100.00%
FEMALES	72.29%	9.46%	8.36%	4.89%	5.35%	100.00%
18-25	58.97%	16.74%	10.92%	8.33%	6.13%	100.00%
26-40	71.82%	10.84%	6.73%	6.14%	4.56%	100.00%
41-65	80.85%	7.45%	4.57%	2.15%	4.85%	100.00%
65+	90.07%	3.10%	1.88%	2.04%	2.91%	100.00%
PRIMARY	75.11%	10.38%	5.06%	2.93%	6.56%	100.00%
SECONDARY	79.10%	9.78%	3.91%	3.50%	3.66%	100.00%
TERTIARY	69.73%	10.98%	8.14%	6.10%	5.40%	100.00%

	The next 1-2 years					
Finland	**Very low probability**	**Low probability**	**Moderate probability**	**High probability**	**Very high probability**	**Total**
MALES	48.85%	20.70%	12.17%	8.07%	10.46%	100.00%
FEMALES	45.56%	19.13%	15.62%	9.55%	11.27%	100.00%
18-25	27.06%	16.27%	22.48%	17.48%	19.51%	100.00%
26-40	41.95%	23.94%	14.94%	9.25%	10.49%	100.00%
41-65	62.47%	17.78%	8.71%	3.57%	7.12%	100.00%
65+	78.12%	10.16%	3.98%	4.02%	3.97%	100.00%
PRIMARY	55.30%	17.20%	8.18%	6.63%	12.29%	100.00%
SECONDARY	57.16%	15.17%	10.49%	7.51%	9.29%	100.00%
TERTIARY	41.48%	22.50%	16.08%	9.65%	11.61%	100.00%

	The next 5 years					
Finland	**Very low probability**	**Low probability**	**Moderate probability**	**High probability**	**Very high probability**	**Total**
MALES	31.18%	17.98%	20.14%	14.92%	16.53%	100.00%
FEMALES	29.47%	16.91%	19.26%	16.37%	19.47%	100.00%
18-25	7.74%	13.24%	20.51%	23.64%	39.07%	100.00%
26-40	26.15%	18.51%	21.39%	17.67%	17.24%	100.00%
41-65	45.27%	18.25%	18.32%	9.73%	7.94%	100.00%
65+	63.60%	18.96%	9.11%	2.95%	6.07%	100.00%
PRIMARY	37.60%	15.10%	21.83%	9.46%	15.38%	100.00%
SECONDARY	36.45%	18.78%	17.32%	9.03%	17.82%	100.00%
TERTIARY	26.67%	16.88%	20.78%	19.47%	18.32%	100.00%

	The next 10 years					
Finland	**Very low probability**	**Low probability**	**Moderate probability**	**High probability**	**Very high probability**	**Total**
MALES	25.08%	13.39%	18.75%	14.31%	29.31%	100.00%
FEMALES	21.76%	12.13%	22.23%	11.64%	33.38%	100.00%
18-25	6.71%	6.76%	11.58%	20.48%	58.46%	100.00%
26-40	19.37%	14.91%	22.11%	12.03%	32.32%	100.00%
41-65	34.59%	13.57%	24.33%	10.48%	16.65%	100.00%
65+	55.94%	11.28%	17.28%	6.91%	9.03%	100.00%
PRIMARY	32.97%	12.99%	18.13%	12.74%	22.79%	100.00%
SECONDARY	29.02%	12.13%	20.70%	8.67%	29.02%	100.00%
TERTIARY	19.86%	13.05%	20.60%	15.15%	33.17%	100.00%

TABLE 7. PROBABILITIES OF CHANGING RESIDENCE BASED ON REGRESSION ANALYSIS FOR PILOT 5 (FINLAND)

4 INTEGRATION OF RELEVANT RRI PILLARS

MOBI-TWIN places the societal dimension of spatial mobility at the core of its research activities. It aligns the project's work with the pillars of Responsible Research and Innovation (RRI) — science education, gender equality, governance, open science, public engagement, and ethics — to ensure that outcomes, outputs, and impacts respond to the needs and values of society (MOBI-TWIN D4.1).

The RRI pillars mainstreamed into the activities described in this report are gender equality, open science, and ethics.

The integration of the gender dimension

In this deliverable, gender is included as a key constraint variable in all pilot areas in order to ensure that the synthetic population weights reproduce as closely as possible the actual gender distribution of each region. Furthermore, the spatial microsimulation models that we developed allow us to conduct analyses based on gender characteristics for mobility outcomes. This enables us to estimate probabilities of moving that are not biased because of gender (exclusion). In later stages of the project, we can further extend and build on that in order to examine in more detail how gender differences interact with mobility intentions and the twin transition, while also drawing on findings from earlier deliverables that have already highlighted the role of gender in mobility.

The integration of open science

As with other project deliverables, the outcomes of this report will be made openly available after approval by the European Commission. This includes the report itself, illustrative R scripts, and synthetic outputs generated by the spatial microsimulation models (in aggregated and anonymised form, where appropriate). All public deliverables, data products, and forthcoming scientific publications related to this work will be deposited in open-access repositories in line with the principles of open science. This ensures transparency, reproducibility, and accessibility for the wider research and policy communities.

The integration of ethics

The work presented in the current deliverable adheres to the ethical standards set out in the Grant Agreement, including Articles 13 (Confidentiality and Security), 14 (Ethics and Values), and 15 (Data Protection), with specific rules detailed in Annex 5. All survey data used in the microsimulation models originate from the MOBI-TWIN survey and EU-SILC, which were collected in compliance with the General Data Protection Regulation (GDPR). This guarantees that personal data are processed lawfully, stored securely, and anonymised prior to use in modelling.

5 Conclusions and Next Steps

This report presented the MOBITWIN model with a particular focus on conceptual, methodological and operationalisation issues as well as examples of model outputs. These outputs include maps of estimated population subgroups, which should be seen as a demonstrator and an example of the potential of the model for the next steps in terms of scenario analysis.

The model outputs presented here will be the basis for further static analysis with the use of EUROMOD, involving the estimation of welfare and fiscal impacts of alternative scenarios (also drawing on the findings of deliverable 3.2) as well as dynamic analysis and agent-based modelling, building on the frameworks developed so far.In particular, the MOBITWIN model can now be further developed and used to assess the social, economic and geographic impacts of the identified scenarios under Task 3.2. For example, the outputs of task 3.2 will be used to develop hypothetical local labour market scenarios (e.g. in terms of new jobs and types of skills and/or in terms of outmigration of particular demographic and socio-economic groups) which will then be analysed with the MOBITWIN model in order to estimate the fiscal and geographical impacts by each pilot region (and within regions, using the spatial microsimulation models).

In addition, the MOBITWIN model will be used to consider changes in spatial patterns of social and spatial inequalities and social and territorial cohesion between and within regions. For example and to that end the MOBITWIN model could be used to estimate inequality indicators before and after migration/mobility as well as the types of areas where migrants with different levels of earned income and other socio-economic attributes may relocate and the potential implications for social and territorial cohesion.

6 REFERENCES

Ajzen, I. (1985). From intentions to actions: A theory of planned behavior. In J. Kuhl & J. Beckmann (Eds.), Action Control: From Cognition to Behavior (pp. 11–39). Berlin Heidelberg: Springer-Verlag. [doi:10.1007/978-3-642-69746-3_2]

Antosz, P., Jager, W., Polhill, J. G., Salt, D., Scalco, A., Alonso-Betanzos, A., Sánchez-Maroño, N., Guijarro-Berdiñas, B., & Rodríguez, A. (2020). SMARTEES simulation implementations, deliverable 7.3. SMARTEES simulation implementations, deliverable 7.3. December [doi:10.18564/jasss.5168]

Ballas, D, Clarke, G P (2009), Spatial Microsimulation, in A.S Fotheringham and Peter A Rogerson (eds.) Handbook of Spatial Analysis, Sage, pp. 277-298

Ballas, D. (2004) Simulating trends in poverty and income inequality on the basis of 1991 and 2001 census data: a tale of two cities, Area 36.2, 146-163.

Ballas, D. & Clarke, G. P. (2000), GIS and microsimulation for local labour market policy analysis, Computers, Environment and Urban Systems, vol. 24, pp. 305-330.

Ballas D, Clarke, G P, Wiemers, E, (2005c) Building a dynamic spatial microsimulation model for Ireland, Population, Space and Place, vol. 11, 157–172.

Ballas, D., Clarke, G P, Dorling D, Rossiter, D. (2007), Using SimBritain to Model the Geographical Impact of National Government Policies, Geographical Analysis 39 (1), pp. 44-77.

Ballas, D., Clarke, G.P., Dorling, D., Eyre, H. and Rossiter, D., Thomas, B (2005a) SimBritain: a spatial microsimulation approach to population dynamics, Population, Space and Place, 11, 13–34.

Ballas, D., Rossiter, D., Thomas, B., Clarke, G., & Dorling, D. (2005b). Geography matters: Simulating the local impacts of national social policies. Geography matters: Simulating the local impacts of national social policies.

Bandura, A. (1962). Social Learning through Imitation. Lincoln, NE: University of Nebraska Press.

Bouman, L., Antosz, A., Jager, W., Polhill, J. G., Salt, D., Scalco, A., Alonso-Betanzos, A., Sánchez-Maroño, N., Guijarro-Berdiñas, B., & Rodríguez, A. (2021). Reports on scenario development and experiments for selected cases. SMARTEES project. Grant agreement No 763912, Deliverable 7.4; October 2021 [doi:10.18564/jasss.5168]

Chapuis, K., Taillandier, P., and Drogoul, A., (2022). Generation of Synthetic Populations in Social Simulations: A Review of Methods and Practices. Journal of

Artificial Societies and Social Simulation 25 (2) 6 <http://jasss.soc.surrey.ac.uk/25/2/6.html>. doi: 10.18564/jasss.4762

Cialdini, R. B., Kallgren, C. A., & Reno, R. R. (1991). A focus theory of normative conduct: A theoretical refinement and reevaluation of the role of norms in human behavior. Advances in Experimental Social Psychology, 24, 201–234. [doi:10.1016/s0065-2601(08)60330-5]

Clarke, G.P. (1996). Microsimulation for Urban and Regional Policy Analysis. Pion, London.

Festinger, L. (1957). A Theory of Cognitive Dissonance. Palo Alto, CA: Stanford University Press.

Jager, W., Alonso-Betanzos, A., Antosz, P., Bouman, L., Guijarro-Berdiñas, B., Polhill, J. G., Rodríguez-Arias, A., Salt, D., & Sánchez-Maroño, N. (2024). Simulating the role of norms in processes of social innovation: Three case studies. Journal of Artificial Societies and Social Simulation, 27(1), 6. [doi:10.18564/jasss.5168]

Jager, Wander, Antosz, Patrycja, Bouman, Loes, Li, Teng, Polhill, J. Gareth, Szczepanska, Timo and Wang, Shaoni (2025) 'HUMAT: An Integrated Framework for Modelling Individual Motivations, Social Exchange and Network Dynamics' Journal of Artificial Societies and Social Simulation 28 (1) 4 <http://jasss.soc.surrey.ac.uk/28/1/4.html>. doi: 10.18564/jasss.5611

Kavroudakis, D, Ballas, D, Birkin, M, (2013), A spatial microsimulation approach to the analysis of social and spatial inequalities in Higher Education attainment, Applied Spatial Analysis and Policy, vol. 3, pp. 1-23.

Li, T., & Jager, W. (2023). How availability heuristic, confirmation bias and fear may drive societal polarisation: An opinion dynamics simulation of the case of COVID-19 vaccination. Journal of Artificial Societies and Social Simulation, 26(4), 2. [doi:10.18564/jasss.5135]

Lovelace, R, Ballas, D (2013), `Truncate, replicate, sample': A method for creating integer weights for spatial microsimulation, Computers, Environment and Urban Systems, volume 41, pp. 1-11.

Maslow, A.H. (1954). Motivation and personality. New York, USA: Harper and Row.

Max-Neef, M. (1992). Development and human needs. In: P. Ekins & M. Max-Neef (Eds.), Real-life economics: Understanding wealth creation (pp. 197-213). London, New York: Routledge.

McCann, P. (2013). Modern Urban and Regional Economics. Oxford University Press.

Mitton, L, Sutherland, H. and Weeks, M. (eds) (2000) Microsimulation modelling for Policy Analysis: Challenges and Innovations. Cambridge: Cambridge University Press

Orcutt G (1957) A new type of socio-economic system. The Review of Economics and Statistics 39(2):116-123

Panori, A., Ballas, D., Psycharis, Y. (2017). SimAthens: A spatial microsimulation approach to the estimation and analysis of small area income distributions and poverty rates in the city of Athens, Greece, Computers, Environment and Urban Systems, Volume 63, 2017, Pages 15-25, ISSN 0198-9715

Pavlov, I. P. (1927). Conditioned Reflexes. Oxford: Oxford University Press.

Petty, R. E., & Cacioppo, J. T. (1986). Communication and Persuasion: Central and Peripheral Routes to Attitude Change. Berlin Heidelberg: Springer.

Redmond, G., Sutherland, H. and Wilson, M. (1998) The Arithmetic of Tax and Social Security Reform: A User's Guide to Microsimulation Methods and Analysis. Cambridge: Cambridge University Press

Schelling, T. C. (1971). Dynamic models of segregation. Journal of mathematical sociology, 1(2), 143-186.

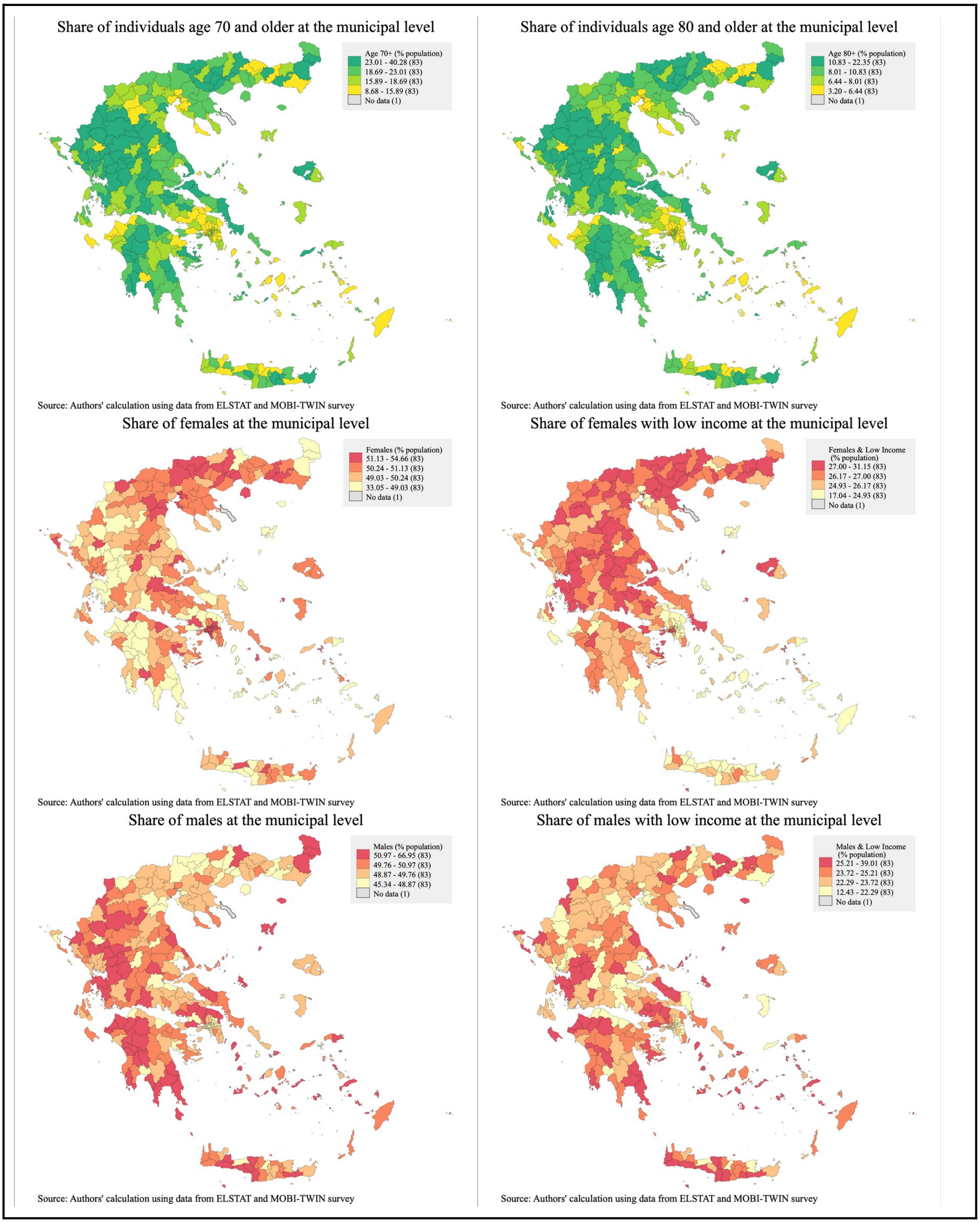
Skinner, B. F. (1953). The Behavior of Organisms. New York, NY: Appleton-Century-Crofts.

Sutherland, H., & Figari, F. (2013). EUROMOD: the European Union tax-benefit microsimulation model. International journal of microsimulation, 6(1), 4-26.

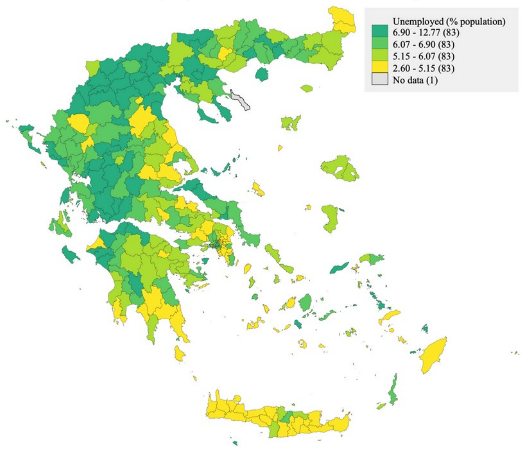
Williamson, P., Birkin, M. and Rees, P. (1998) 'The estimation of population microdata by using data from small area statistics and samples of anonymised records', Environment and Planning A, Vol. 30, pp. 785–816

Wilson, A. G., & Pownall, C. E. (1976). A new representation of the urban system for modelling and for the study of micro-level interdependence. Area, 246-254.

ANNEX 1: TITLE OF ANNEX

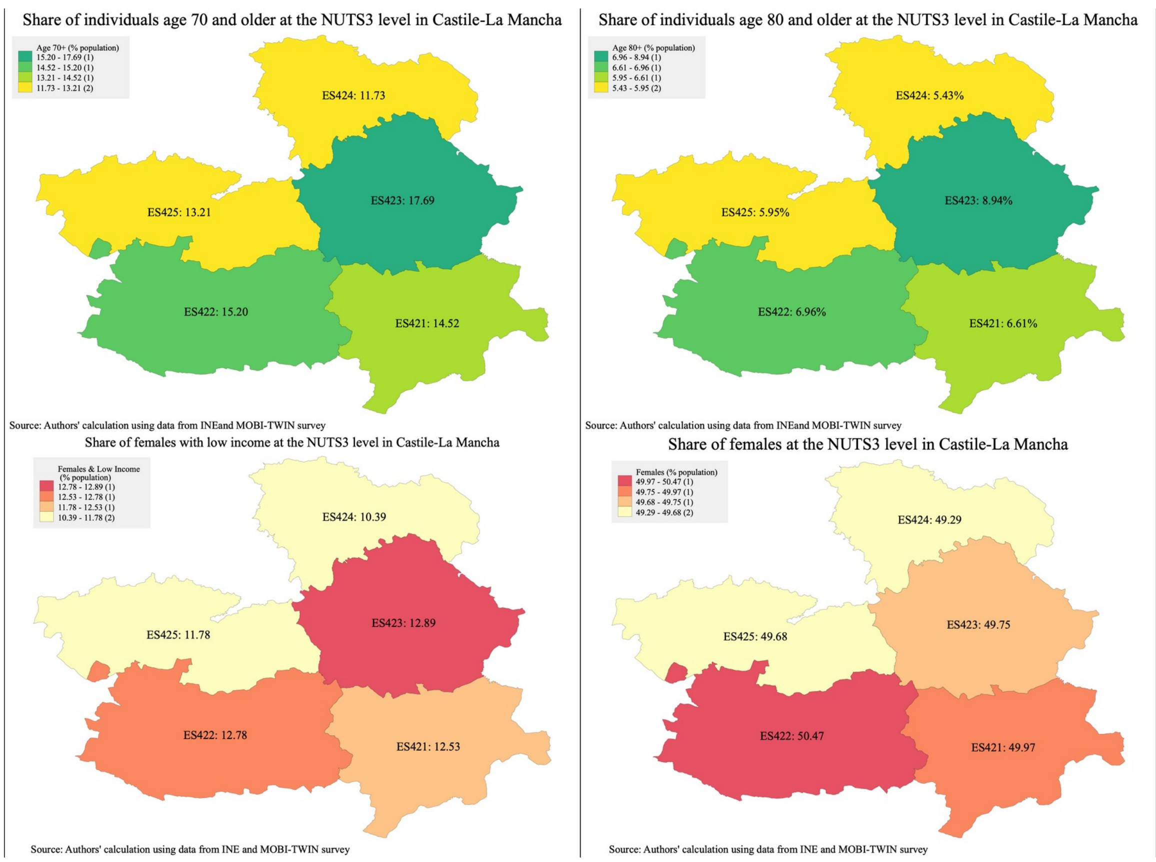

Share of individuals age 70 and older at the municipal level
Age 70+ (% population)
23.01 - 40.28 (83)
18.69 - 23.01 (83)
15.89 - 18.69 (83)
8.68 - 15.89 (83)
No data (1)
Source: Authors' calculation using data from ELSTAT and MOBI-TWIN survey
Share of individuals age 80 and older at the municipal level
Age 80+ (% population)
10.83 - 22.35 (83)
8.01 - 10.83 (83)
6.44 - 8.01 (83)
3.20 - 6.44 (83)
No data (1)
Source: Authors' calculation using data from ELSTAT and MOBI-TWIN survey
Share of females at the municipal level
Females (% population)
51.13 - 54.66 (83)
50.24 - 51.13 (83)
49.03 - 50.24 (83)
33.05 - 49.03 (83)
No data (1)
Source: Authors' calculation using data from ELSTAT and MOBI-TWIN survey
Share of females with low income at the municipal level
Females & Low Income (% population)
27.00 - 31.15 (83)
26.17 - 27.00 (83)
24.93 - 26.17 (83)
17.04 - 24.93 (83)
No data (1)
Source: Authors' calculation using data from ELSTAT and MOBI-TWIN survey
Share of males at the municipal level
Males (% population)
50.97 - 66.95 (83)
49.76 - 50.97 (83)
48.87 - 49.76 (83)
45.34 - 48.87 (83)
No data (1)
Source: Authors' calculation using data from ELSTAT and MOBI-TWIN survey
Share of males with low income at the municipal level
Males & Low Income (% population)
25.21 - 39.01 (83)
23.72 - 25.21 (83)
22.29 - 23.72 (83)
12.43 - 22.29 (83)
No data (1)
Source: Authors' calculation using data from ELSTAT and MOBI-TWIN survey

Share of unemployed individuals at the municipal level



Source: Authors' calculation using data from ELSTAT

TABLE A1.1 SOCIO-DEMOGRAPHIC PROFILE OF GREECE



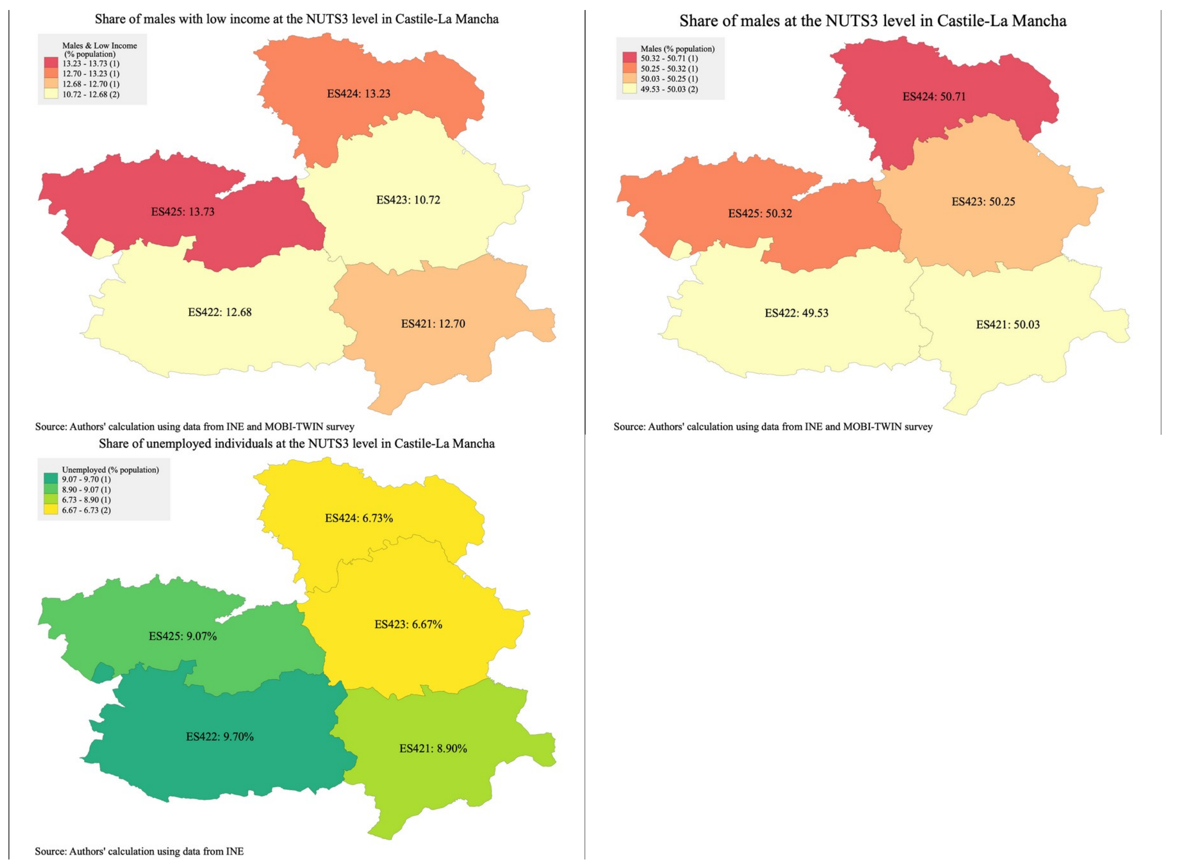
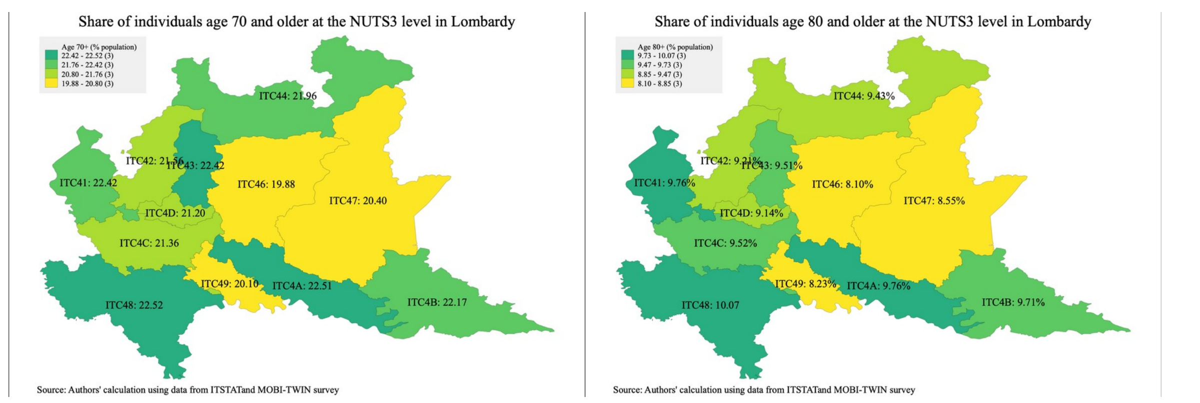


TABLE A1.2 SOCIO-DEMOGRAPHIC PROFILE OF CASTILE-LA MANCHA



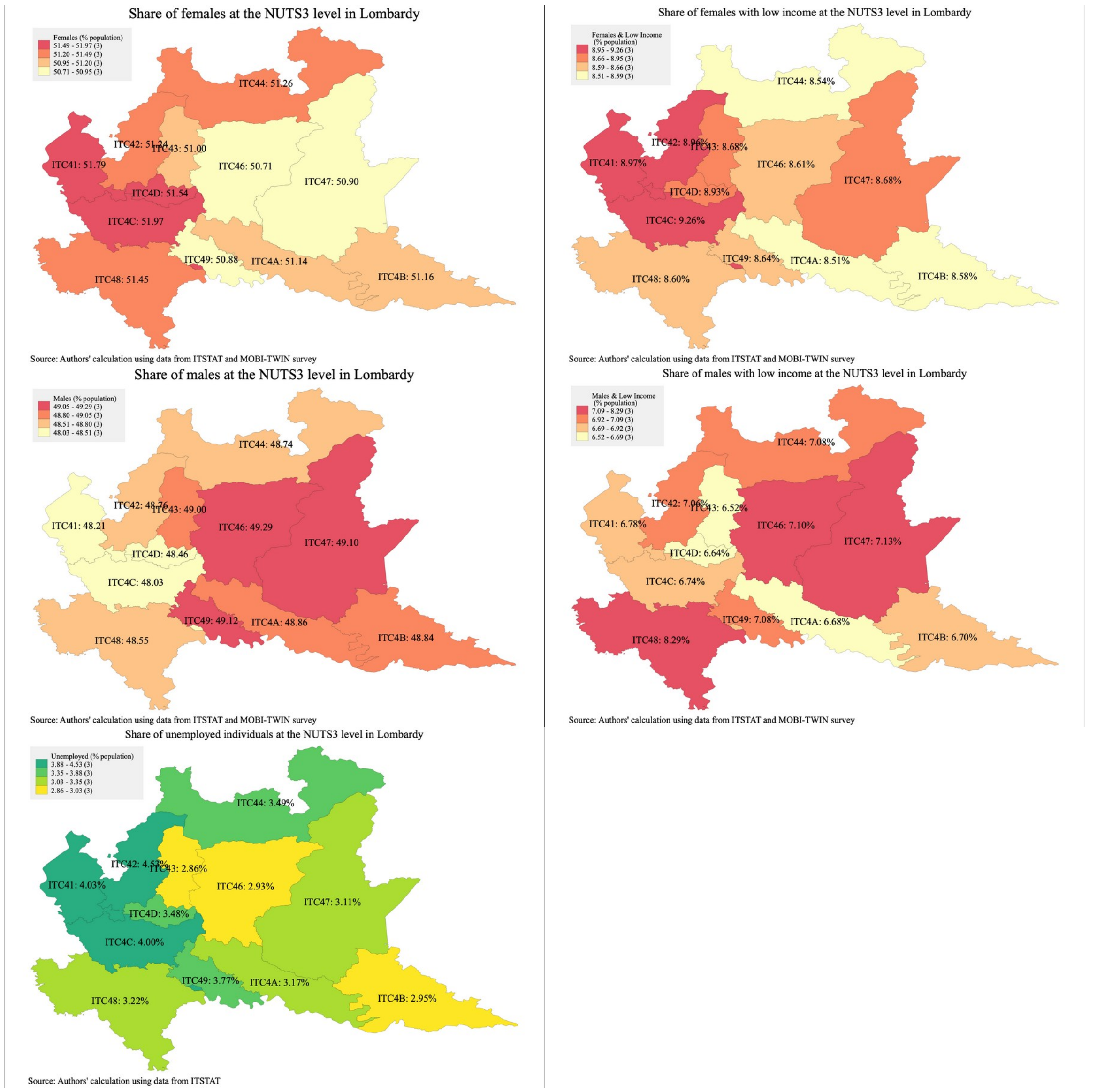
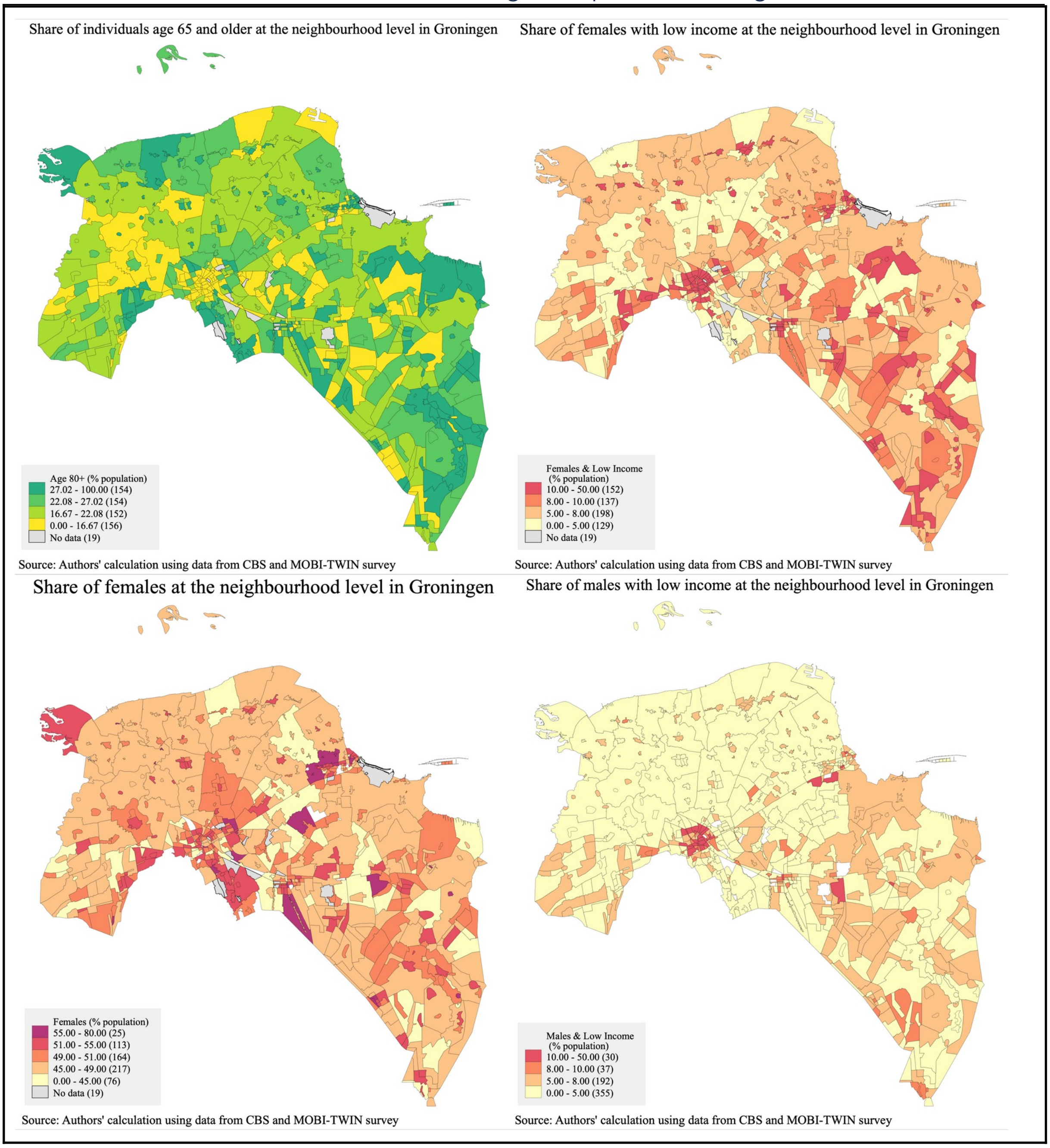


TABLE A1.3 SOCIO-DEMOGRAPHIC PROFILE OF LOMBARDY


Share of individuals age 65 and older at the neighbourhood level in Groningen
Age 80+ (% population)
27.02 - 100.00 (154)
22.08 - 27.02 (154)
16.67 - 22.08 (152)
0.00 - 16.67 (156)
No data (19)
Source: Authors' calculation using data from CBS and MOBI-TWIN survey
Share of females with low income at the neighbourhood level in Groningen
Females & Low Income (% population)
10.00 - 50.00 (152)
8.00 - 10.00 (137)
5.00 - 8.00 (198)
0.00 - 5.00 (129)
No data (19)
Source: Authors' calculation using data from CBS and MOBI-TWIN survey
Share of females at the neighbourhood level in Groningen
Females (% population)
55.00 - 80.00 (25)
51.00 - 55.00 (113)
49.00 - 51.00 (164)
45.00 - 49.00 (217)
0.00 - 45.00 (76)
No data (19)
Source: Authors' calculation using data from CBS and MOBI-TWIN survey
Share of males with low income at the neighbourhood level in Groningen
Males & Low Income (% population)
10.00 - 50.00 (30)
8.00 - 10.00 (37)
5.00 - 8.00 (192)
0.00 - 5.00 (355)
Source: Authors' calculation using data from CBS and MOBI-TWIN survey

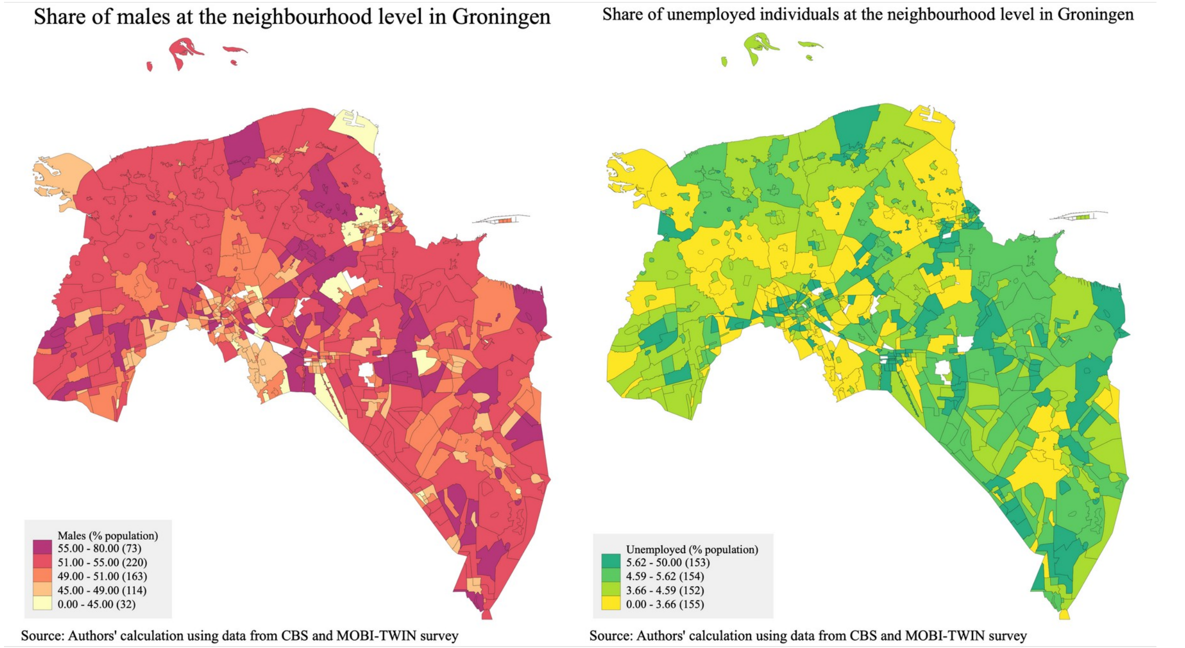
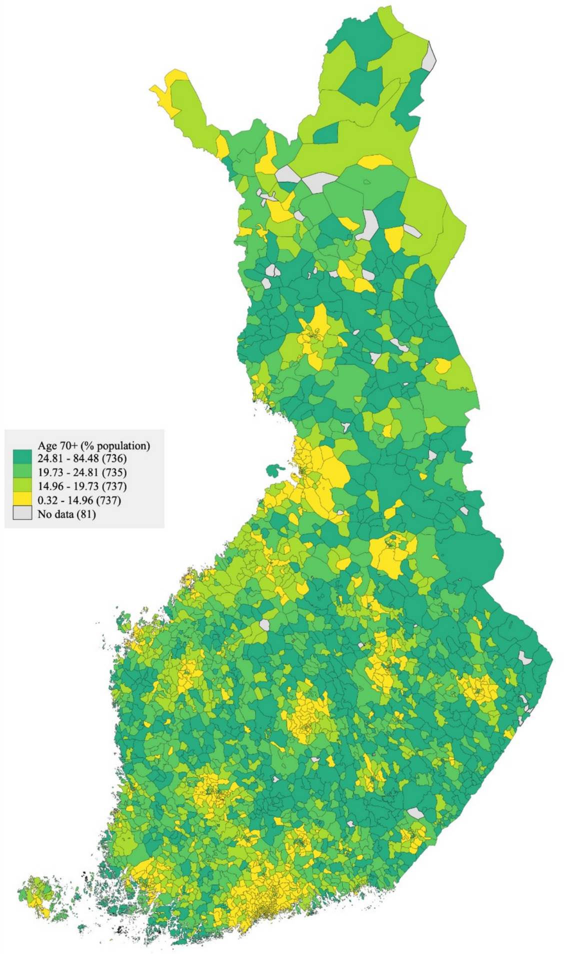


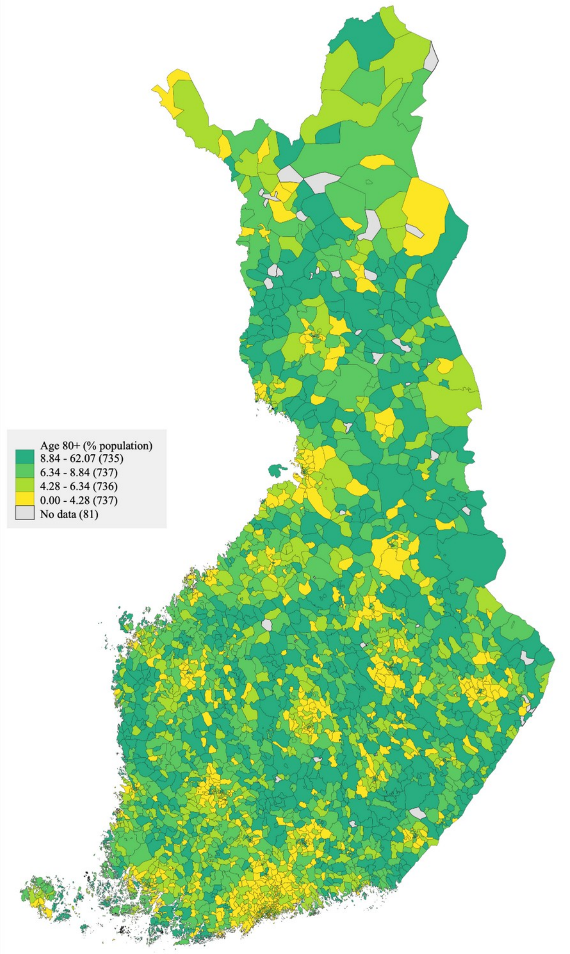
TABLE A1.4 SOCIO-DEMOGRAPHIC PROFILE OF GRONINGEN

Share of individuals age 70 and older at the postal code level

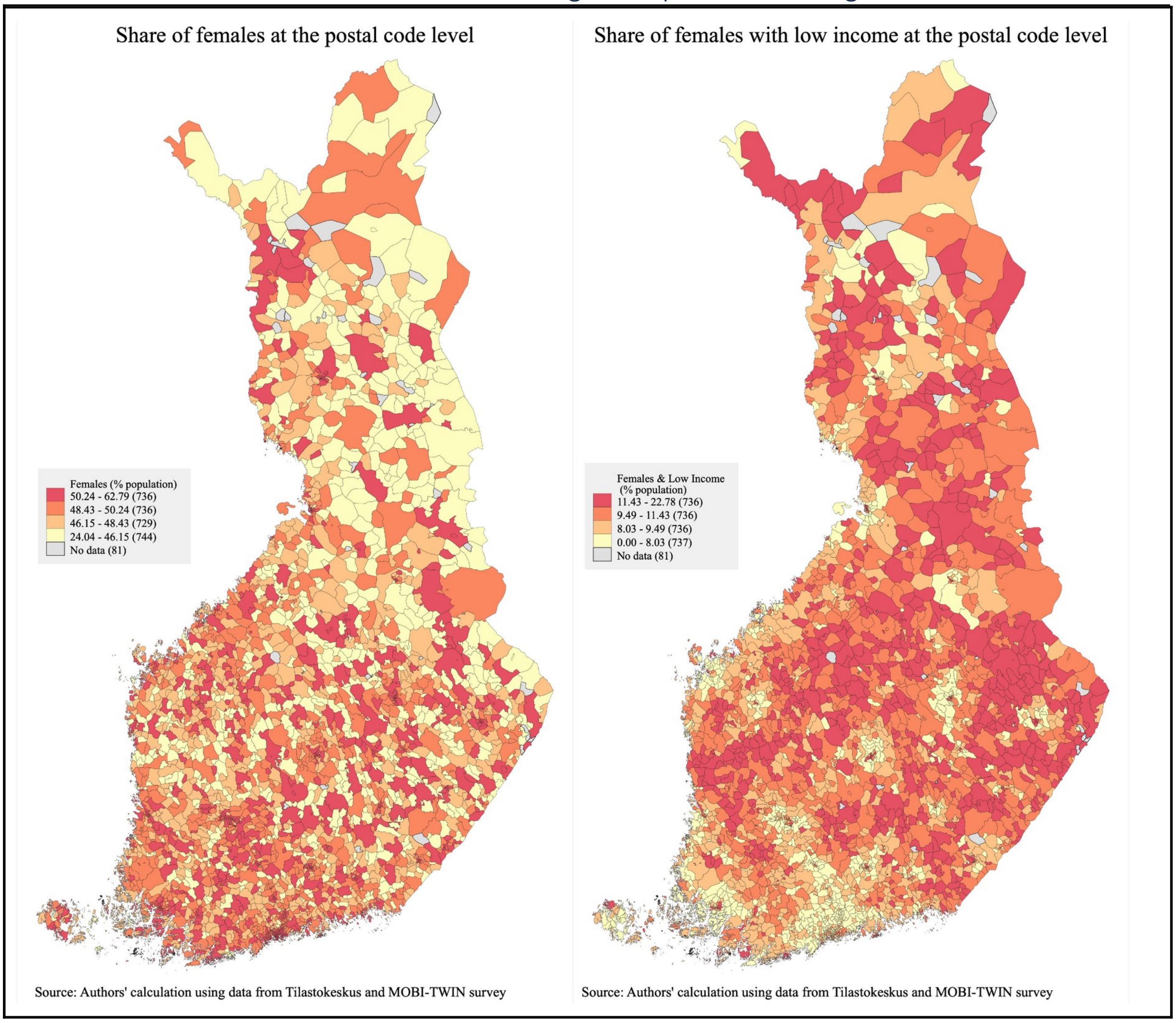


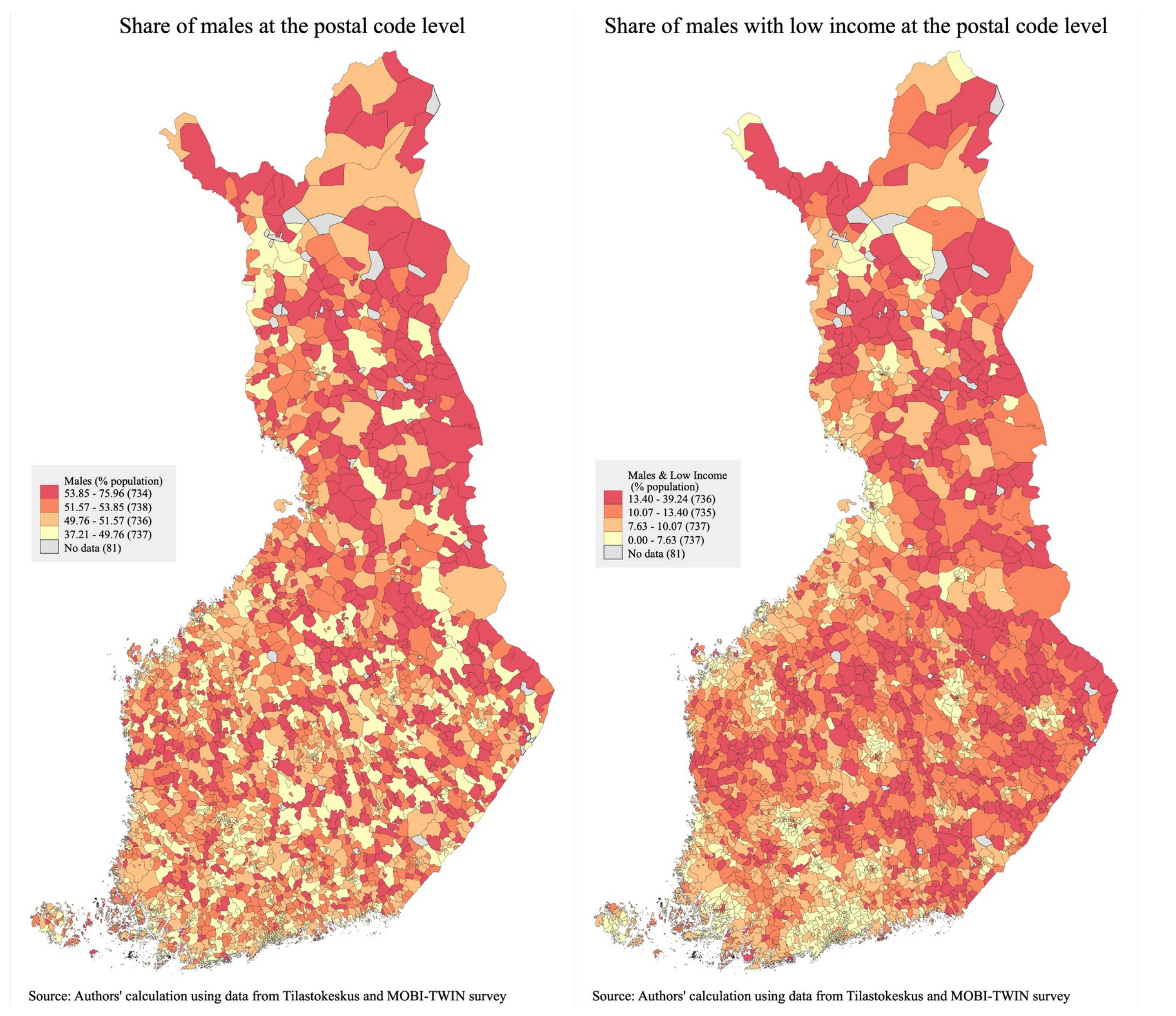
Source: Authors' calculation using data from Tilastokeskus and MOBI-TWIN survey

Share of individuals age 80 and older at the postal code level

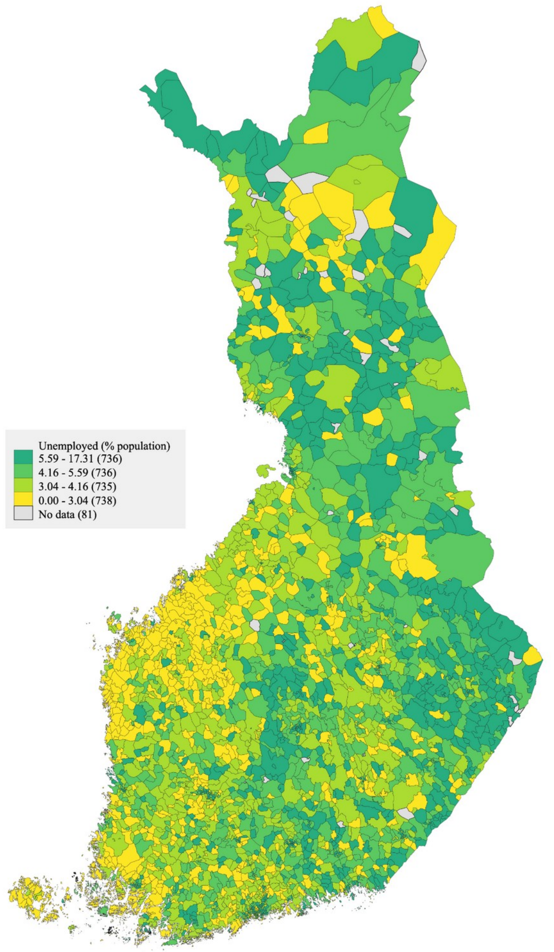


Source: Authors' calculation using data from Tilastokeskus and MOBI-TWIN survey


Share of females at the postal code level
Females (% population)
50.24 - 62.79 (736)
48.43 - 50.24 (736)
46.15 - 48.43 (729)
24.04 - 46.15 (744)
No data (81)
Source: Authors' calculation using data from Tilastokeskus and MOBI-TWIN survey
Share of females with low income at the postal code level
Females & Low Income (% population)
11.43 - 22.78 (736)
9.49 - 11.43 (736)
8.03 - 9.49 (736)
0.00 - 8.03 (737)
No data (81)
Source: Authors' calculation using data from Tilastokeskus and MOBI-TWIN survey


Share of males at the postal code level
Males (% population)
53.85 - 75.96 (734)
51.57 - 53.85 (738)
49.76 - 51.57 (736)
37.21 - 49.76 (737)
No data (81)
Source: Authors' calculation using data from Tilastokeskus and MOBI-TWIN survey
Share of males with low income at the postal code level
Males & Low Income
(% population)
13.40 - 39.24 (736)
10.07 - 13.40 (735)
7.63 - 10.07 (737)
0.00 - 7.63 (737)
No data (81)
Source: Authors' calculation using data from Tilastokeskus and MOBI-TWIN survey

Share of unemployed individuals at the postal code level



Source: Authors' calculation using data from Tilastokeskus

TABLE A1.5 SOCIO-DEMOGRAPHIC PROFILE OF FINLAND

ANNEX 2: PROBABILITIES OF CHANGING RESIDENCE

	<2 years horizon				5+ years horizon			
Spain	short_mob	short_unsure	short_immob	Total	long_mob	long_unsure	long_immob	Tot
MALES	5.20%	17.20%	77.60%	100.00%	24.80%	22.40%	52.80%	100.
FEMALES	9.68%	15.25%	75.07%	100.00%	23.17%	21.41%	55.43%	100.
18-25	14.08%	32.39%	53.52%	100.00%	52.11%	35.21%	12.68%	100.
26-40	12.12%	20.20%	67.68%	100.00%	27.78%	24.24%	47.98%	100.
41-65	3.92%	10.46%	85.62%	100.00%	15.69%	16.99%	67.32%	100.
65+	0.00%	0.00%	100.00%	100.00%	6.25%	25.00%	68.75%	100.
PRIMARY	10.00%	16.67%	73.33%	100.00%	23.33%	13.33%	63.33%	100.
SECONDARY	6.64%	11.85%	81.52%	100.00%	18.01%	19.91%	62.09%	100.
TERTIARY	8.29%	18.57%	73.14%	100.00%	27.43%	23.71%	48.86%	100.

	<2 years horizon				5+ years horizon			
Finland	short_mob	short_unsure	short_immob	Total	long_mob	long_unsure	long_immob	Tot
MALES	6.75%	11.51%	81.75%	100.00%	20.63%	28.17%	51.19%	100.
FEMALES	8.88%	13.82%	77.30%	100.00%	26.97%	27.96%	45.07%	100.
18-25	12.90%	33.87%	53.23%	100.00%	54.84%	32.26%	12.90%	100.
26-40	8.54%	15.85%	75.61%	100.00%	29.27%	26.83%	43.90%	100.
41-65	7.41%	8.15%	84.44%	100.00%	17.04%	28.15%	54.81%	100.
65+	3.33%	3.33%	93.33%	100.00%	10.00%	26.67%	63.33%	100.
PRIMARY	7.41%	11.11%	81.48%	100.00%	25.93%	33.33%	40.74%	100.
SECONDARY	6.62%	11.26%	82.12%	100.00%	20.53%	27.15%	52.32%	100.
TERTIARY	10.00%	15.50%	74.50%	100.00%	29.00%	28.00%	43.00%	100.

	<2 years horizon				5+ years horizon			
Greece	short_mob	short_unsure	short_immob	Total	long_mob	long_unsure	long_immob	Tot
MALES	9.50%	12.93%	77.57%	100.00%	27.70%	28.50%	43.80%	100.
FEMALES	9.56%	22.13%	68.31%	100.00%	35.79%	23.22%	40.98%	100.
18-25	19.80%	28.71%	51.49%	100.00%	66.34%	22.77%	10.89%	100.
26-40	9.48%	21.57%	68.95%	100.00%	35.95%	29.41%	34.64%	100.
41-65	6.50%	10.84%	82.66%	100.00%	17.96%	24.15%	57.89%	100.
65+	6.67%	0.00%	93.33%	100.00%	6.67%	13.33%	80.00%	100.
PRIMARY	0.00%	18.75%	81.25%	100.00%	18.75%	31.25%	50.00%	100.
SECONDARY	8.16%	16.33%	75.51%	100.00%	27.04%	25.00%	47.96%	100.
TERTIARY	10.32%	17.82%	71.86%	100.00%	33.77%	26.08%	40.15%	100.

	<2 years horizon				5+ years horizon			
Italy	short_mob	short_unsure	short_immob	Total	long_mob	long_unsure	long_immob	Tot
MALES	4.12%	13.56%	82.32%	100.00%	22.03%	20.82%	57.14%	100.
FEMALES	4.95%	16.22%	78.83%	100.00%	24.32%	21.85%	53.83%	100.
18-25	8.26%	24.77%	66.97%	100.00%	57.80%	31.19%	11.01%	100.
26-40	8.20%	23.05%	68.75%	100.00%	33.20%	28.52%	38.28%	100.
41-65	2.40%	10.13%	87.47%	100.00%	13.07%	18.13%	68.80%	100.
65+	0.00%	3.42%	96.58%	100.00%	1.71%	6.84%	91.45%	100.
PRIMARY	3.41%	9.09%	87.50%	100.00%	14.77%	11.36%	73.86%	100.
SECONDARY	2.87%	10.44%	86.68%	100.00%	20.10%	20.10%	59.79%	100.
TERTIARY	6.48%	20.73%	72.80%	100.00%	28.24%	24.87%	46.89%	100.

	<2 years horizon					5+ years horizon			
Netherlands	short_mob	short_unsure	short_immob	Total		long_mob	long_unsure	long_immob	Tot
MALES	5.41%	9.46%	85.14%	100.00%		21.62%	22.97%	55.41%	100.
FEMALES	11.89%	14.69%	73.43%	100.00%		32.87%	25.87%	41.26%	100.
18-25	23.33%	30.00%	46.67%	100.00%		63.33%	26.67%	10.00%	100.
26-40	14.06%	14.06%	71.88%	100.00%		42.19%	28.13%	29.69%	100.
41-65	5.62%	10.11%	84.27%	100.00%		17.98%	21.35%	60.67%	100.
65+	0.00%	2.94%	97.06%	100.00%		2.94%	26.47%	70.59%	100.
PRIMARY	0.00%	0.00%	100.00%	100.00%		16.67%	16.67%	66.67%	100.
SECONDARY	4.72%	8.49%	86.79%	100.00%		12.26%	27.36%	60.38%	100.
TERTIARY	15.24%	18.10%	66.67%	100.00%		46.67%	22.86%	30.48%	100.

TABLE A2. PROBABILITIES OF CHANGING RESIDENCE BY GROUPING TIME HORIZONS TO BELOW TWO YEARS AND ABOVE FIVE YEARS.

MOBI-TWIN

TWIN TRANSITION AND CHANGING PATTERNS OF SPATIAL MOBILITY: A REGIONAL APPROACH

The Partners of the MOBI-TWIN Consortium:

Name	Country

White Research SRL	Belgium
Aristotelio Panepistimio Thessalonikis	Greece
Rijksuniversiteit Groningen	The Netherlands
Universitat de Barcelona	Spain
Helsingin Yliopisto	Finland
Fondation Europeene de la Science	France
Politecnico di Milano	Italy
Erevnitiko Panepistimiako Institouto Periferiakis Anaptixis	Greece
Acceligence Ltd.	Cyprus

CONNECT WITH MOBI-TWIN FOR MORE UPDATES:

LinkedIn, Twitter, Website