Unay-Gailhard, İlkay; Baqueiro-Espinosa, Omar

Article — Published Version

Adaptation of a microsimulation model at the municipality level: demographic and employment evolution in the Altmark region of Germany

SIMULATION

Applications

Simulation

Simulation: Transactions of the Society for Modeling and Simulation International
2015, Vol. 91(12) 1088–1103

DOI: 10.1177/0037549715615680
sim.sagepub.com
SAGE SCS

Adaptation of a microsimulation model at the municipality level: demographic and employment evolution in the Altmark region of Germany

İlkay Unay-Gailhard and Omar Baqueiro-Espinosa

Abstract
Adapting microsimulation models to specific regions is a challenging task. This is especially true when aiming to simulate very low regional levels, such as municipalities, owing to anonymity restrictions or simply nonexistence of data at the required levels. In this study, we present a process to define the dynamics of the PRototypical policy Impact on Multifunctional Activities in rural municipalities (PRIMA) conceptual microsimulation model for the Hohenberg–Krusemark municipality network in the Altmark region of Germany. A parameterization process and the prototypical simulation results of the model are given along with a discussion on the validation of the simulation results by local level stakeholders. The presented technical aspects of the model provide the evolution of the demographic structure and the employment status for individuals living in a set of interconnected municipalities in rural areas. Additional insight was gained by comparing the demographic and employment properties of the simulation results with real data. The comparison allows for the observation of how correctly the model replicates past trends in the municipality network. The model's assumptions are based on the results of a probability table created for the years 2000–2009, and the simulation results are presented for the years 2000–2020.

Keywords
agent-based simulation, demographic and employment evolution, Germany, rural areas, municipality

1. Introduction

Policy implementation for rural areas is of high importance within the European Union (EU). From the discussions on the general Common Agricultural Policy (CAP) reforms to changes in how rural policies are implemented, many measures are migrating from top-bottom to bottom-up approaches. The current observed trend is towards the adoption of a bottom-up approach for policy implementation.[1] That is, the participation of local stakeholders in the formulation and implementation of policy programs and projects is becoming more important.

Parallel to this observed trend, there is an increasing demand for the availability of tools that allow for the testing of potential policies at the local level. In our study, this demand is addressed within the framework of the PRIMA (PRototypical policy Impact on Multifunctional Activities in rural municipalities) EU project,[2] where a strongly data-driven microsimulation model has been developed and described in the study of Lenormand et al.[3]

The main objective of this paper is to present the parameterization process and prototypical simulation results of the PRIMA microsimulation model along with a discussion of the simulation results' validation by local level stakeholders. The presented technical aspects of the PRIMA model give the evolution of rural municipalities in terms of the structure of populations and employment under different policies. Previous work presented by Huet and Deffuant[4] and Lenormand et al.[3] described the conceptual PRIMA model and offered an adapted version for a rural

Department of Structural Development of Farms and Rural Areas, Leibniz Institute of Agricultural Development in Transition Economies (IAMO), Germany

Corresponding author:
İlkay Unay-Gailhard, Department of Structural Development of Farms and Rural Areas, Leibniz Institute of Agricultural Development in Transition Economies (IAMO), Theodor-Lieser-Str.2, D-06120 Halle (Saale), Germany.
Email: unaygailhard@iamo.de

area within France. The present paper builds on that work and describes the adaptation of the PRIMA model to the Altmark region in Germany.

Microsimulation models have had increasingly used in social sciences for more than 10 years.[5] These types of models focus on the representation of individuals (such as persons or households) and on the development of the system they inhabit. Such development is represented as probabilistic rules that are applied to select the actions taken by the modeled individuals. Microsimulation models are related to agent-based models (ABM) in that their aim is to observe the behavior of a complete system as a result of the micro behavior of individuals. Indeed, the fine line separating ABM and microsimulation models may be only the agents' decision-making mechanism (behavior rules in ABM or probability distributions in microsimulation). The high data requirement for these types of models presents a challenge when trying to adapt a conceptual model to different regions. This is especially true when modeling low-level geographical units, such as municipalities, owing to anonymity restrictions or simply nonexistence of data at the required levels.

Our paper aims to illustrate the application of a microsimulation model at the municipality level by providing an initial concept of the PRIMA model and by giving an overview on the adoption process. These two base elements of our model are presented in Section 2, with special regard to the composition of individuals' properties and their transitions among economic status. In Section 3, the analysis of a selection of data extracted from the German Labor force Survey (LFS)[6] and used for the adaptation of the model is described in detail within the parameterization process, including the discussion of insights gained after analyzing the data. Section 4 gives an overview of the study region of Altmark and presents the simulation results using demographic and employment trends between the simulated and collected data. Finally, a concluding section is offered, where the lessons learnt and possible improvements for the microsimulation adaptation are noted.

2. Conceptual PRIMA microsimulation model

The PRIMA model simulates virtual individuals living in a set of interconnected municipalities in a rural area. The dynamics of the model include demographic change (such as births, deaths, marriage, migration, and divorce), economic status change (such as being a student, a worker, unemployed, and in retirement), and change of jobs (between a set of defined job types). The simulation evolves at yearly steps, with a starting year of 2000.

2.1. Structure of PRIMA model adaptation in Germany

Figure 1 depicts the major elements that compose the model as well as its structure. The main unit of the model is the individual, which represents a person living in the

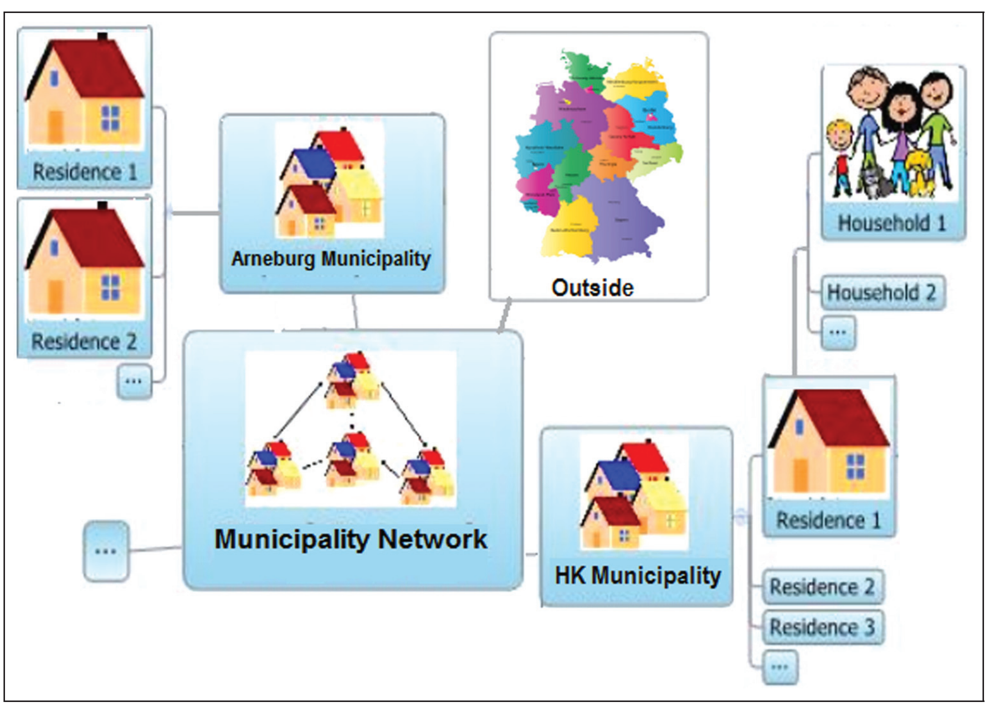

Figure 1. Main structure of the model. HK: Hohenberg–Krusemark.

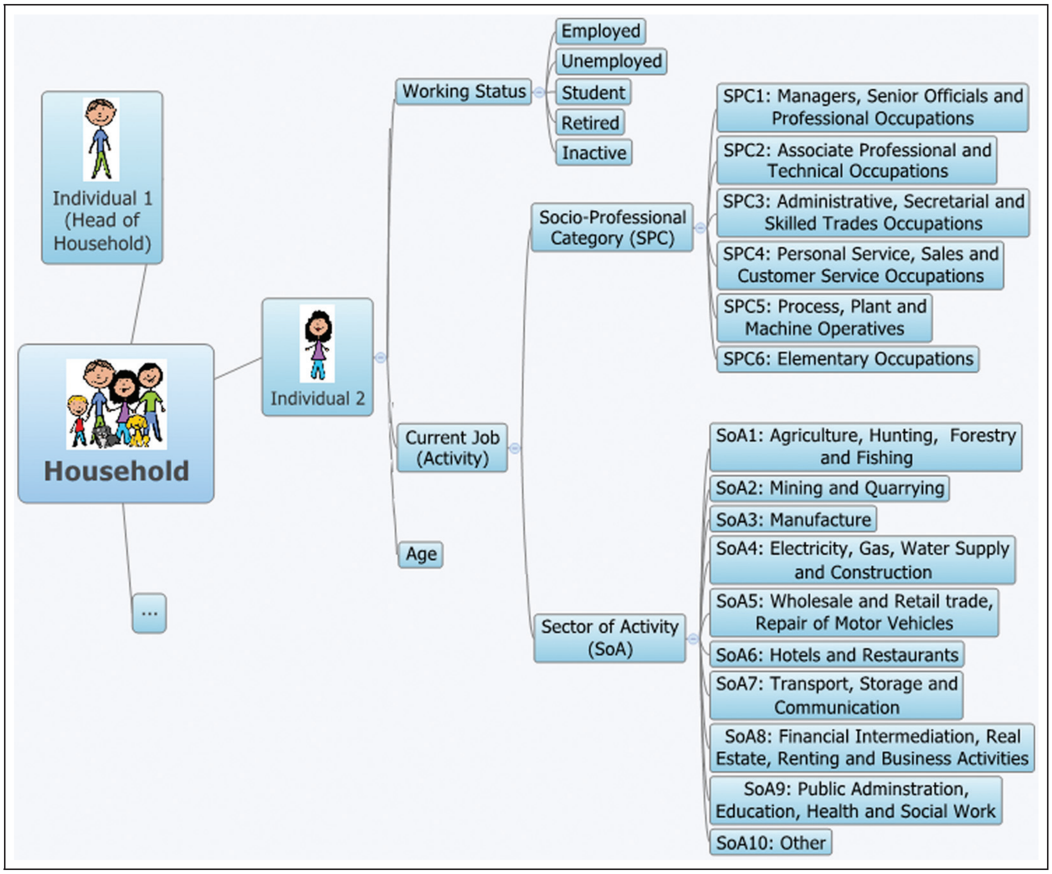

Figure 2. Composition of a household and of individual's properties in the Altmark region.

modeled municipality. Municipalities are composed of residences. A group of individuals form a household, and a set of zero or more households can occupy a residence. A connected group of municipalities is considered as a municipality network, which defines the region of study. A special type of municipality, called the outside municipality, is used to represent locations that are not considered as part of the studied municipality network but that have some effect on the development of the network.

Each individual is characterized by a set of properties, such as age, economic status, and current activity. An activity is defined by the socio-professional categories (SPCs) and the sector of activity in which the individual is working. An example of such a category for the Altmark region adaptation is presented in Figure 2.

The dynamics of the simulations are mainly driven by the decisions made by each individual. These decisions concern decisions made at the individual level related to search for a job; entering the labor market; changing jobs; changes in working status, such as retiring; becoming inactive; or becoming unemployed. Additionally, individuals perform non-job-related actions, such as searching for a partner and household evolution (splits, quitting parenting, death, etc.).

Because each transition is driven by probabilistic decisions, adapting the model to a particular region requires the specification of each probability value (a distribution or a single value) relevant to the adapted region. The following section describes the adaptation process for a particular set of transitions in the model.

2.2. Microsimulation adaptation process

Adapting the PRIMA model to a particular region involved three main steps: first, the specification of input parameters defining the initial state of the simulations; second, defining the probability distributions that drive the dynamics during the simulation; and last, including any additional perturbation (or external event) that could be considered during the simulation time but that is not part of the model dynamics.

In this paper, we describe the process for deriving the probability distributions used as drivers of the simulation. The dynamics of the PRIMA model are mainly related to individuals living in the region. Such dynamics are transitions between individuals' states occurring at each time step, as observed in Figure 3.

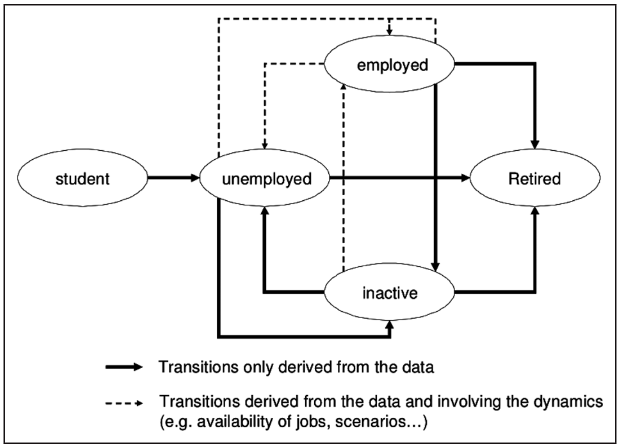

Figure 3. Individual transitions among economic status.
Source: Huet and Deffuant, 2011.

Some of these transitions (denoted by a continuous line in the figure) are fully driven by a probability distribution that characterizes when the transitions occur. Other transitions (denoted by a dotted line in the figure) also depend on the state of the model at the time when the decision is made. As an example, an individual transition from employment to retirement would only be ruled by the probability distribution, whereas a transition from unemployment to employment is partially ruled by a probability distribution but depends on the availability of jobs. The definition of these probability distributions must both have a meaningful interpretation (they must make sense in the context of the decision being made) and be obtainable from the available regional data.

In addition to the transitions in economic status, supplementary probability distributions must be specified for a complete definition of the model dynamics. The complete list of probabilities is shown in Table 1. The table provides the name of a given transition rule and a description of the use of the rule. Not every transition is driven by a stochastic process; for example, the aging rule is deterministic throughout the simulation as every individual becomes older as the simulation moves forward.

The PRIMA model contains additional dynamics not considered here. These dynamics (e.g., ''change of head of household'' or ''merge'' two households) are derived either from some of the distributions mentioned here or from parameters defined during a calibration phase (owing to the lack of statistical data for the parameterization). The detailed descriptions of those dynamics are beyond the scope of this paper.

Each distribution is defined as a probability table containing the choices available for each dynamic (e.g., the available SPC for the previous year student dynamics) given one or more independent variables. The distributions were derived from the LFS. The tables were mainly derived using yearly data at the country level. A separate analysis considering only individuals living in rural areas was performed. This allowed us to determine if the probabilities derived from rural areas differed significantly from the complete dataset.

3. Parameterization process

3.1. Deriving dynamics from the German Labor Force Survey

3.1.1. The concept of socio-professional categories. The SPCs can be classified in several ways. In our research, we use eight categories that correspond to the one digit occupational classification of the International Standard Classification of Occupations (ISCO). The classified occupations are legislators and managers; professionals; technicians; clerical workers; service sector workers; skilled agricultural, forestry and fishery workers; craft workers; machine operators; and elementary occupation workers. This categorization allows us to differentiate individuals by their level of skill. In addition to these occupations, we categorize individuals by their labor force status, such as employed, non-employed and student.

3.1.2. Labor supply in rural Germany. The sample includes individuals living in rural areas, who reported their labor status as employed, unemployed or a student in the previous year and were employed in the current year. Employed individuals and students may also become unemployed. To include only individuals living in rural areas, we considered the LFS variable DEGURBA, which groups individual residences into three classes: densely populated, intermediate and thinly populated area. By considering this classification, we analyzed only the individuals with residence on the ''thinly populated area'' group. This class is defined as local areas with less than 100 km^2 and with a population density of less than 100 inhabitants per km^2.[7]

3.1.3. Probability of working in one occupation. To observe an individual's participation in one profession, our analysis focuses on the transition trends grouped by SPC. For each of these transitions, a probability table was derived from the LFS data. From the model perspective, these probability tables describe the probability that an individual looks for a job in a given SPC. This must be differentiated from the probability that an individual obtains a job in a given SPC. The former is defined by the willingness of change, whereas the latter reflects the actual changes that occurred in the region. This contrast is important for the derivation of the data because the LFS captures only the actual changes between years. For this reason, additional assumptions were made to use the LFS data to construct tables to use as job searching probabilities for the model.

Table 1. List of modelled individual dynamics.

Age increment (in 1 year).
Age at which the individual will die.
Age at which the individual will become a student
Age at which the individual will enter the job market
Age at which the individual will retire.
Choice of Socio-professional Category (SPC) that a person who simply stopped being a student will look for
Choice of SPC that an unemployed individual will look for
Choice of SPC an employed individual will look for (includes the probability of an employed individual looking for a job)
If an individual will become unemployed in this year.
If an individual will become inactive if currently employed
If an individual will become inactive if currently unemployed
If an individual will stay inactive if currently Inactive
An individual creates a new household
An individual moves to another residence
An individual joins a new partner
An individual splits from a “couple” household
An individual changes residence to another municipality
An individual changes residence to another municipality

3.2. Results of probability tables for rural Germany

The probability tables are summarized by age groups and SPCs for the study years 2002 to 2009. The extraction of the data from the LFS was performed using the R statistical analysis program (R Development Team, 2010). The utilized LFS variables and R code are presented in Appendixes 1 and 2, respectively. To observe that the probabilities derived from the rural areas differed significantly from the complete dataset of Germany region, the findings are presented in two columns: considering only individuals living in rural Germany and all individuals. Additionally, Table 2 provides the proportion of transition from student, employed and unemployed states to employed status by age groups.

3.2.1. First-time presence on the labor market. The first analysis focuses on the individuals entering the labor market for the first time. The obtained results (named $ST_{i,c}$) contain the total number of individuals who were students in the previous year that are employed or unemployed in the current year; these results are grouped by ages (i) and by the eight defined SPCs (c). To calculate the proportion, each value is divided by the total of individuals in each age group. Thus, the proportion ($P_{i,c}$) of individuals of age i who started to work on SPC c is obtained as

$$P_{i,c} = \frac{ST_{i,c}}{\sum_{c} ST_{i,c}}$$

Table 3 shows that approximately one-third of individuals who start working earlier in their life (age 15–19) work as craft and machine operator workers in all Germany. Examples for this category are construction, electrical, wood-working workers and industrial machinery operators. In the same age group, the secondly preferred occupations are in the service sector. As individuals’ ages increase (age 20–29), the proportion of individuals entering the job market as professionals and technicians increases; at the same time, the share of craft workers decreases. This trend is similar to when considering only individuals living in rural areas. This could be interpreted as individuals who start working at later ages spend more time on their education[8] and prefer to perform, on average, less physically demanding jobs.

Another SPC result concerns the low proportion of individuals entering the labor market as agricultural workers in rural and all Germany areas. The occupations in this category are classified into three minor groups by the ISCO[9]: market gardeners, mixed crop and animal producers. This shows that the agriculture and fishery sector have a comparatively low attraction rate for individuals who were students in the previous year that enter the labor market. The majority of the students prefer to work as an agricultural worker as their first job between the ages of 15 to 24. For older students, the proportion of them entering the job market as agricultural workers decreases.

In Table 2, we observe that the flow to employment from student status increases until age 34 in both datasets.

Table 2. Proportion (%) of transition from three initial states to employed status by age groups.

	Student		Employed		Unemployed	
Age groups	Rural Areas	All Germany	Rural areas	All Germany	Rural areas	All Germany
15–19	19	17	38	31	39	36
20–24	25	27	59	60	41	38
25–29	32	41	46	56	37	36
30–34	58	50	38	43	34	31
35–39	45	45	37	40	34	30
40–44			34	37	27	28
45–49			33	33	28	26
50–54			30	28	24	21
55–59			31	24	19	14

Source: Calculations based on German Labor Force Survey, average of years 2002–2009.

Table 3. Proportion (%) of students' transitions to employed status by occupation.

	Age groups, rural areas			Age groups, all Germany		
Socio Professional Categories (SPCs)	15–19	20–24	25–29	15–19	20–24	25–29
Manager	0	0	3	0	1	2
Professional	1	10	43	1	12	45
Technician	19	29	22	18	26	21
Clerical worker	11	24	10	13	22	11
Service sector worker	20	14	10	22	21	9
Agricultural worker	4	3	2	2	2	0
Craft worker	35	12	7	31	9	6
Elementary w.	10	7	3	12	7	4

Source: Calculations based on German Labor Force Survey, average of years 2002–2009.

Table 4. Proportion (%) of employed individuals who are looking for another job by occupation in rural Germany.

	Age groups, rural areas					
Socio Professional Categories (SPCs)	20–24	25–29	30–34	35–39	40–44	45–49
Manager	9	3	4	2	2	1
Professional	2	5	2	3	2	2
Technician	5	5	3	3	2	2
Clerical worker	3	4	4	3	3	4
Service sector worker	7	5	5	5	6	5
Agricultural worker	7	6	4	5	3	5
Craft worker	8	9	7	6	5	5
Elementary worker	16	10	8	11	10	10

Source: Calculations based on German Labor Force Survey, average of years 2002–2009.

In rural Germany, we observe substantial differences in the proportion of transition to employment between two age groups: 31% for ages 25–29 and 58% for ages 30–34. By focusing our analysis on these differences, we found that the transition as craft workers for the ages 30–34 was significantly higher and that the transition as a professional and clerk worker for the ages 25–29 have a significantly lower proportion in the rural areas.

From a modeling point of view, the proportions observed for the year 2002 will most likely not reflect the dynamics in 2008 and later, which were characterized by recessions. As a consequence, we decided to use the proportions obtained for the year 2002 for simulations before 2008 and the proportions calculated for the year 2009 when simulating subsequent years.

3.2.2. Employees' job search behaviors. In the second part of the analysis, for each age and SPC, we clustered our sample by individuals employed and looking for another job. This first subset, named $ELO_{i,c}$, gives us the total number of employed individuals who are looking for another job in the current year, grouped by their ages and occupations. We then obtained a second subset with individuals who are employed in the current year also grouped by SPC and year ($E_{i,c}$). Finally, for each age group and professional status, we calculate the proportion of employed individuals who are looking for another job from the total number of employed individuals in each group:

$$P_{i,c} = \frac{ELO_{i,c}}{E_{i,c}}$$

Because it is not possible to know the distribution of the desired SPCs, we use the obtained proportions as a proxy of the probabilities of desired SPCs by employees.

Our results show significant differences in the proportion of employees' job search behaviors across their SPCs in both rural and all Germany areas. As observed in the Table 4, in rural Germany, the proportion of looking for another job for the high skills occupation employees (i.e., managers and professionals) is lower than the proportion for the low skills occupation employees (i.e., elementary occupations, craft workers) by the average of all study years and age groups.

Similar to rural areas, trends of high proportions of job searching for low-skill-occupation employees were found at the country level. One explanation for this may be found in a link between the availability of temporary and low-paying jobs for low-skill-occupation employees. In the literature, while studies[10–12] have found a negative correlation between wage levels and job search decisions, a study on skill levels in Britain[13] found that skilled workers search more than do unskilled workers.

Table 2 shows that the proportion of employed individuals searching for a new job is higher for younger populations (except for ages 15–19, which has a small survey sample) in both datasets. This result is consistent with studies on labor force transition,[8,14] which could be explained by the fact that youths tend to change job more frequently at the beginning of their career to find the best occupation that matches their skills, which is so-called job shopping.[15] For ages 20 to 24, searching for another job is more frequent for elementary occupations, crafts and manager category workers in rural and all German areas. As expected, the proportion of individuals looking for a job decreases with age. However, our analysis shows a lower proportion for job seekers in rural areas with ages between 25 and 44. This may reflect the rural labor market conditions. In good market conditions, lower unemployment rates can increase the probability of finding good job offers; the probability of job search increases for employees.[12] Some studies[13,16] have found a negative correlation of local unemployment on the desire of the employees' job searches.

The comparative results show that the proportion of job searchers aged 25–44 is lower in rural areas for all studied SPCs. In particular, relative to country level sample, elementary occupation employees have the lowest proportion of job searchers in rural areas. Some examples of this category are agricultural and forestry sector workers, laborers in mining, construction and transport sectors. Furthermore, the probability tables indicate that using the data for all Germany to drive the dynamics of the model could be inaccurate in two ways: first, the job mobility of the population aged 25–44 may be overrated, and second, the job mobility of the older population may be underestimated.

3.2.3. Unemployed to employed status. The last part of the rural labor force analysis results (named $UN_{i,c}$) gives us the total number of previously unemployed individuals with a job in the current year. These results are grouped by age (i) and by the eight defined SPCs (c). To calculate the proportion, each value is divided by the total of individuals in each age group. Therefore, the proportion ($P_{i,c}$) of previously unemployed individuals of age i who work in SPC c is obtained as

$$P_{i,c} = \frac{UN_{i,c}}{\sum_{c} UN_{i,c}}$$

The aggregated results for all the studied years show that the highest proportions of individuals transitioning from unemployment to employment are in the categories of craft workers, elementary occupations and service sector workers, and these transition possibilities add up to more than 50% in each age group. An interesting fact shown by the obtained data is that the SPC of skilled agricultural and fishery workers shows the lowest proportion of employed individuals, with a proportion of less than 5%. It is also observed that the SPC technician and clerical worker remain stable between the defined age groups with proportions of approximately 15% and 10%, respectively.

Concerning the year trend analysis, the variation between years 2002 to 2009 is high in the age group 15–19. Additionally, we observed high variations for the transition to service sector from unemployment status. During the reviewed years, the variation in the number of unemployed individuals who become employed exhibits an inverse V shape for the majority of the SPCs. This shows that there is a time-dependent trend that can be grouped into two periods: the first period from 2002 to 2006, where the trends are mainly positive, and from 2006 to 2008, where the trend produces peaks for low-skill occupations. This heterogeneous transition to employed status could be explained by the increase in employment possibilities after the year 2006. This result is relevant to a study[17] that found considerably more job creation between 2006 and 2008 in Germany and explained using the consequence of the changes mainly resulting from the regulatory framework and collective bargaining system.[18]

Additionally, Table 2 shows a decreasing proportion of unemployed to employed status by age in both datasets. Less than 25% of unemployed individuals enter the labor market in a mature population (aged 50 and older). This trend is mostly caused by a reduced attraction to the service and craft sectors.

Comparative results show that the degree of urbanization does not affect the observed proportions on a very significant scale. This decreasing trend by age reassures our assumption to use the age of individuals as a factor of the decision to shift from unemployment to employment in the microsimulation model. Additionally, a year trend analysis suggests that using data for the years 2006 to 2009 to drive the dynamics of the model can cause a distorted result.

4. Prototypical simulations

4.1. Altmark region

The Altmark region, is one of the most sparsely populated regions in Germany and is located in the North of Saxony-Anhalt, west of the Elbe River and between the cities of Hamburg and Magdeburg. The region consists of the Stendal and Salzwedel districts, which are subdivided into 224 municipalities. By administrative reform in 2009, small administrative units have merged, and newly created units have bigger administrative consortia.

The Altmark region adaptation focused on the Hohenberg–Krusemark municipality network, which was selected as the municipality representative of the region (Figure 4).

The adaptation, simulation and result analysis were thus focused on extracting useful data to characterize the selected municipality network. Study focus on commuting patterns in order to give insights into the rural areas dynamics and to understand decision-making process. Based on the methodology of the previous study of Gargiulo et al.,[19] the Hohenberg–Krusemark municipality network was built on the commuting behavior of individuals (Figure 5).[1]

Situated in the district of Stendal, this municipality network is characterized as having only out-commuters. This means that there are no people living in other municipalities working in Hohenberg–Krusemark. The total number of out-commuters for this city is 239; the majority of connected municipalities are major cities near to the region, such as Magdeburg (10 commuters), Gardelegen (13 commuters), Osterburg (24 commuters) and the city of Stendal (83 commuters). In addition, people commute to the municipality of Arneburg (23 commuters). The commute network shows 44 additional commuters who commute to other undefined municipalities in the Stendal District (40 commuters) and in the Salzwedel District (4 commuters). The remaining commuters (10) work in other unknown districts in the same state of Saxony-Anhalt.

For modeling purposes the municipalities in the network are classified in two categories; the municipalities inside the network (Osterburg, Gardelegen, Arneburg, Stendal) are those that should be fully considered as part of the municipality network in the model. The municipalities outside the network (Magdeburg) are those that, although still considered for the model, were not fully parameterized and were only used as external driving forces. Note that the two network nodes that represent undefined municipalities in Stendal and Salzwedel districts have been omitted.

The selected municipalities were included in the model for the simulation by defining the municipalities' properties (individuals, households, jobs and working locations). For the prototypical simulations, the year 2000 was chosen as the starting point of the simulation. This starting year was selected for two main reasons: (a) the availability of data for this year was relatively good because the German microcensus regional dataset was available for 2000; (b) it allowed the use of a future date dataset from the microcensus (2006) and from other sources.

4.2. Simulation results for the Altmark region

The results are reported considering a set of municipalities as a merged group. Grouping the data for several municipalities provided the advantage of simplifying the communication of the simulation results to local stakeholders. Additionally, the availability of real data (from national surveys or other statistical databases) was oftentimes available only at aggregated levels. Thus, a comparison of the

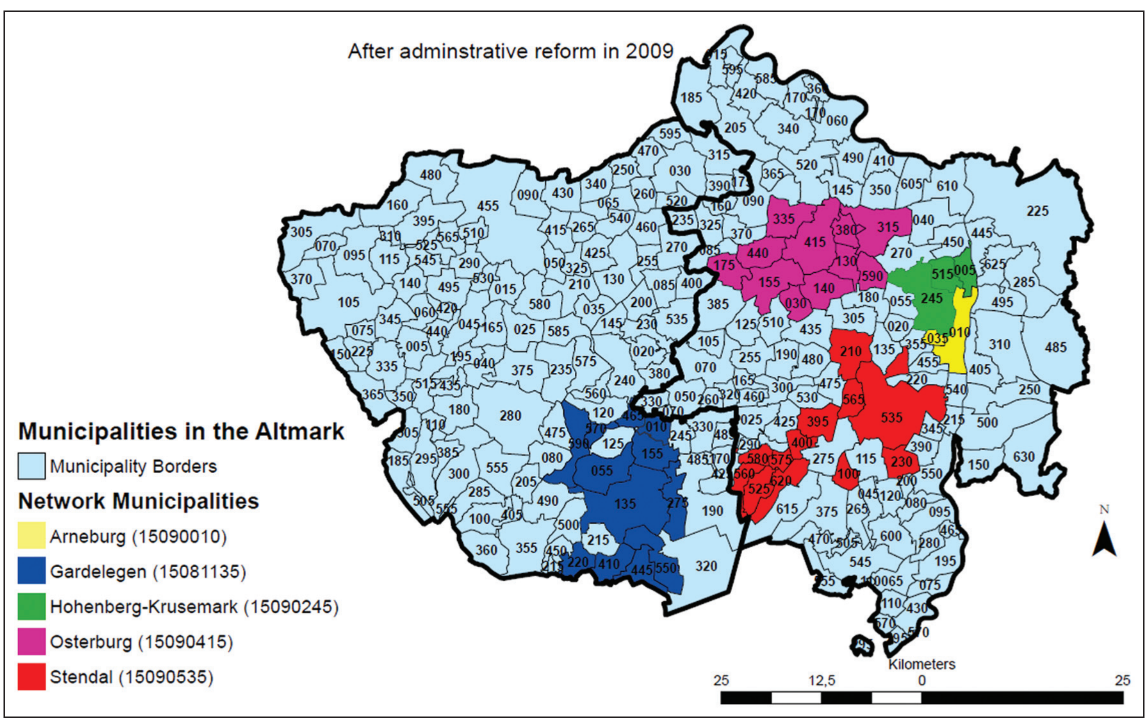

Figure 4. Altmark region, Hohenberg–Krusemark municipality network.

Note: Municipalities in the Altmark region are represented by municipality codes. The colored mapping indicates Arneburg, Gardelegen, Hohenberg-Krusemark, Osterburg and Stendal administrative units that have merged based on the 2009 administrative reforms. These five municipalities are denoted as the Hohenberg-Krusemark municipality network. Color version available online.

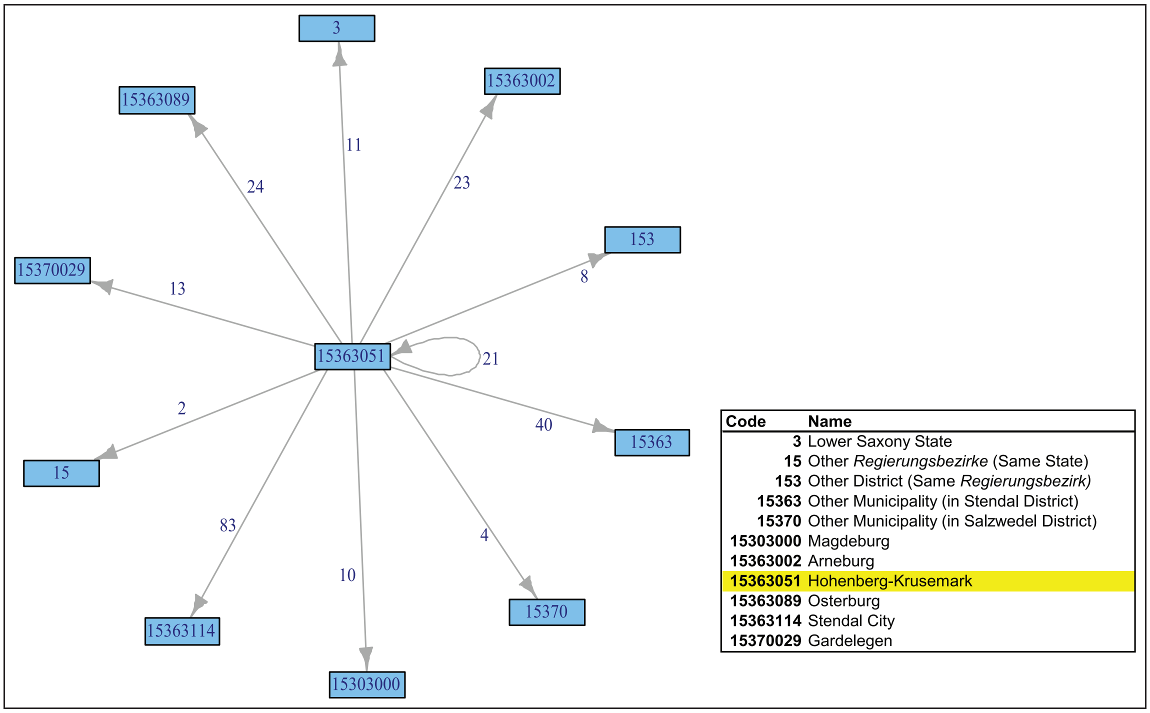

Figure 5. The Hohenberg–Krusemark commute network in 2000.

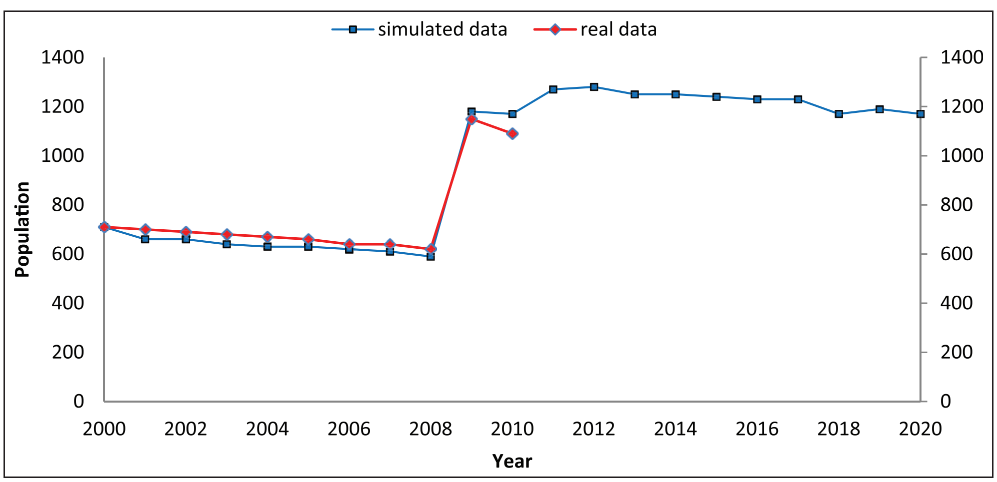

Figure 6. Population trends in the Altmark region, Hohenberg–Krusemark municipality network, Germany (2000–2020).

real data with the data obtained from the simulations was only meaningful at the aggregated level. In the case of the Altmark region, one municipality was selected for reporting. This also reduced the amount of data being presented to stakeholders, while avoiding the issue of loss of precision when grouping several municipalities.

4.2.1. Demographic evolution. The results show the evolution of the demographic properties of the model with the comparison of real data obtained from the statistical offices. The comparison allows one to observe how correctly the model can replicate the past trends in the region. Adjustment was performed in the years 2009 and 2010 by including the population of some municipalities into the population count of the analyzed municipality. This was done to consider a set of municipality boundary modifications realized in the region within those years; this made possible the comparison with the collected data.

Figure 6 gives the general population trend in Altmark within real and simulated data.

Additionally, Figure 7 shows the population evolution trends by age structure for Hohenberg–Krusemark.

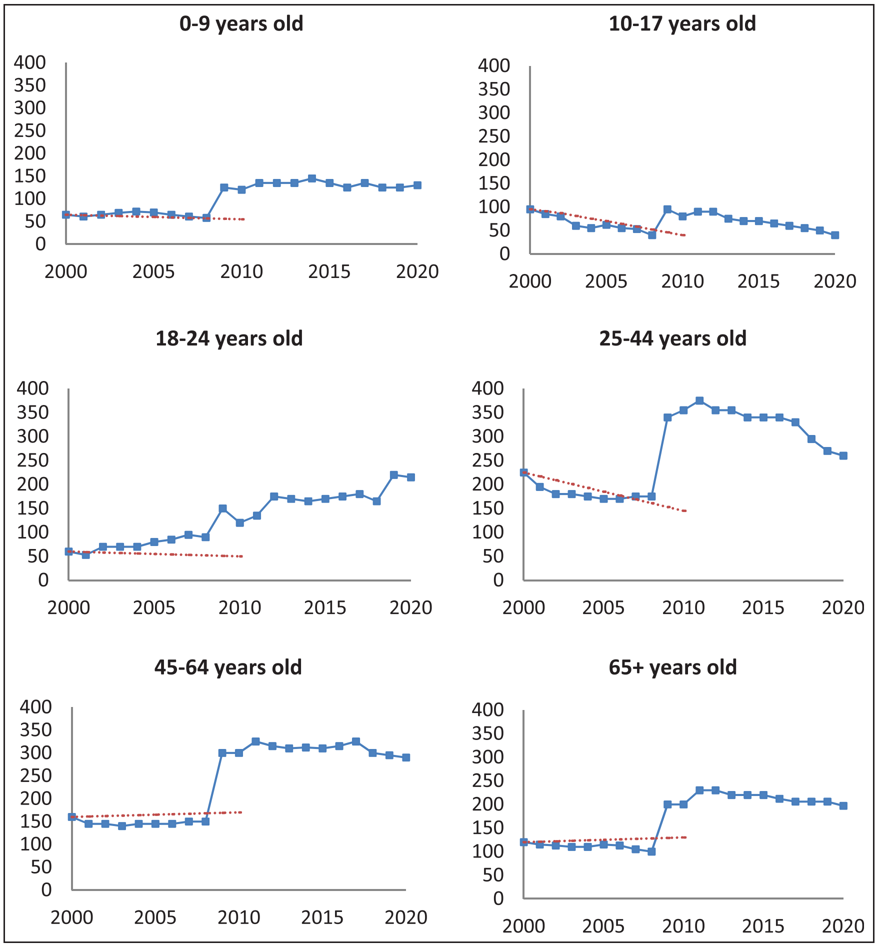

Figure 7. Comparison of population evolution trends by age structure between simulated and collected data, Altmark region, Hohenberg–Krusemark municipality network, Germany (2000–2020).

In the case of the Hohenberg–Krusemark dataset, the model overestimates the number of individuals within the ages of 18 to 24. This discrepancy was discussed with experts during a stakeholder workshop, where it was concluded that the error was because the model does not allow students to move out of the region. Specifically, in the case of the studied region, individuals aged 16 to 24 tend to leave the region to pursue higher education.

In addition, for the Hohenberg–Krusemark region, the model underestimates the number of individuals aged 45 to 64 (and older than 65). This was explained (also during a stakeholder workshop) as the model overestimating the mobility capability of mature individuals. The main cause of the overestimation is the model assumption that a household will always migrate when the head of household finds a job. It was decided that the model needed to consider the residence ownership factor (as individuals owning a house will more likely commute instead of emigrate) and an aversion to move owing to age (because age would affect an individual's willingness to emigrate).

The consulted experts stressed that young individuals are forced to emigrate owing to the lack of affordable housing because the region is part of a protected natural area; thus, residential prices are comparatively high.

In addition to the two previous indicators, birth rate and death rate indicators were used for comparison. Figure 8 depicts the data for the number of deaths.

For the Hohenberg–Krusemark municipality network, the total number of deaths was available at the municipality level; thus, a more detailed comparison was possible. However, for the simulation, death rate data were only available for the second half of the simulated years (2006 to 2010). Although the simulation results are close to the real data, a trend cannot be inferred. The discrepancy between the real and simulated data (and the variations in both) is attributed to the small numbers involved in the comparison. Given that the number of deaths at the reported level is lower than 10, stochastic variations owing to the simulation prevent the replication of the real data.

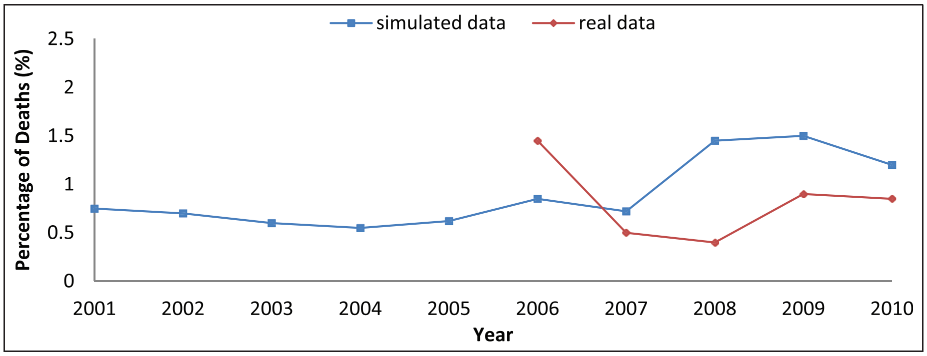

Figure 8. Comparison of death rate evolution trend between simulated and collected data, Altmark region, Hohenberg–Krusemark municipality network, Germany (2000–2010).

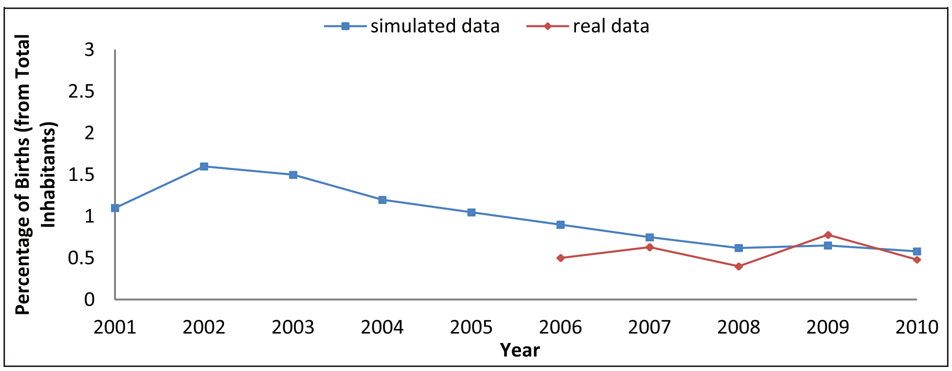

Figure 9. Comparison of birth rate evolution trend between simulated and collected data, Altmark region, Hohenberg–Krusemark municipality network, Germany (2000–2010).

The evolution of the birth rate during the simulation time compared with the collected data is shown in Figure 9.

The simulated birth rate was close to the rate obtained from statistical offices. Nevertheless, the Hohenberg–Krusemark model exhibited a high rate of births at the beginning of the simulation.

A final comparison was performed by comparing the number of individuals living in each household. Such an analysis is omitted from this paper because the data collected from official statistics were unavailable at the municipality level, and thus, a meaningful comparison was not possible. The household structure in rural areas differs significantly to households in urban areas, which were considered in the official data.

Overall, the model is capable of replicating the main demographic trends presented in this section. This fact reinforces the validity of the model, especially considering that the demographic evolution of the model is a function not only of input demographic factors but also of the evolution of employment simulated within the region.

4.2.2. Employment evolution. To evaluate the fitness of the model for the employment trend, a set of indicators were selected. These indicators are employment rate, unemployment rate, and distribution of jobs according to the sector of activity. A different metric was used for each adapted model owing to the type of data available from the statistical offices. In our case study, the data available from the statistical offices allowed for a comparison of the absolute number of employed individuals. These data were available at the municipality level, which allows us to perform a direct comparison. Figure 10 gives the comparison of the employment rate indicators.

As observed from Figure 10, after the year 2008, the model overestimates the number of employed individuals. One explanation for this is the effect of the financial crisis, which the model would not be able to capture given its current assumptions.

The second indicator used for comparison of the employment evolution was the unemployment rate. Similar to the employment indicator, a different metric was used for each analyzed region. Figure 11 depicts the evolution of unemployment.

For the unemployment evolution trend in the Hohenberg–Krusemark municipality network, the simulation data were completely different from the real data. After further analysis, the error was attributed to the

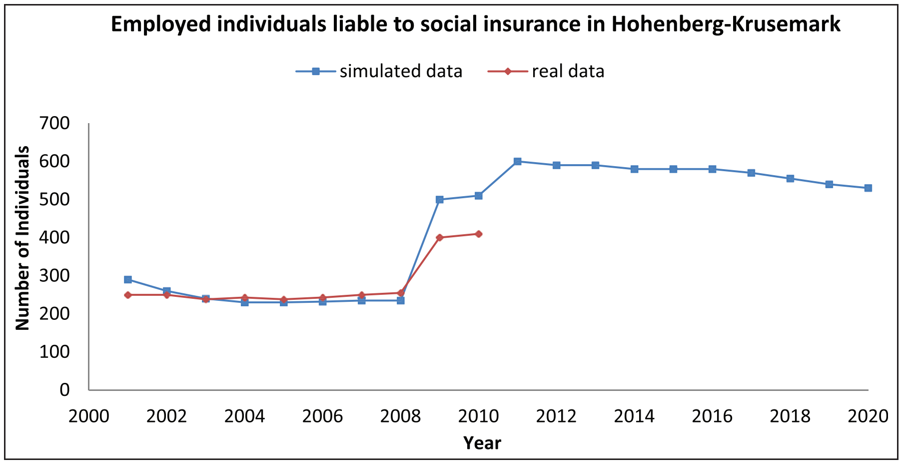

Figure 10. Comparison of employment evolution trend between simulated and collected data, Altmark region, Hohenberg–Krusemark municipality network, Germany (2000–2020).

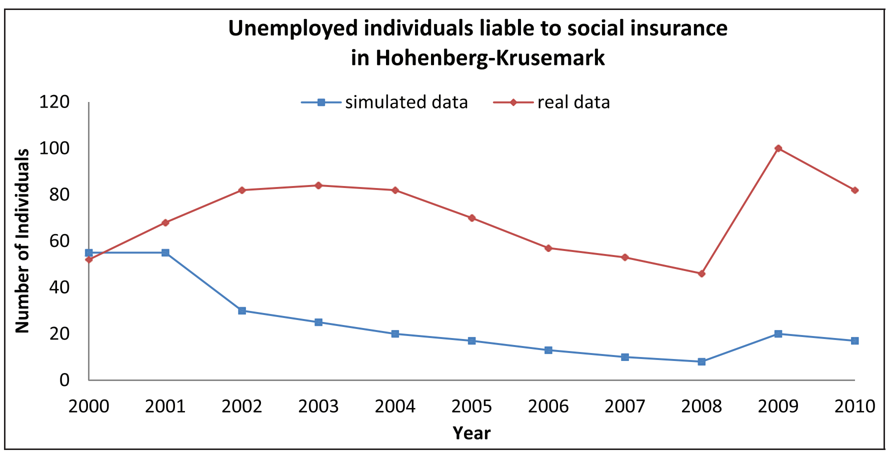

Figure 11. Comparison of unemployment evolution trend between simulated and collected data, Altmark region, Hohenberg–Krusemark municipality network, Germany (2000–2010).

overestimation of mobility of individuals, such as heads of households finding a job in another municipality and moving to the municipality where the job is present.

The country level overview has been refined by Baqueiro-Espinosa and Unay Gailhard[20] on economic inactivity in Germany during the 2008–2009 global economic crises. Authors found that relative to the average of previous years (2002–2007) flow to inactivity from unemployment increased in rural areas.

The analysis of the unemployment data shows that in the case of the adaptation to Germany, the model did not correctly capture the unemployment dynamics. As discussed in the demographic analysis, incorrect assumptions about emigration (e.g., owing to job availability in other municipalities) incorrectly decreased the level of unemployment. This indicated the need to modify such assumptions in the adapted model to represent the unemployment dynamics particular to the region.

Finally, a comparison of the sector of activities was performed using available data from statistical offices. For the Hohenberg–Krusemark municipality network, the sector of activity data was available at the municipality level for several years. The results from this comparison are presented in Figure 12.

As a result, the difference between the simulation data and the real regional data is relatively low, and the real trends observed in each sector are replicated by the simulation data. The only exception is the hospitality sector (Hospitality Industry, SoA7).

5. Conclusion

In this paper, we described a process to define the dynamics of the PRIMA conceptual microsimulation model for the Altmark region in Germany. The adaptation of the model requires defining a set of probability

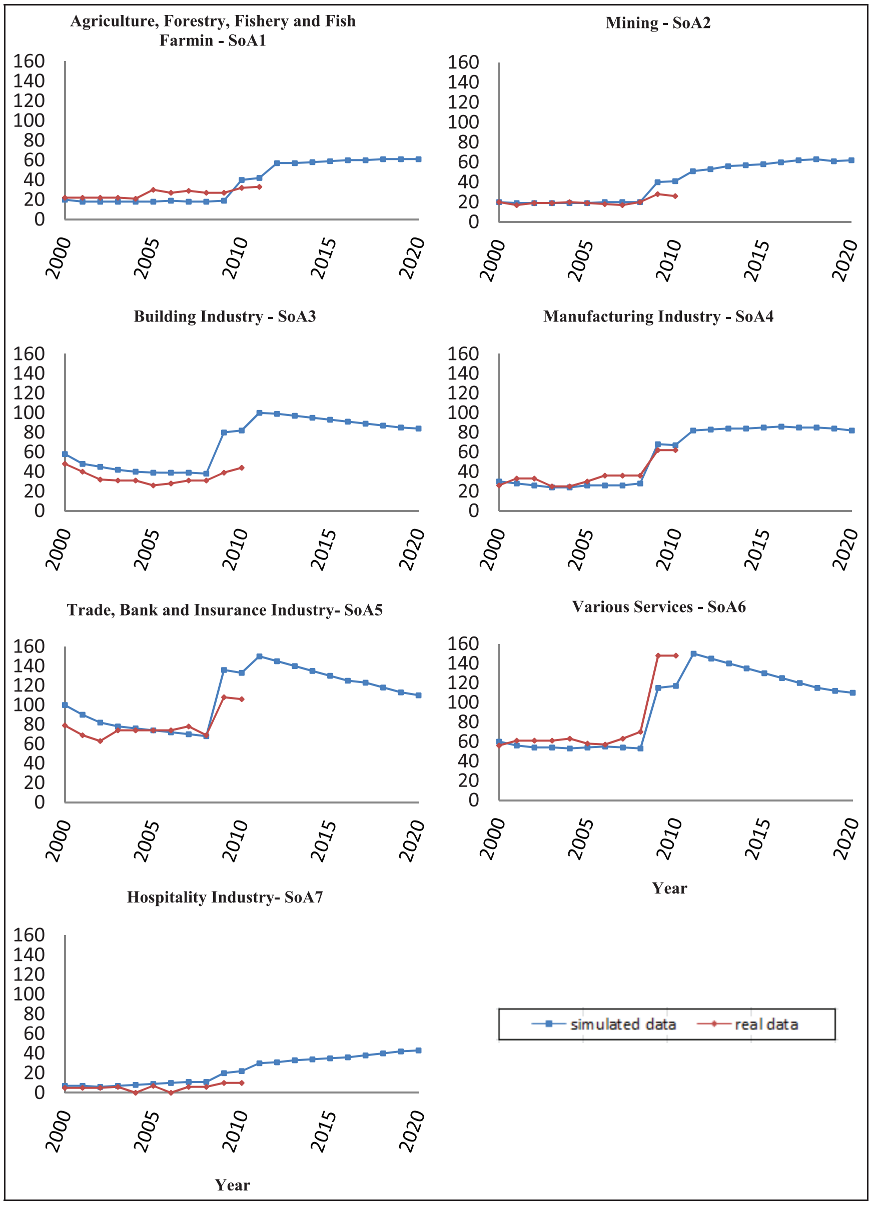

Figure 12. Comparison of sector of activity evolution trend between simulated and collected data, Altmark region, Hohenberg–Krusemark municipality network, Germany (2000–2020).

distributions that drive the choices of individuals and households throughout the simulation. The probability tables were generated from official statistical data derived from the LFS. This data source contains a survey with a sample of the yearly demographic and employment information of individuals and households. To use the available data for the model, some of the model's underlying assumptions needed to be modified. Nevertheless, the presented analysis shows that the modified assumptions provide a sound basis to drive the dynamics of the model.

To improve the derivation of the microsimulation distribution tables, the data from the LFS was analyzed for different years and for a subset of the samples considering the degree of urbanization. On the basis of the constructed assumptions, we separately utilized a comparative approach on rural and all German areas and used eight year results (2002–2009). This was considered necessary to better understand the dataset and to reveal any trends in the data. Indeed, one potential limitation of the microsimulation adaptation is the use of data from a fixed year (or number of years) to represent the model dynamics. The fidelity of such a model would decrease as the simulation continues.

To evaluate the validity of the model, preliminary simulations were run after calibration. The preliminary simulation results show an adequate fit to the statistical data obtained from regional offices. Nevertheless, some of the selected indicators shed light on assumptions that need to be changed or improved for the region. With the help of expert regional stakeholders, a set of plausible explanations for the simulation's disparities were elucidated upon. As a result, a list of refinements to improve the model implementation to better represent each region was proposed. Additional insight was gained when comparing the simulation results with real data for the post-2007 crisis years. Further, it is highly likely that during the crisis period factors that influence transition flow differs in rural and urban regions in Germany.[21]

Based on our simulation findings, it was shown that even though the simulation was able to correctly replicate past trends before the crisis, it was impossible to reproduce post-crisis trends given that the dynamics were strongly affected by external factors not considered in the model. Such an impact should serve as a reminder of the limitation of these types of models.

Adapting microsimulation models to specific regions is a challenging task. This is especially true when aiming to simulate very low regional levels, such as municipalities, given the fundamental data limitations. Nevertheless, the process of parameterization makes it possible to gain an additional understanding of the region. Indeed, in some cases, the best data available may not be as accurate as desired; nevertheless, once the data limitations are understood and acknowledged when using the results of the simulations, the model, albeit imperfect, can be of much help. Indeed, the results from the simulation were used to confront regional stakeholders. Although the data limitations were recognized and discussed, the presented prototypical simulation results showed good replicative validity.

Author Note

In the paper, the authors used German Labor Force Survey data that was obtained from the European Commission, Eurostat, European Union Labor Force Survey annual averages. Eurostat has no responsibility for the results and conclusions in the results of this work.

Funding

This study based on the framework of the multidisciplinary project of PRIMA (PRototypical policy Impact on Multifunctional Activities in rural municipalities) that has been funded by the EU 7th Framework Programme (ENV 2007-1), contract number 212345 (http://cordis.europa.eu/project/rcn/88949_en.html).

Note

1 For a deeper overview on the commuting models in the small geographic units (e.g., municipalities) see the study by Lenormand et al.[3] that focuses on commuting networks on 80 case studies from different regions of the world (Europe and United States).

References

1. Ellis F and Biggs S (2001) Evolving themes in rural development 1950s–2000s. *Development Policy Review* 19(4): 437–448.
2. Turpin N, Laplana R, Kopeva D, et al. (2009) PRIMA Project. Technical Report on Proceedings of the Conference on integrated assessment of agriculture and sustainable development: Setting the Agenda for Science and Policy (AgSAP 2009), Hotel Zuiderduin, Egmond aan Zee, The Netherlands, 10–12 March, pp. 64–65.
3. Lenormand M, Huet S and Gargiulo F (2014) Generating French virtual commuting network at municipality level. *Journal of Transport and Land Use* 7(1): 43–55.
4. Huet S and Deffuant G (2011) An abstract modelling framework implemented through a data-driven approach to study the impact of policies at the municipality level. In: *ESSA 2011 conference, proceedings of the seventh conference of the European social simulation association*, Montpellier, France, 19–23 September 2011.
5. Merz J (1991) Microsimulation – a survey of principles, developments and applications. *International Journal of Forecasting* 7(1): 77–104.
6. Eurostat (2011) Eurostat European Union Labour Force Survey German Annual Averages microdata. 2011.
7. Eurostat (2006) *Labour Force Survey Basic Concepts and Definitions*. Office for Official Publications of the European Communities. Luxembourg.
8. Quintini G, Martin JP and Martin S (2007) The changing nature of the school-to-work transition process in OECD countries. Institute for the Study of Labor (IZA), Discussion paper, no. 2582, January 2007.
9. ILO (2004) International Standard Classification of Occupations ISCO-88. International Labour Organization, Geneva, Switzerland.
10. Black M (1981) An empirical test of the theory of on-the-job search. *Journal of Human Resources* 16(1): 141–151.
11. Hartog J and Van Ophem H (1994) On-the-job search and the cyclical sensitivity of job mobility. *European Economic Review* 38(3–4): 802–808.
12. Ponzo M (2012) On-the-job search in Italian labor markets: an empirical analysis. *International Journal of the Economics of Business* 19(2): 213–232.
13. Pissarides C and Wadsworth J (1994) On-the-job search: some empirical evidence from Britain. *European Economic Review* 38(2): 385–402.
14. Campbell C (1997) The determinants of dismissals, quits and layoffs: a multinomial logit approach. *Southern Economic Journal* 63(4): 1066–1073.
15. OECD (1983) *Employment Outlook*. Paris.
16. Van Ours J (1990) An international comparative study on job mobility. *Labour* 4(3): 33–56.
17. Eichhorst W, Marx P and Thode E (2009) Arbeitsmarkt und Beschäftigung in Deutschland 2000-2009., Institue for the Stuy of Labor (IZA), Reseach Report no. 22.
18. IZA (2010) Labour market and employment in Germany 2000–2009. *IZA COMPACT*. January/February 2010.
19. Gargiulo F, Lenormand M, Huet S, et al. (2012) Commuting network model: Getting the essentials. *Journal of Artificial Societies and Social Simulation* 15(2): 6.

20. Baqueiro-Espinosa O and Unay Gailhard I (2011) Labour force transitions in Germany: adopting micro-simulation model to the study the impact of policies at municipality level. EU FP7 PRIMA Project Working Paper. Halle (Saale): IAMO, 2011.
21. Unay Gailhard I and Kataria K (2014) Economic crisis and labour force transition to inactivity: a comparative study in German rural and urban areas. *Studies in Agricultural Economics* 116(1): 25–32.
22. Lenormand M, Huet S, Gargiulo F, et al. A universal model of commuting networks. *PLoS ONE* 7(10): e45985.
23. Eurostat (2010) EU Labour Force Survey database User Guide, Luxembourg, European Commission. Available at: http://epp.eurostat.ec.europa.eu/portal/page/portal/employment_unemployment_lfs/documents/EULFS_Database_UserGuide_2012_0.pdf (accessed 27 August 2014).

Author biographies

İlkay Unay-Gailhard is research associate at Leibniz Institute of Agricultural Development in Transition Economies (IAMO) in Halle, Germany and a PhD candidate in Martin Luther University, Halle, Germany.

Omar Baqueiro-Espinosa is a former research associate at Leibniz Institute of Agricultural Development in Transition Economies (IAMO) in Halle, Germany.

Appendix 1

Table A1. Used Labor Force Survey (LFS) variables.

Variable Name	Description
DEGURBA	Degree of urbanization
COEFF	Yearly weighting factor for annual averages of quarterly data
WSTA1Y	Situation with regard to activity one year before survey
AGE	Individual age (expressed in age ranges)
MAINSTAT	Main status (optional; not available for certain countries)
ILOSTAT	Working status, as defined by the International Labor Organisation (ILO) and EU regulations.
LOOKOJ	Looking for another job
ISCO1D	Occupation (ISCO 1 digit) from ISCO-88
ISCO3D	Occupation (ISCO 3 digit) from ISCO-88
ISCOPR1D	Occupation in previous job (ISCO at 1-digit level delivered as 3-digit variable to distinguish group 9 and not applicable) ISCO – 88 (COM)
ISCOPR3D	Occupation in previous job (ISCO at 3-digit level delivered as 3-digit variable to distinguish group 9 and not applicable) ISCO - 88 (COM)
YEARPR	Year in which person last worked

Source: Eurostat, 2010.[23]

Appendix 2

Table A2. Used R codes to calculate the first-time presence on labor market from Labor Force Survey (LFS).

(1) library(survey): load survey library
(2) dat < - read.csv ("…") : read dataset
(3) x < -na.omit(dat[(dat$WSTAT1Y==3) & (dat$ILOSTAT==1 | dat$ILOSTAT==2), c("AGE", "ISCO1D", "COEFF")]): obtain the set of all individuals who are students the previous year and are registered as a "employed or unemployed" in the current year (ILOSTAT=1,2)
(4) mysv < -svydesign(id=~1, data=x, weights=~x$COEFF) : create survey design using the data weights (COEFF)
(5) result < -svytable(~AGE + ISCO1D,mysv): create Age and ISCO 1 digit occupation cross tabulation grouping individuals in the subset "x".
(6) result < - result[,-1]: eliminate the "Armed Forces", as shown in the first column
(7) result < - result[,-10]: eliminate the results for "non-available" and "army forces", as shown in the last, tenth column.
(8) result < -prop.table(result, 1): obtain tables of proportions and eliminate the non-available results (999), as shown in the last, tenth column
(9) write.table(result): obtain required result

Table A3. Used R codes to calculate the employees' job search behaviors from Labor Force Survey (LFS).

(1) library(survey): load survey library
(2) dat < - read.csv ("…"): read dataset
(3) x < - na.omit(dat[(dat$ILOSTAT==1) & (dat$LOOKOJ==1), c("AGE", "ISCO1D","COEFF")]): Obtain the set of all individuals who are employed (ILOSTAT==1) and looking for another job in the current year (LOOKOJ==1).
(4) mysv < -svydesign(id=~1, data=x, weights=~x$COEFF): create survey design using the data weights (COEFF)
(5) result < -svytable(~AGE + ISCO1D,mysv): create Age and ISCO1 digit occupation cross tabulation, grouping individuals in the subset "x".
(6) result < - result[,-1]: eliminate the armed forces, as shown in the first column
(7) x2 < -na.omit(dat[(dat$ILOSTAT==1), c("AGE", "ISCO1D","COEFF")]): obtain the second set, called "x2": set of all individuals that are employed (ILOSTAT==1) in the current year.
(8) mysv < -svydesign(id=~1, data=x2, weights=~x2$COEFF): create survey design using the data weights (COEFF)
(9) result2 < - svytable(~AGE + ISCO1D,mysv): create Age and ISCO 1 digit occupation cross tabulation grouping individuals in the subset "x".
(10) result2 < - result2[,-1]: eliminate the responses of "Armed Forces", as shown in the first column
(11) vec < - result
(12) names < - dimnames(result) for (i in names$AGE) {for (j in names$ ISCO1D) {vec[i,j] < - (result[i,j]/result2[i,j]) } }: for each age group (AGE) and professional status (ISCO 1digit) in the resulting table, calculate the number of employed individuals who are looking for another job to the total number of employed individuals.
(12) write.table(vec): obtain required result

Table A4. Used R codes to calculate the unemployed to employed status from Labor Force Survey (LFS).

(1) library(survey): load survey library
(2) dat < - read.csv ("…"): read dataset
(3) x < - na.omit(dat[(dat$WSTAT1Y==2) & (dat$ISCO1D!=999) & (DEGURBA. > 1), c("AGE", "ISCO1D","COEFF")]): obtain the set of all individuals who are unemployed in the previous year (WSTAT1Y==2) and that have a professional status different than "not applicable=9" in the current year.
(4) mysv < -svydesign(id=~1, data=x, weights=~x$COEFF): create survey design using the data weights (COEFF)
(5) result < -svytable(~AGE + ISCO1D,mysv): create Age and ISCO 1 digit occupation cross tabulation grouping individuals in the subset "x".
(6) result < - result[,-1]: eliminate the responses of "Armed Forces" status, as shown in the first column
(7) result < - result[,-10]: eliminate the non-available results (999) as shown in the last, tenth column
(8) result < -prop.table(result, 1) : obtain tables of proportions and eliminate the non-available results (999) as shown in the last, tenth column
(9) write.table(result): obtain required result