A Methodology Approach to Delineate Functional Economic Market Areas with an Iterative Three-Step Spatial Clustering Procedure

Christian A. Oberst
University of Muenster

christian.oberst@gmail.com

CAWM Discussion Paper No. 55

Latest update 2012-03-14

Abstract

This paper proposes an iterative three-step spatial clustering procedure to define Functional Economic Market Areas (FEMAs) with an evolutionary computational approach using flow data on economic linkages. FEMAs are needed as basic observation units in disaggregated economic data analysis, since those have to be taken at the spatial level at which the markets operate. Only then can analyses provide accurate and consistent results and allow useful interpretations of variables and the measurement of spillover effects between markets. Therefore FEMAs should, besides their use as analysis regions, serve as basic areas for regional policy or coordination of these. Although functional markets are particularly used in regional science, the proposed concept with the spatial clustering procedure is transferable to other economic fields like business management and marketing research, network or competition analyses. The presented methodology approach is a generalized and ubiquitous, disjunctive and contiguous delineation procedure based on a joint application of A Multidirectional Optimum Ecotope Based Algorithm (AMOEBA) clustering procedure and an interaction indicator on flow data. The methodology will be illustrated with real-world applications on German commuting data. Also presented is a possible way of computation and illustration of fuzzy market borders (differentiated border densities) as an extension to the procedure

JEL: R12, R23, J61, M3, Z0

Keywords: AMOEBA, functional regions, fuzzy set theory, market delineation, regionalization, regional labor markets

.

Introduction and objectives

This paper proposes an iterative three step spatial clustering procedure to delineate ubiquitous, disjunctive and contiguous Functional Economic Market Areas (FEMAs) in an evolutionary computational approach over the spatial neighborhood. Functional delineated regional markets are widely used by governmental agencies, academic institutions and private market analysts. Common alternative names in regional science literature are functional (urban) regions and, in correspondence to the delineation issues, related terms like travel-to-work areas, regional and local labor or housing markets, as well as planning regions.[1] Generally FEMAs are territorial units defined as economic relations that constitute a functional subdivision of a larger territory, instead of reflecting administrative boundaries, historical events or on similarity of characteristic clustered markets (homogeneity). The most frequent application in regional science is the definition of regional labor markets. The aim of FEMAs is to serve as basic observation units for regional economic analyses and policy (see OECD 2002), or at least to identify zones for policy cooperation that may not represent political boundaries. However, the provided delineation methodology is, with minor customization, transferable to other economic fields like business management, network and competition analyses. While, for instance, in regional economics FEMAs are used to assess employment effects of infrastructure projects and large business developments or the distribution of governmental subsidies, they might be applied in competition analyses to identify the relevant markets, could be used in business management for the allocation of customer service consultants, sales representatives and storage facilities, and have been applied in organizational and technological network analyses (e.g. databank organization, or institutional subdivisions). It should be emphasized that the suggested spatial clustering methodology is not limited to geographical application; but, but to avoid confusion, the focus in this paper will only be on the regional economic context.[2]

The objective of this paper is to present an efficient, universally applicable and transparent procedure for market delineation on the basis of flow data that represent economic linkages – transparent in the sense of comprehensible and traceable for decision makers while meeting the requirements and complexity of flow data structures. The use of flow data is regarded as necessary to comply with the functional principle. According to the functional principle, defined market borders are based on economic activity in form of linkages and interconnections. This stands in contrast to the homogeneity principle, after which markets are formed by similarity of indicators. The flow data can be on any type of economic interaction, e.g. commuting or migration flows, trade, communication and traffic flows, service connections or search patterns, as well as newspaper and advertisement distribution or financial flows (see Karlsson and Olsson 2006, quoted from Vanhove and Klaasen 1987) FEMAs are needed since economic analyses have to be taken at the spatial level at which the markets subjects operate and not at the level of political jurisdictions, because only then can disaggregated data analyses provide accurate and consistent results and allow the useful interpretation of economic variables and indicators as well as the measurement of true spillover effects between markets. In addition to their use as observation and analysis regions, FEMAs should serve as basic areas for regional policy or at least as cooperative areas of local policies as, for example, for public infrastructure and services, business and land development. As mentioned, economic analyses based on local administrative boundaries are inappropriate for many purposes, as markets

[1] See among others e.g.: OECD (2002): Redefining territories: The functional regions, Coombes and Bond (2008): Travel-to-Work Areas, Casado-Díaz (2000): Local Labor Market Areas in Spain, Jones (2002): The Definition of Housing Market Areas and Strategic Planning.
[2] Non-geographical spatial neighborhoods are for example family relations or common research priorities based classification codes as JEL of published papers, etc.

take little account of such boundaries (see West Midlands Regional Assembly 2006). The market situation is dependent not only on the local supply and demand relation within a locality, but also on linkages of supply and demand relations in surrounding areas. In other words the situation in neighboring districts affects the local market conditions and the same applies vice versa. This context can be referred to Cournot's fundamental definition of a market area, "*A market for a good is the area within which the price of a good tends to uniformity, allowance being made for transportation costs*" (Jones 2002 p. 552, there in turn from Stigler and Sherwin 1985, p. 555). In contrast, administrative regions are usually developed historically and represent governmental and administrative structures rather than actual existing economic linkages. It should be noted that the delineation procedure is, in practice, restricted on the availability of data and thus often based on the lowest administrative spatial reference level which units are aggregated to FEMAs.[3]

From Cournot's market definition the main characteristic of FEMAs can be derived, that within a market the prices for comparable goods or services tend towards economic equilibrium on the issue on which the markets were defined. Initially from this statement, one might conclude that price adjustments and reaction sequences are the best indicator. Theoretically, this is most likely the case; but, in reality, price reactions are not immediately apparent and their comparability is limited due to a high heterogeneity of goods and services and imperfect market transparency. This applies especially when aggregated data are used for the above mentioned issues on labor and housing markets.[4] It is, therefore, not surprising that regional market definition based on price data can hardly be found in the regional science literature. One exception is Bode (2008) who argues for the delineation of metropolitan areas that commuting intensity indicator is too narrow and uses fraction (sic) of land prices instead as indicator for a variety of economic interactions and interdependencies (see Bode 2008, although the delineation of metropolitan areas isn't comparable to FEMAs, because there are not characterized by ubiquitous, disjunctive and contiguous). However, the use of price data and not price reactions, whether as fractions, shares, or differentials, does not correspond to the functional principle as described above or to Cournot's market definition, which allows price differences in the amount of transportation costs (including also non-monetary costs as time).

To illustrate the advantage of markets defined according to the functional principle on economic linkages towards the homogeneity principle (clustering by similarity), it should be noted that functional defined areas allow for the division of labor between locations within a defined market. Assume, for example, a region in which most business takes place in the central district, while districts of the surrounding belt take on the residential function. Their characteristics are quite different, although they are strongly linked and dependent on each other. Another example for market segmentation within a FEMA is the classic suburban differentiation with centrally located rental city apartments in multifamily houses and larger owner-occupied homes in the less dense neighboring areas, the local industry mix with highly specialized services and headquarter functions in the highly dense core area, and land-intensive production in the less dense neighboring belt areas that offer lower land prices but still in reach for the agglomeration advantages and infrastructure of the core region.

Further characteristics of the intended delineation results are disjunctives (not overlapping) and geographically contiguous markets that completely cover the larger territory (ubiquitous), to ensure an unambiguous assignment

---

[3] If microdata with spatial information are available there are applicable as well.

[4] To substantiate this it is referred to the study of Michels and Rusche (2008) in where is concluded for regional housing markets that bi-directional migration linkages are more efficient than price data in a small-scale context.

of spatial units for decision makers. Here, however, the AMOEBA (A Multidirectional Optimum Ecotope Based Algorithm, see Aldstadt and Getis 2006) step enables irregular or amoeba-like shaped markets to occur. The intended disjunctive characteristic is later released to some extent in an extension proposal to present market borders with differentiated densities by applying the fuzzy set theory. The suggested procedure is an evolutionary computational approach which works without any prior core-definition or other artificial restrictions of market size and distance limitations and will be calculated simultaneously for all sample units. If the flow data is present as point-related data, it needs to be converted into areal data (polygons) or the default AMOEBA step has to be modified. An alternative approach to the proposed concept of FEMAs in regional science is to model space by spatial dependency parameters with spatial lags of the dependent and/ or independent variables, and/ or a spatial error parameter (see Anselin 1988, Elhorst 2010). The advantage of the suggested non-parametric methodology approach to delineate functional economic market areas towards spatial dependency parameters is that no definition on market areas are imposed apriori. This allows getting around the specification problem, by which the modeling choice is highly susceptible to error. Moreover, there is little way to correct for this error afterwards.

Another alternative is the recently with considerable more attention regarded tool-box of network analysis. One remarkable analysis which corresponds to the later in an application example used commuting flows between German districts is Patuelli et el. (2007). Both of these approaches might be proficient for pure statistical data analyses, but are rather unsuitable for direct local policy implementations.

The outline of the paper is as follows:

- Section 2.2 gives a short overview of related research work in which light the presented approach was developed, with a special emphasis on the wealth of literature on this topic written in German, only limited amounts of which can be found in international literature reviews and whose definitions serve as comparisons for the application example.
- Section 2.3 describes the iterative three-step spatial clustering procedure, starting with an overview of the three main steps together with a short description of the required data input. Afterwards the three steps will be successively discussed in detail as follows:
    - 1st the AMOEBA step used to create a pool of potential markets;
    - 2nd the calculation of the proximity measure and decision variable for the cluster algorithm; and
    - 3rd the simultaneous selection/ definition of FEMAs.
- In section 2.4, real-world applications on German commuting patterns are given, first on a subset of the State of North-Rhine-Westphalia for which some possible variations within the procedure are discussed (including the fuzzy set illustration).
- In section 2.5 follows then the nationwide application, which results are compared to proposals in related research.
- Section 6 concludes with a summary and outline of future development.

Related Research

There is a wealth of literature providing definitions of FEMAs which, however, is scattered across linguistic and scientific disciplines and has been discussed under various names. The objective of the following remarks is to shed light on the discussion of current delineation procedures and to show their diversity. There is first a special emphasis given to the German discussion in which context the methodology was developed and whose works are often absent in international reviews. Afterwards, recognition is given to the recent international literature on regional market delineation, focusing only on the most recent publications, since the interest on the relationship between locations and the formation of markets in which areas reflect economic activity can be traced back to the central place theory by Christaller (1933) and the nature of economic regions by Lösch (1938), for which a complete overview would be beyond the scope of this paper.[5] The overview on German scholarship is separated from the international, due to its quite independent character, the stronger influence on the development of the methodology as benchmark, and for clarity reasons.

The presented delineation procedure has its origin in a series of research work which aimed to define housing market areas for Germany, for which an earlier development stage of the presented methodology was applied on "family migration" patterns of 413 German districts.[6] With this, the approach of previous studies by Michels and Rusche (2008) and Rusche (2009) to delineate housing markets areas on the basis of "family migration" flows was continued, following and referring to the argumentation of Jones (2002) to use migration patterns to define regional housing market areas.[7] In the comparative case study of Rusche (2009) of 13 German regions, the joint application of AMOEBA (see Aldstadt and Getis, 2006) is suggested with an interaction indicator on flow data. This was initially developed in Michels and Rusche (2008) where it was concluded that bi-directional migration interdependency is more efficient than areal price data to define housing markets on the spatial reference level of districts. As it will be discussed in the methodology section, a programmatically implementation with a selection step and some modification of the interaction indicator were necessary for a nationwide disjunctive and ubiquitous delineation.

The placement of the spatial clustering procedure in the German discussion on functional regional markets, which corresponds to the intended characteristics of disjunctive, contiguous and ubiquitous areas, is given with the following brief overview on well-known definitions of functional markets of Germany. There are the three well known labor market definitions on commuting patterns for Germany:

- Kropp and Schwengler (2011) defined functional labor market regions with a graph theory approach, resulting in 50 regional labor markets and 105 local labor markets in a differentiation, respectively
- Eckey et al (2006) achieved with a factor analysis with an oblique rotation 150 regional labor markets

[5] Descriptions of prior publications which are commonly referred to in literature reviews of related research work can be found in the later mentioned overviews; this applies for example to the work of Fox and Kumar (1965), Goodman (1970), Smart (1974), Slater (1976), Ball (1980), Dawson et al (1986) or Boatti (1988), and van der Laan (1998) with daily urban system for the Netherlands, or Pacinelly (1998) with the "sisteme locali di lavour" for Italy.

[6] Setting: queen neighborhood definition, adj.ifi as decision variable with retrospective separation of markets with negative adj.ifi, no fuzzy illustration, and a maximum of calculated iterations. The meaning of this setting is discussed in the following sections. For details see Michels et al (2011), in German.

[7] For "family migration" are used in the three precursor studies of Michels et al (2011), Rusche (2009), Michels and Rusche (2008) the migration flows of the age groups of under 18 years and between 30 and 50 years, which are assumed to be an appropriate proxy for family household migration and thus housing market orientated nearby migration;

- BMWI (2011) uses the Regional Aid Map (GRW) mainly for analyses of regional development, which is defined more under administrative and political requirements, and methodologically based on a threshold method by the work of Klemmer (1976); it consists of 270 regional labor markets[8]

Besides these functional labor markets, the following definitions should be mentioned:

- Michels et al (2011), as mentioned before defined with a precursor procedure of the later presented iterative three-step spatial clustering procedure on the basis of family migration patterns, resulting in 171 regional housing markets
- BBSR (2010) with 96 planning regions, which are the observational and analytical framework of the Federal Regional Planning and defined on a mixture of an administrative and functional concepts which are based on planning regions of the states and commuting patterns between districts[9]
- In Schlömer (2009) are calculated with a cluster analysis 94 migration regions, but without precise clarification if this can be understood as functional regional markets (see Schlömer 2009, p. 140)

An excellent methodological overview of the German discussion on FEMAs is given in Kropp and Schwengler (2008), in which several delineation techniques (threshold method, variants of cluster analysis and factor analysis) are described and applied on commuting patterns. This work is supplemented by the recent publication of Kropp and Schwengler (2011) suggesting a graph theory approach.[10] Further definitions like agglomerations, metropolitan area concepts, the European functional urban regions or other spatially selective delineations like the BBSR spotlight regions (housing markets) are due to a lack of comparability with the intended delineation characteristics here not listed.[11] Only a few metropolitan concepts are used later in a numerical quality comparison. The later presented spatial clustering procedure results in 110 FEMAs and matches according to the number of regional markets, at most the 105 local labor market regions by Kropp and Schwengler (2011), and on a alterative data set focusing only on commuting across districts results into 145 FEMAs which matches more than 150 regional labor markets by Eckey et al (2006). However, the basic procedure of the presented methodology is a cluster analysis, but differs from the one in Kropp and Schwengler (2008) and Schlömer (2009) by the evolutionary computation over spatial neighbors.

International overviews on the extensive literature with a wide range of named applications are given (for example in Feng (2009), van Nuffel (2007), Karlsson and Olsson (2006), and Casado-Diaz and Coombes (2004)) which are partly complementary to each other. Below the delineation approaches will be distinguished in three types:

(1) thresholds and criteria methods

(2) cluster approaches

(3) further approaches for complex structures

---

[8] BMWI stands for German Federal Ministry of Economics and Technology, and GRW for Joint Agreement for the Improvement of Regional Economic Structures (*Gemeinschaftsaufgabe Verbesserung der regionalen Wirtschaftsstruktur*).

[9] BBSR stands for the *Federal Institute for Research on Building, Urban Affairs and Spatial Development.* The framework is used since 1981 with the latest recalculation in 1996, see Böltken (1996). In 2008 there was an adjustment resulting from a reorganization of districts in the states of Saxony and Saxony-Anhalt (see BBSR 2010).

[10] A more generalized methodology description on this modularity clustering method can be found in Brades et. al (2008), quoted from Kropp and Schwengler (2011).

[11] For this kind of regions is as an incomplete overview for example referred to the definitions in Antikainen (2005) with the concept of Functional Urban Areas and the ESPON definition for the European Union, e.g. for Germany the urban agglomerations (see also BBR 2000) or in correspondence to labor market regions in Germany in Eltges (2008) respectively, Dijkstra (2009) for metropolitan areas of the European Commission: DG Regio, Bode (2008) as well as BBR and IKM (2008) and Rusche and Oberst (2010) for alternative metropolitan area definitions in Germany, OMB (2000) for the definition of the metropolitan statistical areas (MSA) in the USA, Hirschle and Schürt (2008) for the BBSR spotlight regions in Germany, or the Larger Urban Zones of the Eurostat's Urban Audit (Eurostat 2010).

and supplemented by a fourth category of more political-administrative motivated and institutional definitions.

1. Thresholds and criteria methods work according to the principle that when a certain threshold is exceeded, or a set of criteria is fulfilled, the corresponding spatial units are combined to a common market area. To this group of delineation methods belongs the definition of Travel-To-Work Areas (TTWA) of the UK Office for National Statistics, whose latest review can be found in COOBES AND BONDES (2008). The TTWA are defined on several criteria in a descending priority, namely:

- self-containment of flows to maximize (autonomy);
- areas within a size range (homogeneity);
- reasonably recognizable boundaries (coherence); and
- alignment with administrative boundaries (conformity).

The basic definition procedure can be found in Coobes et al (1986), in which the regional markets are called official statistical reporting areas. The TTWA delineation procedure is adopted for several countries (see Sforzi et al. (1997) for Italy, Casado-Díaz (2000) for Spain, Andersen (2002) for Denmark, Papps and Newell (2002) for New Zealand and Watts (2004) for Australia), and an analog application on housing markets in the UK is well prepared in Jones (2002). A further work on regional housing markets is Royuela and Vargas (2007) for the region of Catalonia in Spain, in which it is referred to Jones (2002) and Coombes et al (1986), and migration and commuting data are compared. For their application, Royuela and Vargas found that the commuting-defined areas are more homogenous in terms of prices than migration-defined ones, but arguing that both definitions are valid.

Another frequently quoted work on threshold methods is Karlsson and Olsson (2006), in which three commuting-based delineation concepts are applied to Swedish commuting data. First are the "local labor markets" which are delineated by one-way commuting flows that satisfy conditions of existing linkage to a central place, or neighboring central places calculated by differentiated breakpoints. Second, the "commuting zones" are calculated among municipalities on commuter flows in both ways. The third is labeled accessibility approach and is based on potential interaction between locations and calculated by the commuting time. Karlsson and Olsson (2006) also provides a theoretical model for delineating commuting regions based on Beckmann's (1996) spatial equilibrium in labor markets.

2. Cluster analyses are common techniques for the definition of regional markets. In general, cluster analyses are characterized by their proximity measure and fusion algorithm. The proximity measure describes the connection between spatial units, and the algorithm is the procedure in which the spatial units are merged (see Kropp and Schwengler 2008, p. 28). The often- mentioned criticism of cluster analysis – that once merged markets cannot be dissolved even when no longer optimal – is true for many procedure, but does not have to be. For the spatial clustering procedure as suggested here, it is possible that once-merged markets in a further iteration can certainly be dissolved again. An alternative sophisticated cluster approach for market delineation is, for example, the evolutive approach by Flórez-Revuelta et al. (2006) based on evolutionary computation, which means that it involves combinatorial optimization and an iterative progress. The markets are formed by merging spatial units to maximize a fitness function that measures the aggregated interaction under the constraints of inter-region separation and minimum size and is applied to define local labor markets for the Valencia region on commuting flows. A disadvantage is that their evolutive approach does not guarantee that the results would remain unaltered in different trials and are, therefore, to some extent random. Almost the same author group suggested in Flórez-

Revuelta et al. (2008) a memetic algorithm as extension to the evolutive algorithm with a slightly but significant improvement.

Cörvers et al (2009) delineate functional regions for the Netherlands with a further cluster analysis approach by using a Markov analytic functional distance approach. There the interaction matrix (in their case also commuting flows) is transformed by applying the so-called Ward clustering procedure into a mean first passage time matrix (MFPT) which represents the functional distances between each unit of the spatial reference area (see Cörvers et al. 2009, in particular pp. 21f., 29 and the cited methodology of Ward 1963. The authors argue to test the delineation results against administrative definition of regions on a set of relevant criteria to indicate that regional disparities on income level, housing prices, employment and unemployment rate are lower. However these tests contradict the functional principle, which approval of the division of labor and market segmentation within FEMAs is exactly an advantage of a functional towards homogeneity delineations.

Although cluster analysis approaches can already reach a high level of methodology and computational complexity, under the name of delineation approaches for complex structures (3) are grouped procedures which are primarily related to the representation of complex flow structures, namely are listed the approach with graph theory, factor analysis, degree of polarization and fuzzy-set. The mentioned work of Kropp and Schwengler (2011) combines the graph theoretical approach with the traditional threshold methods and a hierarchical clustering.[12] The graph theory or concept of dominant flows goes back to Nystuen and Dacey (1961) and is often subject in network analysis. It assumes that a number of nodes (spatial units) are connected by different levels of inflows and outflows. In several stages are then dominant flows that exceed the certain threshold singled out, and over hierarchical levels integrated into a threshold method. From a pool of possible solutions, whose size depends on the threshold, the optimal definition is chosen on a quality criterion with an algorithm over several iterations. Kropp and Schwengler (2011) suggest and refer to the Q modularity of Newman (see Newman and Girwan 2004). Other methods for complex flow structures are factor analytic delineation procedures, where areas with similar flow structures are combined and thereby also indirect linkages considered. An application example is the mentioned German labor market regions of Eckey et al (2006); however, according to Kropp and Schwengler (2008) the application of factor analytic procedures on German commuting data produce delineation results which are more homogenous in size, but less effective in capturing the linkages.

In Van Nuffel (2007), using the analysis of Flandern in Belgium, is argued that, for regions with complex commuting structures, a functional delineation does not provide the appropriate geographical structuration. Instead is suggested the degree of polarization combined with a three-step procedure of multiple linkages, cluster and skewness analysis to determine threshold values. The three-step procedure should reduce inherent arbitrariness. The areas are labeled according to the degree of polarization of their outgoing flows and do not belong to the attributes (see Van Nuffel 2007, p. 510). In the same direction of complex labor market structures aims the work of Feng (2009), who provides a fuzzy-set solution of TTWA by allowing for overlapping markets, arguing that disjunctive and ubiquitous functional market definition are imperfect in the sense that there are always commuting trips crossing the defined boundaries. Although a fuzzy set definition does not meet the described intended characteristics of disjunctive, contiguous and ubiquitous areas, it is listed here because the

[12] Kropp and Schwengler (2011) refer to similar approaches by Rabino and Occelli (1997), Haag and Binder (2001); Gorman et al. (2007).

fuzzy concept can easily be integrated as an extension to the suggested spatial clustering procedure. A first possible form of this is illustrated later for the application example.

A more **institutional international survey** of more or less FEMAs is given in OECD (2002), also it is to point out that not all of the listed delineations correspond to the intended characteristics and various definition of metropolitan areas, agglomerations are listed. An older attempt of an international summary can be found in Eurostat (1992) which provides a summary of methods which should be overhauled by now, but is still often quoted in recent literature (e.g Coobes and Bondes, 2008), in particular because the description of general principles.

This overview on related research showed the variety of approaches, from simple threshold methods to complex structure analysis; and in the end, many of the complex procedures are still based on some kind of pre-defined or calculated threshold values which seem somewhat arbitrary. The overview could easily be extended by adding further country applications and prior work, which can be found in the referred literature reviews.

The following iterative three-step spatial clustering procedure should supplement these delineation approaches, characterized by the particular use of spatial information and the use of flow data to be in accordance with the functional principal. The iterative three-step spatial clustering procedure uses the cluster analysis techniques with a proximity measure and fusion algorithm and can be classified as evolutionary computation with combinatorial optimization and iterative progress and is therefore most related to the work of Flórez-Revuelta et al (2006, 2008), but likely with a more easily comprehensible methodology for non-specialists. However, none of the listed procedures can be theoretically identified as clearly superior to any of the other methods, and since theoretical derivations for the appropriate procedure are hardly possible, the suitability is ultimately displayed in the result and by its acceptance in practice.

Methodology[13]

The basic idea of the procedure is to maximize the share of intra-market flows on total flows within an overall space. This objective will later be used to assess the quality of the delineation result. An overview of the general procedure is displayed in Figure 1. The three basic steps follow each other within an iteration and the procedure iterates until it converges to a delineation result. The procedure is performed simultaneously on all underlying spatial observation units (e.g. districts). The economic linkages have to be provided in a matrix and the spatial neighborhood has to be specified (default is the queen criterion). After the following brief overview, a detailed description for each of the three basic steps follows.

The **first step** is the identification of potential markets by using the combinatorial clustering procedure AMOEBA (A Multidirectional Optimum Ecotope Based Algorithm) by Aldstadt and Getis (2006). Iteration, one step of the AMOEBA on each defined core is processed. The cores are in the first iteration all units of the spatial reference level and in the following iteration the preliminary defined market areas. The resulting set of potential markets is referred to as FEMA-Pool$_i$, with i as index for the iteration. In the later following application example on commuting flows between 54 NRW districts, the starting set of cores of FEMA-Pool$_{i=0}$ is 54:1, containing 54 markets of 1 unit. After that the FEMA-Pool$_{i>0}$ dimension strongly depends on the defined neighborhood, e.g. in the later following application example for the 54 NRW districts with queen continuity results in a FEMA-Pool$_1$ of 6,337:14, for four nearest neighbors in 655:7 and ten nearest neighbor 42,608:13. The default neighbor in the procedure is as in Aldstadt and Getis (2006) a row-stochastic first-order queen-contiguity with 1 if markets i, j, (i $\neq$ j) share common borders and 0 otherwise. Other neighbor criteria like k-nearest neighbors, or staggered distance matrices are conceivable.

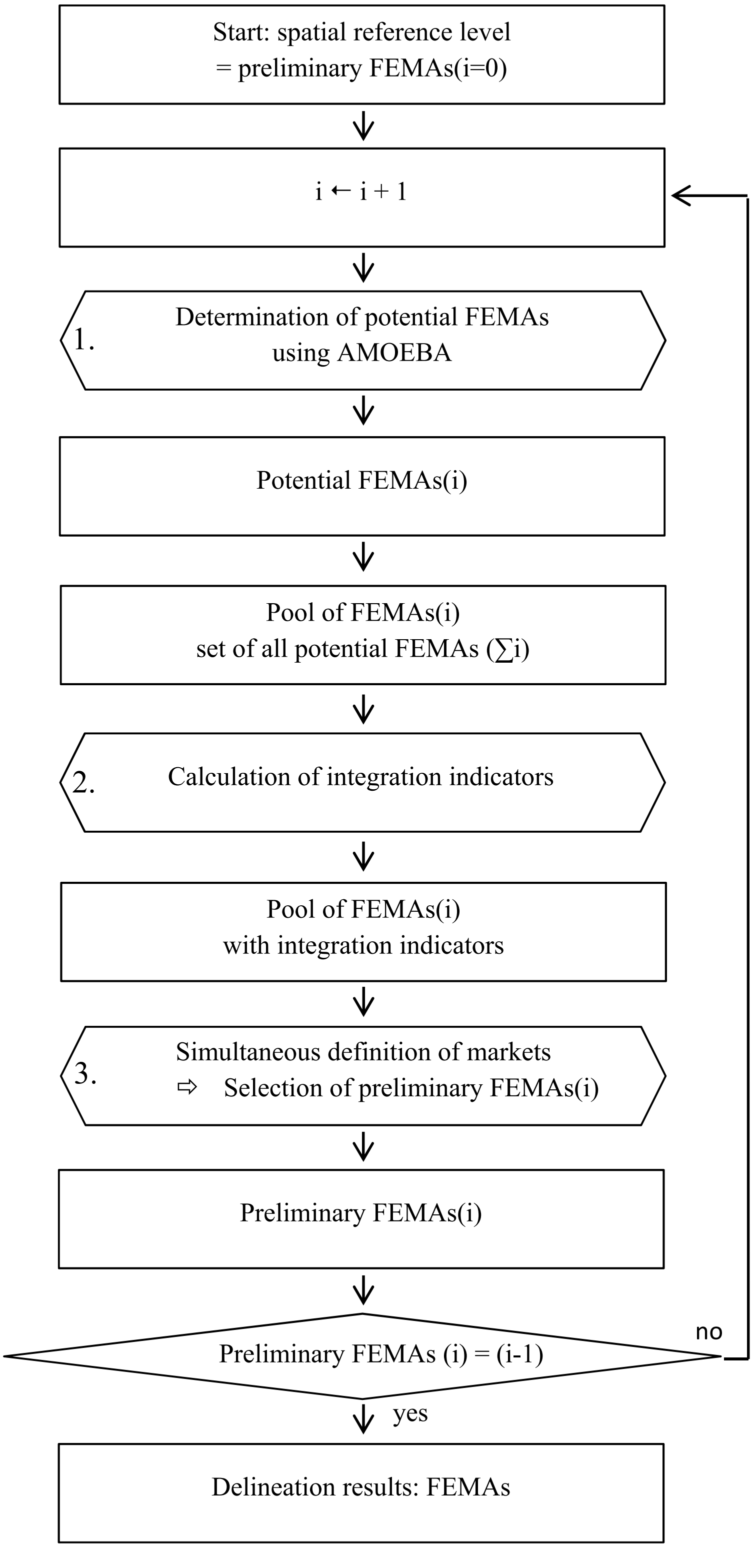
The **Second step** is to calculate the proximity measure or decision variable for the fusion algorithm of new markets, which is labeled intra flow indicator and is calculated for any of the potential markets of the FEMA-Pool(i). In this paper two indicators are used: One is named the basis intra-flow indicator (base.ifi), which is the originally measure "*Wanderungsverflechtungskennzahl*" [English: migration interdependence ratio] of Rusche (2009). The base.ifi turned out to be not sufficient as proximity measure and can only be used to illustrate the market integration. However, it presents the base term of the used specially developed second indicator and used proximity measure for the delineation procedure, which is labeled adjusted intra flow indicator (adj.ifi).

The Third step is the simultaneous determination of the preliminary FEMAs, which is the selection of optimal potential markets out the FEMA-Pool$_i$. The preliminary FEMAs$_i$ are the delineation result after the i^{th} iteration.

If the preliminary FEMAs$_i$ differ from the previous one (FEMAs$_{i-1}$) than the three steps are repeated in a further iteration. Over the iterations the FEMA-Pool$_i$ accumulates (with no duplicate entries of a potential market) and the trend correction part in adj.ifi has to be calculated for each iteration. The procedure continues until the delineation results converge to a stable results what is achieved when the preliminary FEMAs of Iteration i equals the one of iteration i-1.

[13] The procedure is implemented in R, see Development Core Team (2011). The delineation procedure itself, illustration and analysis are based on functions and classes out of the following packages/libraries and their associated packages: spdep (Bivand 2011), classInt (Bivand 2009), rgeo (Bivand and Rundel 2011), rgdal (Keitt et al. 2011), maptools (Lewin-Koh and Bivand 2011), RColorBrewer (Neuwirth 2011), gpclib (Peng 2010), MASS (Venables and Ripley (2002).

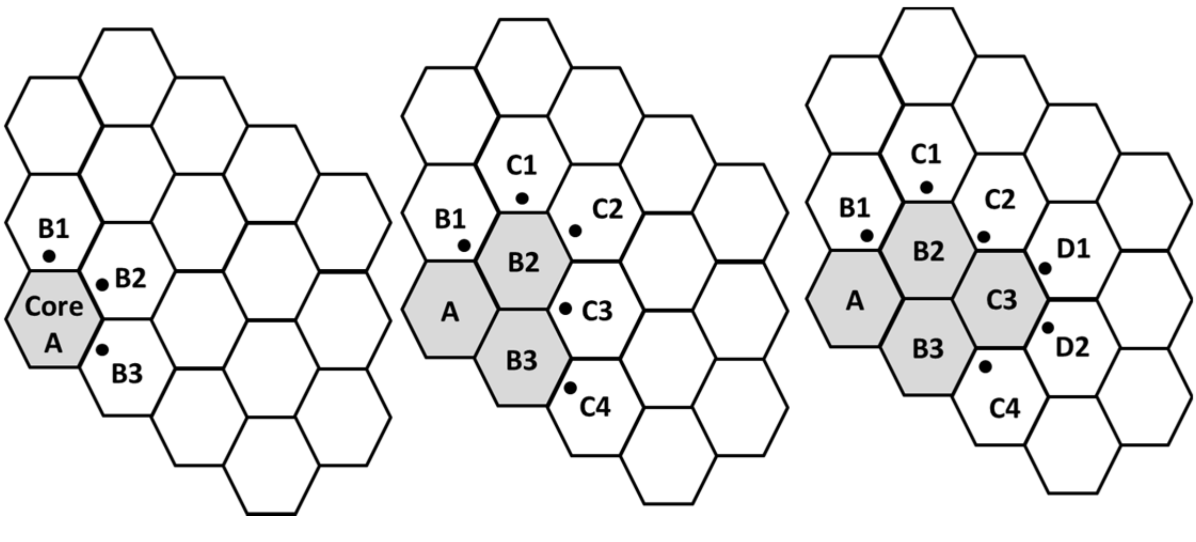
Figure 1: Overview delineation procedure



Step 1: AMOEBA iteration to determine potential FEMAs (see Aldstadt and Getis 2006, pp. 326-333). Originally it was used to create spatial weight matrices and to identify spatial cluster with areal spatial units. To the knowledge of the author, Michels and Rusche are the first suggesting to use the multidirectional sequence part of AMOEBA in a delineation procedure on flow data (see Michels and Rusche 2008), and Rusche 2009). However for this initially by Aldstadt and Getis computed Getis-Ord local statistic G^*_i (Ord and Getis 1995), or mentioned alternatives of local Moran's I_i (Anselin 1995) and spatial scan statistics (Kulldorf 1997) are no longer applicable, and another indicator is needed for flow data. As such, Rusche (2009) suggested a migration interdependence ratio here labeled as base.ifi, and Michels et al (2011) the corresponding trend adjusted variation labeled adj.ifi. The adopted part of the AMOEBA procedure step will be described in general without further specification of the proximity measure.

The combinatorial AMOEBA procedure identifies optimal value clusters starting from a predefined core by taking the neighborhood structure of each spatial unit into account. At first the indicator is computed for the core itself as initial value. Then the indicator variable is computed for each possible market that contains the core and all combinations of its contiguous neighbors. This step is illustrated in Figure 2 in which for an ecotope A with three neighbors (B1, B2, B3) eight possible markets need to be evaluated in the first iteration.[14]

Figure 2: Step 1, AMOEBA



The general formula of all possible combinations ($\Omega(c)$) of n neigbors and therefore potential FEMAs are given in formula 1.

$$\Omega(c) = \binom{n}{n} + \binom{n}{n-1} + \binom{n}{n-2} + \ldots + \binom{n}{n-n} \qquad (1)$$

The exponential increase of possible combinations by each additional neighbor can be easily illustrated by using the formula: for one neighbor it results in two combinations, for two neigbors in four (1+2+1=4), and for three in eight (1+3+3+1=8). For ten neighbors this sums up to 1,024 combinations, for 15 to 32,768, and for 20 to 1,048,576.

---

[14] Three combinations of two (A-B1,A-B2,A-B3), three combinations of three (A-B1-B2,A-B1-B3,A-B2-B3), as well as the one combination of four (A-B1-B2-B3) and the one with itself (A).

The combination of the spatial units with the highest indicator value is merged as new market, when it exceeds the value for the core. If a spatial unit is not merged with any other it will be referred to as a self-sufficient or autonomous market. In the illustration it is assumed that the combination of A-B2-B3 has the highest value and thus is merged. While the spatial unit B1 not considered in the original version by Aldstadt and Getis (2005) is excluded for the following analysis, here it will be allowed to serve in the next iteration as one of the neighbor units.[15] In the next iteration the indicator variable is then computed for all combinations with neighbors of the before-merged market. In the illustration the new core A-B2-B3 has five neighbors: Again B1 as well as C1, C2, C3 and C4. An Ecotope with 5 neighbors results in 32 potential market combinations. Combination A-B1-B2-C3 is assumed to have the highest indicator value and therefore merged. This new ecotope has six neigbors, still B1, C1, C2, C4 and newly added D1, D2, and it results in 63 additional combinations. The algorithm iterates as long as no new combination achieves a higher value than in the previous iteration. With this short description the advantages of the AMOEBA for the delineation procedure are revealed: Any path dependency is avoided, by comparing all possible value-increasing combination with no (pre-) specified direction of the algorithm. Due to this characteristic amoeba-shaped markets can occur, but geographically contiguous functional regional markets are ensured.

In **step 2: proximity measure calculations,** the intra-flow indicators for the potential FEMAs are computed. The intra-flow indicator describes the connection between spatial units and expresses the basic idea of the delineation procedure. The consideration is that, from a starting spatial unit with each added neighbor, the indicator increases until it reaches a maximum with the highest possible share of internal total flows within a market, the FEMA. Beyond this point of expansion any additional units will lead to a decrease of the indicator, because the spatial units are less integrated (or more independent) to the spatial units within the aggregated market than those among themselves.

First the initial basis intra flow indicator (base.ifi) is presented, since it is also the base term of the afterwards suggested trend-adjusted intra flow indicator (adj.ifi). The base.ifi is defined as the proportion of internal on overall (internal + external) flows of a potential market.

$$ifi.base_{M=\bigcap k_{i=1}, k_{i=2}\ldots} = \frac{\sum I_i + \sum O_i}{\sum I_j + \sum O_j} \qquad (2)$$

with
- I = inflows,
- O = outflows,
- i = index for all spatial units (k) within the market M; i = 1, 2…, Ω(M)
- j = Index for all spatial units (k) of the larger overall space; j = 1, 2,…, Ω(S); Ω(M) ≤ Ω(S);
- M = Set of spatial units included in the market, M = $\bigcap k_i$
- Ω(M) = No. of spatial units within the market

The value range of ifi.base is zero to one, one if all flows are within the market and zero if all flows are external flows going beyond defined market boundaries. Although every inflow from A to B is an outflow from B to A ($\sum O_i$) and the sum of inflows equals the sum of outflows ($\sum I_i = \sum O_i$) for spatial units within a market, this is

[15] The computational efficiency loss through this is tolerable, but for samples with higher numbers of neighbors it might be recommended to exclude once not considered spatial units from the further algorithm.

mostly not the case for market external flows and thus total flows, which is revealed by unequal column and row totals of the flow matrix ($\sum I_i \neq \sum O_i$).

The simple and easily understandable proximity measure base.ifi indicates directly the share of intra-market flows. However, for an effective operating procedure, the proximity measure of the algorithm has to be independent from the number of spatial units of a market, but in most regional geographic applications the base.ifi tends to increase with each additional included spatial unit and even slightly with distant spatial units. If there is a systematic dependence of the base.ifi on the amount of included spatial units, which is labeled as spatial trend in analogy to trend components in time series analysis, it needs to be corrected. If such a spatial trend is present, the algorithm on the basis of base.ifi probably will not stop the expansion at its optimal level, and may even convert to a result. An intuitively explication for the necessity of a correction term is that, when all spatial units of the overall space are merged to one market, it results in a base.ifi of one, which is the maximum value. In any way, a spatial trend will lead to a longer range of the first merged markets and crowds out smaller markets. Literally formulated, larger markets swallow smaller markets in their neighborhood. Within the delineation procedure, this effect exacerbates with the third step of the delineation methodology, in which the preliminary FEMAs are chosen according to the value of the proximity measure.

The phenomenon of dependency within a data-generating process is of a general nature and, on a disaggregated data level, it is common that nearby spatial units are highly correlated. In analogy to autocorrelation of time series analysis where values are correlated with their own time delayed values, the correlation of neighboring spatial units is called spatial autocorrelation, and in network analysis and for flow data it is called network autocorrelation. In the present case it results in high amount of flows between nearby spatial units. This relation will be displayed in the application examples with a preceding summary of the underlying matrix to each procedure. For German commuting data it shows, for example, that the amount of flows in the first circle of neighborhood is very high and then falls rapidly with each additional circle of neighborhood level. That near things are more related than distant things is a common observation and corresponds to First Law of Geography (see TOBLER 1970, p. 236). However, although the procedure builds to some extent on this relation with its evolutionary progress over spatial neighbors it leads to a non-stationary data process that has to be corrected.

The idea behind the adjustment and the adj.ifi is to eliminate the spatial trend. An additional spatial unit is included in the market area if the improvement on the intra-market share of flows is above the pure trend related increase. In analogy to time series analysis the adjustment is done by subtracting the significant trend from the base.ifi values for potential markets in the FEMA-Pool(i). The trend parameters are estimated by ordinary least squares in each iteration. The calculation formula for adj.ifi is shown in formula (3), with the trend parameters α for the intercept and β for the slope, whereby just β is a crucial parameter for the delineation procedure and α a pure level effect, M is the set of spatial units which are included in the market, Ω(M) the number of included spatial units, i the index for the corresponding iteration, $\overline{ifi.base}_i$ the average ifi.base in FEMA-Pool$_i$, and Ω(c) depicts the number of potential markets.

$$adj.ifi_{M,i} = ifi.base_{M,i} - (\alpha_i + \beta_i \Omega(M_M)) \tag{3}$$

with:
$$\beta_i = \frac{\sum_{c=1}^{\Omega(c)} (\Omega(M)_{c,i} - \overline{\Omega(M)_i}) \cdot (ifi.base_{c,i} - \overline{ifi.base_i})}{(\Omega(M)_{c,i} - \Omega(\overline{M})_i)^2}$$

$$\alpha_i = \overline{ifi.base_i} - \beta_i \; \overline{\Omega(M)}_i$$

For the adj.ifi any additional inclusion of a spatial unit will be “punished“ by subtracting the trend influence. The trend parameters are determined over the complete FEMA-Pool(i) and can differ slightly between iteration. If insignificant or irrelevant, the adj.ifi reduces to ifi.base.

Plots for values for base.ifi and adj.ifi values show that the results on adj.ifi correspond better to the idea of the delineation procedure. A disadvantage of the adjustment is that the indicator loses its intuitive understandability. For the quality of the global delineation result the weighted average of base.ifi seems appropriate, because it is nothing else than the share of intra-market flows. However, it needs to be seen together with the indicated number of FEMAs.

The **step 3** is the simultaneous **determination of preliminary FEMAs** by selecting them from the FEMA-Pool. These markets are formed or selected according to the height of the proximity measure; and once assigned spatial units are blocked within the iteration to form any other FEMA, so each spatial unit can just be assigned once. Among the competing potential markets, the delineated market with the highest contribution to an improvement of the intra-market share of flows is chosen.

Application Example 1: State of North Rhine-Westphalia

The application example is calculated on the German commuting dataset on employees subjected to social insurance contribution from the Federal Employment Agency in Germany for the year 2006 at district level (“*Landkreise und kreisfreie Städte*, 413 units, see Bundesagentur für Arbeit 2010) First the delineation analysis is performed on a subset that comprises the 54 districts of the State of North Rhine-Westphalia (NRW)[16], an area with rather complex commuting structures which is chosen for illustration purposes. Afterwards nationwide delineation results on the basis of all 413 districts are presented.

The procedure is applied for two variants of commuting flow matrices: matrix M1 considers district internal commuting flows and thus the districts self-sufficiency as degree of autonomy, and the other matrix M2 with a diagonal of zeros considers only flows between districts. Summaries of both matrices are given in Table 1. They show that in M1 5.2 million flows are considered of which about 65% (3.4 million) remain within districts. These 3.4 million district internal flows are not included in M2, which considers only the remaining 1.85 million flows across district borders. NRW districts are contiguous in average to 5.1 neighboring districts towards about 1.3 million employees commute, which is about 25% of the flows in M1 and 70 % in M2, or cumulatively regarded 90% and 70% respectively.[17] Within the second neighborhood, which is the district itself, its neighbors and the neighbors of neighbors 97% or 92% of commuting flows take place, and in the third neighborhood this are 99% or 97%, respectively. A note for the computational efficiency of the procedure is that, within the third

[16] NRW (North Rhine-Westphalia) is with 18 million residents the most populated German state, and has about 5 million commuters. It’s characterized through its heterogeneity by different landscapes of the historical regions Rhineland, Westphalia, Lipper Land and further subdivisions, as well as the polycentric Ruhr Area which overlaps between historical and administrative regions.
[17] As previous under methodology discussed are neighbors defined on the queen continuity.

neighborhood, a district has in average 14 neighbors and for the processing power more important in maximum 23 neighbors which enables 8,338,608 possible market combinations. Although matrices that contain less than 15% zero values[18] cannot really be described as sparse matrices, the summary indicates that commuting is mainly taken place in the close geographical neighborhood and therefore particularly well suited for the evolutionary spatial clustering procedure. The sample size with 1.3 million flows on 137 linkages[19] in the first neighborhood can be regarded as sufficiently large.

Table 1: Summaries for the two commuting matrices for 54 NRW districts of 2006

	M1: sum	M1: in %	M1: in cum %	average no. of links	Most Connected
Total Flows	5,258,149	100%			
within flows	3,405,099	65%	65%	1	1
flows to 1st neighb.	1,302,830	25%	90%	5.1	11
flows to 2nd neigb.	394,041	7%	97%	10.8	19
flows to 3rd neigb.	107,275	2%	99%	14	23
flows to all further units	48,904	1%	100%		
zero flows	376	13%			

	M2: Sum	M2: in %	M2: in cum %	average no. of links	Most Connected
Total Flows	1,853,050	100%			
within flows	0	0%	0%	1	1
flows to 1st neighb.	1,302,830	70%	70%	5.1	11
flows to 2nd neigb.	394,041	21%	92%	10.8	19
flows to 3rd neigb.	107,275	6%	97%	14	23
flows to all further units	48,904	3%	100%		
zero flows	430	15%			

Overviews of the two delineation procedures on M1 and M2, using the adj.ifi as proximity measure, common boundary (queen continuity) as neighborhood definition, and a disjunctive definition are given in Tables 2 and 3 with the procedural output tables. Both applied procedures result for the subset of 54 NRW districts 16 delineated FEMAs, which, however, are shaped differently as later is illustrated and discussed. For the delineation based on M1, 86% (4,494,930) of commuting flows take place within the 16 FEMAs^{M1}, of which, however, 65% (3,405,099) are internal values as listed in Table 2 and thus within the markets per definitions. From the 1,853,050 flows across district borders are therefore 59% (1,089,831) captured within the FEMAs$_{M1}$. The ratio of 59% for the FEMAs^{M1} is slightly higher than for the FEMAs^{M2}, which only capture only 54% (1,006,127) of total flows between districts. Based on this indicator the result FEMAs^{M1} are superior to FEMAs^{M2}. However as noted on the TTWA typical criteria of homogeneity, the range of included districts show that the FEMAs^{M1} are more heterogeneous in size FEMAs^{M2}. Besides the information on structure and quality, are in Tables 2 and 3 indicators for the distribution of the proximity measure adj.ifi and the processing power

[18] In Matrix M2 54 entries of 2,916 are per definition zero.
[19] 54 districts with in average 5.1 neighbors.

provided.[20] One is the maximum of nb counts which shows that for the processing power decisive maximum number of newly added neighbors to each predefined core within the iteration. For the delineation process on M1 were in total 40,078 potential FEMAs regarded with markets that in maximum included 23 districts, but of which only 14 neighbors were newly considered in the iteration to a core of 9 districts. The procedure based on M2 features a smaller pool of potential FEMAs and a stronger spatial trend. The latter seems plausible due to the not considered self-sufficiency of districts. In both procedures is the major part of assignment already done in the first iteration.

Table 2: Summary delineation process on M1 NRW subset

Setting		**M1 NRW (internal values), adj.ifi, queen, no fuzzy**		
Iteration		Iteration 1	Iteration 2	**Iteration 3=4**
structure	no. of FEMA	18	17	**16**
	min no of units FEMAs	1	1	**1**
	mean no of units in FEMAs	3.3	3.5	**3.8**
	max no of units in FEMAs	8	10	**10**
quality	FEMA internal flows	4,364,772	4,438,655	**4,494,930**
	FEMA external flows	893,377	819,494	**763,219**
	total flows	5,258,149	5,258,149	**5,258,149**
	share of internal flows	83.01%	84.41%	**85.49%**
	min base.ifi	0.5962	0.5962	**0.6397**
	average base.ifi	0.8088	0.8117	**0.8267**
	max base.ifi	0.9452	0.9452	**0.9452**
proximity measure	min adj.ifi	-0.0889	-0.0938	**-0.0462**
	average adj.ifi	0.1125	0.1071	**0.1136**
	max adj.ifi	0.2497	0.2459	**0.2377**
	trend parameter intercept	0.6627	0.6653	**0.6745**
	trend parameter slope	0.0112	0.0124	**0.0114**
	R^2 Trend	0.0977	0.6118	**0.6538**
	adj.R^2 Trend	0.0975	0.6117	**0.6538**
processing	dimension FEMA-Pool	6,337	26,158	**39,982**
		14	21	**23**
	max. of nb counts	11	14	**14**
	calculation time (min)	1	4	**4**

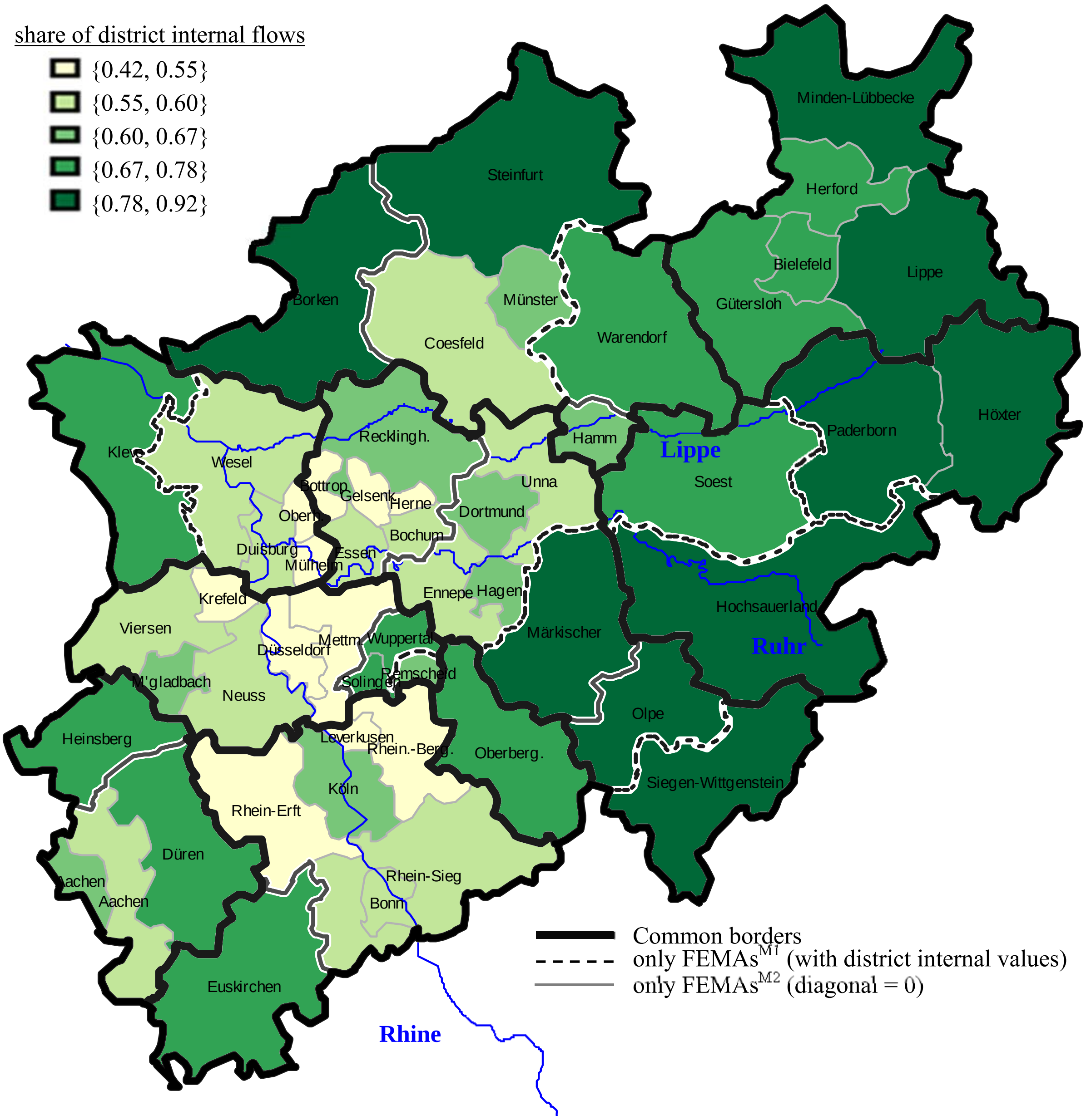
[20] R version 2.13.2 (2011-09-30) with 2.67 GHz processor and 8,00 GB RAM.

Table 3: Summary delineation process on M2 NRW subset

Setting		M2 NRW (diagonal of zeros), adj.ifi, queen, no fuzzy	
Iteration		Iteration 1	**Iteration 2=3**
structure	no. of FEMA	16	**16**
	min no of units FEMAs	21	**1**
	mean no of units in FEMAs	3.5	**3.7**
	max no of units in FEMAs	6	**6**
quality	FEMA internal flows	929,539	**1,006,127**
	FEMA external flows	923,511	**846,923**
	total flows	1,853,050	**1,853,050**
	share of internal flows	50.16%	**54.30%**
	min base.ifi	0	**0**
	average base.ifi	0.3748	**0.365**
	max base.ifi	0.7774	**0.7774**
proximity measure	min adj.ifi	-0.0878	**-0.1625**
	average adj.ifi	0.1781	**0.1195**
	max adj.ifi	0.5061	**0.4749**
	trend parameter intercept	0.0419	**0.1275**
	trend parameter slope	0.0459	**0.035**
	R^2 Trend	0.4197	**0.76**
	adj.R^2 Trend	0.4196	**0.76**
processing	dimension FEMA-Pool	6,337	**26,936**
		14	**21**
	max. of nb counts	11	**14**
	calculation time (min)	1	**4**

The geographical shapes of the two delineation results are mapped together in Figure 3, where the underlying districts are colored according to the district internal flows. For a quality analysis the results will be afterwards illustrated separately in Figure 4, where the FEMAs are colored according to the proximity measure adj.ifi. and the more intuitively interpretable base.ifi, respectively. In general the breaks of the coloring interval correspond to quintiles and a darker color indicates higher integration intensity. In Figure 3 the common market borders, which result in both delineation procedures, are drawn with big black lines and are clearly in the majority. Market borders only for FEMAs^{M1} are drawn with dashed black lines on white ground, and market borders that only results for FEMAs^{M2} are drawn with grey lines on white ground. For orientation reasons are drawn the three main rivers in the state in blue lines, Rhine, Ruhr and Lippe.

Figure 3: Delineation results for NRW



Although the great majority of borders are defined commonly, identically delineated are only the FEMAs in the northeast around *Bielefeld* and the single district FEMA of *Oberbergischer Kreis* in the central south. With the occurring differences in Figure 3, several effects can be observed. For the interpretation it has been emanated from a change of the underlying flow matrix from M2 to M1, by additionally considering district internal flows. Figuratively this is shown with the elimination of the grey lines and the addition of the dashed black lines. The first and most obvious effect is that FEMAs^{M2} consisting of large-area districts with high shares of district internal flows are separated and form FEMAs^{M1} at a smaller scale likely with a single district. This effect will be labeled as autonomy of large-area districts and regards the TTWA criteria of sufficient self-containment of flows. This effect is clearly visible in the east of the state, which are rather rural areas. A second effect is that smaller FEMAs^{M2} (often single district) contiguous to bigger, strongly interlinked FEMAs with low shares of district internal flows are merged together. This is labeled as big-neighbor-market-effect and is, for example, the case for *Euskirchen* and *Heinsberg* in the southwest, as well as for the districts *Wuppertal and Solingen* toward

Düsseldorf, although *Remscheid* remains as a single district FEMA with only a medium level internal district flows. The opposite is shown for *Kleve* in the west, however, with a reasonable high share of district internal flows (0.77). Furthermore, larger shifts of market borders have been observed. One for the Ruhr Area with the FEMA(s) *Essen-Bochum* and *Dortmund-Hagen*, where, by considering district internal flows, the former two FEMAs are merged to one by only excluding the district *Märkischer Kreis.* As kind of the big-neighbor-market-effect, the shift results due to the consideration of district autonomy, the two big neighbors are merged. To this much larger area the additional contribution of *Märkischer Kreis* is then too small. Thus the assignment in the contiguous southeast region changes, because if the *Märkischer Kreis* is excluded from the FEMA along the Ruhr, it merges with *Olpe*, and in turn *Siegen-Wittgenstein* is excluded. For itself the three districts of the southeast region qualify for the first described effect of autonomy, in which, however, the merge of *Märkischer Kreis* with *Olpe* is inconsistent and needs to be examined more closely in the following quality analysis. The second larger shift can be observed for the large-area districts in the northwest, where *Borken* is added to the FEMA with *Münster*, as in the big-neighbor-market-effect, but on the other side *Warendorf* is excluded and merges with *Hamm,* with which the linkage was not strong enough without district internal flows and is, therefore, questionable as well. This presentation invalidates a possible initial assumption that the consideration of district internal flows would yield to a higher amount of delineated market due to the self-efficiency and autonomy of some districts. The effect is, indeed, true but counteracted by the big-neighbor-market effect and so a reduction or increase of defined markets depends on which effect prevails.

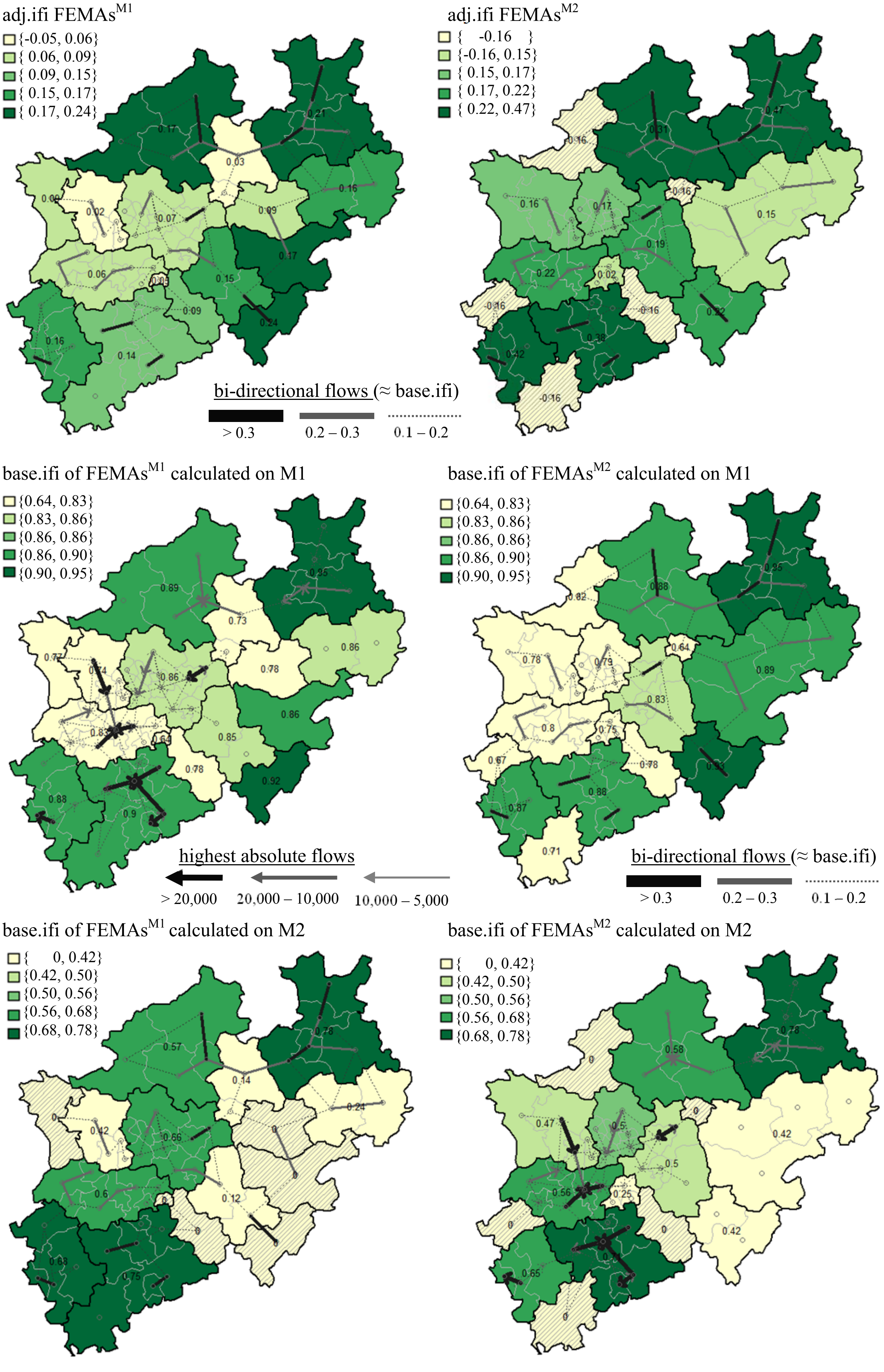
The quality of market reflection is discussed with the maps in Figure 4. On the left side are mapped the FEMAs^{M1} and on the right side the FEMAs^{M2}. In the first row the FEMAs are colored according to the proximity measure adj.ifi and the strongest relative bi-directional commuting flow connections are drawn, which are calculated as the base.ifi but without taking the district internal flows into account. The big black lines represent linkages with values over 0.3 (16 linkages), medium grey lines for values between 0.2 and 0.3 (30) and small light grey lines for values between 0.1 and 0.2 (82). However, the adj.ifi of the two definitions are not comparable to each other, due to different trend corrections. Therefore, in the second and third rows are the base.ifi mapped for the FEMA, which are comparable to each other and intuitively interpretable with their range from 0 to 1. For an appropriate comparison are, in the second row for both definitions, the base.ifi's calculated on the M1, and in the third row for both definitions on M2. The breaks of the coloring class intervals are defined each time with quintiles on values for the delineation result on the original underlying matrix. In two of the four maps are additionally drawn the strongest absolute commuting flows with big black arrows for over 20,000 commuters (9 flows), medium grey arrows for 10,000 to 20,0000 commuters (14) and small light grey arrows for 5,000 to 10,000 commuters (66).[21] Since for single district FEMAs the base.ifi calculated on M2 is per definition zero and the adj.ifi is the negative sum of the trend parameter (here -0.16), these markets are marketed with diagonal lines. For the district names it is referred to Figure 3. So the first row in Figure 5 with adj.ifi is the used proximity measure of the delineation procedure shown, the second row with base.ifi on M1 illustrates the self-containment of FEMAs, and the third row with base.ifi on M2 the inter-linkage between districts within a market.

[21] Thus 89 highest flows out of 2916 (= 54^2) connection or 2486 none zero flows. The distribution and choice of displayed interval is illustrated in the attachment.

Starting with the interpretation of the drawn flows, the maps in Figure 4 show that, generally, both FEMA definitions capture the most relevant commuting flows within the defined market. The borders of FEMAs^{M1} cross only once a bi-directional flow with a value over 0.3 (in the south-east) and four times with flows between 0.2 and 0.3. The border crossing in the southeast of the one district FEMA *Siegen-Wittgenstein* seems plausible and acceptable due the high degree of autonomy (0.92). The same applies for *Soest* and *Hochsauerlandkreis*, all concerning the autonomy effect of large-area districts. The other three crossings of defined borders raise questions. First the two at *Warendorf* in the north, which forms a relatively low integrated FEMA with *Hamm* and lies between two highly interlinked FEMAs with *Münster* and *Bielefeld.* The second is the crossing between *Märkischer Kreis* and *Hagen.* Note that both questions have already been raised in the previous discussion on effects. For the borders of FEMA^{M2} no relative bi-directional flow over 0.3, and only one with a value between 0.2 – 0.3 crosses a market border, which concerns again the district *Warendorf*. This assessment on bi-directional flows basically can be confirmed with the highest absolute flows, where of FEMAs^{M2} no border is crossed by a flow over 10.000, and of FEMAs^{M1} only one. This one indicates that the district *Warendorf* is absolutely stronger linked to the western FEMA with *Münster* and less to *Hamm* or the easterly FEMAs with *Bielefeld*. This conclusion is supported by adj.ifi's, which indicate a relatively weak integration for the FEMA^{M1} *Hamm-Warendorf* and high adj.ifi's for the FEMA$_{M2}$ with *Münster-Warendorf*. The illustration of delineation quality, together with the descriptive flows, suggest hence that *Warendorf* is located between the two highly integrated FEMAs with *Münster* and *Bieldefeld*, linked to both of them, but is rather assigned to the FEMA *Münster*. The intermediate location is probably amplified by the wide area of the district. Ahead in text, the reassessed FEMA with *Münster-Warendorf* is also the result of the nationwide analysis on both matrices. The apparently false assignment in the subset delineation might be caused by border effects, since the FEMA with *Münster* is strongly linked to districts in the contiguous state of Lower Saxony and to the neighborhood to *Hamm*, which is a specific case.

The maps show for the district *Hamm* and as well for *Remscheid* a quite difficult market assignment, because both are with a medium high share of district internal commuting (0.64) rather not self-sufficient and with relevant links to all of their surrounding districts, which are in highly integrated FEMAs. Although they are related to their neighbors, no linkage is sufficient to form a common market, so is Hamm located between *Warendorf* which should be accounted to the well-integrated FEMA with *Münster* (0.88), the FEMA with *Dortmund-Hagen-(Essen)* (0.86 or 0.83), and the single district FEMA *Soest* (0.78) or a larger rural FEMA in the east (0.89), respectively. Although *Soest* has, for a large-area rural district, a surprisingly low degree of autonomy, the linkage to *Hamm* is not sufficient to from a common market and if *Soest* is rather considered to their easterly neighboring districts. An analog interpretation can be made for Remscheid, with the districts *Wuppertal* and *Solingen* to the FEMA with *Düsseldorf* (0.83), FEMA *Köln-Bonn* (0.9) in the south and *Oberbergischer Kreis* (0.78). Both are, together with *Warendorf,* likely intermediate or overlaying markets, which will be viewed in detail later with the fuzzy set extension.

Figure 4: Quality of delineation results for NRW



Furthermore, low adj.ifi's and base.if's result for the FEMA^{M1} with Duisburg-Wesel at the upper Rhine section. This might be caused by three reasons: first, by relevant linkages with easterly districts of the neighboring FEMAs with *Essen*; second, the high absolute flow to Düsseldorf which is, however relatively less important for both; and third, the connection of Wesel to Kleve which is less relevant for Duisburg. While the region between Duisburg and Essen along the Ruhr is a (well known) potential region of overlay markets, the flow between *Düsseldorf-Duisburg* might be interpreted as a kind of market openness. In all cases, the addition of the districts to a market is not high enough to raise the integration of the whole market, as it is on a small scale for Kleve to the FEMA with *Duisburg*. *Kleve* is relatively important to the district Wesel which is within the FEMA *Duisburg*, but the linkage is not important enough to the whole region. The same can be seen for Hamm to Unna. That this question of overlays and degree of openness occurs for the polycentric area along the rivers Ruhr and Rhine in NRW, often referred and partly cooperated in as metropolitan region RheinRuhr, was assumable because it is well known in the planning literature and regional policy practice (see e.g. Blotevogel 2005, pp. 47-70, or Blotevogel 1998). It is indicated by rather low adj.ifi and base.ifs on M1, but high base ifi's on M2 tantamount to highly relevant bi-directional and absolute flows. However, the separated markets of the lower Rhine area (FEMA with *Köln-Bonn*) and the Rhine Area toward the Ruhr Area are clearly indicated by the market integration indicators as well as descriptive drawn flows.

The district autonomy effect is shown in Figure 4 with low adj.ifi and base.ifi on M2 for the larger markets, but high adj.ifi and base.ifi on M1 for the smaller areas as well as for the larger area on M1. For the big-neighbor-market-effect, it is shown with base.ifi of FEMAs^{M2} on M1 first, the differences in integration of the big market and contiguous single district FEMA, and second, that the added districts do not have the degree of autonomy as the identified single district FEMAs (excepted *Borken,* which is certainly discussable). Alone on their linkages the added districts do not raise the proximity measure in the procedure for the potential FEMA, because it is for the highly integrated big market on the whole not important enough. However, together with the relatively high share of internal flows, the added districts increase the integration. That this effect relates to the whole area can be well seen by the example of *Heinsberg,* which is relatively highly linked to two districts within the FEMA with *Aachen*, but the strongest absolute flow goes to the other FEMA with *Möchengladbach*. This makes it another interesting viewing region for the following analysis of overlaying markets and degree of openness. One self-evident interpretation of the big-neighbor-market effect is that, for the added region, the continuous larger and already highly integrated market is of importance, while for the larger market the added district is of less importance. The added districts can, to some part, already be seen as sub-market and are therefore also illustrated for the later following nationwide delineation result. Besides the discussed overlaying intermediate markets, the degree of border openness, and the sub-market character of added single-district FEMAs^{M2} to big neighbor markets, another form of sub-markets and way of differentiation patterns within the delineated FEMAs is indicated by the drawn flow. Particularly visible for the FEMA with *Düsseldorf* where stronger linkages are indicated for the west part (*Krefeld-Viersen-M'gladbach*), the center (*Düsseldorf-Neuss-Mettmann)* should to some extent be open to the north to *Duisburg*, and a lesser linkage with the eastern part (*Wuppertal and Solingen).* For the FEMA of the Ruhr Area such sub-areas are indicated as well, namely with the FEMA a western part with *Essen,* an eastern part with *Dortmund*, and southeastern sub-market with *Hagen,* as well as probably the eastern contiguous FEMA with *Duisburg*. Thinkable is also a retroactive adjustment of the numerical optimal market definitions, by separating FEMAs with negative adj.ifi or very low values (e.g.

outliners). Depending on the delineation purpose and used data input (e.g. type as M2) this might be useful, but it contradicts with the objective to minimize arbitrariness by choosing subjective parameters and correction.

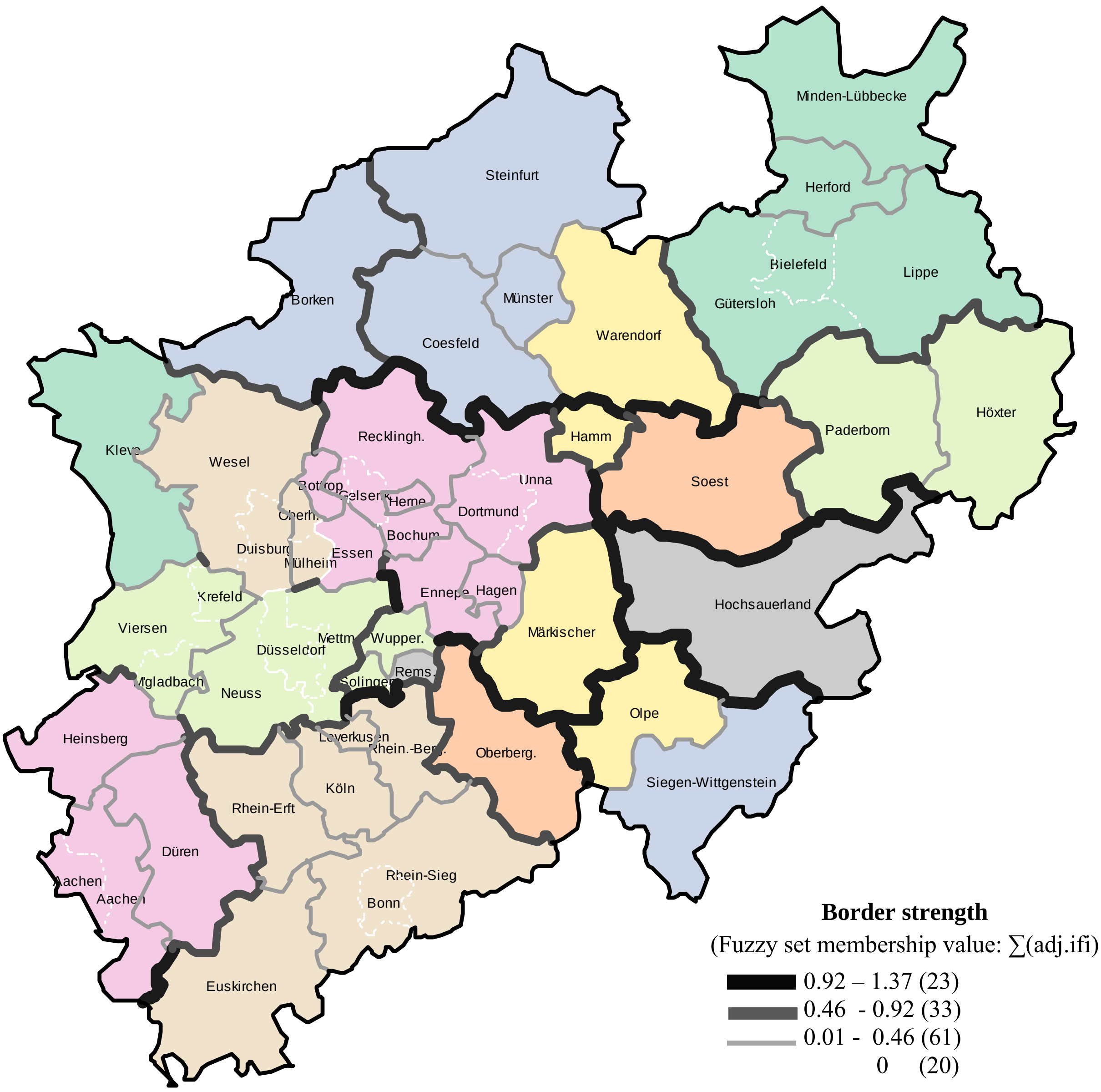
The maps in Figure 4 confirm the conclusion which was made on the numeric output table that the applied delineation procedure captures the majority of commuting linkages, but also illustrates that there a still relevant commuting trip between the *FEMAs,* especially in areas with complex commuting structures as the RhineRuhr area. However, it also indicates overlays and intermediate market situations for larger rather rural districts as, for example, *Warendorf* or *Heinsberg*. Therefore, in Figure 5 a fuzzy set delineation result according to the idea of FENG (2009) is provided as a possible methodology extension. To enrich the delineation procedure with a fuzzy set (overlapping) output of FEMAs the presented delineation procedure has to be slightly modified. First, a fuzzy membership function has to be established that allocates to each element a certain degree of border strength. As elements are here chosen the district borders of which the NRW subset consists 137. Further, it was chosen to omit the third step in each iteration so that, for each spatial unit, there exists one optimal corresponding FEMA to which a value (e.g. adj.ifi or base.ifi) can be assigned. In the ideal case the process results for each district to the same FEMA results, which would then be identical to the original result.

In Figure 5, one thinkable fuzzy extension is illustrated on the previous presented delineation procedure on M1 without step 3. The districts are colored to distinct between the previously disjunctive delineated FEMAs^{M1}. The presented fuzzy membership function is the accumulation of adj.ifi over each of the 54 individual calculated FEMAs to the corresponding district borders. The resulting values for the set of 137 borders range between 0 and 1.37 and are drawn for three equal classes plus a forth class for zero values. Borders that are assigned to the class of highest values are drawn with big black lines, the second in medium grey lines and the third in thin light grey lines. The value of 0 occurs for 20 borders to which no FEMA-border is defined. These are separated by white lines and would also remain for alternative membership function. Further more sophisticated fuzzy membership functions are thinkable.[22] The fuzzy presentation of functional markets delivers a more realistic and accurate presentation of market borders, especially in polycentric regions with complex flow structures. A disadvantage is that the geographic presentation can be challenging and suffers in clarity and simplistic view of maps. Moreover, it is confirmed that disjunctive and ubiquitous functional market definitions are imperfect, but so are the fuzzy set definitions. Both differ in their degree of abstraction from reality.

Generally the fuzzy delineation result confirms clearly the functional market borders along the river Lippe and the definition of the two larger FEMAs in the north as discussed before. The autonomy of the larger rural districts in the East and the three FEMAs with *Köln-Bonn*, *Aachen* and *Oberbergischer Kreis* in the south-west is confirmed as well. In contrast are the markets along the Ruhr and upper Rhine area (without *Köln-Bonn*) which illustrate more permeable borders.

[22] Some alternative calculated fuzzy membership functions are presented in an analog way in the attachment.

Figure 5: Fuzzy delineation results for NRW based on M1



For the individually discussed assignment, the fuzzy set illustration shows that the overlay of *Warendorf* is less than it had seemed in the previous quality analyses and clearly supports the argumentation to assign *Warendorf* to the FEMA with *Münster*. Further, it shows the high autonomy of *Borken*, which was before identified as a sub-market of the big-neighbor-market-effect to the FEMA with *Münster*. Confirmed as well is the classification of *Hamm* as an intermediate market, and even somewhat indicates an openness towards *Soest* and *Dortmund-Unna* than *Warendorf*. For the larger, more open market border for the Ruhr and upper Rhine area that consists of FEMAs with *Düsseldorf Duisburg*, *Essen*, *Dortmund*, *Hagen*, as well as *Kleve* and *Remscheid*, more permeable sub-markets are indicated. Only the named sub-market of *Wuppertal-Solingen* and *Remscheid* is shown more autonomy than it was concluded before. For the southwest, it is shown that the district *Märkischer Kreis* is rather a sub-market of the southern sub-market of the Ruhr Area with *Hagen*, and that the two districts of *Siegen-Wittgenstein* and *Olpe* form a common market. The fuzzy delineation result also confirms the previously concluded openness of the FEMA with *Aachen* towards the eastern part of the FEMA with

Düsseldorf over the district *Heinsberg*. As certainly a common market are shown the districts of 20 borders (white lines), which are assigned to one market from every starting point as one area and correspond therewith to the ideal case of a functional area. These are *Bielefeld-Gütersloh-Lippe*, *Dortmund-Unna*, the center of the Ruhr-Area with *Recklinghausen-Gelsenkirchen-Bochum*, Duisburg-Wesel, *Aachen* city region *(*since 2009 one district*)*, *Bonn-Rhein-Sieg,* as well as the western and middle submarkets in the FEMA with *Düsseldorf,* where the openness of *Düsseldorf-Duisburg* occurs in all individual 54 optimal FEMAs, but not if the highest selected ones are chosen. This can be explained in that this definition was blocked in the procedure, because one of the districts was previously assigned to a more highly integrated definition. The same can be seen for *Essen-Mülheim-Oberhausen*, so that the interpretation of a market overlay seems the most plausible one. Both the western and eastern neighbors of *Essen* are strongly connected to *Essen* but not to each other, so that the area gets split.

For the first robustness check, slightly modified procedures were applied to test the impact of the chosen proximity measure and neighborhood criteria. First, in alternative to the adj.ifi as a proximity measure the original suggest base.ifi is applied. The results differed in so far that the procedure resulted in just three markets on M1, one with the two *Aachen* districts, a second with seven districts of *Ostwesfalen-Lippe* (FEMAs^{M1} with *Bielefeld* and *Höxter-Paderborn*), and a third for all other 45 districts; and only two markets on M2 where the two *Aachen* districts are merged with the other 45 districts. These results are obviously unsuitable, and the procedure on all 413 German districts is likely to collapse by reaching computational limitations. As neighborhood criteria the queen continuity was set as a default by the AMOEBA. Procedures alternatively using k-nearest neighborhood criteria converge with an increasing number of k (considered neighbors) to the delineation results based on the queen continuity, but are less efficient on the delineation principle to maximize the share of market internal flows and considered potential market.[23]

Another aspect is the stability over time and randomness in the definition caused by insufficiently large cases in input data sets. A variation over time is omitted here –first, because the application example is applied on a total survey consisting of over five million flows on 54 districts; second, the emphasis is on the illustration of the procedure; and third, due to the high data cost. For a regional market definition the need for a time variation increases when subgroups (for example by age, education, or on a smaller reference level) are viewed, and the amount of flows per district is lower. The regional markets may differ over time, as in the case of commuting due to changes in transportation costs and needs; but, on the basis of the used spatial reference level of German districts, these changes can be assumed as rather long-term developments. For a conducted time variation it can referred to the application on migration data in Michels et al. (2011) which found for the period 2000 to 2008 only minor variations. Another check for the robustness of the procedure is to compare it with the results that occur by randomizing the neighbor as in a permutation test. The delineation procedure on M1 on randomized neighborhood criteria in ten trials did not exceed a total share of market internal flows of 70% by 65% per definition with district internal flows with a number of markets between 35 and 48, in comparison to the presented results with 85.5% to 16 FEMAs. In a nationwide analysis the differences are even more apparent, since the distance between randomly assigned neighbors is higher and the share of districts within a common (poly-centric) region is less. So it is very likely that the districts are not randomly assigned within the procedure.

[23] Tabulated numerical summary of a selection of processed alternative delineation procedures analogues to table 3 are given in the attachment.

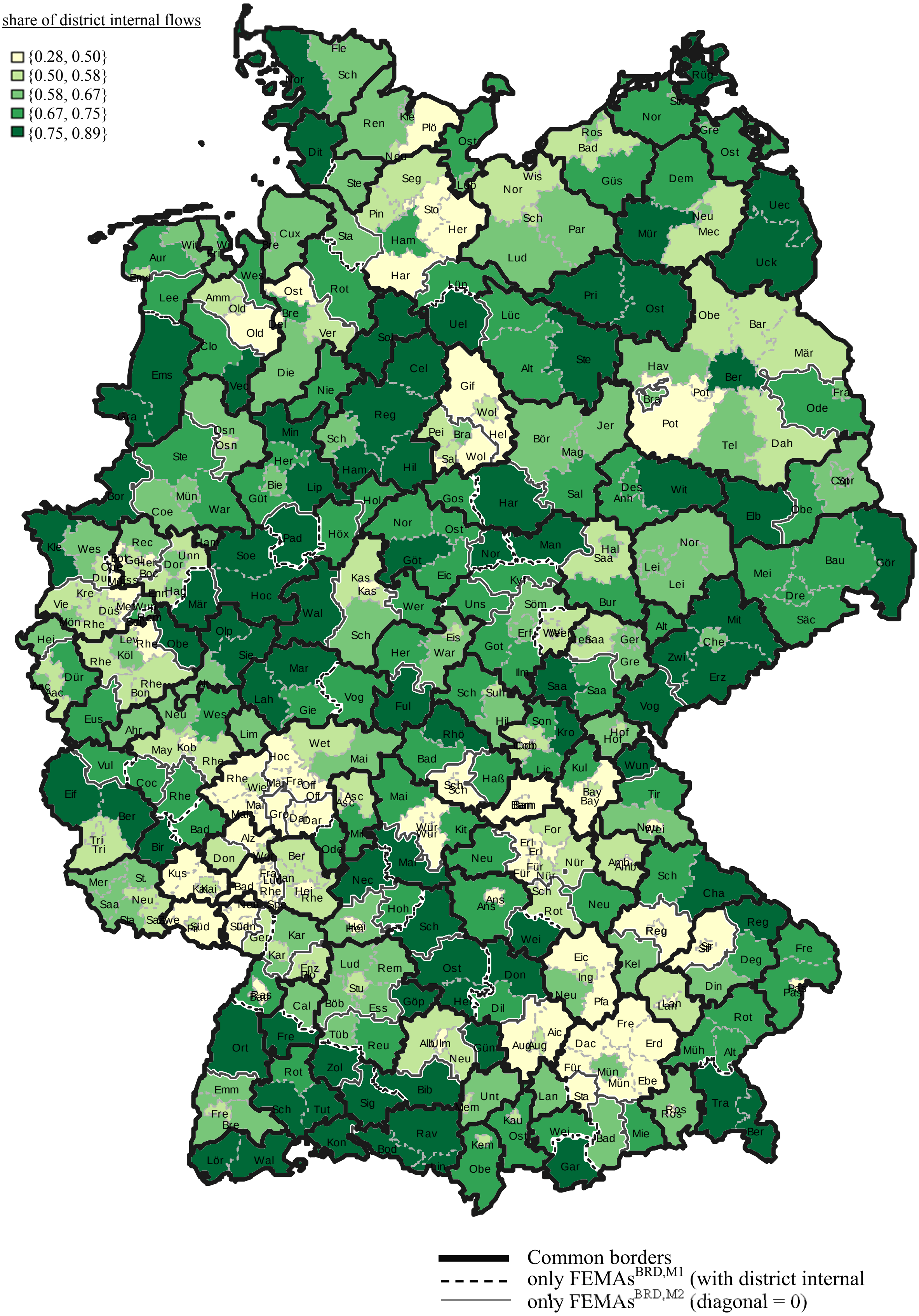
Application Example 2: Germany

In the following, the nationwide delineation results are presented and discussed to classify the results toward common functional concepts for Germany. The FEMAs for Germany on both types of matrices for the year 2006 are presented in Figure 6. Although the numerical quality criteria showed the results on matrix M1 as superior, results on both matrices are presented due to, among others, their interpretation as sub-markets and larger cooperation areas (big-neighbor-market effect). The delineation procedure on the nationwide matrix of M1 results in 110 FEMAs^{BRD,M1} and captures 86.4% (23.1 million) of commuting flows and 63.2% (6.2 million) of cross- border commuting flows, respectively. When only considering district external flows with a matrix type M2, the procedure results in 145 FEMAs^{BRD,M2} and captures 53.4% (5.3 million) of cross-border flows, but with smaller and more homogenous market areas.[24] Toward the previously mentioned possible initial assumption, that including district internal flows leading to smaller market extensions, the opposite is shown. The designated big-neighbor-market outweighs the district autonomy effect. With calculation time that took several days and a FEMA-pool dimension of over half a million potential markets with up to 30 included districts, the nationwide application reaches the computational limitations of the current software implementation. The main driver is the number of newly considered neighbors in an iteration to form potential FEMAs, caused by the combinatory within the AMOEBA step. This performance issue could, in future research, be addressed with parallel computing (multi-task processing)[25] restrictions on possible considered additional neighbors within an iteration (for example, with a combination of the queen and k-nearest neighbors or stages of maximums commuting distances) to apply the procedure on very large data sets and structures with high numbers of neighbors.

The nationwide results are mapped in Figure 6 for both matrix types in the same way as before. A comparison of the underlying data matrices yields that the delineation procedure on M1 leads to larger areas for population and economic centers, both in monocentric regions as well as polycentric regions. This applies not only for the largest well-known metropolitan regions as *Berlin*, *Hamburg*, *Frankfurt*, *Stuttgart* and *München*, but can also be seen for smaller centers as, for example, *Münster-Osnabrück*, *Regensburg* and *Kassel.* Most of the times the big-neighbor-market effect goes along with the autonomy effect of large-area districts by taking one of the large-area districts out of that FEMA^{BRD,M2} and reassigning and separating the remaining districts.

[24] A detailed tabulated numerical summary of the delineation processes can be found in the attachment.
[25] The Software RHIPE would be an obvious choice to continue with the current program implementation in R, which allows R users to analyze larger complex data with parallel computation across subsets. See for further details: www.rhipe.org.

Figure 6: Delineation results for Germany



A detailed discussion of individual assignments and fuzzy set extensions for the NRW subset is omitted here. Instead, the $FEMAs^{BRD}$ are compared to alternative regional market definitions that are basically based on commuting patterns, in a first step to comparable ubiquitous functional and administrative concepts, and later in addition to non-ubiquitous concepts. The alternative concepts are listed and compared in Tables 4-6 with their number of defined regional markets, the corresponding number of district borders defined as market borders, the distribution of the number of included districts within a market, and the flows captured with defined market borders on both types of matrices. Further is listed the distribution of base.ifi on M1 and, as an example, the annual employment growth rate between 1999 and 2006 with the associated Moran's I as an indicator for spatial dependency and the internal versus external variation with the standard deviation. The alternative spatial reference levels are the functional concepts of 50 labor market regions (LMR KS) and 105 local labor market regions (LLMR KS) both by Kropp and Schwengler (2011), the 96 planning regions (*Raumordnungsregionen*, ROR) by BBSR (2011), the 147[26] (150) local labor markets (LMR Eckey) by Eckey et al. (2006), as well as the administrative concepts of 39 NUTS-2 regions of Eurostat, as well as classification the 413 districts and the aggregation of the whole nation (overall).

In addition some metropolitan and agglomeration concepts are considered which are not directly according to the ubiquitous characteristic. These are the eleven European Metropolitan Regions (EMR) defined by their own assignment by district decision makers, the 55 metropolitan regions (Metros) of the European Commission DG Regio, and the 19 agglomeration (Agglo).[27] In order to adjust them to a ubiquitous characteristic, two versions are considered: one where the remaining periphery is merged to one area, and another one where each remaining district forms a periphery market on its own (EMR*, Metros*, Agglo*). The last could serve since there are mostly large area districts with likely high degrees of autonomy. By the comparison it has to be noted that the FEMAs are specifically delineated with data from 2006, while the input data of the alternative delineation differs and includes longer periods, so the FEMAs are slightly favored. There are also other differences possible regarding the reference level (e.g. municipalities) and restrictions (e.g. no markets across state borders are allowed for the planning regions).[28] However, since the alternative definitions ultimately propose an assignment on a district level which is basically based on commuting patterns, they are compared to each other.[29]

[26] The original definition by Eckey at al. (2006) accounts for 150 labor market regions, but due the merger of districts as part of district reforms in the States of Saxony-Anhalt and Saxony (nationwide that lead to a reduction from 439 districts to 413) the labor markets needed to be reassigned and some were merged through this too.

[27] For the EMR see BBR and IKM (2008) and Rusche and Oberst (2010) for the used geographical form, for Metros see Dijkstra (2009) and for the Agglomeration Eltges (2008) f.

[28] Same accounts of course as well for the quality criterion, so when Kropp and Schwenger (2008, 2011) optimize their delineation on the Q modularity measure of Newman (Newman and Girvan 2004), their own procedure will likely perform better on this. Same is true here for the base.ifi.

[29] The 171 housing market regions of Michels et al. (2011) and the 94 migrations regions in Schlomer (2009) are not included in this comparison, since they are based on migration data.

Table 4: Number of districts and borders of regional market definitions

definitions	no.	number of included districts					borders	
		max	mean	median	min	sd		
Overall	1	413	413	413	413	-	**0**	0%
NUTS-2	39	23	10.6	11	1	5.4	**386**	36%
LMR KS	50	40	8.3	5	1	8.2	**366**	34%
ROR	96	9	4.3	4	1	1.9	**598**	56%
LLMR KS	105	18	3.9	3	1	3.4	**566**	53%
FEMAsM1	110	14	3.8	3	1	2.7	**564**	53%
FEMAsM2	145	9	2.8	3	1	1.6	**670**	62%
LMR Eckey	147	12	2.8	2	1	2.1	**671**	62%
districts	413	1	1	1	1	0	**1074**	100%
EMR	11+1	33	21	19	13	7.3	**284**	26%
EMR*	195	33	2.1	1	1	4.9	**617**	57%
Metro	55+1	15	3.4	3	1	2.6	**452**	42%
Metro*	282	15	1.5	1	1	1.5	**846**	79%
Agglo	19+1	17	8.1	8	3	4.2	**263**	24%
Agglo*	279	17	1.5	1	1	2.1	**902**	84%

Table 5: Market internal flows of regional market definitions

definitions	no.	flows within markets/ regions			
		matrix M1		matrix M2	
Overall	1	26,738,130	100,0%	9,841,082	100.00%
NUTS-2	39	23,075,221	86.30%	6,178,173	62.80%
LMR KS	50	24,360,648	91.10%	7,463,600	75.80%
ROR	96	21,856,378	81.70%	4,959,330	50.40%
LLMR KS	105	22,889,535	85.60%	5,992,487	60.90%
FEMAsM1	110	23,111,799	86.40%	6,214,751	63.20%
FEMAsM2	145	22,155,690	82.90%	5,258,642	53.40%
LMR Eckey	147	22,168,087	82.90%	5,271,039	53.60%
districts	413	16,897,048	63.20%	0	0.00%
EMR	11+1	24,508,365	91.70%	7,611,317	77.30%
EMR*	195	22,408,656	83.80%	5,511,608	56.00%
Metro	55+1	22,534,438	84.30%	5,637,390	57.30%
Metro*	282	20,985,149	78.50%	4,088,101	41.50%
Agglo	19+1	24,059,322	90.00%	7,162,274	72.80%
Agglo*	279	21,058,214	78.80%	4,161,166	42.30%

Table 6: Characteristics of regional market definitions

Definitions	base.ifi(M1)					example employment growth rate 99-06			
	max	mean	median	min	sd	Moran's I	Pr(>|t|)	mean internal sd	external sd
overall	100.00%	100.00%	100.00%	100.00%	-	-	-	6.30%	-
NUTS-2	94.80%	86.10%	87.50%	68.20%	5.90%	0.52	0	4.60%	3.80%
LMR KS	95.00%	88.20%	88.70%	79.40%	3.90%	0.62	0	4.00%	4.90%
ROR	92.30%	81.80%	83.10%	58.20%	6.70%	0.58	0	4.00%	4.80%
LLMR KS	95.00%	82.70%	83.00%	67.50%	5.60%	0.53	0	3.90%	5.10%
FEMAsM1	94.30%	81.50%	84.60%	41.80%	9.20%	0.48	0	3.60%	5.20%
FEMAsM2	92.60%	79.10%	80.50%	41.80%	8.70%	0.51	0	3.50%	5.80%
LMR Eckey	92.10%	81.00%	80.70%	64.20%	5.20%	0.39	0	3.60%	5.80%
districts	89.00%	62.00%	63.00%	28.10%	12.70%	0.54	0	-	6.30%
EMR	95.40%	91.30%	91.60%	84.80%	2.70%	0.2	0.06	5.10%	2.70%
EMR*	95.40%	67.10%	67.60%	36.00%	12.70%	0.52	0.05	5.00%	6.10%
Metro	92.70%	79.30%	80.60%	59.90%	7.50%	0.12	0.05	3.70%	4.40%
Metro*	92.70%	69.40%	71.30%	36.00%	12.30%	0.55	0	3.70%	6.20%
Agglo	94.30%	87.40%	87.10%	80.70%	3.20%	0.27	0.01	4.30%	3.00%
Agglo*	94.30%	66.90%	68.60%	36.00%	13.10%	0.56	0	4.20%	6.20%

The definition of 50 LMR KS (labor market regions by Kropp and Schwengler) captures with over 90% most of the commuting flows in 2006. The FEMAs^{BRD,M1} capture slightly less with 86%, but allow for 60 more markets and are thus almost as efficient but on a much smaller scale. The definition of 147 LMR Eckey is rather similar to the FEMAs^{BRD,M2}, but in comparison slightly better in performance. NUTS-2 regions and ROR have a considerably lower share of captured flows than the next on number of markets following functional market concept (for example LMR KS with 50 markets captures 91.1/ 75.8%, while the 39 NUTS-2 regions capture only 86.3/ 62.8%). Based on this criteria using data for the year 2006, the 110 FEMAs^{M1} capturing 86.4/ 63.2% can be seen as superior to the 105 LLMR KS that capture only 81.7/ 50.4%. The three functional concepts--50 LMR KS, 110 FEMAs^{BRD,M1} and 147 LMR Eckey--are superior to the other ubiquitous concepts in this interpretation. However, the FEMAs^{BRD,M1} and LMR Eckey are more homogenous in size, but the 50 LMR KS capture more flows. In contrast to Kropp and Schwengler (2008), the interpretation here is that FEMAs^{BRD,M1} and LMR Eckey are more effective in capturing, since there need far fewer regional merges of districts to obtain a comparable degree of integration.

As Royuela and Vargas (2007) argue that, despite differences for their Catalonian housing markets, their definitions on both commuting and migration data are valid, here it is concluded that the 50 LMR KS, the 110

FEMAs^{BRD,M1} and the 147 LMR Eckey are all valid and represent different zooms for disaggregated analysis of the German labor markets. The LMR with 50 regional markets are rather large commuting regions, and even in an overall perspective the regional development can be well recognized in its entirety. As it is shown in Table 3, the range and variation of included districts to a market is high, while there is a more homogeneous blend on the base.if indicator. Using the example for the employment growth rate, their variation within the markets (mean internal sd) is higher and between the markets (external sd) lower than for the FEMAs^{BRD,M1} and LMR Eckey. The FEMAs^{M1} are one zoom, and the LMR Eckey a second zoom further disaggregated on the local labor markets. Their range of included districts into a market declines, but on the variation of FEMAs^{M1} base.ifi is higher. On the other hand, in the example on employment growth, the variation of FEMAs^{M1} within the markets is as low as for the LMR Eckey by a lower external variation. From this level, a more detailed illustration of labor markets should be done, and differentiation of the market borders (as in the earlier-presented fuzzy extension or market segments as labor market – for example by level of education, professions, or age groups) should be used. For the discussed definitions, this and their deviations and inaccuracies are shown with the following graphical intersection analysis of LMR K/S, FEMAs^{M1} and LMR Eckey.

The comparison with metropolitan concepts shows that those cover already with 22.4 million (83.8%; 56.0%, EMR*), 21.0 million (78.5%; 41.5%, Metros*), and 21.1 million (27.8%; 42.3%, Agglo*) the most of the 26.7 million. or 9.8 million commuting flows, respectively. As to size, the EMR correspond rather to the LMR K/S, and Metros* to FEMAs^{M1} and LMR Eckey. The indicator of Moran's I show for all functional and administrative a highly significant and relevant spatial dependency. To avoid the need of quite complex spatial parameters in econometric estimation on a disaggregated data level, only the not ubiquitous metropolitan concepts are approximately appropriate, and there rather the Metros. With almost 21 million (78.5%) of their counterpart Metros* shows that they cover already the major of all commuting flows of 26.7 million.

The application example is concluded with a graphical comparison in an intersection analysis by measuring overlaying borders.[30] The largest overlay can be found between the FEMAs^{BRD,M1} with the LLMR K/S which assign to 85% same markets borders. Note that the LLMR are a differentiation of LMR K/S. Of 1047 district borders, 485 are assigned commonly as market borders, 429 commonly as district borders within markets, and there are only 79 borders defined as borders in FEMA^{M1} but not as local labor markets, and 81 the other way around. This overlay is even larger than the one between FEMAs^{BRD,M1} and FEMAs^{BRD,M2} that show only an overlay of 80%, which is the same level of overlay as M1 has with the common borders of the regional labor markets of Kropp and Schwengler (2011) and Eckey (2006).[31]

The analysis of overlays in the graphical illustration is limited to the three definitions of the 50 LMR of Kropp and Schwengler (2011), the 110 FEMAs$_{M1}$ and the 147 LMR of Eckey, which result in seven possible overlays and thus a quite complex illustration. Besides the different combinations of overlay, for illustration reasons again the highest absolute flows are drawn within the map and the districts colored to the shares of district internal flows. The detailed design of the graphical overlay illustration in Figure 7 is described together with the results. Out of 1047 borders between German districts, 739 are defined as market borders in at least one of the three definitions, the other 308 are in all three definitions accounted to one region and drawn by dashed white lines. Of

---

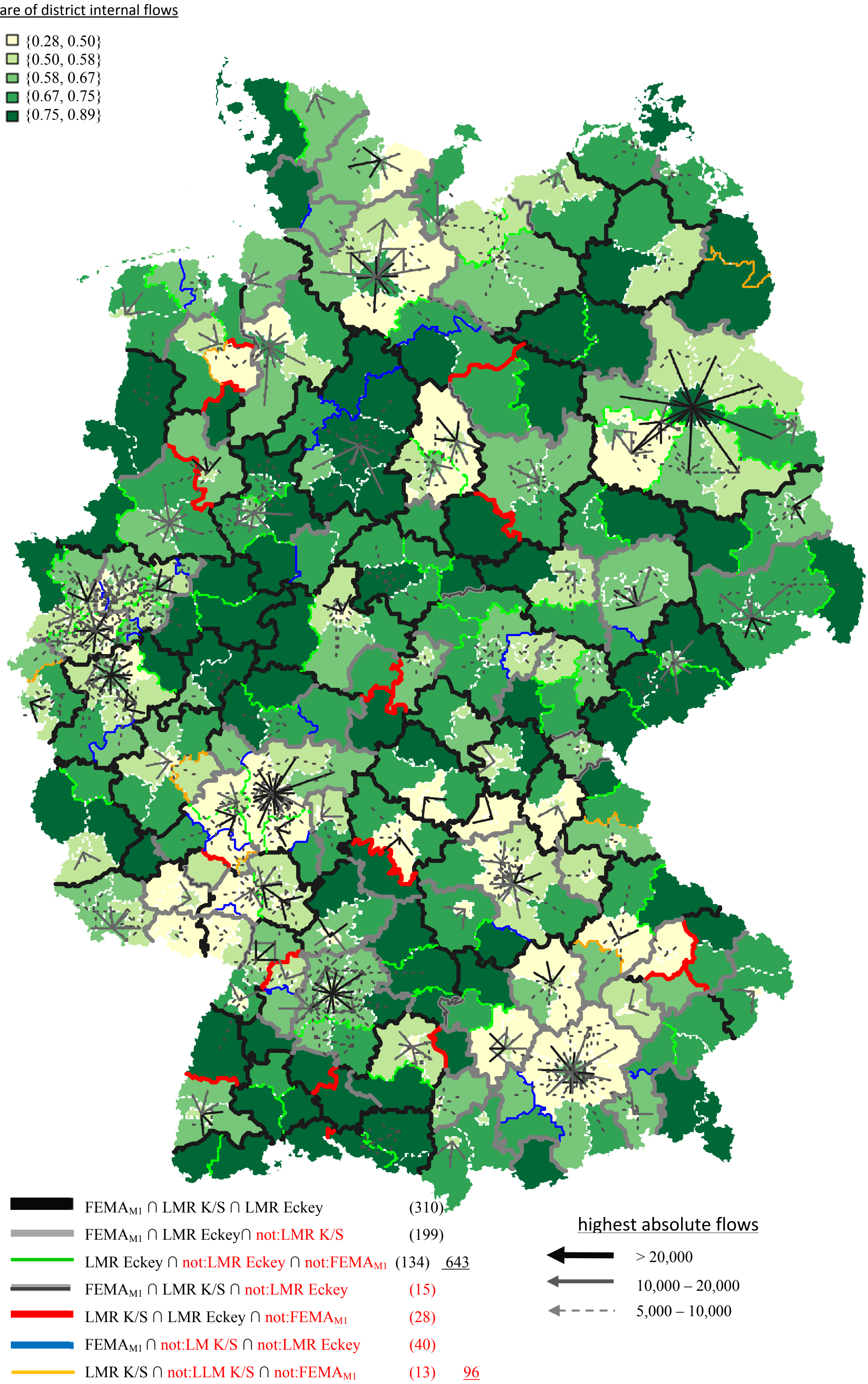
[30] An alternative way would be the spatial overlay index (*Räumliche Überlagerungs Kennziffer*) by Rusche (2009), calculated by the amount of congruent districts divided by amount of districts in definition one, minus the not congruent districts divided by the amount of districts in the compared definition from; which was used to compare labor and housing markets.
[31] A detailed overview is listed in the appendix.

739 defined market borders, 310 (42%) are commonly in all three definitions. These borders are drawn with big black lines and should be fairly uncontroversial. Borders that occur only in the smaller-scale delineations of the 110 FEMAs^{M1} and 147 LMR of Eckey are drawn with big grey lines and are the case for 199 (27%) borders. Together with the drawn highest absolute flows the scale interpretation is clearly visible for the greater area of *München* (in the southeast), for which the FEMAs^{M1} and Eckey definition identifies several independent sub-centers within the larger area of LMR K/S of *München.* The same applies among others for *Stuttgart, Heilbronn, the Rhein-Ruhr area,* and *Hamburg*. The even finer delination can be made with the LMR by Eckey. These additional borders that only occur in the Eckey definition are drawn with green lines, quantifies in 134 (18%) borders and concerns mainly rural area districts and polycentric city regions as the RheinRuhr, RheinMain (with *Frankfurt*), or *Wolfsburg-Braunschweig*. These 643 (87%) assigned borders fit into the interpretation of the three zoom levels to explain the differences in geograpical shape and numbers of markets.

The remaining 96 borders do and have to be seen as deviations and inaccuracies. Of those borders, 28 occur in the LMR K/S and in LMR of Eckey, but not in the FEMAs$_{M1}$ and are drawn with red lines. In the illustration it can be seen that most affected are large area districts within a FEMAS^{BRD,M1} with an higher share of internal flows contiguous to densly interlinked markets. Only a few separation are on a larger scale, as for example the merge of the LMR Münster and LMR Osnabürck to one FEMA[32], or LMR Karsruhe and LMR Pforzheim in the southwest. The 15 borders defined in LMR K/S and FEMA^{BRD,M1}, but not in LMR Eckey are likely inaccurancies and are drawn with thin black lines on a grey ground. Furhermore 13 borders are definied only for the 50 LMR K/S, which are marked with thin orange lines and 13 borders result only for FEMAs^{M1}, which are marked with thin blue lines. These borders also can be interpretated as inaccurancies in one of the definitions. A note on this is that the FEMAs^{M1} shape of the *Hannover* and *Munich* market is clearly smaller, and the NRW subset discussed intermediate market situation of *Hamm* and the overlay of *Essen* as overlaying markets might be random to some extent as to which side these districts are accounted.

[32] The common market might be anecdotally justified with the Airport Münster/Osnabrück (FMO).

Figure 7: Overlay alternative market definitions



Summary and future Development

The special characteristic of the presented methodological approach is that no prior core definition is necessary. Such preliminary determinations are quite common in related research and might be in some cases legitimate (for example, centers of labor markets in metropolitan areas where commuting patterns are aligned to an employment centers). In other cases such as the residential location decisions, they are already not justified from a theoretical point and are arbitrary. Considering urban developments like post-suburbanization and raising polycentric city regions or rural areas without dominating cores, this assumption becomes more and more doubtful. Within the presented procedure are no artificially-set restrictions like threshold values or neighborhood limitations, as often in alternative delineation procedures, sometimes hidden behind complex calculation methods. This kind of restriction might be only applied if computational limitations are reached, although the possibilities to improve the computational implementation are far from exhausted and are open to further research. In contrast to some other cluster analyses, the procedure avoids path depend since it is possible that, within the procedure, existing clusters are resolved to allow newer and "better" clusters, even possibly that the number of markets raises again as it arose in some of the randomized neighborhood applications. Most especially, the AMOEBA step ensures that geographically contiguous markets occur and allows that irregular shaped (amoeba-shaped) markets are defined. The outcome of the simultaneous definition of markets in Step 3 is that, to each delineated market, potential neighbor markets are formed as counterweights and ensures a disjunctive and ubiquitous (nation-wide) definition of markets. Alternatively to that can serve the presented fuzzy set extension. The presented iterative three-step procedure to delineate FEMAs is, with its iterative approach, under explicit consideration of the geographical neighborhood structure the appropriate delineation method, especially for cases where near linkages are the theoretically basis of a FEMA, as for regional labor and housing markets. An appropriate regional delineation helps to limit the problem of the Modifiable Areal Unit Problem, but since it is often based on an administrative spatial reference level it can't completely eliminate it.

In future work the presented methodology should be utilized together with other recent approaches, like the graph theory approach of Kropp and Schwengler (2011), the particular application of a factor analyses by Eckey et al. (2006), other cluster analysis techniques, in particular the evolutionary ones of Flórez-Revuelta et al (2008, 2006), and traditional methods as the threshold procedure. This should be implemented in a common software environment for more transparency in delineation procedures. For the comparison of published regional market definition are always two causes possible for variations, one can be found in the methodology, the other in the data input. A simplified comparability seems especially useful since no delineation procedure is theoretically superior, and all are more or less numerical optimization procedures. In this common software environment a set of various quality indicators such as the base.ifi or the Q modularity measure of Newman (Newman and Girvan 2004) should also, be integrated.

With regard to expectations, it is important to know that a perfect regional mapping is not possible, because the one region does not exist. A real economic market region is multilayered, complex, and heterogeneous, and most of the time can be divided into further sub-markets. Functional delineated markets deliver a simplified picture of the true markets, and this needs to be chosen depending on the research question and policy issue. Even fuzzy set definitions, are in the end, imperfect. The presented delineation method is one procedure which can be applied for these definitions and is especially useful when spatial neighbors should explicitly be used in the delineation process on flow data. For defined regional markets in Germany, the application examples show that the

FEMAs^{M1} are a compromise between the large area scaled labor market definition of Kropp and Schwengler (2011) and the smaller scaled labor markets of Eckey et al. (2006) as both the numerical comparison and mapped flows show.

The quite detailed and intensive discussion on effects, overlays, and openness is seen as useful to determine what is behind the numeric optimization procedure. It should help for the acceptance of policy decision makers and can lead to derive different individual regional cooperation strategies besides the numerical suggested optimal definition, although the presented results and analysis of the example have to be seen primarily as an example, and less as a regional market proposal. Some research questions could require some reassignments, for that the discussion should have shown starting points. It might even indicate new worthwhile areas of comparison – for example, on the economic development of *Hamm* and *Remscheid* which are both in a position of an intermediate market. The advantage of the methodology procedure should, in future research, also be proven on alternative applications in other research fields besides regional development, such as organizational subdivision, transport systems and markets and more.

To alleviate doubts on functional delineation in areas with complex flow structure with imperfect definitions, if disjunctive or fuzzy, it should be noted that these imperfect structures are not random and still a better spatial reference level for evaluations and analysis than administrative districts or even overall national averages. In any regional definition, there will always be commuting trips between designed areas, and in the close neighborhoods they are likely to be relevant. In order to choose the appropriate definition, it is necessary to know what stands behind the numerical optimization by transparent procedures, but also not to be arbitrary by c hosen parameters of the analyst. Both apply for the presented methodology.

Acknowledgements

I thank, Norbert Hiller, Harald Hecking Srinivas Parinandi and the participants of the 13th INFER Annual Conference (London) for their very helpful comments and suggestions.

A precursor procedure was applied on migration data to define Housing Market Regions in Germany (see MICHELS/ OBERST/ HILLER, 2011) and discussed at a workshop on the same topic in Münster February 2011, both within a research project commissioned and financed by the WÜSTENROT STIFTUNG. This research provided a valuable foundation for the presented methodology approach to delineate Functional Economic Market Areas. The present paper

References

Aldstadt J, Getis A (2006) Using AMOEBA to Create a Spatial Weights Matrix and Identify Spatial Clusters. Geographical Analysis 38:327-343.

Anselin L (1995) Local Indicators of Spatial Association-LISA. Geographical Analysis 27:93–115

Anselin L (1988) Spatial Econometrics: Methods and Models. Kluwer Academic Publishers Dordrecht

Antikainen J (2005) The Concept of Functional Urban Area, Findings of the ESPON Project 1.1.1. Informationen zur Raumentwicklung 7.2005:447-452.

Ball R M (1980) The Use and Definition of Travel-to-Work Areas in Great Britain: Some Problems, Reg. Studies 14:125-139. DOI:10.1080/09595238000185121

BBR Bundesamt für Bauwesen und Raumordnung; IKM Initiativkreis Europäische Metropolregionen in Deutschland (2008) Regionales Monitoring 2008. Bonn. http://www.region-frankfurt.de/media/custom /1169_2602_1.PDF?1225098167. Accessed 10 March 2012

BBR Bundesamt für Bauwesen und Raumordnung (2000) Raumordnungsbericht 2000, Bonn, Berichte Band, http://www.bbsr.bund.de/cln_032/nn_22548/BBSR/DE/Raumentwicklung/RaumentwicklungDeutschland/Raumordnungsberichte/ROB2000/ROB2000__pdf,templateId=raw,property=publicationFile.pdf/ROB2000_pdf.pdf. Accessed 10 September 2011

BBSR Bundesinstitut für Bau–, Stadt– und Raumforschung (2010): http://www.bbsr.bund.de/cln_016/nn_340582/BBSR/DE/Raumbeobachtung/Werkzeuge/Raumabgrenzungen/Raumordnungsregionen/raumordnungsregionen.html?__nnn=true. Accessed 10 September 2011

Beckmann J M (1996) Spatial Equilibrium in Labor Markets. In: Van den Bergh JCJM, Nijkamp P, Rietveld P (eds) Recent Advances in Spatial Equilibrium Modeling, Methodology and Applications. Springer, Berlin Heidelberg New York, pp 111–117

Blotevogel H H (2005) Vom Ballungsraum zur Metropolregion – Selbstfindung an Rhein und Ruhr. In: Meffert H, Steinbrück P (eds) Trendbuch NRW, Perspektiven einer Metropolregion. Verlag Bertelsmann Stiftung, Gütersloh, pp. 47-70

Blotevogel H H (1998) Europäische Metropolregion Rhein-Ruhr, Theoretische, Empirische und Politische Perspektive eines Neuen Raumordnungspolitischen Konzeptes. Institut für Landes- und Stadtentwicklungsforschung des Landes Nordrhein-Westfalen, Dortmund

BMWI Bundesministerium für Wirtschaft und Technologie (the German Federal Ministry of Economics and Technology) (2011), Gemeinschaftsaufgabe. http://www.bmwi.de/BMWi/Navigation/Wirtschaft/Wirtschaftspolitik/Regionalpolitik/ gemeinschaftsaufgabe,did=163706.html. Accessed 20. September 2011

Boatti G (1988) I sistemi urbani, struttura del territorio lombardo. Edilizia Popolare 25:2-27

Bode E (2008) Delineating Metropolitan Areas using Land Prices. Journal of Regional Science 48.1:131-163

Böltken F (1996) Neuabgrenzung von Raumordnungsregionen nach den Gebietsreformen in den Neuen Bundesländern. Arbeitspapier 5/1996 der Bundesforschungsanstalt für Landeskunde und Raumordnung, Bonn.

Casado-Díaz J M (2000) Local Labor Market Areas in Spain: A Case Study. Regional Studies 34.9:843-856.

Christaller W (1933) Die zentralen Orte in Süddeutschland: Eine ökonomisch-geographische Untersuchung über die Gesetzmäßigkeit der Verbreitung und Entwicklung der Siedlungen mit Städtischen Funktionen. Gustav Fischer Verlag, Jena.

Coombes M, Bond S (2008) Travel-to-Work Areas: the 2007 Review. Office for National Statistics, UK.

Coombes M G, Green A E, Openshaw S (1986) An Efficient Algorithm to Generate Official Statistical Reporting Areas: The Case of the 1984 Travel-to-Work Areas Revision in Britain. Journal of the Operational Research Society 37.10:943-953.

Cörvers F, Hensen M, Bongaerts D (2009) Delimitation and Coherence of Functional and Administrative Regions. Regional Studies 43.1:19-31.

Dawson J A, Findlay A M, Sparks L (1986) Defining the Local Labor Market: An Application of Log Linear Modeling to the Analysis of Labor Catchment Areas. Environment & Planning A 18:1237-1248.

Dijkstra L (2009) Metropolitan Regions in the EU. Regional Focus 01/2009. http://ec.europa.eu/regional_policy/sources/docgener/focus/2009_01_metropolitan.pdf. Accessed 10 September 2011

Eckey H-E, Kosfeld R, Türck M (2006) Abgrenzung deutscher Arbeitsmarkt-regionen. Raumforschung und Raumordnung, 4:299-308

Elhorst J P (2010) Applied Spatial Econometrics: Raising the Bar. Spatial Economic Analysis 5.1:5-28

Eltges M (2008) Das Ruhrgebiet – eine regionalwirtschaftliche Analyse. Informationen zur Raumentwicklung 9/10:535-547

Eurostat (2010) European Regional and Urban Statistics Reference Guide, eurostat Methodologies and Working Papers. http://epp.eurostat.ec.europa.eu/cache/ITY_OFFPUB/KS-RA-10-008/EN/KS-RA-10-008-EN.PDF. Accessed 20 September 2011.

Eurostat (1992) Study on Employment Zones. Eurostat (E/LOC/20), Luxembourg

Feng Z (2009) Fuzziness of Travel-to-Work Areas. Regional Studies 43.5:707-720

Flórez-Revuelta F, Casad-Díaz J.M. Martínez-Bernabeu L (2006) An Evolutive Approach for the Delination of Local Labor Markets. Lecture Notes in Computer Science 4193/2006:342-351

Flórez-Revuelta F, Casad-Díaz J M, Martínez-Bernabeu L, Gómez-Hernández R (2008) A Memetic Algorthm for the Delineation of Local labor Markets. Lecture Notes in Computer Science, 5199/2008:1011-1020

Fox K A, Kumar T K (1965) The Functional Economic Area: Delineation and Implications for Economic Analysis and Policy. Papers and Proceedings of the Regional Science Association 15:57–85

Goodman J F B (1970) The Definition and Analysis of Local Labor Markets: Some Empirical Problems. British Journal of Industrial Relations 8:179–196.

Hirschle M, Schürt A (2008): Suburbanisierung … und kein Ende in Sicht? – Intraregionale Wanderungen und Wohnungsmärkte. Informationen zur Raumentwicklung 3/4:211–227

Jones C (2002) The Definition of Housing Market Areas and Strategic Planning. Urban Studies 39.3:549-564

Karlsson C, Olsson, M (2006) The Identification of Functional Regions: Theory, Methods, and Applications. Annals of Regional Science 40.1:1–18.

Klemmer P (1976) Problem einer arbeitskräfterelevanten Typisierung von Regionen. In: Kommission für wirtschaftlichen und sozialen Wandel (eds.), Band 78, Teil B, Göttingen, pp. 175–274.

Kulldorf M (1997) A Spatial Scan Statistic. Communications in Statistics - Theory and Methods 26:1487–1496.

Kropp P, Schwengler B (2011) Abgrenzung von Arbeitsmarktregionen – ein Methodenvorschlag. Raumordnung und Raumforschung 69:45-62.

Kropp P, Schwengler B (2008) Abgrenzung von Wirtschaftsräumen auf der Grundlage von Pendlerverflechtungen, Ein Methodenvergleich. IAB-Discussion Paper 41/2008, Research Institute of the German Federal Employment Agency.

Lösch, A (1938) The Nature of Economic Regions. Southern Economic Journal 5.1:71-78.

Michels W, Rusche K (2008) Abgrenzung von Wohnungsmarktregionen mit Hilfe von Arbeitsmarktverflechtungen. Materialien zum Siedlungs- und Wohnungswesen, Band 43 - Münster

Michels, W, Oberst, C, Hiller, N (2011): Wohnungsmarktregionen in Deutschland – Abgrenzung und Struktur funktionaler Wohnungsmärkte, Materialien zum Siedlungs- und Wohnungswesen 45, Münster, also online publication on the website of the Wüstenrot Stiftung. http://www.wuestenrot-stiftung.de/pdf/Wohnungsmarktregionen.pdf. Accessed 20 September 2011

Newman M E J, Girvan M (2004) Finding and Evaluating Community Structure in Networks. Physical Review E 69.2:26113(15)

Nystuen J D, Dacey M F (1961) A Graph Theory Interpretation of Nodal Regions. Papers and Proceedings of the Regional Science Association 7.1:29-42

OECD Organisation for Economic Co-operation and Development (2002) Redefining Territories: The Functional Regions. OECD Publication, Paris

OMB Office of Management and Budget (2000) Standards for Defining Metropolitan and Micropolitan Statistical Areas. Federal Register 65(No. 249), December 27, 2000, pp. 82228–82238

Ord J K, Getis A (1995) Local Spatial Autocorrelation Statistics: Distributional Issues and an Application. Geographical Analysis 27:286–306

Papps K, Newell J O (2002) Identifying Functional Labor Market Areas in New Zealand: A Reconnaissance Study Using Travel-to-Work Data. Discussion Paper 443, Institute for the Study of Labor (IZA), Bonn.

Patuelli R, Reggiani A, Gorman S P, Nijkamp P, Bade F-J (2007) Network Analysis of Commuting Flows: A Comparative Static Approach to German Data. Networks and Spatial Economic 4:313-331.

Royuela V, Vargas M (2007) Defining Housing Market Areas using Commuting and Migration Algorithms. Catalonia (Spain) as an applied case study. http://www.ub.edu/irea/working_papers /2007/200707.pdf. Accessed 20 September 2011

Rusche K (2009) Abgrenzung von Wohnungsmarktregionen mithilfe von Wanderungsverflechtungen: eine vergleichende Fallstudie. Raumordnung und Raumforschung 1:34–43

Rusche K, Oberst C (2010) Europäische Metropolregionen in Deutschland – eine regionalökonomische Evaluation. Raumforschung und Raumordnung 68:243-254

Schlömer C (2009) Binnenwanderungen in Deutschland zwischen Konsolidierung und neuen Paradigmen. Bundesinstitut für Bau–, Stadt– und Raumforschung (Hrsg.), Berichte Band 31, Bonn.

Sforzi F. Openshaw S, Wymer C (1997) Le procedura di identificazione dei sistemi locali del lavoro (The procedure to identify local labor market area). In Sforzi, F (ed.) I sistemi locali del lavoro 1991, pp. 235-242. ISTAT, Rome

Slater P B (1976) A Hierarchical Regionalisation of Japanese Prefectures using 1972 Inter-Prefectural Migration Flows. Regional Studies 10:123-132

Smart M W (1974) Labor Market Areas: Uses and Definition. Progress in Planning, 2.4:238–253

Stigler G, Sherwin R (1985) The Extent of the Market. Journal of Law and Economics 28:555-585

Tobler, W. R. (1970) A Computer Movie Simulating Urban Growth in the Detroit Region. Economic Geography 46.2:234-240.

Van der Laan, L (1998) Changing Urban Systems: An Empirical Analysis at Two Spatial Levels. Regional Studies 32.5:235–247

Van Nuffel (2007) Determination of the Number of Significant Flows in Origin-Destination Specific Analysis: The Case of Commuting in Flanders. Regional Studies 41.4:509-524

Ward J H (1963) Hierarchical Grouping to Optimize an Objective Function. Journal of the American Statistical Association 58:236–244

West Midlands Regional Assembly (ed., 2006): Study into the Identification and Use of Local Housing Market Areas for the Development of the Regional Spatial Strategy. Report of Initial Findings, C3205, Birmingham

Appendix

Appendix 1 Robustness check

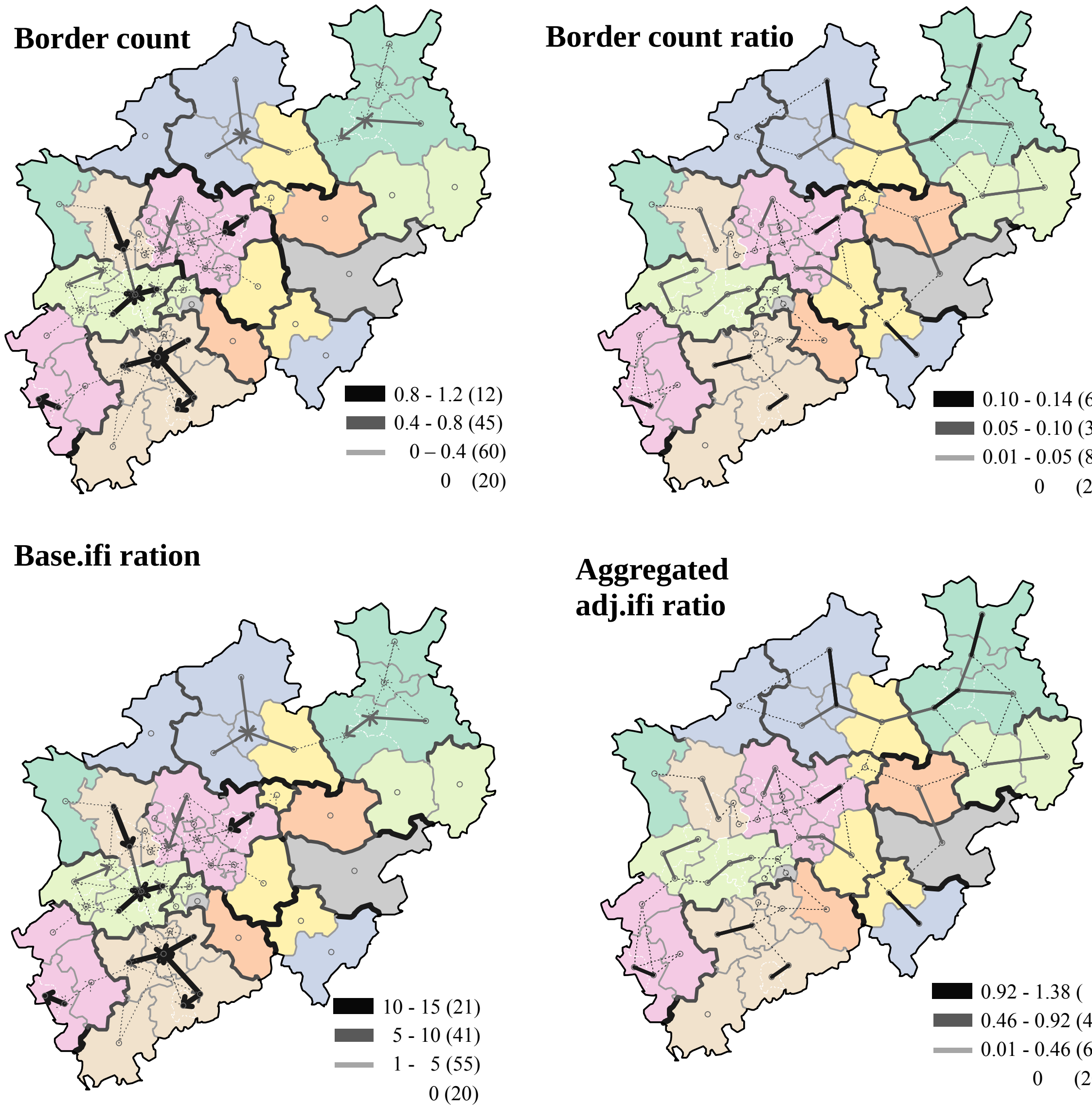
Delineation Procedure	On M1 with adj.ifi & queen	On M2 with adj.ifi & queen	On M1 with adj.ifi & 10nn	M2 adj.ifi &10nn	On M1 with adj.ifi & 4nn	On M2 with adj.ifi & 4nn
Iterationen	4	3	5	4	4	5
No. of FEMA	16	16	15	16	26	21
FEMA internal flows	4,494,930	1,006,127	4,489,692	1,006,127	4,134,045	761,252
FEMA external flows	763,219	846,923	768,457	846,923	1,124,104	1,091,798
Total flows	5,258,149	1,853,050	5,258,149	1,853,050	5,258,149	1,853,050
Share of internal flows	85.49%	54.30%	85.39%	54.30%	78.62%	41.08%
min no of units FEMAs	2	2	2	2	2	2
mean no of units FEMAs	3.8	3.7	3.9	3.7	2.5	2.8
max no of units FEMAs	10	6	9	6	5	5
min ifi.base	0.6397	0.0000	0.4360	0.0000	0.4360	0.0000
average ifi.base	0.8267	0.3650	0.8034	0.3650	0.7708	0.3164
max ifi.base	0.9452	0.7774	0.9452	0.7774	0.9236	0.7774
min adj.ifi	-0.0467	-0.1642	-0.2736	-0.1198	-0.2359	-0.0920
average adj.ifi	0.1132	0.1180	0.0688	0.1591	0.0733	0.1244
max adj.ifi	0.2372	0.4736	0.2139	0.5125	0.2516	0.4569
Dimension FEMA-Pool	40,078	27,044	242,111	172,604	1118	1173
	23	21	25	22	11	12
Trendparameter Intercept	0.6750	0.1293	0.7001	0.0836	0.6483	0.0284
Trendparameter Slope	0.0114	0.0349	0.0096	0.0362	0.0236	0.0637
R² Trend	0.6494	0.7526	0.4398	0.6118	0.2542	0.5689
adj.R² Trend	0.6494	0.7526	0.4398	0.6118	0.2536	0.5685
Cal. time in min	13.3	8.2	105.8	69.0	0.5	0.5
maximum of nb_Counts	14	14	16	15	7	7

Delineation Procedure	On M1 with base.ifi & queen	On M2 with base.ifi & queen	On M1 with base.ifi & 10nn	On M2 with base.ifi & 10nn	On M1with base.ifi & 4nn	On M2 with base.ifi & 4nn
Iterationen	6	7	6	7	6	10
No. of FEMA	3	3	2	2	8	8
FEMA internal flows	5,161,455	1,756,356	5,210,191	1,805,092	4,692,641	1,465,526
FEMA external flows	96,694	96,694	47,958	47,958	565,508	387,524
Total flows	5,258,149	1,853,050	5,258,149	1,853,050	5,258,149	1,853,050
Share of internal flows	98.16%	94.78%	99.09%	97.41%	89.25%	79.09%
min no of units FEMAs	2	2	7	7	2	2
mean no of units FEMAs	18.0	18.0	27.0	27.0	6.9	6.9
max no of units FEMAs	45	45	47	47	19	27
min ifi.base	0.8435	0.6125	0.9615	0.8352	0.7080	0.0000
average ifi.base	0.9314	0.8061	0.9782	0.9106	0.8776	0.5446
max ifi.base	0.9892	0.9707	0.9948	0.9860	0.9615	0.8688
min adj.ifi	0.8435	0.6125	0.9615	0.8352	0.7080	0.0000
average adj.ifi	0.9314	0.8061	0.9782	0.9106	0.8776	0.5446
max adj.ifi	0.9892	0.9707	0.9948	0.9860	0.9615	0.8688
Dimension FEMA-Pool	34,414	108,090	451,189	1,423,903	3836	2961
	50	50	52	52	28	36
Trendparameter Intercept	0.7248	0.3215	0.7261	0.1367	0.7283	0.2457
Trendparameter Slope	0.0078	0.0164	0.0069	0.0223	0.0093	0.0223
R² Trend	0.6078	0.7099	0.7268	0.8010	0.5551	0.7585
adj.R² Trend	0.6078	0.7099	0.7268	0.8010	0.5550	0.7585
Cal. time in min	13.9	50.1	307.1	2005.1	1.3	1.6
maximum of nb_Counts	4	5	7	7	8	8

Appendix 2: Alternative fuzzy set memberhsip functions

The other fuzzy membership functions are presented analog. In the second function, border counts for each resulting board are counted as one. This is probably the most simple membership function and it reaches a maximum of 15 counts. However, it is very likely that the number of times the region can be considered within the 54 FEMAs influences the count. This will be approximated and corrected by the number of borderlines that exists for the neighboring two districts of a border. For example, for the border between the districts *Münster* and *Steinfurt* it is seven, since *Münster* has three neighbors and borders and *Steinfurt* has four. The ration is then calculated by dividing the counted value for each line through the number of border lines. This correction is also done for the membership function of accumulated adj.ifis, and for an analog function that accumulates base.if.

Alternative fuzzy set illustrations



Appendix 3: Summary delineation process for Germany

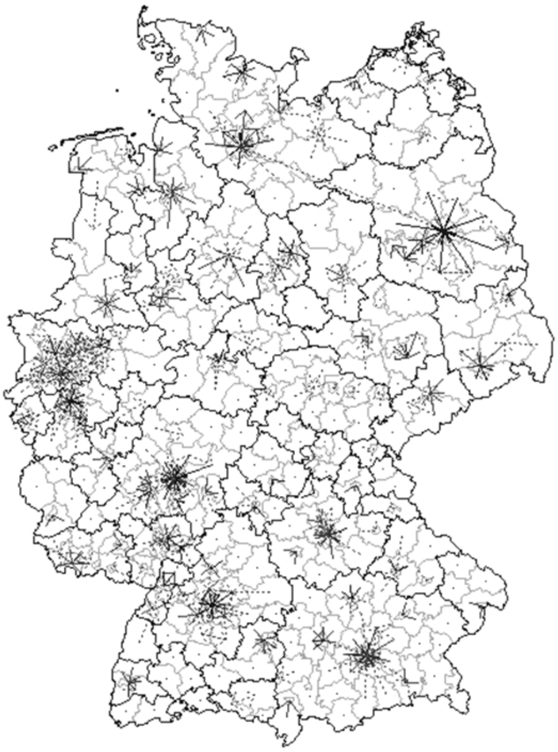
Delineation	M1 BRD (internal values), adj.ifi, queen, no fuzzy				M2 BRD (diagonal of zeros), adj.ifi, queen, no fuzzy		
Iteration	Iteration 1	Iteration 2	Iteration 3	Iteration 4	Iteration 1	Iteration 2	Iteration 3
No. of FEMA	134	124	110	110	154	145	145
FEMA internal flows	22,481,648	22,820,040	23,111,799	23,111,799	5,041,248	5,258,642	5,258,642
FEMA external flows	4,256,482	3,918,090	3,626,331	3,626,331	4,799,834	4,582,440	4,582,440
Total flows	26,738,130	26,738,130	26,738,130	26,738,130	9,841,082	9,841,082	9,841,082
Share of internal flows	84.08%	85.35%	86.44%	86.44%	51.23%	53.44%	53.44%
min no of units FEMAs	1	1	1	1	1	1	1
mean no of units FEMAs	3.4	3.6	3.9	3.9	2.9	3.0	3.0
max no of units FEMAs	10	10	14	14	9	9	9
min ifi.base	0.4398	0.4175	0.4175	0.4175	0.0000	0.0000	0.0000
average ifi.base	0.7875	0.8001	0.8145	0.8145	0.3558	0.3764	0.3764
max ifi.base	0.9271	0.9271	0.9425	0.9425	0.7499	0.7969	0.7969
min adj.ifi	-0.2338	-0.2821	-0.3047	-0.3078	-0.0793	-0.1014	-0.1001
average adj.ifi	0.0974	0.0790	0.0733	0.0702	0.2055	0.2041	0.2054
max adj.ifi	0.2448	0.2175	0.1972	0.1941	0.5862	0.5718	0.5731
Dimension FEMA-Pool	43398	290790	496931	553324	43398	147195	152968
	15	25	27	30	15	23	23
Trendparameter Intercept	0.6657	0.6903	0.7153	0.7184	0.0371	0.0632	0.0617
Trendparameter Slope	0.0079	0.0093	0.0069	0.0069	0.0422	0.0383	0.0384
R^2 Trend	0.0240	0.3727	0.3166	0.3336	0.2686	0.6166	0.6107
adj.R^2 Trend	0.0240	0.3727	0.3166	0.3336	0.2686	0.6166	0.6107
Calculation time in min	24	230	774	366	39	69	77
maximum of nb_Counts	12	16	16	16	12	14	14

Appendix 4: Numerical overlay of functional regional market definitions

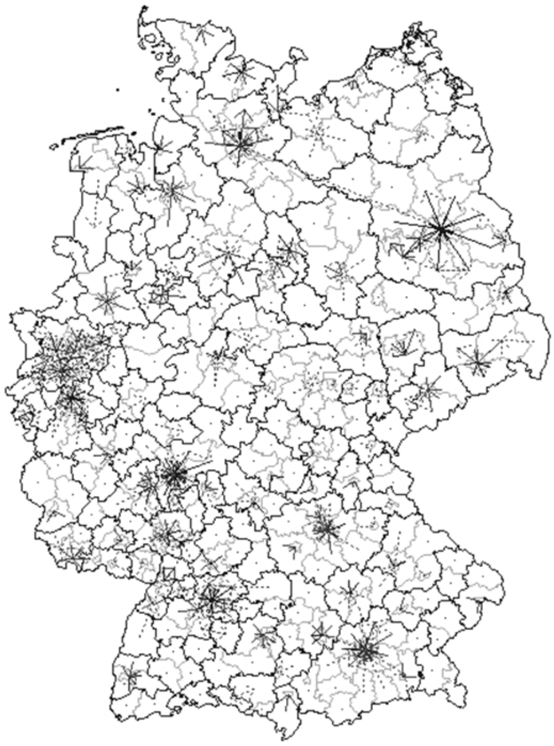
Def. 1	Def. 2	Borders in			borders commonly			borders not in common		
		Def. 1	Def. 2	common	total	assigned	not assigned	total	Def.1 not Def. 2	Def. 2 not Def. 1
FEMAs$_{M1}$	RLM K/S	564	366	73.9%	794	325	469	280	239	41
FEMAs$_{M1}$	LLM K/S	564	566	85.1%	914	485	429	160	79	81
FEMAs$_{M1}$	LMR Eckey	564	671	79.8%	857	509	348	217	55	162
FEMAs$_{M1}$	ROR	564	598	75.6%	812	450	362	262	114	148
FEMAs$_{M1}$	FEMAs$_{M2}$	564	670	80.6%	866	513	353	208	51	157
FEMAs$_{M2}$	RLM K/S	670	366	64.1%	688	325	363	386	345	41
FEMAs$_{M2}$	KS LAM	670	566	75.2%	808	485	323	266	185	81
FEMAs$_{M2}$	Eckey	670	671	69.9%	751	509	242	323	161	162
FEMAs$_{M2}$	ROR	670	598	65.7%	706	450	256	368	220	148
FEMAs$_{M1}$ ∩ FEMAs$_{M2}$	RLM K/S	564	366	70.4%	756	306	450	318	258	60
FEMAs$_{M1}$ ∩ FEMAs$_{M2}$	LLM K/S	564	566	74.5%	800	428	372	274	136	138
FEMAs$_{M1}$ ∩ FEMAs$_{M2}$	Eckey	564	671	73.3%	787	474	313	287	90	197
FEMAs$_{M1}$ ∩ FEMAs$_{M2}$	ROR	564	598	71.7%	770	429	341	304	135	169
FEMAs$_{M1}$	RLM K/S ∩ RLM Eckey	564	338	73.7%	792	310	482	282	254	28
FEMAs$_{M1}$	RLM K/S ∩ RLM Eckey ∩ ROR	564	294	70.9%	762	273	489	312	291	21
FEMAs$_{M1}$	LLM K/S ∩ RLM Eckey	564	516	80.6%	866	436	430	208	128	80
FEMAs$_{M1}$	LLM K/S ∩ RLM Eckey ∩ ROR	564	415	78.1%	839	372	467	235	192	43
FEMAs$_{M2}$	RLM K/S ∩ RLM Eckey	670	338	65.4%	702	318	384	372	352	20
FEMAs$_{M2}$	RLM K/S ∩ RLM Eckey ∩ ROR	670	294	61.5%	660	275	385	414	395	19
FEMAs$_{M2}$	LLM K/S ∩ RLM Eckey	670	516	77.1%	828	470	358	246	200	46
FEMAs$_{M2}$	LLM K/S ∩ RLM Eckey ∩ ROR	670	415	72.0%	773	392	381	301	278	23
FEMAs$_{M1}$ ∩ FEMAs$_{M2}$	RLM K/S ∩ RLM Eckey	564	338	70.9%	762	295	467	312	269	43
FEMAs$_{M1}$ ∩ FEMAs$_{M2}$	RLM K/S ∩ RLM Eckey ∩ ROR	564	294	68.3%	734	259	475	340	305	35
FEMAs$_{M1}$ ∩ FEMAs$_{M2}$	LLM K/S ∩ RLM Eckey	564	516	76.0%	816	411	405	258	153	105
FEMAs$_{M1}$ ∩ FEMAs$_{M2}$	LLM K/S ∩ RLM Eckey ∩ ROR	564	415	75.1%	807	356	451	267	208	59

Appendix 5: Geographical illustratiation alternative functional regional market definitions

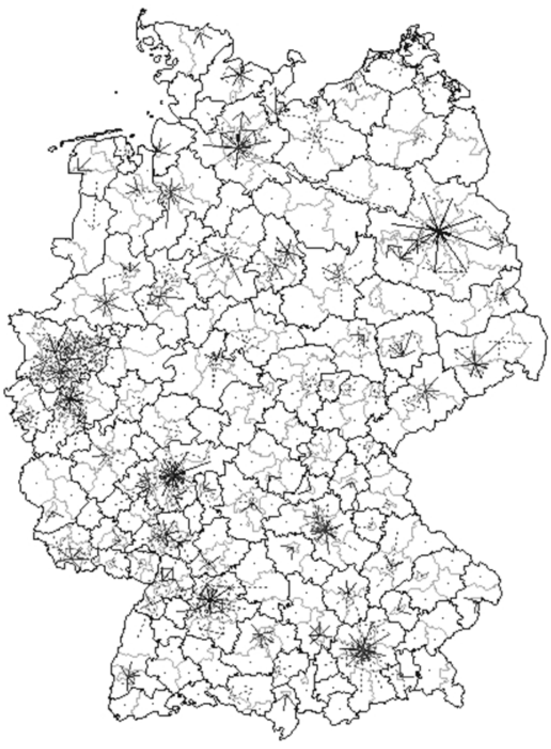
LMR Kropp/Schwengler (50)



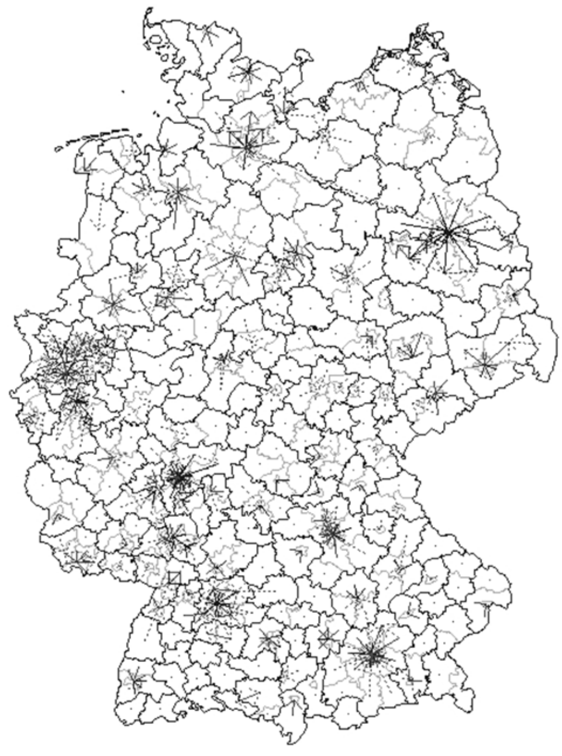
LLMR Kropp/Schwengler (105)



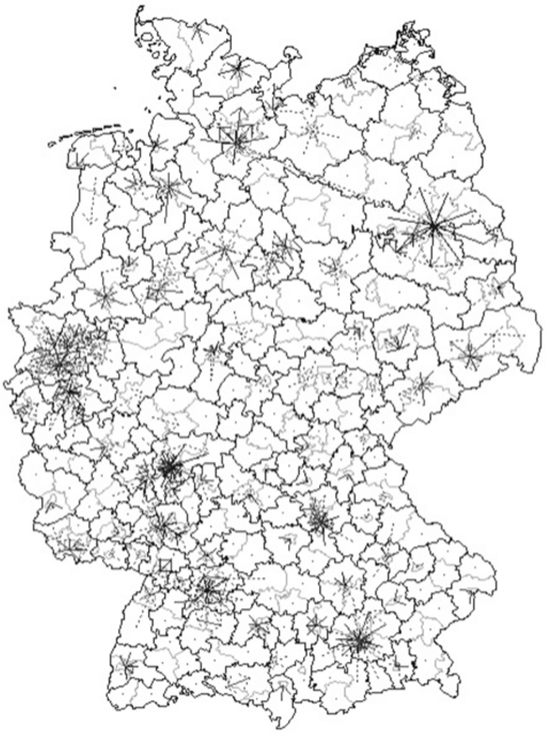
FEMA M1 (110)



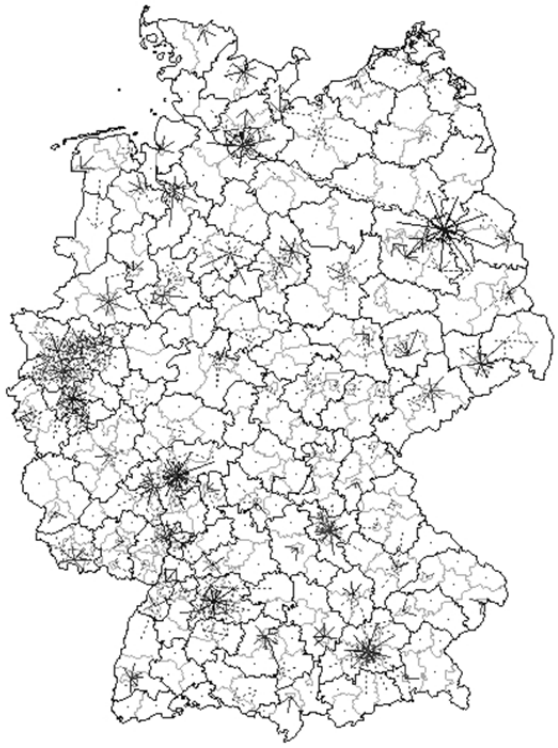
LMR Eckey (147*)



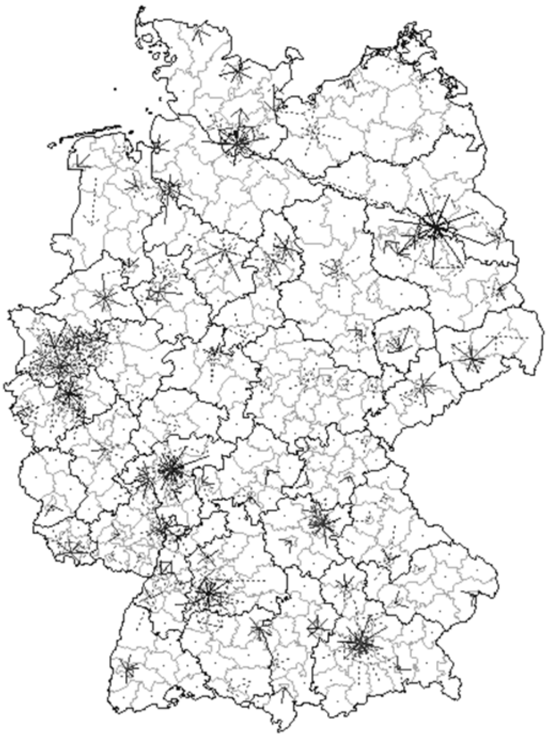
FEMA M2 (145)



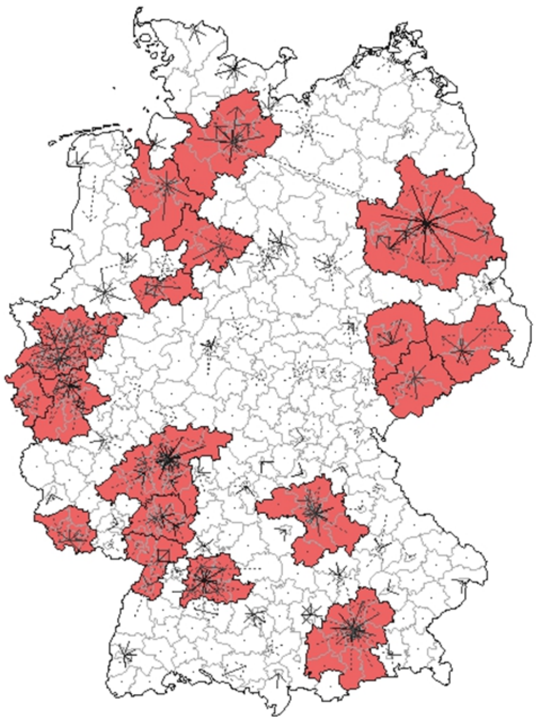
ROR (96)



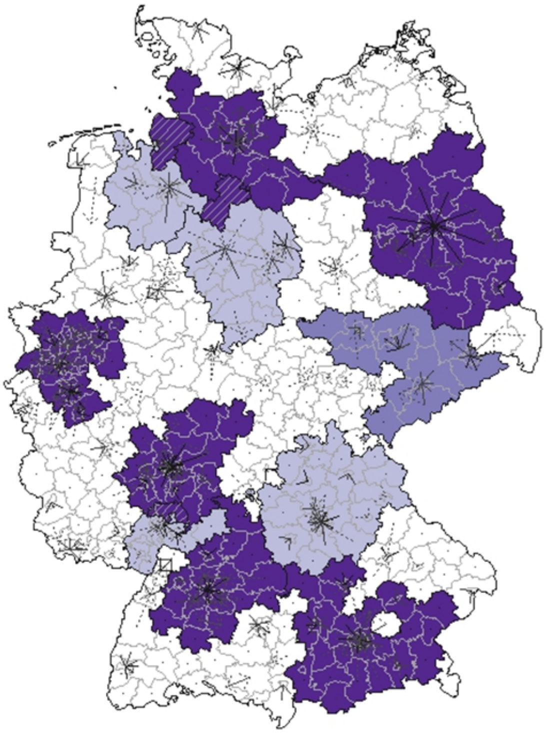
NUTS-2 (39)



Agglomeration (19+1, 279)



EMR (11+1, 195)



Metro (55+1, 282)