Study on comparable Labour Market Areas
- Final research report –

Framework contract n° :	**6001. 2008.001 - 2009.065**
Specific contract n°:	**50405.2010.004 – 2011.325**
Authors:	**Mike Coombes, José M. Casado-Díaz, Lucas Martínez-Bernabeu, Florabela Carausu**
Date of submission:	**17 October 2012**
Version:	**V1.0**

Table of contents

1. Introduction

Overview of the project

In 2007, when the implementation of the NUTS Regulation was reviewed by Commission Regulation (EC) No 105/2007, National Statistical Institutes (NSIs) proposed an investigation of alternative classifications to the administrative levels below NUTS for the EU management of territory. As a result, Eurostat committed itself to exploring functional regions and the possible application of that concept to the entire EU. The first step was then to investigate with the help of the research community the potential value-added and feasibility of, and best practice for, a consistent EU-wide definition of labour market areas.

The report that follows is the final output from research activities undertaken by, and on behalf of, Eurostat in this direction. It is a result of the work of Eurostat Unit E4, coordinated by Mr. Oliver Heiden, and the external contractor DevStat – Servicios de Consultoría Estadística with its associated researchers, Prof. Mike Coombes from Newcastle University, and Prof. José Manuel Casado and Dr. Lucas Martínez from the University of Alicante.

The purpose of the "Study on comparable Labour Market Areas" is to explore the possibility of a consistent statistical classification of the whole EU territory, defined on a functional basis. To be specific, the objectives are to: (1) outline the state-of-art of applied sciences in the field of LMAs; (2) compare the LMA concepts recognised and implemented in each Member State; (3) draw conclusions on relevant best practice; (4) explore the added value of a common definition for the entire EU; and (5) identify possible ways and means of harmonising LMA definitions across the EU.

What makes this study different from the previous ones[1] is the intention to cover all the territory of EU 27 Member States (MSs) while also empirically testing a proposed method for the delineation of LMAs so as to formulate proposals for a possible EU wide harmonised grid of comparable LMAs. This empirical research is to reflect the assessment of LMAs' potential applications in the policy fields of the EC, but will also rely on the involvement of MSs in the collection of comparable information and the evaluation of intermediate results of the study.

[1] The topic was previously investigated by other international organisations and DGs of the European Commission (EC). Two decades ago, Eurostat and Newcastle University carried out a study of principles upon which to base definitions of LMAs to be used in a policy context. In 2001, OECD undertook a cross-national survey in order to examine the relevance of functional delineation of regions on the basis of travel-to-work. OECD's study offers a clear view of different existing concepts of LMA in different states, although it does not include all EU countries. More recently, DG REGIO expressed a strong interest in LMA statistics and Eurostat carried out a survey of the LMA definitions in Member States.

Structure of the Final Research Report

Apart from this introductory section the report includes two main chapters which are the product of two of the activities of the project:

- **Assessment of LMAs' potential applications in policy fields of the EC**
  This chapter focuses on the need for comparability of areas in territorial policy analyses and makes a review of the main policy domains of the EC that could benefit from the appropriate statistics at sub-national level.
- **Recommendations on the establishment of an EU-wide harmonised grid of comparable LMAs**
  This chapter is concerned with empirical tests of the feasibility of establishing for the whole EU a harmonised grid of comparable LMAs, and with making recommendations on the means and resources necessary for future research to achieve this objective.
  Each chapter is completed by specific appendices.

This Final Research Report also includes in an Annex some reports produced earlier in this project. **Annex 1** includes a study reviewing LMAs in the applied social sciences, plus a study comparing existing national LMA definitions.

2. Assessment of LMAs' potential applications in policy fields of the EC

2.1. Introduction and context

This chapter is the product of one activity from a research study on the potential for creating a common European definition of Local Labour Market Areas (LMAs). These functional areas are an alternative to local and regional administrative areas for statistical and policy purposes. The increasing interest in LMAs reflects the fact that administrative boundaries are frequently the result of historical circumstances, rather than of present day territorial reality. The current debate on future regional policy in the EU suggests the need to strengthen Cohesion Policy, but this would not be helped if it was limited to using administrative boundaries. This chapter is set the task of investigating whether analysing statistics on LMAs might improve the implementation of the Europe 2020 Strategy by giving more appropriate insights into the performance of regions and the impacts of policy.

Policy researchers have seen welcome increases in the available data on regions and smaller areas in recent years. One result is that it is no longer simply the case that they must use administrative areas for their analyses, as it had been when the NUTS system identified three levels of broadly comparable sets of administrative areas within each Member State (MS). Now there is often data available for more than one set of areas at a sub-national level that is appropriate for the targeting and/or monitoring of key policies of the EU. The choice between sets of boundaries can help improve policy-making, but it draws attention to what is known as the MAUP (Modifiable Areal Unit Problem). Put most simply, the results of any territorial analysis will partly depend on the areas used for that analysis: a change in areas used will produce different results.

The relevance of the MAUP for policy analyses involving comparisons amongst regions and cities was examined in detail by ESPON[2] and their conclusion was that the fact that changing the areas used while analysing the same data produced "[s]uch contradictory results could, quite naturally, be very disturbing for the decision maker" (ESPON 2006). Two relevant illustrations of the MAUP have also been provided here in the Appendix to this chapter (section 2.5):

[2] ESPON (2006) *Final Report of the Project ESPON 3.4.3 "The Modifiable Areas Unit Problem"*

- Official statistics on the earnings rate (€) in the ten largest cities of France are compared and the rankings of the cities are shown to vary markedly depending on which type of official boundaries is used.

- Unemployment statistics in the UK for both the official labour market areas and a set of administrative areas are analysed to assess how much difference there is in the places identified as being in the highest unemployment areas on each basis.

The fact that the results when comparing areas are sensitive to the choice of areas that are used for those comparisons prompts the question of whether there is then a set of areas which ***should*** ideally be used. The basic answer to this question is that there is no one set of areas which is the ideal for all types of analyses, but that the most appropriate set of areas will depend on the purpose of the analysis concerned. To give an example: the ideal areas for analyses related to the increasingly vital issues around fresh water availability would probably be defined as river basins. Such areas would provide more accurate analyses of the supply of and demand for fresh water and so enable more appropriate policy targeting (eg. to identify where there is the most urgent need for new infrastructure or other policy actions).

The focus for this research is on sub-national areas which are appropriate for the analysis of statistics in the socio-economic realm of territorial policies. In this field the significance of the choice of areas for key analyses is increasingly recognised. Perhaps the single most important policy indicator in this policy field is *per capita* average value of Gross Domestic Product (GDP) but it has been shown[3] that when using administrative areas "regional economic indicators, such as GDP per capita, are frequently distorted" (ESPON 2007:23). The reason is that the income generated in one area, such as a city, may be largely consumed by households in other areas. The equivalent case in the fresh water policy field would be to consider as separate the areas where the water is used and the area where it falls and may be stored. Pursuing this analogy further, what is need in the socio-economic policy field is a set of areas which are equivalent to the water catchment and consumption areas.

The best documented examples of the need to look beyond administrative areas have often used the metaphor of catchment areas when emphasising that data for a city should be analysed as part of a larger area which also includes the commuter catchment area of that city. The most dramatic cases[4] are often provided by capital

[3] ESPON (2007) *Final Report of the ESPON Project 1.4.4 "Preparatory Study on Feasibility of Flows Analysis Final Report"*

[4] European Commission (2007) *Growing Regions, growing Europe, Fourth report on economic and social cohesion*, Luxembourg, Office for Official Publications of the European Communities

cities where "GDP is overstated relative to that produced by residents by between 4% and 76%" (European Commission 2007:15). One particular example is Brussels where the "Brussels-Capital Region (NUTS 1, 2 and 3) can be presented as one of the wealthiest in Europe (ranking twentieth at NUTS 3 level as regards its GDP/head), whereas its available income per inhabitant is in fact today one of the lowest of the three Belgian Regions, since almost 6/10 of the workers who contribute to the creation of its GDP reside in the other two Regions of the country. The political impact of such situations, especially in a federal country like Belgium where there is no financial cross-subsidisation between the Region's budgets, is easy to imagine" (ESPON 2006:139).

The recognition of the fundamental linkages between a city and its 'catchment area' is far from new in Europe or many other more developed parts of the world where it has led to many studies to define metropolitan regions, culminating most recently in that[5] from the OECD (2012). The common logic to these definitions is that it is not sufficient to simply look beyond the administrative boundary of the city to include adjacent heavily urbanised areas – the basis for the definition of a conurbation – because the hinterland of a city will extend further to embrace rural areas along with some other urban areas. This has been termed a 'functional region' and in most such definitions great significance is placed on the pattern of commuting to assess the extent of the functional linkages between cities and their surrounding areas. This dependence upon commuting data means that, in almost all cases, such region definitions are a form of LMA.

For this research however, the relevance of LMAs to analysis in socio-economic policy fields is seen as a general principle, rather than an issue limited to major cities and their immediate regions. Those factors which underlie the functional linkages binding together metropolitan regions, such as the long-term increases in personal mobility and the re-location of employment sites, have also created dispersed patterns of commuting across the many varied types of territory found in the EU. The primary implication is that the need here is for LMA definitions which are not limited to metropolitan regions but that embrace all areas. Less obviously perhaps, the second implication is that the notion of a single centre and its catchment area may not be as appropriate in areas which are more distance from the major cities (and in fact the increasingly widespread recognition that major urban regions are becoming more polycentric structure may mean that even there, a definition of LMA, which presumes a single centre and its hinterland, may not be ideal).

[5] OECD (2012): *Redefining "Urban". A New Way to Measure Metropolitan Areas*, OECD Publishing, Paris

The need for a set of consistently defined LMAs covering the whole territory has for some time been recognised in several MSs. Earlier in this research a survey of MSs was undertaken and some of findings from the information collected have already been provided in the report included here as Annex n° 1 but here it is appropriate to reflect on the evidence about LMAs that have been defined across whole countries for reporting and analysis of socio-economic regional statistics. This survey information has both updated and extended in coverage an earlier review[6] for the OECD that had shown, for example, that the defined LMAs have been invoked when identifying eligible areas for the allocation of European funds in the cases of Finland, France, Germany, Italy and the United Kingdom. The following are a small selection of key points which are mainly drawn from the new survey information:

- Germany uses four indicators (including unemployment rate and wage rates) for their national set of LMAs to guide policy actions to improve regional economic structures.

- Italy uses the boundaries of their national set of LMAs in analyses[7] to define industrial districts (ISTAT, 2006).

- France has incorporated a national set of LMAs into its standard processes for the dissemination of socio-economic statistics at various spatial scales.

- A national set of LMAs has been defined and then regularly updated in the UK for over four decades and over that period its uses have ranged from the targeting of funding for industrial development to providing a template for the new areas needed for sub-regional analysis of housing supply and demand mismatches.

- In both the Czech Republic and Estonia their national sets of LMAs have been used regional and local planning by, for example, using their evidence on the daily activity spaces of people to help improve public transport provision.

- Finland has recently referred to its set of LMA boundaries when revising their local government structure, with one aim being that the new municipalities would more closely reflect functional areas.

This summary of some current uses of LMAs, in those countries where they have been defined as a national set of functional regions, reveals widespread recognition of their value. The diversity of uses which has been highlighted is 'over and above' that of their basic purpose of the reporting and analysis of labour market statistics. The

[6] Cattan N (2001) *Functional regions: a summary of definitions and usage in OECD countries* DT/TDPC/TI(2001)6 OECD, Paris

[7] ISTAT (2006) *Distretti industriali e sistemi locali del lavoro 2001*, Rome, Istituto Nazionale di Statistica

reason why areas specifically designed to be LMAs have proved to have wider uses is that across many socio-economic issues the labour market is the most crucial domain in which the life chances of people are shaped. It is through commuting that most people access the work that brings them income, so patterns of commuting provide a window on the geography of opportunity. It is for this reason that LMAs are widely seen as appropriate areas for spatial policy-related analyses of so many socio-economic challenges to the quality of life of contemporary Europeans.

2.2. The case for cross national delineation of LMAs

The preceding section of this chapter emphasised the need for comparability of areas in territorial policy analyses. The arguments presented were reinforced by citing evidence that numerous MSs have not only accepted the need for specially-defined LMAs but then also used their national set of LMA boundaries for other purposes. The value of appropriately defined LMAs is that they allow meaningful comparisons to be made between all the diverse parts of a territory. This value would be all the greater for comparisons between the highly diverse parts of the territory of the EU because the need for area comparability is accentuated when the analyses are both within and between countries.

In relation to socio-economic policy agendas, the need for area comparability in any EU-wide analyses is especially acute. The uneven impact of the recent steep growth of unemployment has only heightened the importance of regional policy targeting, and the issue of cross-national comparability is always a concern in these analyses. One of the most recent international studies tackling a related challenge was the joint research by the OECD and European Commission (OECD, 2012) that was aimed at establishing a new set of urban area boundaries. Although the coverage of these boundaries ranges from metropolitan areas down to smaller urban area populations (as low as 50,000 people), they do still exclude less urbanised areas. In their focus on urban areas, these new definitions are adopting an approach that has been pursued – with varying levels of success – by many academic studies over past decades, as well as by others in the statistics and policy fields more recently.

Of particular note in this connection are efforts at the European level such as those of the Urban Audit which was defined[8] as a "joint effort by the Directorate-General for Regional Policy (DG Regio) and Eurostat to provide reliable and comparative information on selected urban areas in Member States of the European Union and

[8] European Commission (2004): *Urban Audit. Methodological Handbook*, Luxembourg, Office for Official Publications of the European Communities

the Candidate Countries" (European Commission, 2004:5). Of most relevance here is what the Urban Audit terms a larger urban zone (LUZ) which is intended to reflect a functional region that is centred on the designated urban area. Due in part to the distinct process adopted in creating the Urban Audit data infrastructure, the basis of the LUZ definitions used to vary between countries. Although the definitions were indeed often derived by reference to commuting patterns around cities and towns, it was not possible to consider them as a comparable set of cross-national definitions of LMAs.

One directly relevant assessment[9] highlighted the "very large heterogeneity in the national approaches used to define LUZ" (ESPON 2010:40). This study by ESPON then proceeded to develop its own approach to defining a new set of functional regions extending across a large number of MSs, but once again the basic strategy was to first identify a set of urban cores and then to designate their hinterlands using information on commuting flows. The recent combined efforts of the EU and the OECD (OECD, 2012) have resulted in the definition of a new set of metropolitan areas based on common principles. This exercise has answered the concerns expressed above regarding comparability issues but the nature of the method used remains the same. It is notable that this urban-centred strategy has become so familiar that in many studies there is little effort devoted to justify it, beyond simple statements such as that there is a wide range of economic flows which tend to be orientated around larger urban areas. The evidence of some long-term trends that might be termed counter-urbanisation, as well as evidence of polycentric trends in modern urban systems, is in effect being ignored by exclusively urban-centred approaches to region definition.

The consequence has been that cross-national region definitions have to date produced sets of boundaries that fail to include large parts of the territory analysed. For example the recent research by OECD (2012) defined urban-centred regions that covered less than 40% of the population of both Slovenia and the Slovak Republic. Yet clearly there are equally important policy challenges associated with creating appropriate conditions for socially and environmentally sustainable growth in areas without highly concentrated populations.

In the EU, the main published source for cross-national regional comparisons is the *Eurostat regional yearbook* which includes information for areas at two levels in the NUTS hierarchy of statistical reporting units. According to the regulation on which NUTS definitions are based, the lowest level of aggregation (NUTS 3) in this EU harmonised hierarchy of regions should have between 150,000 and 800,000

[9] ESPON (2010) *Final Report of the Project* ESPON 2013 Database

residents each, and this has led to them becoming a *de facto* option for policy analyses concerned with socio-economic issues at the regional scale in the EU. However the emphasis placed on administrative boundaries in the area definitions of the NUTS hierarchy has led to the criticism that the result is that "they introduce a confusion of scales, especially at NUTS 3 level…in countries with large NUTS 3 divisions like France or Spain, the three categories of spatial structure (intra-urban, peri-urban, rural) are always "mixed" [but in the]…countries with smaller NUTS 3 units like Germany or Belgium, each category of spatial structure can be isolated" (ESPON, 2006:86).

Thus the conclusion here is that although the NUTS cross-national region definitions do cover the whole EU territory as is necessary for policy purposes, they do not provide a suitably comparable set of region definitions. Of particular importance here is that they are 'generalist' in their purpose, with a default emphasis on local government boundaries rather than making any pretence at providing functional region definitions as would be the case of a set of well-defined LMAs.

In direct contrast are the increasing number of metropolitan region definitions that seek to identify functional regions around major centres, whether these be based on the NUTS 3 areas[10] as in Dijkstra (2009) or the smaller units used by OECD (2012). Here the problem is that the focus on major urban areas has enabled a greater comparability of boundary definition but only at the cost of not covering the whole territory of interest. Although most metropolitan region definitions share the basic motivation here of identifying clusters of commuting patterns so as to demarcate sub-national economic regions, this research has set itself the more demanding task of defining a set of comparable LMAs which includes not only metropolitan regions but also all other types of area found in the territory of the EU.

2.3. How a set of LMAs could impact on policy

The previous sections of this chapter emphasised that LMAs are widely seen as very appropriate areas for spatial policy-related analyses at sub-national levels, so this section considers the main policy areas of the EU that require sub-national scale analyses for effective monitoring of their outcomes.

The multilevel model of governance promoted in Europe over the past decades has evolved to extend the policy scope of the European Union, although the EU continues to function according to the general principle of subsidiarity. Most policy domains of the EU are a shared responsibility between the European level and

[10] Dijkstra, L. (2009) *Metropolitan regions in the EU*, Regional Focus n° 01/2009

Member States, while others benefit from coordination at the EU level. These latter include the economic and employment policies. Recent EU policy strategies prioritise more integrated development along with decentralised management and implementation of policies. Assessing the social, economic and environmental sustainable outcome of EU policies at sub-national level requires appropriate areas for intervention and monitoring.

The main features of LMAs are their functional coherence, coverage of both urban and rural territories, and linkage between areas as shown by the interaction revealed in commuting patterns. These features are directly relevant to the following policies: employment policy, regional policy, rural development policy and transport policy. These main policy domains are in turn related to youth policy and environment policy. As a result, this section will consider how a consistent set of EU LMAs could support the implementation of these policies.

The overarching EU strategic document, Europe 2020 Strategy, provides the starting point for this discussion of relevant policies. The approach here recognises that policy making, monitoring and evaluation require appropriate statistical data, and so, for each of the policies identified, an overview of the key indicator statistics at sub-national level is presented. These discussions of individual policies then lead to an assessment of the possible relevance of LMAs for the Eurostat "vision" towards improving the efficiency of European statistics.

i. Europe 2020 Strategy

Europe 2020 Strategy, designed as successor to the Lisbon strategy, aims to turn the EU into a smart, sustainable and inclusive economy delivering high levels of employment, productivity and social cohesion. There are three primary priorities:

- **Smart growth**: developing an economy based on knowledge and innovation;
- **Sustainable growth**: promoting a more resource efficient, greener and more competitive economy;
- **Inclusive growth**: fostering a high-employment economy delivering social and territorial cohesion.

Five measurable EU **headline targets** have been established for 2020.

Headline targets:

- 75% of the population aged 20-64 should be employed;
- 3% of the EU's GDP should be invested in R&D;
- The '20/20/20' climate/energy targets should be met (including an increase to 30% of emissions reduction if the conditions are right);

- The share of early school leavers should be under 10% and at least 40% of the younger generation should have a tertiary degree;
- 20 million less people should be at risk of poverty.

Source: COM(2010) 2020 final, EUROPE 2020 – A strategy for smart, sustainable and inclusive growth
http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=COM:2010:2020:FIN:EN:PDF

To monitor the progress towards Europe 2020 strategy headline targets, eight main indicators and three sub-indicators[11] have been developed. Headline targets relate to the EU level but link to targets set for each MS that take account of their different circumstances.

Among the tools and initiatives put in place at EU level for measuring the progress towards the achievement of EU 2020 Strategy goals is the 2012 Annual Growth Survey[12] which should support economic and employment policy coordination with a focus on growth and competitiveness for today and tomorrow, thereby tackling unemployment and the social consequences of the crisis. Bringing the focus down to these labour market issues means that LMAs can be a critically important scale for targeting policy interventions and subsequently for monitoring the impact of those policies.

Moreover the labour market is a central component of a local economy, and as such allows analyses to link two Europe 2020 goals together, smart growth (which is primarily related to labour demand) and inclusive growth (which highlights 'supply side' social issues). Indeed if the local economy is drawn quite broadly – as with city regions – it is likely that the environmental externalities linked to economic development will be largely contained within the same boundary: thus the same areas can also be appropriate for analyses of the other goal of the Europe 2020 Strategy, sustainable growth. In this way the LMA might provide a suitable areal unit for a sub-national 'triple bottom line' monitoring of the progress towards the objectives of the Europe 2020 Strategy.

ii. Future Cohesion Policy and Territorial Agenda 2020

Despite some successes in previous programming periods, inequalities still exist among EU regions and so the **EU Cohesion Policy 2014-2020**[13] proposal will be

[11] http://epp.eurostat.ec.europa.eu/portal/page/portal/europe_2020_indicators/headline_indicators
[12] The Annual Growth Survey sets out what the Commission believes must be the EU's priorities for the coming 12 months in terms of economic and budgetary policies and reforms to boost growth and employment.
[13] COM(2011) 615 final/2

"strongly linked to the Europe 2020 strategy, focusing on results [and] monitoring progress" towards two goals: Investment in Growth and Jobs and European Territorial Cooperation. A crucial feature here is that economic and social cohesion at the European level is seen to call for a stronger focus on the territorial impact of EU policies. The territorial development tools proposed include one termed Integrated Territorial Investments (ITI). ITIs involve several key elements but the main interest here is in the requirement[14] to designate a territory to be the focus of their actions. By delegating policy delivery to sub-national actors this approach applies the subsidiarity principle to the Territorial Agenda 2020, while placing cohesion centrally on the political stage:

"*The multiple challenges confronting Europe – economic, environmental and social – show the need for an integrated and territorial place-based approach to deliver effective response.*"[15]

Discussions of the place-based approach feature references to functional areas along with diverse aspects of local economies which contribute to place competitiveness. In this way the policy is being conceived as ideally delivered within areas that are each functional economic areas, and as such could be identified as labour market areas.

iii. European Employment Policy and European Youth Policy

MS governments have the primary responsibility for employment policies but the EU increasingly sets common objectives and also analyses the measures taken at national level. As noted earlier, in the Europe 2020 strategy there are headline employment and unemployment targets. Two related strategies are the European Employment Strategy (EES) and the Youth Employment Strategy.

The common priorities and targets of the EES are set through the Employment Guidelines (Council Decision 2010/707/EU of 21 October 2010 on guidelines for the employment policies of the Member States). These guidelines recognize that employment policies link economic strategies to social inclusion concerns, and this linkage is illustrated by the following guidelines agreed by national governments.

Guideline 7: Increasing labour market participation of women and men, reducing structural unemployment and promoting job quality

http://ec.europa.eu/regional_policy/sources/docoffic/official/regulation/pdf/2014/proposals/regulation/general/general_proposal_en.pdf

[14] EC, EU Cohesion Policy 2014-2020 Factsheets: Integrated Territorial Investment http://ec.europa.eu/regional_policy/what/future/proposals_2014_2020_en.cfm

[15] Ibidem

Guideline 8: Developing a skilled workforce responding to labour market needs and promoting lifelong learning
Guideline 9: Improving the quality and performance of education and training systems at all levels and increasing participation in tertiary or equivalent education

All these guidelines recognize the need to link the economic development issues of labour demand to the more social issues of labour supply. LMAs are the local expression of this matching process and thus are an ideal framework for assessing its success or failure.

Regional Europe 2020 indicators have been published[16] at the NUTS 2 level by DG Regio. Indicators for additional employment analysis were specified in the Council Decision 2010/707/EU. The two main sources of employment data are the EU Labour Force Survey and the National Accounts, but data availability limits some indicators monitoring at the national level. At the same time, there are indicators from the Eurostat Regional Statistics[17] database which allow analyses of NUTS2 or even NUTS3 regional levels, as the following small sample illustrates.

EU Labour Force Survey, annual data • Persons in employment in age groups 15-64, 15-24, 25-54, 55-59, 55-64, 60-64, 65-69 and 20-64 as a proportion of total population in the same age group. Breakdown by sex. • Total hours worked divided by the average annual number of hours worked in full-time jobs, calculated as a proportion of total population in the 15-64 age group. Breakdown by sex. • Gender segregation in occupation/sectors, calculated as the average of national share of employment for women and men applied to each occupation/sector (ISCO classification/NACE classification)
Eurostat, European System of Accounts • Growth in GDP per person employed and per hour worked
Eurostat, Structure of Earnings Survey and national sources • Difference between men's and women's average gross hourly earnings as a percentage of men's average gross hourly earnings (for paid employees)
LMP database, Eurostat • N° of participants in regular activation measures (LMP categories 2-7)/in assisted activation programmes (LMP sub-category 1.1.2)/n° of recipients of support (LMP categories 8-9) divided by the n° of persons wanting to work (ILO unemployed plus labour reserve)

[16] http://ec.europa.eu/regional_policy/information/brochures/pages/country2012/index_en.cfm
[17] For more details, see http://epp.eurostat.ec.europa.eu/portal/page/portal/region_cities/metropolitan_regions/data_metro/database_sub3

Job Vacancies Database and EU LFS, Eurostat • Ratio between the total number of the stock of vacancies compared to the total number of unemployed (v/u ratio)

Source: EC, DG Employment,Social Affairs and Equal Opportunities (2010): *Indicators for monitoring the Employment Guidelines including indicators for additional employment analysis, 2010 compendium*

Clearly these labour market indicators are prime candidates for analysis at the LMA scale to enable monitoring of the interaction between local dynamics, broader shifts in the economy and the policy interventions to tackle uneven development.

Unemployment among young Europeans is one of the major problems facing the new EU Youth Strategy for 2010-2018 adopted by the EU Council in November 2009. There are in fact two broad objectives set for this new framework:
(1) more and equal education and labour market opportunities for young people and
(2) active citizenship, social inclusion and solidarity of young people.

EU Youth Strategy (2010-2018) covers fully eight fields of action, so a very broad assessment of the situation of young people was needed in a Dashboard of youth indicators[18]. All the EU indicators on youth are fully complementary to the headline targets of the Europe 2020 strategy and its initiative Youth on the Move[19] that tackles employment and mobility issues.

Certainly for the employment-related youth indicators, the conclusions drawn above about the wider Employment Policy apply here equally, not least because the data sources are mostly the same. In fact the relevance of the local scale is all the greater for young people who tend to find their first employment opportunities very locally (unless they are such 'high flyers' that they are not the concern of EU Youth Strategy). As a result policy interventions need to be tightly targeted at the localised concentrations of youth problems.

iv. European Transport Policy and European Environmental Policy

ECs most recent transport strategy (Transport 2050, White Paper 2011: Roadmap to a Single European Transport Area – Towards a competitive and resources efficient transport system) seeks to develop "a competitive transport system that will

[18] http://epp.eurostat.ec.europa.eu/portal/page/portal/employment_social_policy_equality/youth_policy/dashboard_youth

[19] Commission Staff Working Document on EU indicators in the field of youth, SEC (2011) 401 final

increase mobility … [but also] cut carbon emissions in transport by 60% by 2050" (CE, 2011). It sets different goals for different types of journeys: within cities, between cities, and long distances. Targets such as reducing by 50% urban travel by conventionally fuelled cars are directly related to the commuting patterns which are fundamental to the definition of LMAs.

The transport sector is not only important due to accessibility issues involved in the discussions above about employment opportunities, but also because of its major contribution[20] to greenhouse gas emissions. This issue has considerable sub-national variation due to contrasts between rural and urban areas generally, as well as those between different urban regions depending on their compactness as well as their public transport provision. As a result there would be very good reasons for LMA scale analysis of progress against the objective of efficient mobility[21]that is seeking to improve access to jobs and thus to have a positive impact on employment rates.
Eurostat disseminates a variety of transport indicators at the regional level, and this partly reflects the fact that EU cohesion funding for transport projects has been very substantial. However the datasets of most interest here – unlike the statistics on road transport of goods, or those on maritime and air transport – are not derived from data collected[22] under legal acts.

v. Common Agricultural Policy

The reform of the Common Agricultural Policy (CAP) post 2013 is currently on the political agenda of the EU and as yet there are only proposed[23] objectives for the future, which are:

- Viable food production;
- Sustainable management of natural resources and climate action;
- Balanced territorial development.

These new main objectives link to the EU 2020 Strategy 2020 and also the emerging emphasis on territorial development in seeking smart, sustainable, and inclusive growth.

[20] Eurostat, Statistics in Focus n° 42/2009, Highlights of the Panorama of Transport.
[21] Commission Staff Working Paper, Analysis associated with the Roadmap to a Resource Efficient Europe, Part II, accompanying the document Roadmap to a Resource Efficient Europe, SEC (2011) 1067 final
http://ec.europa.eu/environment/resource_efficiency/pdf/working_paper_part2.pdf
[22] http://epp.eurostat.ec.europa.eu/cache/ITY_SDDS/en/reg_tran_esms.htm
[23] Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions, The CAP towards 2020: Meeting the food, natural resources and territorial challenges of the future, COM(2010) 672 final
http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=COM:2010:0672:FIN:en:PDF

These proposals increase the importance placed on rural development, while also reinforcing the need for a common monitoring and evaluation system

> *(a) to demonstrate the progress and achievements of rural development policy and assess the impact, effectiveness, efficiency and relevance of rural development policy interventions, (b) to contribute to better targeted support for rural development, and (c) to support a common learning process related to monitoring and evaluation.*[24]

The initial indications [25] are that the list of common indicators will cover rural development issues in ways that parallel several indicators discussed earlier, such as GDP per head and employment rates. There will also be indicators related to major sustainability concerns, such as water quality and biodiversity. As noted in relation to policies for economic development generally, these sustainability indicators consider the externalities which can be the outcomes of some development strategies, and most externalities of this kind tend to be localised in their impacts. Thus a rather local scale is appropriate for monitoring progress towards more balanced territorial development.

At present many rural development statistics are provided by Eurostat for the NUTS 3 regions and new urban/rural typology[26] (nb. it is notable that the relevant labour market statistics mostly derive from the LFS, but this is not a guaranteed data source at the regional scale [27]). An additional reason for a sub-national scale of rural development policy comes from the growing importance of community-led local development. Based on the experience of LEADER strategies for rural development, the trajectory of policy development has been towards a locally based approach. One other aspect which has even greater relevance to the LMA scale of interest here is the growing recognition of the need to structure policy activity within functional regions which combine both rural and urban areas.

---

[24] Proposal for a Regulation of the European Parliament and of the Council on support for rural development by the European Agricultural Fund for Rural Development (EAFRD), COM (2011) 627 final/2
http://ec.europa.eu/agriculture/cap-post-2013/legal-proposals/com627/627_en.pdf

[25] http://ec.europa.eu/agriculture/cap-post-2013/monitoring-evaluation/workshop-03-2012/technical-paper_en.pdf

[26] See http://epp.eurostat.ec.europa.eu/cache/ITY_OFFPUB/KS-HA-10-001-15/EN/KS-HA-10-001-15-EN.PDF

[27] The transmission of labour market data at NUTS 3 level has no legal basis. However many countries transmit labour market NUTS 3 figures to Eurostat on a voluntary basis, under the understanding that they are not for publication with such detail, but for aggregation in few categories per country, i.e. metropolitan regions and urban-rural typology. Most of the NUTS 3 data are based on the Labour Force Survey (LFS), however, some countries transmit data based on registers, administrative data, small area estimation and other reliable sources. For more details, see http://epp.eurostat.ec.europa.eu/cache/ITY_SDDS/en/urt_lmk_esms.htm

vi. Relevance of LMAs to the Eurostat vision for the next decade

A major statistical policy document on Eurostat's "Vision" was released in the Communication from the Commission to the European Parliament and the Council on the production method of EU statistics (COM(2009) 404 Final)[28] and this represents a major updating of the European statistical approach.

One of the key issues facing the European Statistical System (ESS), as identified in Eurostat's Vision, are the dramatic changes that affect the ESS environment, such as:

- New information requirements in terms of quantity and quality; in particular, the integration of different data sources and domains is crucial to satisfy important information needs;
- Enterprises and citizens require that the regulatory environment is simplified and the response burden is decreased;
- New ICT tools are available for the production and dissemination of data.

In this context, Eurostat's vision for the next decade proposes strategies to cope with this changing environment through two dimensions:

- Horizontal integration across statistical domains, by both the NSIs and Eurostat:
  - For instance, by linking micro-data from different sources including administrative data or alternative sources;
  - By establishing legal acts that cover a variety of statistical domains.
- Vertical integration of the process from data source to final statistical product:
  - By establishing joint structures, tools and processes;
  - By applying the "European approach to statistics" which recognises that European aggregates do not need full national datasets (because modeling or sampling at the EU level could provide some of the data).

The overview of the policies that may benefit from using LMAs noted that some relevant indicators and statistical variables are not available at an appropriate scale, but it also argued that LMAs could support a more integrated approach to these sectoral policies which would then better contribute to overall EU strategy targets. These potential benefits are clearly relevant to the Eurostat vision and its emphasis on data integration and a more flexible response to the needs of policy users.

In addition it is clear that producing a consistent set of LMA boundaries across the EU would be an example of collaboration between NSIs and Eurostat which is part of the Vision. The possibility of consistent LMA definitions will be greater with the availability of data from the 2010/11 Population Censuses (collected within the framework of a Regulation of the European Commission), along with statistics that

[28] Available at: http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=COM:2009:0404:FIN:EN:PDF

for some countries come from sample surveys or administrative sources. In this way compiling the commuting datasets needed for the LMA definitions would itself involve horizontal integration, as well as the vertical integration of national and Eurostat statistical practices.

Although more integrated activity between national and EU levels can be welcomed as an example of the Eurostat vision in practice, the key potential benefit of the LMAs is in providing multi-sectoral statistics to support evidence-based in numerous policy areas. A possible additional advantage would be that the LMAs could be used as the basis for new statistical products of value to policy makers. For example, there could be a typology of LMAs; the potential value of typologies for policy was very recently illustrated in European Commission (2010).

2.4. Conclusions

The results of statistical comparisons of areas are sensitive to the choice of areas that are used for these analyses. It is not possible to say that one set of areas ***should*** always be used, because the most appropriate areas will depend on the purpose of the analysis. Whereas analyses related to fresh water availability should probably use data for river basins, the concern here is with suitable sub-national areas for social and economic policies. Administrative areas have often been used, but there is increasing recognition that 'functional regions' are more appropriate because they minimise the distortion of area comparisons. Almost all functional region definitions are based on commuting patterns and so are labour market areas (LMAs).

A large proportion of MSs recognize the need for a set of consistently defined LMAs covering the whole of their territories. Yet the importance of area comparability becomes still greater when the analyses are both within and between countries. The pressing need for comparable areas for cross-national analyses related to social and economic policy issues has, for example, prompted the definition of the OECD's new metropolitan areas and the EU's LUZ entities in the Urban Audit that in their latest version have converged with the OECD's definitions to give place to the new metropolitan areas presented in OECD (2012). These are identifiable as forms of LMA *but* they only cover large urban areas, whereas for this study the relevance of LMAs to analysis in socio-economic policy fields is not seen as being limited to major cities and their immediate regions.

The use of LMAs in a range of MSs extends well beyond the reporting and analysis of labour market statistics. The reason why areas specifically designed to be LMAs have proved to have wider uses is that across many socio-economic issues the labour market is the most crucial domain that shapes the life chances of people: indeed it is

by commuting that most people access work which brings them income. Moreover the spatial pattern of commuting provides an insight into the geography of opportunity more generally.

It was this understanding which underpinned the section of this report considering ways in which a consistent cross-national set of LMAs could support the delivery and monitoring of EU policies, and in particular the Europe 2020 Strategy and a set of related thematic policies that have sub-national implementation or monitoring. These policies have indicators that emphasise competitiveness and social cohesion, and this combination of concerns makes LMAs a key scale because it is within labour markets that the benefits of growth can reach the unemployed. It is also within such wider LMAs as city regions that environmental externalities of economic development occur, thereby making the same spatial scale relevant to sustainability issues in growth policies. The conclusion was then that LMAs are possible areal units for a sub-national 'triple bottom line' for monitoring progress against the objectives of the Europe 2020 Strategy.

In moving on 'from theory to practice' it is not straight-forward to draw conclusions in the form of a specification for a set of LMAs for the EU policy uses reviewed here. A major aspect of this uncertainty is the size of LMAs: the environmental concerns of sustainability may be best addressed over relatively large areas (as noted above), while social cohesion issues call for a narrower spatial focus because marginalised groups have restricted mobility levels, requiring a localised response to their needs. In fact this dilemma may be at least partly resolved by the practical considerations of obtaining the indicator data for EU policy analyses. Many of the necessary measures rely on sources such as the Labour Force Survey which are not available for very small areas.

In reality it would probably be inappropriate to seek to prescribe strong guide-lines for the definition of LMAs across the whole EU territory when creating acceptable boundaries from remote Scandinavia to urbanised Andalusia is a challenge at the research frontier. For this reason, it was appropriate that the empirical component of this study has proceded without fixed guide-lines to limits its experimentation, instead drawing primarily on experience of 'what works' in existing national definitions of LMAs.

Appendix to chapter 2 - Assessment of LMAs' potential applications in policy fields of the EC

This appendix provides two brief demonstrations of the policy-related issues that motivate this research. They use different policy-relevant indicators for the ranking of areas, and they are applied in different Member States. It is not suggested that these two worked examples are actual cases of existing policy indicator analyses; the sole aim here is to provide analyses that are representative of the *general principles* with which this research is concerned.

The two cases each use different sets of areas to analyse the same data for:

- rankings of large French cities in terms of earning rates
- rankings of all parts of the UK in terms of unemployment rates.

In both cases, the results obtained using the existing sets of nationally-defined LMAs are compared with the results from using one or more set of local government areas.

French city earnings rates

The first step in this case of the sensitivity of area comparisons to the areas used was to identify the ten largest cities in France. In fact this step itself could provide an example of the sensitivity of results to the areas used, but here the 'traditional' areas for identifying cities were used – viz: the set of local government areas which can be officially designated as "cities" – and in the case of France these are the Communes (LAU2).

Table A1 shows the ranking of the 10 largest Communes by population (2008), ranging from Paris down to Lille (which is just over a tenth of the size of the capital). Immediately to the right of the city names in Table A1 are the average hourly € earning rates (2009) for each of the Communes, and the **emboldened** figures next to these are the rankings of the cities on those rates (where **1** = the highest rate). There is no prior assumption here that the ranking of cities on their earning rates should match that on their size, instead the interest is in how far this ranking on city earning rates varies when different boundary definitions of the cities have been used to report the values from the same 'raw' data on how much people earned.

Moving across the columns in Table A1 to the right, the first of the non-Commune areas examined are the Départements (a higher level set of administrative areas). Whereas all other cities are then 'represented' by an area which is not only larger but also has a name that makes no reference to the enclosed city (eg. Lille is in the Nord

Département) the size of the Paris Commune results in it being a Département on its own. Table A1 thus shows the same € rate for the Commune and the Département in the case of Paris, but the ranking of every other city has changed due to the different way in which their € rates are affected by the change to the areas used for the analysis. These two set of € values have a very similar range (from just over 12 to around 13.5 in the nine non-capital city cases), and there is not an immediately 'obvious' way of choosing between them as ways to compare cities, because both sets of boundaries very largely derive from decisions made a very long time ago when the current realities of economic geography could not have even been imagined.

Comparing the earnings rates using the Communes with those for Départements finds that in the majority of cases the former values were higher than the latter – suggesting that € rates tend to be higher in urban cores than in the nearby areas – ***but*** the crucial point here is that there is a significant minority of cities where the opposite pattern is the case. The second and third cities in the size ranking illustrate this variation: Marseilles had only the sixth highest Commune rate but using the higher Département rate promotes it to third equal with Lyon (which has a lower rate at the Département level than for its own Commune).

Table A1 also reports the Région level rates in the same way, because these too are official statistical reporting areas of no little policy interest. Similar shifts in city rankings can be observed, whether these comparisons are made against Communes or against Départements. At this higher level Paris becomes part of Île de France, whose average € rate is lower than that of the urban core, although its top ranking is unaffected. More notable is the fac that at the Région level two of the cities are grouped together: the second largest Commune – Marseilles – is in the same Région (Provence-Alpes-Côtes d'Azur) as the smaller Nice and this, of course, results in the two cities being ascribed the same Région earnings rate. The average € rate for this relatively affluent Région is second only to that of Île de France among those that are shown in Table A1. While for all the other eight cities the Région € rate is lower than that of the urban core Commune, for these two port cities the opposite is true. The policy implications are potentially severe, with the prospect that these cities may benefit from any assistance that is targeted narrowly at low earning rate cities depending very substantially on the areas that are used for such an analysis.

Table A1 shows finally the results using the Zone Emploi which are areas specifically designed to be LMAs for presenting and analysing policy-relevant data at this scale. It is noticeable that on this basis the € rate for Nice sees this city back near the foot of the ranking. Of these ten Zone Emploi earnings rates, four are closest to the Commune rate for the same city, four closest to that of the Département, and two

closest to the rate for the respective Région. To summarise then, this example has shown that a key statistical value:

- is very sensitive to the choice of administrative areas used for the analyses
- is higher for some cities, but lower for others, when larger areas are used
- yields substantially different rankings of cities due to this area sensitivity
- ranks cities differently if LMAs are used to any administrative area ranking.

UK local unemployment rates

In this example, the general principles illustrated in the case study of France (above) are explored in a more explicitly policy-driven example in the UK. In each analysis, areas are ranked in descending order of their unemployment rates, and then the cumulative proportion of the national population that they include is calculated while proceeding down the area unemployment ranking. The objective is to thus identify the '10% of areas with the highest unemployment rates' on each area basis. As with the case in France, the rankings compared here use either local government areas or the nationally defined LMAs known as TTWAs (Travel-to-Work Areas).

Whereas in the French case there are several scales of local government areas that could be compared, in the UK there is a unitary system of local government ***except*** in some parts of England. As it happens, the areas with a two-tier system tend not to have very high unemployment rates, so the highest 10% unemployment rate areas (Table A2) are in fact exactly the same whether the analysis uses the upper or lower tier of local authority (LA) areas in those parts of England: the highest rate areas are all either in the parts of England that have a unitary system, or are outside England (where a unitary system exists in all three countries). Figure A1 shows the distribution of the highest unemployment rate LAs in UK as at 2011. The map also shows the boundaries of the constituent countries and the regions within England: only the three southern regions outside London do not include at least one of the 10% highest unemployment LAs. The map also names several areas of the areas featuring among the 10% highest rate LAs, including the largest city of Scotland (Glasgow) and also of N. Ireland (Belfast) plus a substantial part of London. Although there is little doubt that unemployment is indeed high in this 'East End' part of the capital, neighbourhood scale data would cast doubt on whether there are not equally severe problems in other conurbations such as that of Manchester. The probability is that these provincial inner urban problems are not do not appear in this analysis because they have been 'averaged out' by the boundaries of their LAs being large enough to also embrace some more affluent neighbourhoods, whereas in London there are relatively small individual LAs in the inner area which cover few if any of the low unemployment neighbourhoods of the capital.

Table A3 and Figure A2 show the equivalent results using TTWAs. To help with the comparison against the results with the LAs, the named areas on the map are those which do not have an equivalent area identified when using the other set of areas. Thus whereas on Figure A1 the named areas included some large city LAs (as has just been discussed), the areas identified in the TTWA analysis include several smaller and more peripheral places like Holyhead and Great Yarmouth (Figure A2). At the same time, Bradford is a large city TTWA which is also included on this basis. Comparing the unemployment rates themselves (Tables A2 and A3) shows that the 'qualifying' rate for TTWAs is lower than that for LAs: LAs need to have a rate of 6% or higher to be among the highest decile of LAs in the UK by unemployment rate, whereas a rate of 5.5% is enough to be included in the ranking when using TTWAs. This is the primary reason why the TTWA ranking includes areas like Great Yarmouth and indeed Bradford, because in fact the LAs covering these places have rather similar boundaries to their TTWAs and so their LA unemployment rates were similar to those for their TTWAs. As can be seen from Tables A2 and A3, these rates are not high enough for these places to be among the highest 10% if the ranking is of LAs. The key explanation lies in the LA ranking including parts of London along with the large cities such as Glasgow (where the TTWA rates are lower): when these large populations are included in identifying the highest unemployment areas, the top 10% threshold is achieved before the analysis reaches the 6% unemployment rate areas.

There is no definitive empirical 'proof' that one set of areas is superior to another for analyses such as those illustrated here, although some academic work has shown evidence to support the theoretical position that LMAs are the appropriate units for comparative analyses of local economies. Here the principal argument refers to the question of which areas are more 'fit for purpose' in policy contexts. To be specific: would it be more appropriate for a policy targeted at the highest unemployment areas in the UK to include a part of London such as Waltham Forest (Table A2) rather than Bradford? To ask this question tends to imply the use of LAs – because otherwise Waltham Forest will not be individually identifiable – and on this basis it certainly has a higher rate than Bradford. However the reason why it does not have its 'own' TTWA, as Bradford, is that there are major commuting flows into and out of Waltham Forest: it is heavily integrated into the London LMA. Consequently a policy response involving, for example, fostering job opportunities in the area is very likely to see its effects greatly dissipated by the jobs being taken by people of other areas. Instead the scale of the unemployment problem affecting Waltham Forest needs to be considered – and addressed – on the appropriate basis which is London-wide. By contrast the unemployment problem in Bradford is much more localised and, when

appropriately addressed within its TTWA boundary, it can be anticipated that a policy intervention will mostly impact on the local residents.

Table A1 Ranking of largest French cities by wage rates, using four different types of area

rank	*Population 2008*	Commune	(€)		Département	(€)		Région	(€)		Zone Emploi	(€)	
1	*2211297*	Paris	18.2	**1**	Paris	18.2	**1**	Île-de-France	15.5	**1**	Paris	16.8	**1**
2	*851420*	Marseille	12.9	**6**	Bouches-du-Rhône	13.2	**3**	Provence-Alpes-Côte d'Azur	12.7	**2**	Marseille-Aubagne	13.1	**4**
3	*474946*	Lyon	13.9	**2**	Rhône	13.2	**3**	Rhône-Alpes	12.4	**4**	Lyon	13.2	**2**
4	*439553*	Toulouse	13.2	**5**	Haute-Garonne	13.3	**2**	Midi-Pyrénées	12.3	**5**	Toulouse	13.2	**2**
5	*344875*	Nice	12.3	**10**	Alpes-Maritimes	12.8	**5**	Provence-Alpes-Côte d'Azur	12.7	**2**	Nice	12.7	**8**
6	*283288*	Nantes	13.3	**4**	Loire-Atlantique	12.2	**8**	Pays de la Loire	11.5	**10**	Nantes	12.3	**10**
7	*272116*	Strasbourg	12.7	**8**	Bas-Rhin	12.4	**7**	Alsace	12.3	**5**	Strasbourg	12.8	**6**
8	*252998*	Montpellier	12.4	**9**	Hérault	12.2	**8**	Languedoc-Roussillon	11.9	**8**	Montpellier	12.8	**6**
9	*235891*	Bordeaux	13.5	**3**	Gironde	12.5	**6**	Aquitaine	12.0	**7**	Bordeaux	12.7	**8**
10	*225784*	Lille	12.8	**7**	Nord	12.1	**10**	Nord-Pas-de-Calais	11.9	**8**	Lille	12.9	**5**

(€) = Salaire net horaire moyen 2009

Sources

population: www.insee.fr/fr/ppp/bases-de-donnees/donnees-detaillees/base-cc-evol-struct-pop/base-cc-evol-struct-pop-08.zip

salaire: www.insee.fr/fr/ppp/bases-de-donnees/donnees-detaillees/base-cc-salaire-net-horaire-moyen/base-cc-salaire-net-horaire-moyen-09.zip

Table A2 10% highest unemployment areas using LAs

LA	unemployment %	cumulative %
Kingston upon Hull	7.9	0.4
Derry	7.9	0.6
Wolverhampton	7.7	1.0
Middlesbrough	7.5	1.2
Birmingham	7.3	2.9
Hartlepool	7.3	3.0
Sandwell	7.2	3.5
Hackney	7.1	3.9
Newham	7.1	4.3
Blaenau Gwent	7.1	4.4
Strabane	7.1	4.4
Belfast	7.0	4.9
Limavady	7.0	4.9
Liverpool	6.8	5.7
Haringey	6.6	6.1
Walsall	6.6	6.5
North East Lincolnshire	6.5	6.7
North Ayrshire	6.4	6.9
Redcar and Cleveland	6.3	7.2
South Tyneside	6.3	7.4
Knowsley	6.3	7.6
Barking and Dagenham	6.3	7.9
West Dunbartonshire	6.2	8.1
Blackpool	6.2	8.3
Tower Hamlets	6.2	8.7
Waltham Forest	6.2	9.1
Merthyr Tydfil	6.1	9.2
Newry and Mourne	6.1	9.4
Glasgow City	6.0	10.4

Table A3 10% highest unemployment areas using TTWAs

TTWA	unemployment %	cumulative %
Newry	7.6	0.2
Strabane	7.6	0.2
Derry	7.5	0.5
Hartlepool	7.0	0.6
Wolverhampton	6.6	1.2
Irvine & Arran	6.3	1.5
Birmingham	6.2	4.1
Middlesbrough & Stockton	6.2	4.9
Liverpool	6.1	6.4
Dudley & Sandwell	6.0	7.2
Holyhead	6.0	7.2
Merthyr Tydfil & Aberdare	5.9	7.4
Hull	5.8	8.2
Grimsby	5.8	8.5
Ebbw Vale & Abergavenny	5.7	8.7
Great Yarmouth	5.6	8.9
Margate, Ramsgate & Sandwich	5.6	9.1
Kirkcaldy & Glenrothes	5.6	9.3
Bradford	5.5	10.1

Source for Tables A2 & A3: average of the 12 months of 2011 claimant count proportions of estimated resident population aged 16-64 [www.nomisweb.co.uk]

Figure A1 10% highest unemployment areas using LAs

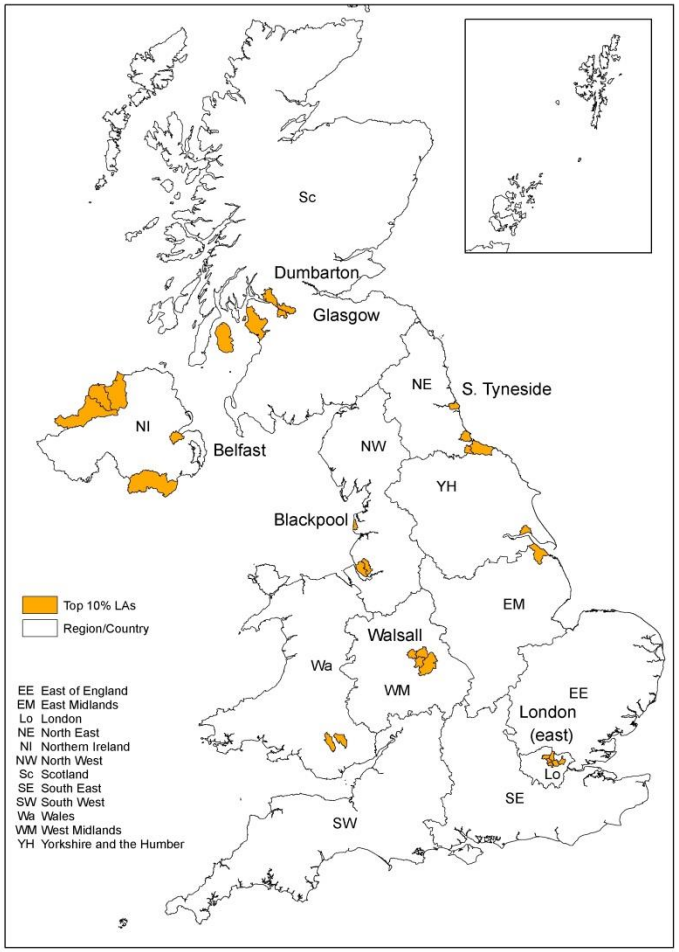
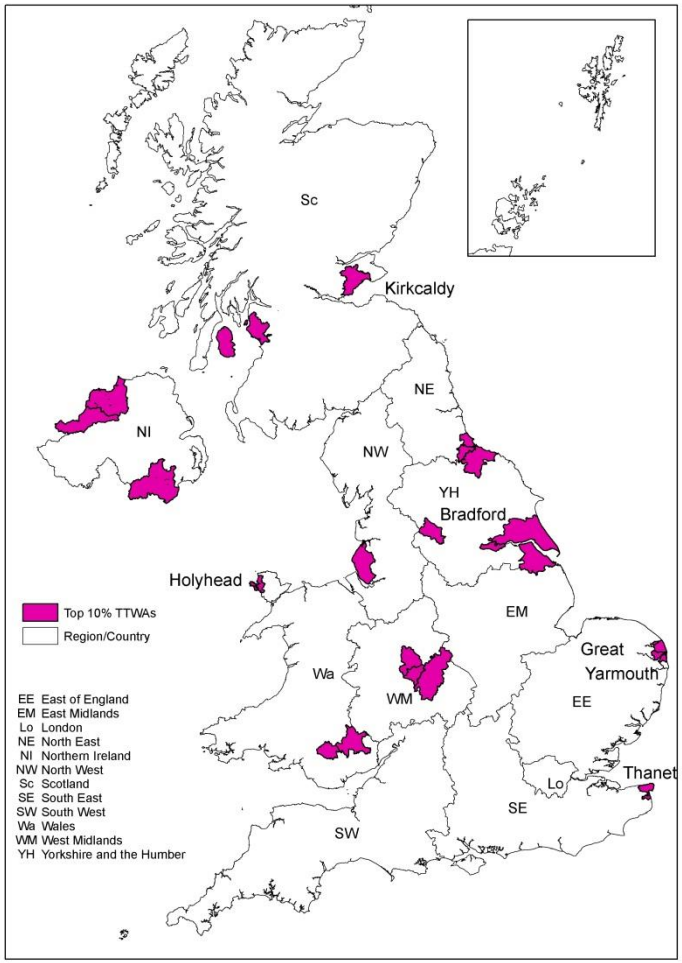


Figure A2 10% highest unemployment areas using TTWAs



3. Recommendations on the establishment of an EU-wide harmonised grid of comparable LMAs

3.1. Introduction

This part of the study is concerned with empirical tests of the feasibility of establishing for the whole EU a harmonised grid of comparable Local Labour Market Areas (LMAs), and with making recommendations on the means and resources necessary for future research to achieve this objective.

This last stage of the study builds upon findings of the earlier research. Chapter 1 of the Interim Research Report (IRR) (here included as Annex I) presented a review of the state-of-art of applied sciences in the field of LMA definitions, showing that most existing concepts of LMAs require all LMAs to be self-contained[29] in their commuting flows. Chapter 2 of the IRR had then shown that the commuting data needed to produce LMA definitions in the EU will very soon be available, mostly from Population Censuses completed in the last few years. The information will be mostly available at LAU2 level, and these zones are suitable 'building blocks' for defining LMAs (as they already are in those countries were LMAs are officially defined).

Chapter 3 of the IRR reviewed LMA definition method best practice and concluded that two official methods (that which defines Swedish LAMs *[Lokala arbetsmarknader]*, and the method defining the TTWAs *[Travel-to-Work Areas]* in the United Kingdom) which deserved empirical evaluation, along with a more 'academic' innovative method (based on genetic algorithms). This empirical evaluation should be undertaken on data for case studies countries providing

- variation in territorial terms found across the EU (e.g. heavily urbanised zones as well as predominantly rural and peripheral areas, including islands)
- variety of LAU2 area sizes (nb. this can strongly impact on some methods)
- some cases with existing official definitions, and some where there are none.

[29] Self-containment is a two-fold variable that comprises supply-side self-containment (the proportion of an area's employed population that works within the area) and demand-side self-containment (the proportion of jobs within an area that are filled by residents of that area). In many of the tables included in this chapter the variable *lower self-containment* is reported: this refers to the lower of the supply and demand-side self-containments and is the best measure of the degree of closeness of the area to being a 'perfect' LMA.

The study had access to national datasets from different sources for three countries[30] that between them covered all three criteria: Spain, Sweden and the UK. Access to data from other countries was facilitated by Eurostat through granting access to SIRE. These additional datasets included the necessary commuting information for the following countries: AT, BE, CH, ES, FR, IE, SE. These datasets presented some limitations; namely, incompleteness due to its coverage of each origin area being limited to the commuting flows to the 30 LAU2 destinations receiving the largest flows.

Glossary

Countries	
xx	the standard EU 2 letter abbreviations are used (eg. ES = Spain)
General	
LMA	(local) labour market area
NSI	National Statistical Institute
MS	Member State
IRR	this project's Interim Research Report (which is included as an Annex to the Report of which this chapter is part of)
Data	
ES01	2001 ES Census commuting data obtained directly from the NSI
SE01	2001 SE Register-based commuting data on Sweden obtained from the web-site of the NSI
SE96/SE06/SE10	1996/2006/2010 SE Register-based commuting data on Sweden obtained from the web-site of the NSI
UK01	2001 UK Census commuting data obtained directly from the NSI
xx(SIRE)	2001/similar commuting data for country xx obtained indirectly (via the Eurostat database SIRE)
Methods	
GEA	experimental method based on grouping evolutionary algorithms
LAM	method that defines official SE labour market areas: *Lokala arbetsmarknader*
TTWA	method that defines official UK labour market areas *Travel-to-Work Areas* (with criteria adapted minimally for transferability)
'Euro'	method that defines official UK labour market areas Travel-to-Work Areas (with criteria changed on a hypothetical basis)

The next three sections of this Chapter explore the definition of LMAs produced when the three alternative methods are applied to commuting datasets for the three MSs:

- Section 3.2 applies the centre-based method developed in Sweden to define the official local labour market areas (LAMs)

[30] The data for UK and Spain was available as a result of the geographical origin of two of the experts involved in the project, while the data for Sweden is publicly available on Statistics Sweden webpage.

- Section 3.3 uses the TTWA method (which does not presume that LMAs all have the same spatial structure) that was devised to define official LMAs in the UK
- Section 3.4 applies an experimental methodology based on the use of grouping evolutionary algorithms (GEA).

These main tests – three methods applied to three national cases (with data on 2001) – are enhanced in a number of ways. Section 3.2 includes data for Sweden from several years so as to assess the sensitivity of LMA definitions to changes in commuting patterns over the longer term. Section 3.3 then widens the test of the 'transferability' of the TTWA method by also applying it to datasets from SIRE.

To build towards a set of recommendations, Section 3.5 starts by comparing the equivalent size and self-containment criteria in existing official methods and thereby develops a hypothetical 'Euro' set of criteria which is then tested in several countries. Section 3.6 then summarises findings from the research by outlining recommendations in a question-and-answer format that hopefully helps to clarify their policy relevance. The chapter also has two appendices. Appendix 1 details the key characteristics of: (a) LAU2 areas which are basic territorial units here, (b) key commuting variables for these analyses, (c) all the 32 alternative sets of LMAs produced in the course of the research. In Appendix 2 the possibility of raising the 'Euro' set of criteria is briefly explored.

3.2. LMAs based on centres

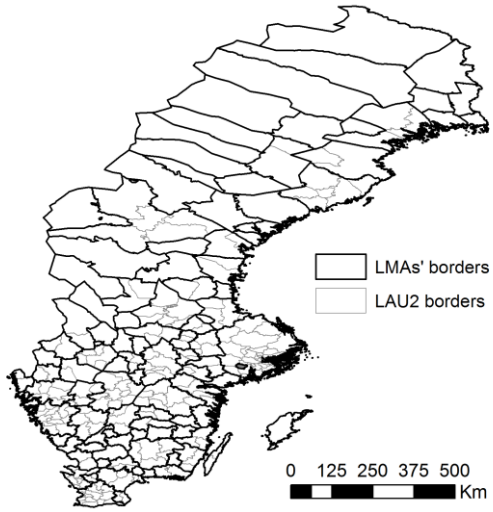
The research starts with a centre-based method because this approach has been the most widely used and recognised. Earlier in this study the LAM method was identified as arguably 'best practice' among these methods. The method has been implemented here as a computerised algorithm that can be applied to suitable data for any country. Comparing the results with the official Swedish definitions confirms that the method has been successfully interpreted by this software. Map 1 shows the results of applying this implementation of the LAM method to 1996, 2001, 2006 and 2010 commuting data (made available via the internet by SCB).

Before moving on to evaluate the results of applying the LAM method to the data for other countries, the data for Sweden provides a very valuable opportunity to assess how stable LMA boundaries are, given persistent changes to commuting patterns. Sweden reports an updated commuting dataset nearly every year, and here the datasets for four

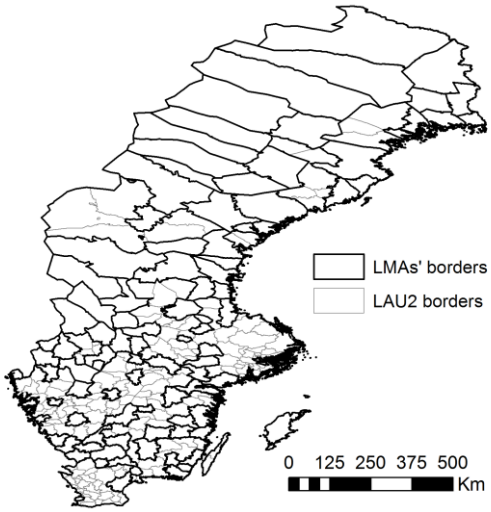
years have been analysed in parallel. Table 1 and Map 1 show the results. Table 1 also shows that the number of local authority areas used for the

Map 1. LMAs from the implementation of the LAM method applied to SE96, SE01, SE06 and SE10

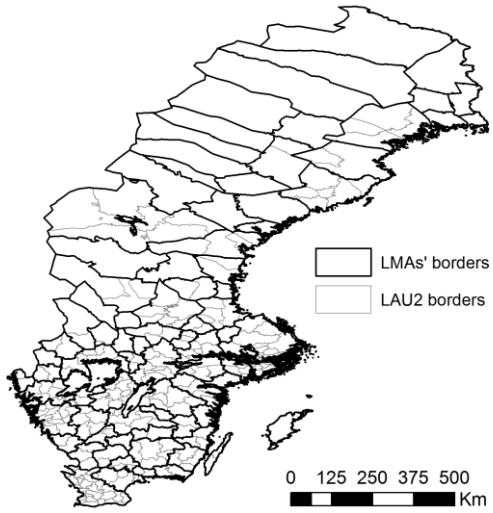
1996



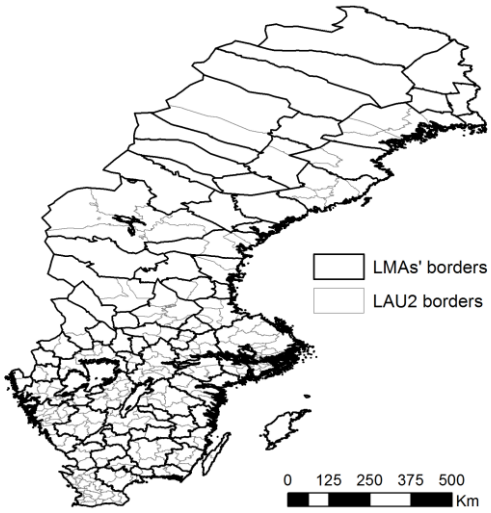
2001



2006



2010



number of separably identifiable LMAs is the gradual increase in longer-distance commuting that gradually erodes the separateness of all localities as labour markets.

The same process has been seen in all modern countries with comparable data available over the medium to long term. Some methods of defining LMAs are more sensitive to

this effect than others, although the level of sensitivity shown in Table 1 was in fact quite closely matched by the level found when the TTWA method was applied to the same datasets. Indeed a method which is ***not*** very sensitive to this

Table 1 Sensitivity of the results from applying the LAM method to SE96, SE01, SE06 and SE10

Year of data	Number of municipalities	Number of LMAs
1996	288	105
2001	289	88
2006	290	79
2010	290	76

effect is failing to reflect the realities of commuting patterns. Hence the conclusions here are that an adequate LMA definition method ***should*** produce results which are time-specific because they are sensitive to change in commuting patterns, which then serves to reinforce the need to update LMA definitions when a new dataset is available (assuming that commuting patterns do continue to change, as they have for decades).

To explore in how transferable this city-based method[31] is to other MSs, the crucial test involves applying it to countries[32] with more complex urban systems, and also datasets with small LAU2 areas – that is, a higher "granularity" – and with boundaries that often cut through current settlements. Spain and the UK both have LAU2 areas with these features, and the results from applying the LAM method to these countries are shown respectively in Maps 2 & 3, while being summarised in Table 2.

The clearest problem found in transference of the LAM method to data for other countries is the low levels of self-containment of the LMAs defined in those countries (Table 2). This problem arises directly from the method not having a self-containment test for valid LMAs; one was simply not necessary in Sweden, partly due to the nature of the LAU2 areas there. By contrast, the same method produces some Spanish LMAs

[31] LAM method has a first step to identify centres (zones where the supply-side self-containment is over 80% and where the maximum outgoing flow to a single destination is under 7.5% of working population, also grouping pairs of zones that have each other as destinations of their largest commuting outflows). The rest of the zones are assigned to the destination of their largest outflow, within a hierarchical process of assignments until all zones are allocated to LMAs (all of which must include a centre): considerable extra detail is provided in the Annex to Chapter 3 of the IRR.

[32] In fact the SCB has applied the LAM method in other Nordic countries (Denmark, Finland and Norway) but these countries do not meet these criteria. Even so the survey conducted by this study revealed that LAU2 areas in Denmark were considered by the Danish NSI to be too different from those in Sweden, resulting on the LAM method not producing results in DK which were likely to be adopted officially.

which are less than 10% self-contained, and in the UK the minimum level is under 20% and more notably perhaps, the median value in the UK is under 60%).

Table 2 LMAs from the implementation of LAM method applied to several countries

	Number LAU2 zones		Number of LMAs	Lower self-containment level		LMA employed population	
	Allocated	Unallocated[33]		Min.	Median	Min.	Median
ES01	8030	78	1536	7.6%	85.2%	1	261
SE01	289	0	88	70.7%	87.5%	1253	14537
UK01	10474	84	681	18.7%	59.0%	464	14133

Notes: The full set of available statistics for each national exercise is available in the Appendix (Table A.6).

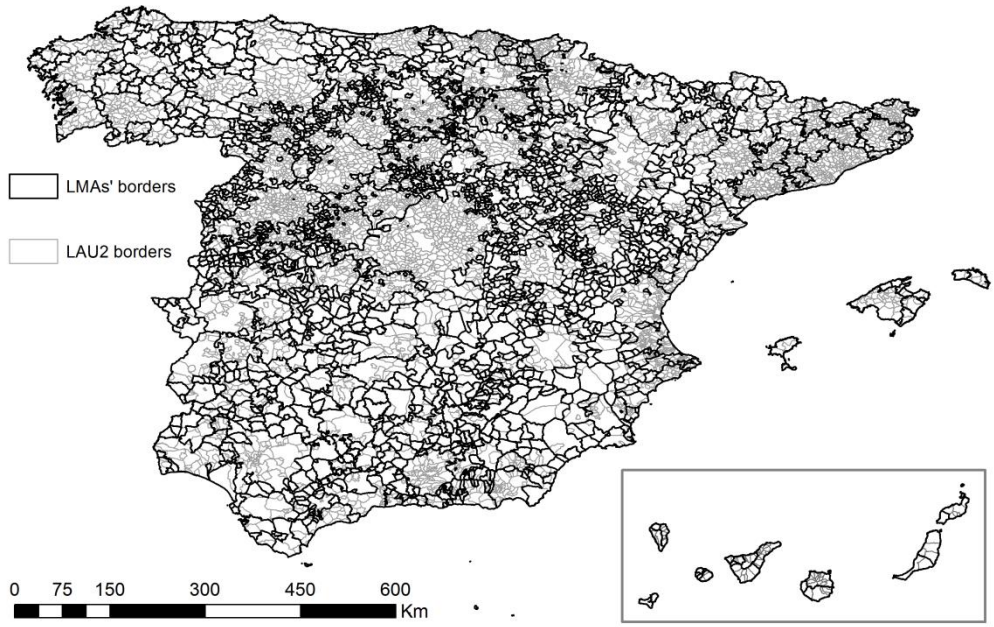
Another feature of the LAM method is the very wide size range of the LMAs it defines. It identifies very large LMAs around the main cities so that in the Spanish case the 20 largest LMAs cover 1876 municipalities and 8.6 million workers yet, at the same time, there is an overabundance of LMAs that each comprises one LAU2 zone in isolation (the median number of employed population in the Spanish LMAs resulting from the application of LAM method is in fact as low as 261). The main reason for the abundance of small sized LMAs is the absence of a minimum size constraint in selecting centres, which is not so relevant in Sweden due to its LAU2 zones' populations (in fact the LMA with lowest size there has 1253 occupied residents). In addition, there is no size test for the final set of LMAs. The problem of small size LMAs from the LAM is not however restricted to the Spanish case because in the UK too some of these LMAs house under 500 occupied residents.

The simultaneous existence of some very large areas is the outcome of the criterion chosen for aggregation: in the LAM method LAU2 areas are grouped simply according to the absolute size of commuting flows with other areas (unlike most other methods, the flows are not expressed as a proportion of all the flows at the origin and/or destination end, because if flow sizes are relative in this way the largest LMAs have less in-built tendency to grow ever larger). In most policy uses, LMAs that are not too variable in size are preferred because then it is more reasonable to see the defined areas

[33] The figures on unallocated zones in the Spanish case reflect the existence of LAU2 zones for which there are no in or out commuting flows; such zones cannot be assigned to any LMA by methods that rely exclusively on commuting flows. In the British case however, the cause is the LAM method relying exclusively on the largest flows of areas which, in some cases, can produce 'loops' [eg. A->B->C->A] which remain unallocable zones if none of the zones qualify as centres.

as comparable. As a result, the tendency of the LAM method to define areas which vary radically in size is a distinct disadvantage.

Map 2. LMAs from the implementation of LAM method applied to ES01[34]



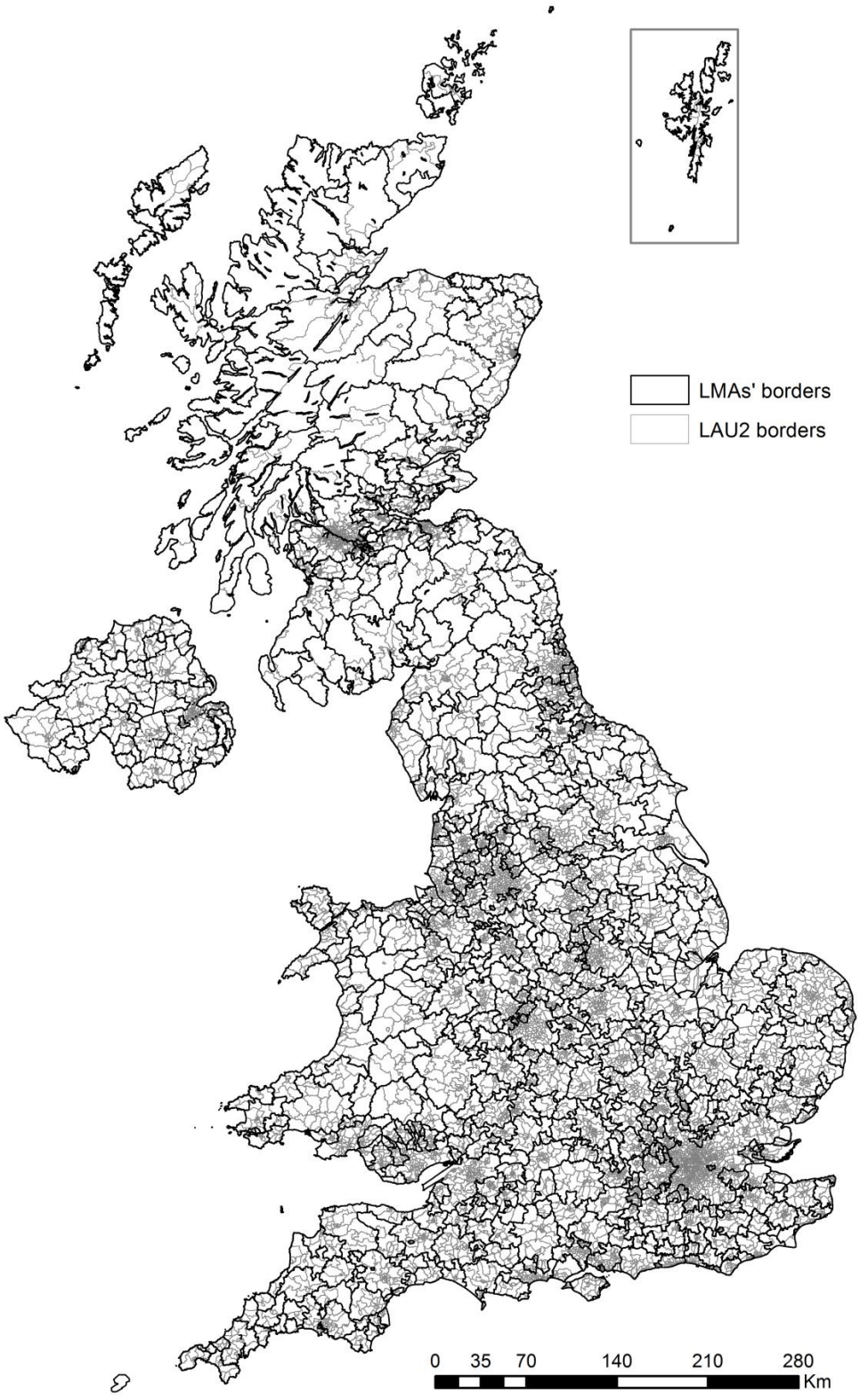
Maps 2 & 3 make it immediately evident that the size of the defined LMAs across both Spain and the UK varies hugely, often over short distances, in a way which was not found in Sweden (Map 1). In modern countries most people can commute some distance to access suitable jobs, with the result that very small LMAs scarcely exist (except perhaps on remote islands). This means that the results here in Spain (Map 2) are implausible in areas such as those immediately surrounding Madrid – the large LMA in the centre of the country – where almost all of the LAU2 areas have remained either unallocated zones or have been defined as LMAs of just one or two very small zones. Similar

[34] Commuting data from Census of Population 2001 were facilitated by the Spanish NSI (INE) to the University of Alicante for research on LMAs (a full explanation of the details can be found in Casado, J.M.; Martínez, L. and Flórez, F. (2010) "Los mercados locales de trabajo españoles. Una aplicación del nuevo procedimiento británico", in Albertos, J.M. y Feria, J.M (ed.) *La ciudad metropolitana en España: procesos urbanos en los inicios del siglo XXI*. Madrid, Thomson-Civitas, pp. 275-313.

problems can also be seen, if to a rather less extent, in between substantial LMAs in the UK (Map 3).

Another problem is illustrated by the large LMAs such as that which covers the Madrid area and that covering London: there are 'enclave' areas which are totally surrounded by the main LMAs. These situations can arise where some LAU2 areas have met the criteria to be LMA centres but then the pattern of their commuting inflows does not

Map 3. LMAs from the implementation of LAM method applied to UK01


LMAs' borders
LAU2 borders
0 35 70 140 210 280
Km

match the criteria of the method for them to be further grouped. These outcomes can be seen to be due to the criteria for identifying centres in the LAM method being based on the situation in Sweden: there the LAU2 areas are whole settlements, perhaps with nearby areas included too, but this is far from always the case elsewhere.

This is a problem which is not restricted to the LAM method because no centre-based method has been able to produce consistently satisfactory results when analysing individual small zones (especially in the UK). In fact this was also shown by the new OECD centre-based city region definitions: this method had to begin with a very major task of GIS-based analysis to group the small zones so that they identify whole built-up areas so to guarantee that the final areas would not to include any enclave areas. Whilst this is indeed one certain way of preventing such enclaves, it is inappropriate when applied as a basis for defining LMAs across all parts of the EU (whether or not these are 'metropolitan'). It clearly prioritises the morphology of urban areas over the functional linkages which are fundamental to LMAs. To put this into clear perspective, there will be some areas which are physically connected and yet are not strongly linked by commuting flows (eg. in the UK, W. Yorkshire and the W. Midlands). In any such case to solve the potential enclave problem by initially grouping all the built-up areas will prevent the commuting analyses defining the most appropriate results in those physical conurbations which embrace several functionally distinct LMAs.

The limitations of the LAM method which have been identified in this part of the report must be weighed against its strengths. Above all perhaps, its nature as a centre-based definition of LMAs gives it familiarity and intuitive appeal. It is also a distinctly simple procedure which makes it readily replicable, an advantage in the policy environment where transparency is valued. Unfortunately this very simplicity is part of the reason for some of the weaknesses identified here, including its poor transferability to MSs where many LAU2 areas are small and/or only cover part of a coherent settlement. Part of the problem is the inherent difficulty of identifying centres with 'building block' zones lacking those inherent characteristics of the LAU2 areas in Sweden which will have been taken for granted when their LMA definition method was created for application to these zones. In much the same way, the method produced Swedish LMAs with suitable characteristics even though it had no explicit self-containment or size minima applied. When the same method was applied to data for other countries such as ES and UK however, the results were not so satisfactory (Appendix, Table A.6).

The final question to ask here is whether these problems with the LAM method could be redressed by adding an extra element to the definitional process to further group areas which were inadequately self-contained (for example). One existing official method which proceeds by iteratively grouping 'proto-LMAs' until they all meet the statistical requirements set is the method to define TTWAs. Experimental research has found that it is indeed possible to take results from the LAM method and input these to the TTWA method so all the final LMAs are of adequate self-containment and size. Without going into the details of the results, it is possible to recognise that solving the problem of the LAM method in this way must have a key disadvantage of negating two of the crucial benefits of that method viz: its simplicity and the primacy of its focus upon commuter flow links to identified centres.

3.3. LMAs defined by a method that does not prescribe a single structure

The previous section concluded that although a centre-based method has a clear advantage in the intuitive appeal of areas orientated around towns and cities, enhanced by the simplicity of the sample method (LAM) tested here, major problems arise when applying the method to other countries. It seems that the centre definition method is not transferable to countries in which the LAU2 boundaries cut through settlements because, for example, they are relatively small. This section of the report considers a method without a prescribed single structure – such as that every LMA should have (at least) one main centre – to test whether that method is more transferable between the countries examined in detail here.

There is one method which already has a record of transferability because it has been adopted and adapted in different countries in several continents: this is the method that has been developed as the way to define TTWAs (Travel-to-Work Areas) in the UK over several decades now. Although earlier forms of the method did initially identify 'job foci' as a type of centres, its evolution since then has resulted in TTWAs that have no presumed structure. The method used now[35] 'explores' the data to seek out clusters of commuting flows of any form. The fact that the TTWAs have had long-term UK policy

[35] In the TTWA method every zone is initially a potential LMA. The method then proceeds by iteratively considering the LMA with the lowest score on the criteria of validity. If that LMA does not fulfill the set criteria it is dismembered and its constituent zones are reassigned to whichever remaining potential LMAs scores highest on the interaction index. The criteria of validity are codified in terms of a trade-off relationship that has a target level for the lower supply and demand-side self-containment values which is relaxed for zones exceeding a minimum size in terms of economically active population (as described in the Annex to Chapter 3 in the IRR).

use shows the method has proved acceptable in policy arenas, despite lacking the ready acceptability that centre-based methods are here credited with. On the same basis it can be argued that the fact that the TTWA method involves a more 'black box' definition process than the simple LAM centre-based method described above has not held back the use of TTWAs in relevant policy fields.

The clarity and relevance of its primary criteria has helped the TTWA method achieve its proven transferability to a diverse range of different geographical circumstances. All the defined TTWAs must possess the appropriate set statistical characteristics, guaranteeing that every one satisfies minima in terms of self-containment and size (with a certain level of 'trade-off' allowed between these two criteria). In any transfer of the method to another context, these criteria can be treated as parameters that are adjusted to make them appropriate to different circumstances. The key objective for this study is to produce ***comparable LMAs across the whole of the EU*** so it is essential that the same criteria values are used in all countries, but exactly what those values should be is yet to be determined.

In this section of the report the TTWA method is applied to several countries using the criteria values derived from those that defined the official 2001-based UK LMAs (Coombes & Bond 2007). In fact the criteria values used here are not precisely those used for the official definitions, because they had used an 'external' data source for the size measure, whereas here all the analyses rely completely upon commuting data. (In the original UK criteria the unemployed in the LMA size measure, adding the counts of those without work to total numbers of people in the commuting data.) As a result, the criteria here are not quite the same as those in the official UK TTWA definitions but instead are a 'scaled down' version[36] to adjust for the lack of unemployed people in the commuting data analysed here.

There is a rather more significant way in which this process of defining LMAs is not identical to the one which produced the official TTWAs in the UK. Before the official boundaries were confirmed there was a limited consultation process on the 'raw' boundaries produced by the computerised analysis of the UK01 commuting dataset. Any

[36] The parameters that were used for the 2001-based official set of TTWAs were adapted in this research to be based on total trip origins rather than total economically active populations: a simple calculation that divided the national number of trip origins by the economically active population gave some guidance as a 'deflator' (0.9395); applying this to both the minimum value 3500 and target value 25000 and then slightly rounding these produces minimum=3300 and target=23500 and so these values are used for TTWA runs here (along with the existing self-containment minimum .6666 and target .7500).

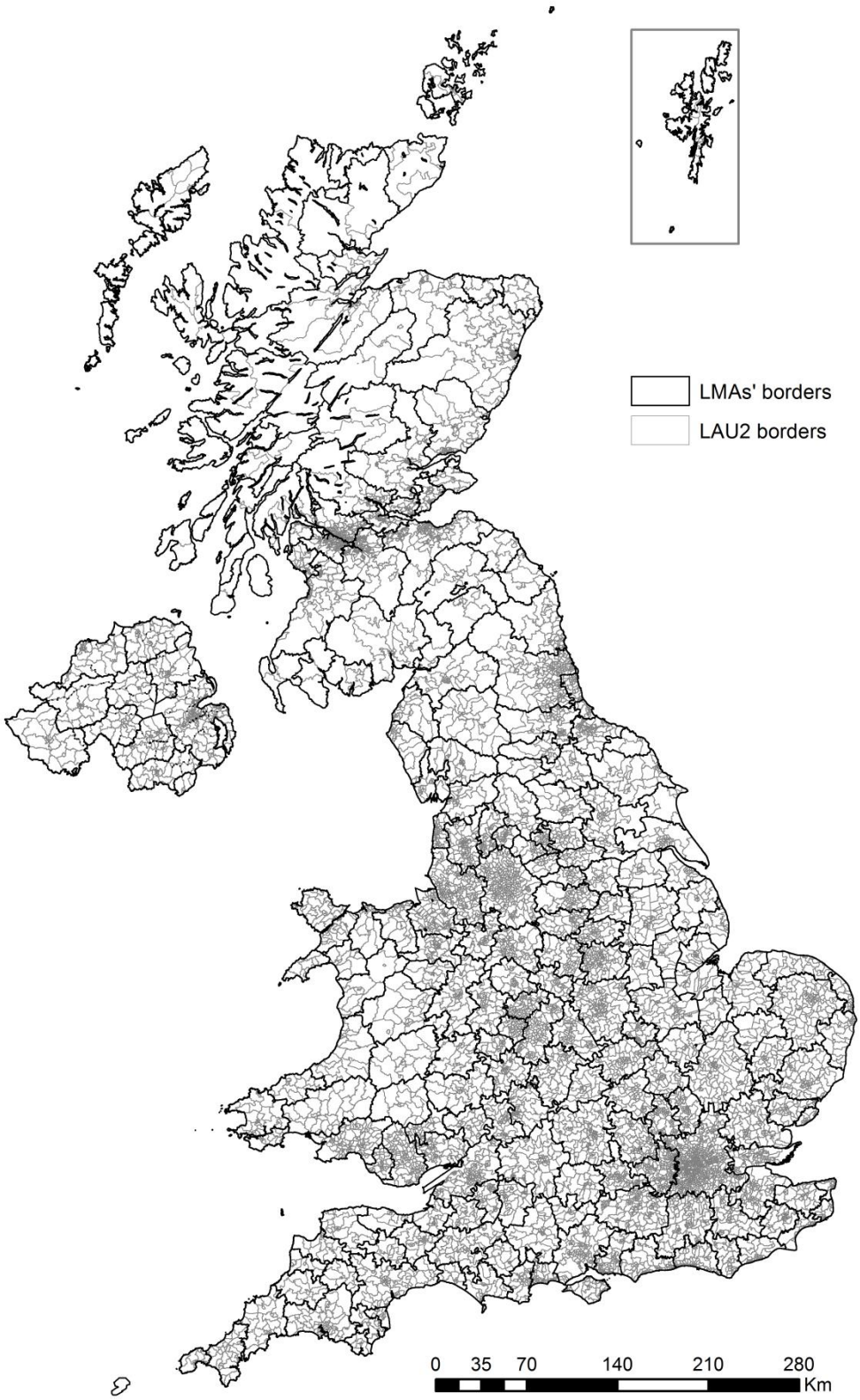
change to the boundaries had to meet rigid conditions to protect the critical statistical characteristics of the definitions, while the process did enable the small number of non-contiguities in the 'raw' boundaries to be resolved. Map 4 shows the LMAs produced here by applying the TTWA method to the UK01 dataset: it is of course 'raw' in that it has ***not*** been through a consultation process, so although the overall match to official TTWAs is of course very close there are some notable differences (which primarily involve there being more separate official TTWAs in some areas).

Map 4 shows the results of this TTWA analysis of UK01 dataset and as such is directly comparable to the results of the LAM analysis of this dataset (Map 3). Probably the most striking difference is that the TTWA method produces LMAs that are remarkably similar in area size in almost all parts of the UK. Most people avoid lengthy commuting trips and this means that localised clusters of flows dominate the commuting patterns. The key feature of the TTWA method is that it does not prescribe any single structure, such as a set of primary centres; instead the method has revealed this spatial structure in the flows themselves. By contrast when same flow dataset was analysed by the LAM method it could not reveal this inherent pattern because its definition procedure was determined to impose a centre-and-hinterland structure that is not always appropriate. This lack of transferability is partly due to the geographical conditions found in the UK (eg. the sheer size of London and the extensive old industrial regions), but also the challenge of defining centres when analysing the very small and seemingly randomly shaped zones in the UK01 commuting dataset.

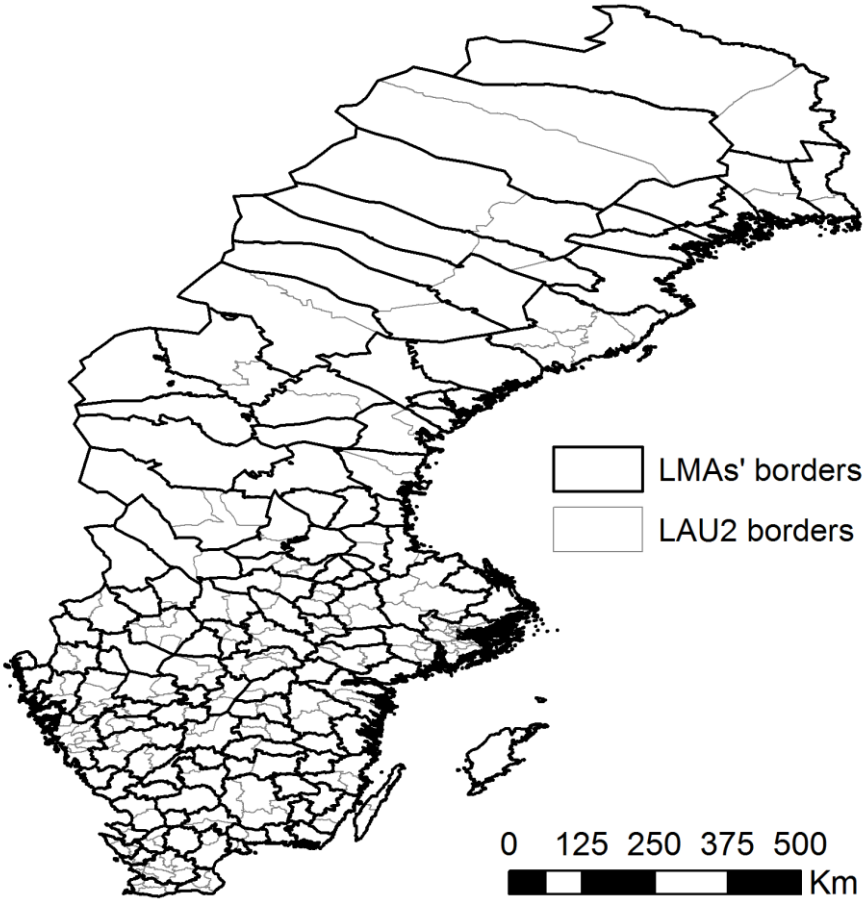
This leads naturally to the question of how transferable the TTWA method is to the Swedish situation for which the LAM method was devised. Map 5 shows the 126 LMAs produced applying the TTWA method to the SE01 data, a notable contrast in numbers to the 88 LAMs. In fact a substantial minority of the LAMs are exactly recreated by the TTWA method and this gives some encouragement that the results may not be wholly unacceptable in Sweden. The most notable difference is that LMAs produced by the TTWA method tend to be smaller than the LAMs in the areas around the largest cities. In these areas the local towns must not qualify as centres on the LAM criteria and thus have to become part of LAMs based on the nearby major cities, whereas the TTWA method is designed to enable any area to remain separable so long as it finally meets the statistical criteria set, and these LMAs must be satisfactory in terms of their size and commuting self-containment.

Of course, all the LAMs will be suited to the purpose they were created for, but the fact that the TTWA method has defined numerous additional statistically robust LMAs does not mean that it its results are inherently either worse or better. It may be the view of Swedish users that these TTWA-defined LMAs – while different to the LAMs – do also represent one 'dimension' of the complex modern labour market in these parts of the country. Such a view would be more evidence that the TTWA method does have a genuine transferability.

Map 4. LMAs from the TTWA method applied to UK01


LMAs' borders
LAU2 borders
0 35 70 140 210 280
Km

Map 5. LMAs from the TTWA method with applied to SE01



The issue of transferability is further evaluated by applying the TTWA method to the ES01 commuting data for Spain. Map 6 shows these results (nb. there is no official set of LMAs to compare this against, in an equivalent way to the above comparison of the Swedish results with the LAMs). There are numerous LAU2 areas which are isolated – not grouped with any others in LMAs – although there are considerably fewer than there were in the results from the LAM method analysis (Map 2). As was mentioned when discussing those earlier results, there are some LAU2 areas with no commuting flows in or out in the ES01 dataset and this makes them unallocable for any method that relies exclusively on commuting data. Setting aside these problematic LAU2 areas, Map 6 suggests that the TTWA method has produced LMAs of rather similar spatial extent

across much of Spain, although the most mountainous areas have larger LMAs (as was found in the north of the UK and also the large swathe of upland areas in the north-west of Sweden).

Map 6. LMAs from the TTWA method applied to ES01

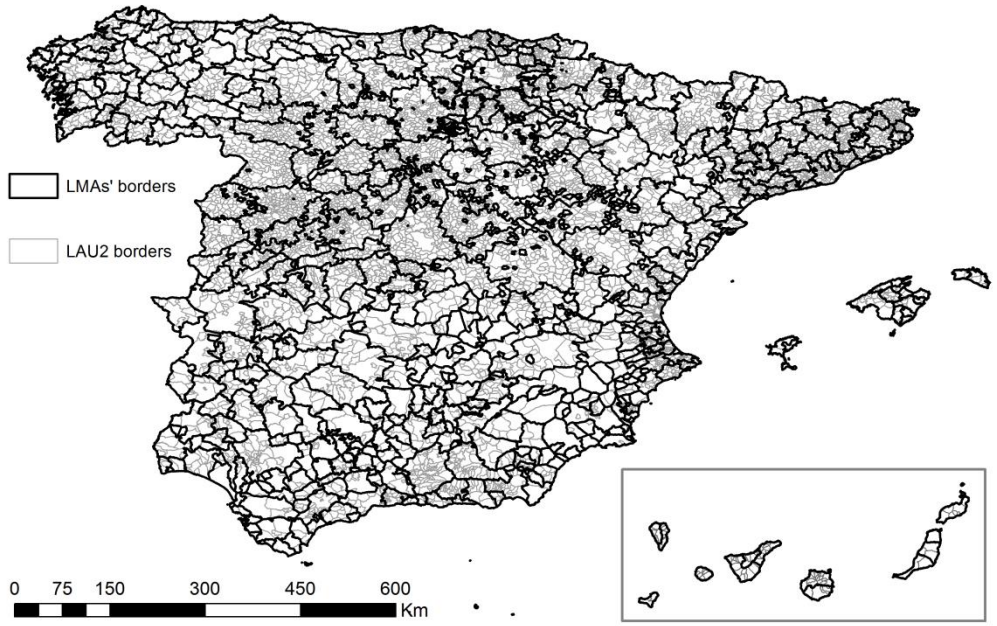


Table 3 provides the vital statistics on the LMAs defined by the TTWA method but, unlike the statistics for the LMAs defined by the LAM method seen earlier (Table 2), there is no need to state the minimum size or self-containment values of the LMAs defined here because the TTWA method ensures that they all meet the minimum values that were set. It is very positive finding that the median self-containment values for both ES01 and SE01 are well above 80% despite the TTWA method having defined large numbers of separable LMAs in both countries. The presence of some larger LMAs in both Spain and (especially) the UK than any found in Sweden clearly reflects genuine geographic realities: in the UK there is the world city of London, and also numerous large conurbations that are a similar size to the Barcelona or even the Madrid metropolitan regions, and as such they are larger than any to be found in Sweden. More surprising is the exact match between the median size of employed population of the Swedish LMAs from both methods considered so far: TTWA (Table 3) and LAM (Table 2). Given the much larger number of TTWA-defined LMAs than LAMs, this can

Table 3 LMAs identified by the TTWA method applied to ES01, SE01 and UK01

	Number of LMAs	LMA min. self-containment (%)		LMA employed population			Number of zones per LMA		
		Median	Mean	Median	Mean	Max.	Median	Mean	Max
ES01	492	85.9	85.0	8438	29947	2260167	8	16.3	156
SE01	126	82.7	82.6	14537	32469	840401	2	2.3	23
UK01	218	76.4	77.7	57819	122129	3376179	33	48.4	727

Note: The full set of available statistics for each national exercise is available in the Appendix (Table A.6)

be seen as further encouraging evidence of strong similarities between TTWA-defined LMAs and the officially-adopted LAMs.

A final test here of the transferability of the TTWA method is to apply it to the SIRE data available for the seven other countries whose commuting datasets are thought (after a briefer than ideal evaluation) to be sufficiently comparable for analysis here. Table 4 shows these results. It is not appropriate to draw too detailed conclusions from such overview statistics, but again it can be argued that the TTWA method has reflected the genuine geographical contrasts between the countries that are covered. In particular, Belgium is the most heavily urbanised of these countries, and it has been defined with LMAs which have the lowest median self-containment value and the highest median size of LMA employed population, just as is to be expected in a country with many closely-spaced substantial urban areas.

The evidence here of a high transferability of the TTWA method to other countries is, of course, rather limited and would benefit greatly from careful examination of the boundaries by NSI representatives of the countries covered. On present knowledge however it seems safe to conclude that the transferability which had already been seen in other countries has not been disproved here. One reason for this high degree of acceptability is that the method was developed to cope with UK LAU2 areas whose boundaries can seem to be almost random (as well as being highly unstable over time). A method which can find deep-seated patterns in the commuting flows between such areas will find little difficulty in analysing LAU2 areas with geographically meaningful boundaries whereas, by contrast, a method designed to be satisfactory where the LAU2 areas are highly suitable 'building blocks' for defining LMAs struggles greatly when transferred to a situation such as that in the UK.

Table 4 LMAs identified by the TTWA method applied to those countries with sufficiently comparable SIRE data c.2001

	Number of LMAs	LMA min. self-containment (%)	LMA employed population	Number of zones per LMA

		Median	Mean	Median	Mean	Max.	Median	Mean	Max
AT(SIRE)	69	77.6	78.2	22860	49224	998073	22	34.4	252
BE(SIRE)	28	76.4	77.4	49604	93760	656537	12	21.0	117
CH(SIRE)	76	79.7	80.5	21076	41971	501292	30	38.1	192
ES(SIRE)	477	82.6	82.8	9289	31179	2217852	8	17.0	216
FR(SIRE)	519	87.4	87.7	24315	57368	4837335	54	70.4	706
IE(SIRE)	38	82.3	82.1	16551	30243	392551	78	90.7	441
SE(SIRE)	124	84.2	84.0	13950	31332	788168	2	2.3	23

Note: The full set of available statistics for each national exercise is available in the Appendix (Table A.7)

Another reason for the relative success of the TTWA method is that its definition criteria have been developed to produce adequate results in the many different geographical circumstances in the UK (from a huge metropolis to remote islands). These criteria in fact feature a trade-off between the size and self-containment objectives and this approach can help with transferability to other very different geographical circumstances. In fact transference to countries beyond EU has usually involved adjustment to the size and self-containment criteria values in the trade-off. That approach is relevant for applications covering a single country, but here the interest is in a consistent application of the same criteria. The conclusion is that the TTWA method has encouraging levels of transferability but the criteria applied here (based on those used in the UK) may need to be adjusted to produce acceptable results in the maximum number of MSs.

3.4. LMAs defined by a method that is not deterministic

In very general terms, all official LMA definition methods, including those in SE and UK, are similar in that they start with all the individual LAU2 areas and then apply specific aggregation criteria to choose which areas to group. These aggregation criteria are measures of interaction and/or dependence derived from the commuting flow data, with the areas to be merged being those with the maximum score on that measure. The process ends when a certain global condition is met. What may not be self-evident is that the aggregation criteria – whatever they are – do not guarantee that the final set of areas is optimal. This is because they are identifying the best possible solution in terms of the immediately available options: the analyses are considering the local properties at that stage of the process, but these may favour an aggregation which may prevent the eventual emergence of the best global solution. This is a problem that is widely recognised in the research literature on solving very complex problems such as

the definition of LMAs (especially when the 'building block' areas run into hundreds of thousands).

The official methods are all deterministic: only one result can be derived from their application to any one dataset, given a certain set of criteria. It is useful here to also examine a non-deterministic approach, and the method[37] chosen is termed the GEA. This method seeks to maximise a global objective function that in effect measures the quality of any possible solution. The procedure can also incorporate some restrictions, and here the applications require that none of the final LMAs include non-contiguities.

The size of the optimisation problem – the best partition of a given territory into LMAs – makes it relevant to use a genetic algorithm. This is a stochastic exploration of the universe of all possible solutions, an intensive trial-and-error procedure that gradually improves the quality of the set of LMAs in terms of the objective function. It is possible to impose different termination conditions, but here the analysis proceeds until no further improvements can be found after 10000 trials. The optimisation strategy involves stochastic aggregation and disaggregation of LMAs – as well as the exchange of LAU2 areas between LMAs – and by allowing some changes which initially are locally sub-optimal the method can reach a better global final result due to the wider exploration of the many alternatives.

To be specific: the method does not stop when minimum self-containment and size criteria are met by all the potential LMAs, but instead the GEA continues exploring other alternatives to find the best definition of LMAs in global terms. To enable these results to be compared with those in previous presented here, the criteria used in the GEA analyses follow those that the TTWA method used. The objective function was defined as the sum of the attraction that in terms of commuting flows links each LAU2 area with the LMA it is part of, for the whole territory. The restrictions that the GEA method also observes are identical to those of the TTWAs in terms of self-containment and minimum size requirements for the final LMAs. In addition to the choice of criteria for the method, when there is also boundary information[38] available then a contiguity constrained version of the GEA method can be used, as is it has been here.

[37] Martínez-Bernabeu L, Flórez-Revuelta F and Casado-Díaz JM (2012) "Grouping genetic operators for the delineation of functional areas based on spatial interaction" *Expert Systems with Applications* 39, 6754--6766.

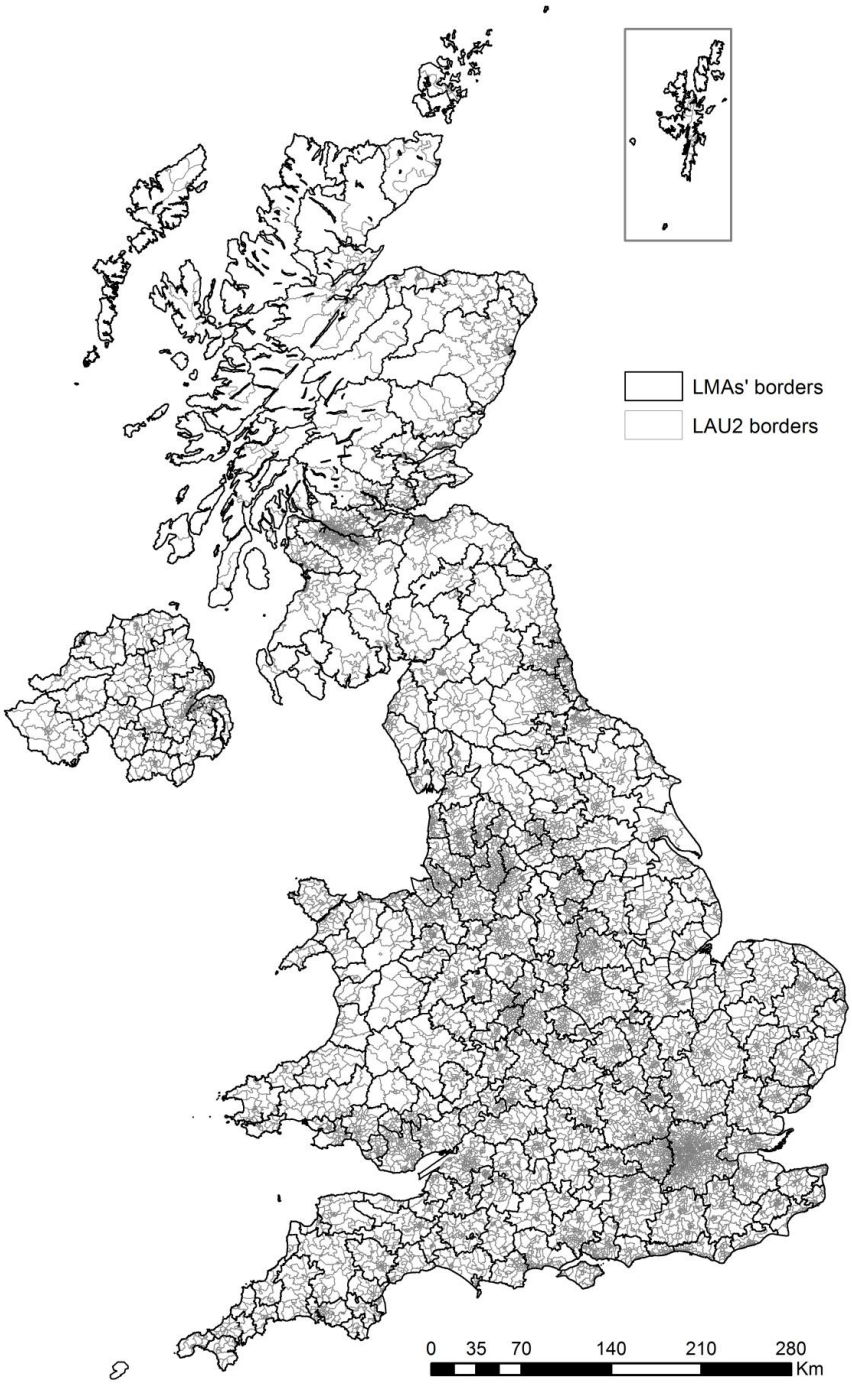
[38] The quality of this information is crucial here: it should ideally reflect not only physical neighbouring but also the existence of bridges linking two non-neighbouring localities (eg. across a bay). In fact the

The two main drawbacks of this procedure are its complexity – which makes it less readily understandable – and that its non-deterministic nature means that in this experimental form it may not always produce the same results when the same criteria are used on the same dataset, due to its inclusion of several stochastic components. This last feature is mainly relevant to the analysis of those areas with a sparse pattern of commuting flows where there is no one clearly dominant pattern. In such cases the GEA method continues exploring the space of solutions and depending on when the process is stopped some LAU2 areas may not always be allocated to the same LMA.

The following maps depict the LMAs definitions resulting from the application of the GEA method to the three main datasets for this study: UK01 (Map 7), SE01 (Map 8) and

datasets used here illustrate this because the UK boundary information was less than ideal so that some of the results there are prone to be inappropriate, essentially due to this dataset ignoring some bridges.

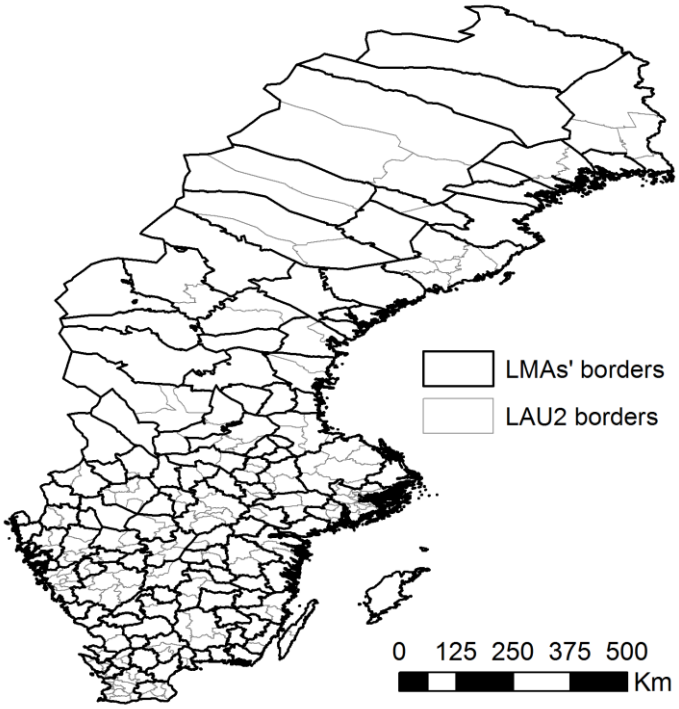
Map 7. LMAs from the GEA method applied to UK01



ES01 (Map 9). The results are broadly in line with expectations: boundaries produced by the GEA method are more similar to those produced by the TTWA method than those by the LAM method with its dependence upon the initial definition of centres. This is largely due to the GEA method using criteria values (self-containment and size) which

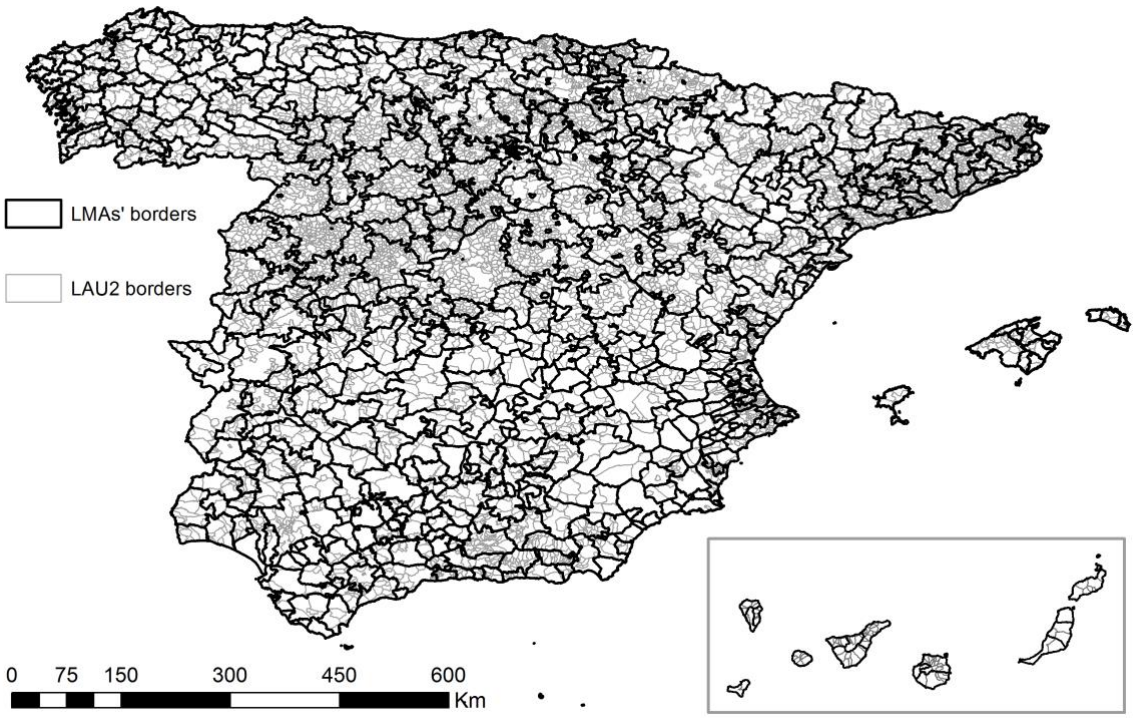
the TTWA method had used. All the same, it is valuable to discuss where the two methods have produced rather different sets of LMAs.

Map 8. LMAs from the GEA method applied to SE01



Map 9. LMAs from the GEA method applied to ES01[39]

[39] In all the cases in this section of the report, LMA definitions include a contiguity constraint; even so, there are some zones in that Spain for which no contiguously grouped solution can be found by GEA.



One relevant criterion to judge the results of different methods by is a count of LMAs defined which meet set statistical criteria. Table 5 shows the GEA optimising approach has defined 20% more LMAs in the UK than the TTWA method found. About half of the 47 'extra' LMAs are around large urban centres: for example in the region including Liverpool as well as Manchester the number of GEA-defined LMAs is more than double the number the TTWA method defined here. Around a third of all the 'extra' GEA LMAs are in the more rural parts of the UK. Perhaps more importantly, the results from GEA are boundaries that seem likely to gain ready acceptance due to closely matching the 'common knowledge' of local geography in most areas. The most surprising result may be the considerable variation between large cities in the how large their GEA-defined boundaries are when compared to those from the TTWA method, because despite the two analyses using the same data and basic criteria the LMAs of some cities are larger in one set of results, while the opposite is true for some other cities.

Table 5 LMAs identified by the GEA method when applied to ES01, SE01 and UK01

Data	No. LMAs	LMA min. self-containment (%)		LMA employed population			Number of zones per LMA		
		Median	Mean	Median	Mean	Max.	Median	Mean	Max
SE01	120	84.4	83.4	14029	34092	840401	2	2.4	23
ES01	583	85.2	84.4	6259	25273	2244969	8	13.8	157
UK01	265	73.6	75.2	50346	100468	3214712	27	39.8	692

Note: The full set of available statistics for each national exercise is available in the Appendix (Table A.6)

Similar results are obtained when comparing the maps of ES LMAs identified using the TTWA method and those using the GEA. The total number of identified LMAs is around 20% higher in the GEA case while the size and self-containment values of the largest LMAs are rather similar. As in the UK case the GEA identifies both smaller and larger LMAs in large cities when compared to the TTWA method and this result may at least partially derive from the fact that GEA is contiguity constrained. What is more certain is that the GEA method is especially active in identifying more LMAs in less populated parts of Spain, and especially around second-rank provincial cities, where the TTWA method tends to define LMAs whose surface area is comparable to that of the main urban agglomerations.

A deeper analysis of the results from GEA shows some suboptimal assignments, primarily a number of LAU2 areas having higher interaction values with some LMAs other than those to which they have been allocated. Such results may be termed 'convenience marriages' and occur more frequently in the GEA method results than they had in the TTWA results as a direct result of the GEA method getting closer to the global solution for a given set of criteria and restrictions: they are a local disadvantage of achieving a global benefit. Although there are some technical alternatives which could deal with this issue, to add these to the method would worsen one of the negative features of the GEA method in a policy context: its complexity.

The factors by which to evaluate methods include the statistical characteristics of the areas produced, and the views of experts in the countries analysed. Among the other relevant considerations for policy-makers could be a preference for sets of areas which are less queried: this is in fact probably one reason why the use of administrative areas has continued for so long. From that perspective – despite the results from the GEA method being promising – the conclusion at this stage seem likely to be that its more technical advantages *may* be outweighed by it still being experimental technology, while its greater complexity also makes it more difficult to explain. The other concern is that its non-determinism at this experimental stage may not be welcomed by the policy community. In the short term then, an alternative possibility that deserves mention is the use of this procedure as part of a final stage in the definition of LMAs through which the contiguity constraint necessary for policy areas could be ensured. To be more ambitious, this step could also provide a degree of optimisation of the 'raw' boundaries in consort with a process of consultation.

3.5. Towards future research options

Section 3.3 concluded by recognising that the TTWA method has a relatively high level of transferability but its criteria have usually been adjusted when the method has been applied to other countries. This section seeks hypothetical 'Euro' standard criteria from an examination of existing official methods of defining LMAs. Subsequently the criteria are applied to the datasets collated for this study.

Comparing equivalent thresholds in existing official methods (which are fully described in the Annex to Chapter 3 in the IRR) reveals the values which are relevant here.

- The method in Sweden has no minimum size, but actually their smallest LAU2 has over 2500 people so that is roughly equivalent to 1000 residents in work
- In the official methods examined, the lowest population minimum was the 1000 jobs in the IT method and this is very similar to 1000 employed residents
- The highest population minimum was the 10,000 people in the CZ method which is roughly equivalent to 5000 residents in work[40]
- The highest self-containment minimum is the 75% in the method of IT
- The lowest self-containment value is the 66.66% minimum used in the UK

How can a hypothetical 'Euro' standard set of criteria be derived from these values? Two potential considerations present themselves. Where an official definition exists, the respective NSI may consider that 'Euro' LMAs should be no bigger than their 'native' LMAs, and similarly that their minimum self-containment should not be higher than has been required of the LMAs they define themselves. However there could also be the opposite response, with the NSI in fact asking "why at the European level did you define smaller and/or less self-contained LMAs than we think make sense?"

Given that it has already been suggested that the TTWA method provides the basis for experiments with different criteria, it is valuable that this method provide a trade-off between the self-containment and size criteria. For each of these, it is necessary to set an absolute minimum and a 'target' that is, in effect, the level above which all higher values are considered equal (ie. the target values are those that are the highest 'minimum' values that are of interest).

[40] The criteria in the DE method are very different to those discussed here, but the method also includes discretionary decisions which can over-rule the criteria, so they are not considered here.

Bringing all these considerations together now allows the thresholds in existing official methods to yield a hypothetical 'Euro'standard set of criteria to test.

- 'Euro' LMAs must have an employed population size minimum of 1000 but they are set a target size of 5000
- 'Euro' LMAs must meet the self-containment minimum of 66.67% while at the same time their target is 75% (where both these values are the lower of their supply and demand self-containments)

It is important to note that this is a purely hypothetical set of 'Euro' criteria which has been derived here purely for experimental purposes[41].

Tables 6 and 7 summarise the main features of the sets of LMAs defined here using these criteria. One preliminary issue to deal with is that of the observable differences between the results in Spain and in Sweden depending on whether the analyses used SIRE data or not [ie. ES01 & SE01 vs. ES(SIRE) & SE(SIRE)]. These differences arise due to the SIRE datasets only including the largest 30 flows from the LAU2 areas covered. The impact of this difference is substantial in Spain where applying the TTWA method to the SIRE data defines 12% less LMAs (that are also less self-contained) in comparison to the LMAs defined using the unrestricted dataset ES01.

[41] In Appendix 2, the alternative of raising the criteria that must be satisfied (therefore the identification of significantly fewer separate LMAs) is briefly explored.

Table 6 LMAs defined by the 'Euro' method when applied to ES01, SE01 and UK01

Data	No. LMAs	LMA min. self-containment (%)		LMA employed population			Number of zones per LMA		
		Median	Mean	Median	Mean	Max.	Median	Mean	Max
SE01	166	80.9	80.4	10174	24645	840401	1	1.7	23
ES01	892	82.9	82.6	3641	16518	2258670	5	9.0	130
UK01	257	75.0	76.5	43312	103596	3310191	25	41.1	706

Note: The full set of available statistics for each national exercise is available in the Appendix (Table A.6)

Table 7 LMAs defined by the 'Euro' method in countries with sufficiently comparable SIRE data c.2001

	No. LMAs	LMA min. self-containment (%)		LMA employed population			No. of LAU2 zones per LMA		
		Median	Mean	Median	Mean	Max.	Median	Mean	Max
AT(SIRE)	80	76.5	77.6	17709	42456	1008059	19	29.6	259
BE(SIRE)	30	76.0	76.4	49604	87509	664508	12	19.6	120
CH(SIRE)	107	78.5	78.9	12049	29811	502635	16	27.1	165
ES(SIRE)	798	79.5	80.1	4572	18637	2220037	5	10.1	177
FR(SIRE)	729	85.4	85.8	13889	40842	4498780	36	50.1	623
IE(SIRE)	55	79.5	78.9	10225	20895	393719	46	62.7	447
SE(SIRE)	166	81.7%	81.4%	9645	23404	788168	1	1.7	23

Note: The full set of available statistics for each national exercise is available in the Appendix (Table A.7).

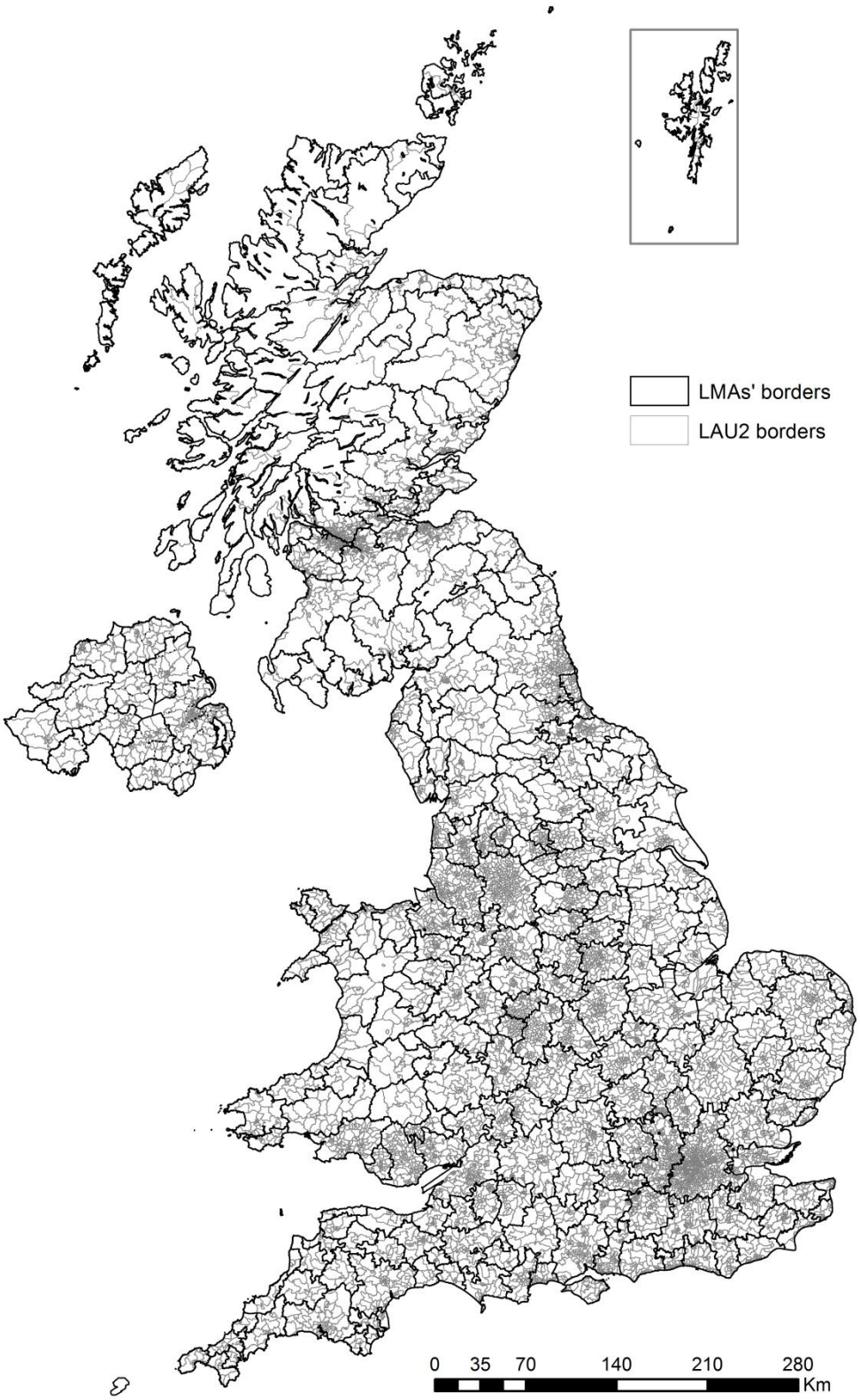
One way to assess the appropriateness of the selected criteria is their ability to define a set LMAs which is 'balanced' throughout the territory. Maps 10, 11 and 12 shows the results produced in the UK, Sweden and Spain. All the maps suggest sets of LMAs which are quite balanced in terms of their size distribution. Where there are much larger areas, these appear to be appropriately reflecting local geographical realities. For example the larger Spanish LMAs either cover metropolitan regions or other areas where there are dense interrelationships in commuting terms, or they are in the most rural areas where the low populations lead to the amalgamation of LAU2 areas in order to meet the population minimum size. There are similar cases in the UK but otherwise there is a strong size consistency, and similar generalisations can be made about the results in Sweden (where the rural areas cover most of the north-west of the country).

An overview at this point involves comparing the results of the 'Euro' analyses with the number of LMAs in countries where there is an official definition of LMAs.

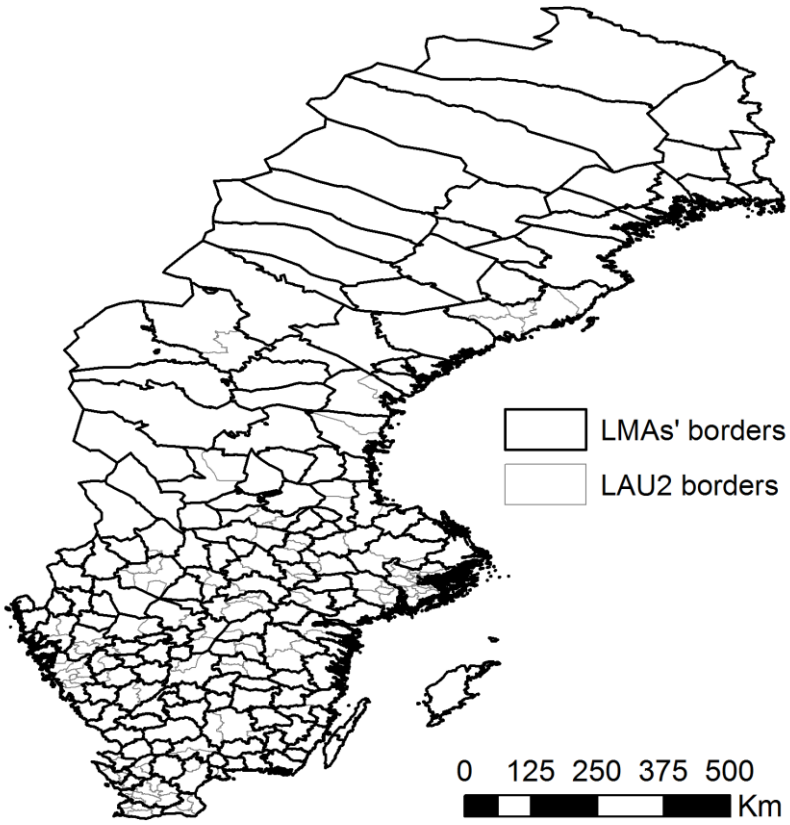
- The number of 'Euro' LMAs identified is much higher in SE compared to the official definition of LAMs and a similar outcome applies to FR; in both cases there are roughly twice as many 'Euro' LMAs as their national equivalents.
- In dramatic contrast, there are slightly fewer 'Euro' LMAs than TTWAs in the UK and the difference is more notable still in BE (47 vs. 30).

These results suggest that the 'Euro' criteria have perhaps found a 'middle way' between the national definitions considered. However the aim of this part of the report was not to propose a definitive set of criteria, but instead to further illustrate the possibility of producing reasonable results in different countries with a consistent application of a common set of criteria.

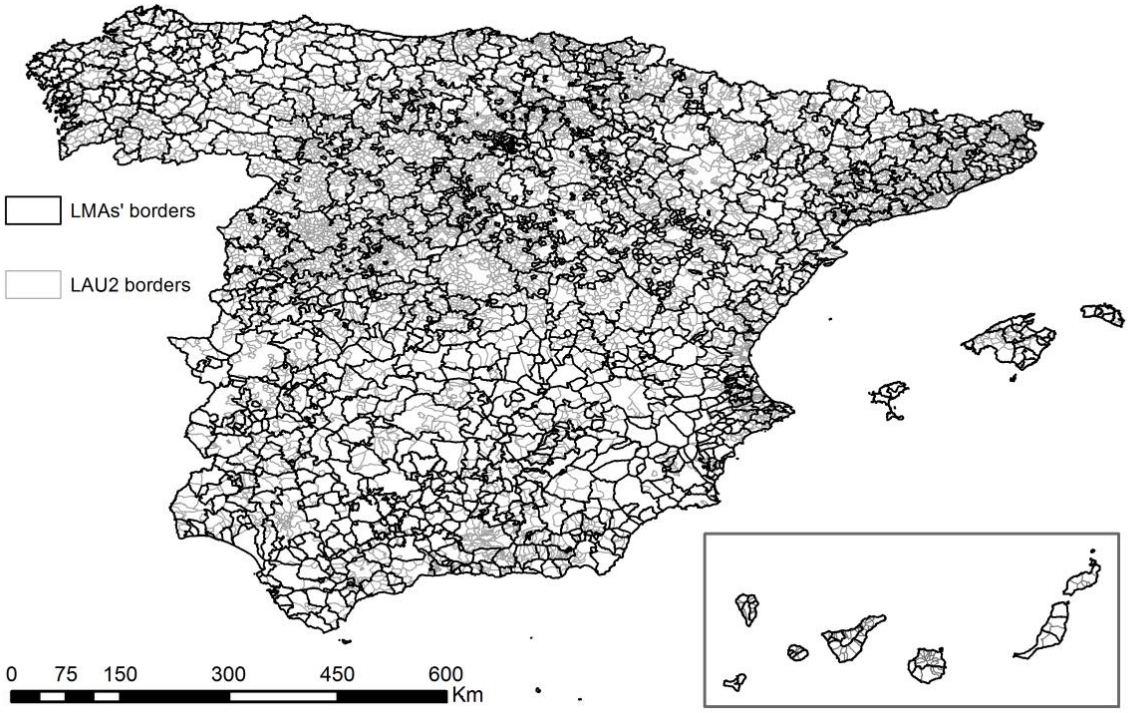
Map 10. LMAs from the 'Euro' method applied to UK01


LMAs' borders
LAU2 borders
0 35 70 140 210 280
Km

Map 11. LMAs from the 'Euro' method applied to SE01

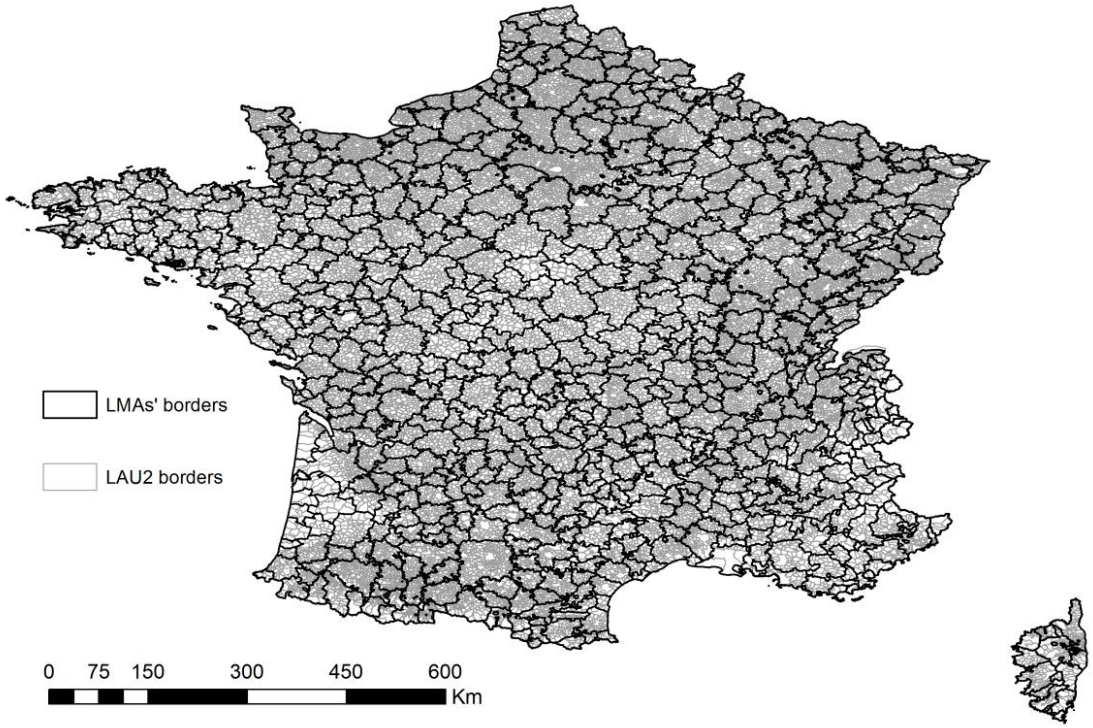


Map 12. LMAs from the 'Euro' method applied to ES01



Map 13 is particularly interesting because it covers France and so can be seen as one example of seeking transferability to another country not previously examined in this report in any detail. The application of the 'Euro' method to FR(SIRE) results in the definition of 729 LMAs, a figure more than double that of the official set of LMAs (*bassins d'emploi*). The increase is distributed in a very balanced way across the whole territory except for the region around Paris. In the national set of LMAs this region has been divided into many *bassins d'emploi* whilst here Map 13 shows the region divided into only four LMAs: more and smaller LMAs are simply not sufficiently self-contained due to the strong commuting flows around the metropolis. The different result with *bassins d'emploi* is in practice the result of the French method allowing for several 'special cases' such as that *grandes communes* like Paris can be the subject of different criteria within the national definition procedure. That approach stands in stark contrast to the consistent application here of the 'Euro' method to all areas of each country, and indeed to several countries simultaneously.

Map 13. LMAs from the 'Euro' method applied to FR(SIRE)



3.6. Summary and recommendations

This last section summarises the conclusions which can be drawn from the research, leading towards some recommendations through a set of questions and answers.

1. ***Is there a need for a grid of comparable LMAs?***

Other parts of this study have shown not only that several MSs find consistently defined national set of LMAs useful for policy analyses, but also that there are EU policies for which consistent cross-national LMA definitions could be very valuable.

2. ***Is the definition of EU-wide LMAs technically feasible?***

The analyses conducted in this study suggest that there are no insurmountable technical obstacles to producing cross-national LMA definitions; in the following questions the issues involved are considered separately in turn.

3. ***What data would be needed?***

The study found that the vast majority of official definitions of LMAs in the EU MSs rely on commuting flows between LAU2 areas and for the analyses reported here the only other information used was GIS-based boundary data. In the vast majority of cases this information derives from Population Censuses and there newer data will be available for most MSs in the next few years. (The equivalent commuting data from the 2001 Census 'round' was collated in SIRE: a repeat procedure should avoid recreating the problem caused by several NSIs supplying datasets subjected to severe reduction – notably only including the largest 30 flows from any LAU2 area – because this prevents consistent analyses being conducted.). There remain some MSs with no commuting dataset and this presents a residual difficulty.

4. ***Can any method work adequately in very varied geographical conditions?***

This study has extended the existing evidence that the TTWA algorithm has a high degree of transferability (as a result of its long-term development dealing with the geographical variety in the UK, and the technical challenges that its LAU2 areas pose). The flexible criteria of the method reveal patterns in the commuting flows which are usually recognisable as territorial reality, whether the LMA has a single centre or not. LMAs in all regions – from metropolitan to remote rural – are defined so that they possess the same minimum statistical characteristics.

5. ***What criteria should be used for EU-wide LMAs?***

As well as the essential requirement that LMAs should all be reasonably self-contained in terms of commuting flows, there is a good case for a size minimum in order to limit the sensitivity of data analysed using LMAs (nb. the sample survey collection process of some relevant datasets might make very small LMAs of minimal value in any case because key policy indicators would be unavailable at that scale). There do not exist 'ideal' minimum values of size or self-containment for LMAs: selecting these values must be resolved empirically by examining alternatives to find a generally acceptable solution in terms of the nature and size of the LMAs that are produced. It was in this experimental mode that a test has been conducted here on an extensive group of MSs using a hypothetical set of 'Euro' self-containment and size criteria. The necessary next step is for further consideration of this issue by Eurostat and the NSIs.

6. ***Can the definitions be entirely automated?***

Most national methods include a final refining step in which 'raw' definitions are evaluated against local knowledge. In this extra step any boundary non-contiguities within the 'raw' definitions can be resolved (nb. those definition processes that are

themselves contiguity constrained tend to produce sub-optimal LMAs). To ensure that a truly consistent EU-wide grid of LMAs definitions is produced it will be necessary for any such adjustments to be restrained so the final boundaries continue to meet the set statistical characteristics, thus avoiding *ad hoc* solutions that undermine the general validity of the definitions. In fact the GEA method used in Section 3.4 can offer one way to ensure the final LMAs have no non-contiguities, as part of a general 'fine tuning' (perhaps linked to a consultation process).

7. *Who should carry out the definitions?*

The familiar alternatives are either parallel national analyses, or a centralised process conducted by Eurostat. In the latter case the NSIs would need to provide commuting data meeting a Eurostat specification and would also be crucial to any consultations. The other alternative is that NSIs conduct the analyses themselves, subject to Eurostat overview to ensure that the common method is applied consistently and to protect the statistical characteristics of the LMAs. Successful precedents exist for both these options so the decision is left to be agreed between the NSIs collectively and Eurostat.

Appendix 1 to chapter 3 - Recommendations on the establishment of an EU-wide harmonised grid of comparable LMAs

Table A.1. Characterisation of the basic zones used in the analysis (i)

	No. LAU2	Population					Area (m²)				
		Min	Max	SD	Mean	Median	Min	Max	SD	Median	Mean
BE[1]	589	84	472071	28706.39	18110	11612	1142297	213750185	**37820201.5**	40100910	51830094.2
BG[2]	5302	0	1165503	18311.37	1427	243	0	492029000	**23980272.8**	15490500	20942204.1
CZ[1]	6249	3	1233211	17574.47	1675	412	422868	496025461	**16689088.3**	7986477	12622469.7
DK	2244	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	**n.a.**	n.a.	n.a.
DE	12229	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	**n.a.**	n.a.	n.a.
EE[1]	227	72	398594	27773.35	5905	1808	1760000	871620000	**143774185.1**	175520000	191331894.3
IE[2]	3441	n.a.	n.a.	n.a.	n.a.	n.a.	50000	127030000	**14011428.2**	19450000	20395884.9
GR[1]	6130	0	789166	13526.32	1784	381	200000	577171000	**24155819.8**	14518500	21526494.8
ES[1]	8112	5	3255944	47368.70	5763	585	25784	1750327196	**92377422.2**	34896870	62220848.6
FR[2]	36682	n.a.	n.a.	n.a.	n.a.	n.a.	40000	18360000000	**143778260.6**	10750000	17250025.1
IT[1]	8094	n.a.	n.a.	n.a.	n.a.	n.a.	150000	1307710000	**49944646.5**	21815000	37232037.3
CY[1]	615	n.a.	n.a.	n.a.	n.a.	n.a.	276339	153834000	**13512659.7**	11718600	15028374.3
LV[2]	118	1364	706413	65679.94	19054	7082	17490684	2524643008	**506206781.0**	369623654	547134034.7
LT	555	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	**n.a.**	n.a.	n.a.
LU[1]	116	293	88586	8982.24	4254	1985	5290000	113360000	**14106227.4**	19675000	22296206.9
HU[1]	3152	12	1712210	31829.39	3182	837	560000	525130000	**36767506.2**	18685000	29513781.7
MT[2]	68	241	22492	4877.34	6073	3994	159823	26599210	**5206010.9**	2789774	4634657.2
NL[2]	431	n.a.	n.a.	n.a.	n.a.	n.a.	4520000	460320000	**68076079.6**	59100000	78256542.9
AT[2]	2357	n.a.	n.a.	n.a.	n.a.	n.a.	358100	466783800	**38524247.7**	24274000	35783686.6
PL[1]	2478	1346	1709781	50645.53	15390	7407	3320000	633700000	**78799993.9**	111825000	126182155.0
PT[1]	4260	n.a.	n.a.	n.a.	n.a.	n.a.	51550	416802430	**34517479.9**	11287720	21618398.4
RO[2]	3180	n.a.	n.a.	n.a.	n.a.	n.a.	1890000	804490000	**55662685.3**	61350000	74965632.1
SI[1]	210	322	276091	21624.83	9678	4708	6934510	555385710	**88364178.0**	64739460	96538140.6
SK[1]	2928	0	112907	5995.67	1848	637.5	357863	359787871	**18735969.3**	11607310	16753298.2
FI[1]	348	115	576632	40575.91	15305	5752	6	17333.89	**1620.1**	699.11	1123.3
SE[1]	290	2549	795163	61429.10	31665	15285	8710000	21891000000	**3078599788.1**	679610000	1669456034.5
UK[2]	10310	n.a.	n.a.	n.a.	n.a.	n.a.	43450	1638643400	**68453662.6**	4984152.5	23868798.7

Notes
([1]) 1 January 2009: zip file all EU27 countries from http://epp.eurostat.ec.europa.eu/portal/page/portal/nuts_nomenclature/local_administrative_units (except Greece, from individual file same address)
([2]) 1 January 2010: individual country file from http://epp.eurostat.ec.europa.eu/portal/page/portal/nuts_nomenclature/local_administrative_units
A major reduction in LAU2 took place in LV 01/07/2009

Table A.2. Characterisation of the basic zones used in the analysis (ii)

Dataset	Working population per zone					Jobs per zone				
	Min	Max	Mean	Median	Std.dev.	Min	Max	Mean	Median	Std.dev.
AT(SIRE)	20	667006	1432	642	14134.0	2	837072	1432	306	18146.4
BE(SIRE)	22	105173	4457	3044	6470.0	5	187872	4457	1554	13313.3
CH(SIRE)	7	165494	1101	376	4288.3	2	315892	1101	169	7629.7
ES(SIRE)	1	1209296	1834	177	17052.3	0	1560221	1834	117	21268.5
FR(SIRE)	0	1535231	814	181	9400.2	0	2330159	814	88	13627.1
IE(SIRE)	0	6819	325	154	499.7	0	20657	325	71	1028.6
SE(SIRE)	1018	345749	13443	6447	25918.2	830	515612	13443	5441	36055.8
ES01	0	1225956	1817	150	17325.8	0	1485561	1817	122	20382.5
SE01	1033	374121	14156	6731	27879.5	825	531912	14156	5953	37315.5
UK01	237	17725	2522	2032	1725.2	61	266442	2522	1386	5222.7

Note
Employed population and *jobs per zone* as per data used in the analyses included in the text, based on commuting datasets from SIRE or NSIs sources (INE, ONS and SCB).

Table A.3. Characterisation of the basic zones used in the analysis (iii)

Dataset	Trip origins per zone					Trip destinations per zone				
	Min	Max	Mean	Median	Std.dev.	Min	Max	Mean	Median	Std.dev.
AT(SIRE)	0	2136	29.2	15	67.0	7	30	29.2	30	2.6
BE(SIRE)	0	580	29.4	16	46.6	9	30	29.4	30	2.4
CH(SIRE)	0	1399	26.7	11	56.9	1	30	26.7	30	6.7
ES(SIRE)	0	4931	18.6	6	81.4	0	30	18.6	19	10.3
FR(SIRE)	0	10374	22.6	7	78.4	0	29	22.6	27	7.8
IE(SIRE)	0	462	12.9	3	31.0	0	30	12.9	10	9.3
SE(SIRE)	4	288	29.9	17	38.7	24	30	29.9	30	0.5
ES01	0	1313	22.6	10	41.5	0	1145	22.6	12	36.1
SE01	14	287	99.4	85	59.5	19	279	99.4	89	45.1
UK01	3	4477	144.0	92	181.1	18	442	144.0	128	73.4

Note
Trip origins/destinations per zone refer to the number of LAU2 zones that act as origins/destinations for an specific LAU2 zone (e.g. if zone A receives 20 workers from zone B and 30 workers from zone C then number of trip origins for zone A is 2). It is noticeable that for SIRE data a maximum of 30 destinations is reported for each LAU2 (a limit that is evident when those data are compared with those provided by national NSIs: INE, ONS and SCB).

Table A.4 Characterisation of the basic zones used in the analysis (iv)

Dataset	Supply-side self-containment (%)					Demand-side self-containment (%)				
	Min	Max	Mean	Median	Std.dev.	Min	Max	Mean	Median	Std.dev.
AT(SIRE)	5.3	94.5	31.6	28.9	13.9	6.8	100.0	62.7	64.0	21.1
BE(SIRE)	9.1	81.1	30.9	26.8	14.2	6.2	100.0	52.9	53.5	16.9
CH(SIRE)	2.4	98.7	34.6	31.3	15.3	5.4	100.0	67.3	69.7	21.9
ES(SIRE)	0.0	100.0	54.7	55.4	18.2	0.0	100.0	79.6	84.5	18.7
FR(SIRE)	0.0	100.0	45.1	43.0	18.3	0.0	100.0	78.5	82.8	17.2
IE(SIRE)	0.0	100.0	27.5	25.0	17.5	0.0	100.0	56.8	56.3	29.3
ES01	0.0	100.0	62.9	65.4	21.0	0.0	100.0	74.1	78.9	19.3
SE01	21.0	95.1	67.6	71.7	18.3	13.4	96.8	76.5	80.1	13.1
UK01	6.6	95.3	25.9	23.3	10.6	0.2	97.4	38.6	37.8	18.7

Note
Supply-side self-containment is the share of employed population that work locally. Demand-side self-containment is the share of jobs that are occupied by workers who reside in the zone.

Table A.5 Characterisation of commuting datasets

	Number of LAU2 zones	*Connectivity*	*Polarisation*
AT(SIRE)	2371	1.23%	8.39%
BE(SIRE)	589	5.01%	11.04%
CH(SIRE)	2896	0.92%	8.11%
ES(SIRE)	8108	0.23%	8.19%
FR(SIRE)	36565	0.06%	6.92%
IE(SIRE)	3558	0.38%	11.60%
SE(SIRE)	289	20.79%	26.64%
ES01	8108	0.28%	12.19%
SE01	289	34.51%	27.34%
UK01	10558	1.38%	13.93%

Notes
Connectivity refers to the number of connections in the commuting matrix (flows≠0) over the total number of possible connection, n(n-1), where n is the number of zones. *Polarisation* refers to the % of zones that are destinations for the largest outflow from 1(+) other zone(s).

Table A.6 Characterisation of the results (i). Based on data provided by NSIs for a selection of countries

Data	Method	Assigned zones		No. LMAs	Global SC (%)	LMA min. self-containment (%)					LMA employed population					Number of zones per LMA				
		Yes	No			Min.	Median	Mean	Max.	SD	Min.	Median	Mean	Max.	SD	Min	Median	Mean	Max	SD
SE96	LAM	288	0	105	93.1	68.5	88.8	87.7	96.2	5.3	1256	11732	36313	963222	103428	1	1	2.7	34	4.1
SE96	TTWA	288	0	136	88.5	70.2	84.0	83.8	96.1	6.2	3332	12749	28036	746138	71229	1	2	2.1	23	2.3
SE01	LAM	289	0	88	93.0	70.7	87.5	87.7	95.3	4.5	1253	14537	46490	1082322	129439	1	2	3.3	35	5.0
SE01	GEA	289	0	120	88.7	68.0	84.4	83.4	95.1	6.3	3368	14029	34092	840401	86321	1	2	2.4	23	2.6
SE01	TTWA	289	0	126	87.4	68.0	82.7	82.6	95.1	6.4	3368	14537	32469	840401	83016	1	2	2.3	23	2.4
SE01	'Euro'	289	0	166	86.6	66.8	80.9	80.4	95.1	7.0	1253	10174	24645	840401	73008	1	1	1.7	23	2.1
SE06	LAM	290	0	79	92.7	79.6	87.4	87.5	95.0	4.1	1248	17016	54187	1117352	141949	1	2	3.7	36	5.3
SE06	TTWA	290	0	118	86.8	68.6	81.8	82.3	94.4	6.1	3354	16756	36278	867188	89351	1	2	2.5	23	2.5
SE10	LAM	290	0	76	92.3	76.0	85.9	86.2	95.1	4.6	1245	16768	57799	1197405	155031	1	2	3.8	36	5.5
SE10	TTWA	290	0	102	87.5	67.5	82.8	82.4	93.5	6.2	3359	19286	43066	929463	103773	1	2	2.8	22	2.8
ES01	LAM	8030	78	1536	94.4	7.6	85.2	83.7	99.0	9.4	1	261.5	9592	2447627	83904	1	1	5.2	469	19.3
ES01	GEA	8030	78	583	90.1	66.7	85.2	84.4	98.9	6.9	3300	6259	25273	2244969	112508	1	8	13.8	157	
ES01	TTWA	8031	77	492	90.7	68.0	85.9	85.0	98.9	6.8	3332	8438.5	29947	2260167	122742	1	8	16.3	156	
ES01	'Euro'	8031	77	892	89.8	66.8	82.9	82.6	98.9	7.1	1002	3641	16518	2258670	91872	1	5	9.0	130	12.6
UK01	LAM	10474	84	681	72.2	18.7	59.0	58.9	96.0	14.6	464	14133	38926	3619455	151000	1	9	15.4	809	36.2
UK01	GEA	10558	0	265	78.6	66.7	73.6	75.2	96.5	6.8	3769	50346	100468	3214712	220046	2	27	39.8	692	52.0
UK01	TTWA	10558	0	218	81.4	66.7	76.4	77.7	96.5	7.4	3769	57819	122129	3376179	261627	2	33	48.4	727	62.6
UK01	'Euro'	10558	0	257	81.0	66.8	75.0	76.5	96.5	7.3	1267	43312	103596	3310191	242733	1	25	41.1	706	59.2

Notes
SD: standard deviation. *Assigned zones* refers to the number of LAU2 zones effectively assigned to a valid LMA (there is a discrepancy in the number of basic zones assigned in each case due to the characteristics of the different procedures). Global SC refers to global self-containment (the percentage of resident workers that occupy a job within the boundaries of the LMA where they reside over the total number of jobs in the country)

Table A.7 Characterisation of the results (ii): based on SIRE data for a selection of countries

Data	Method	Assigned zones		No. LMAs	Global SC(%)	LMA min. self-containment (%)					LMA employed population					Number of zones per LMA				
		Yes	No			Min.	Median	Mean	Max.	SD	Min.	Median	Mean	Max.	SD	Min	Median	Mean	Max	SD
AT(SIRE)	TTWA	2371	0	69	87.9	66.9	77.6	78.2	95.8	6.1	4078	22860	49224	998073	124175	2	22	34.4	252	38.3
AT(SIRE)	'Euro'	2371	0	80	88.1	66.8	76.5	77.6	94.5	6.2	1306	17709	42456	1008059	117617	1	19	29.6	259	37.6
BE(SIRE)	TTWA	589	0	28	85.0	67.9	76.4	77.4	90.9	6.2	6262	49604	93760	656537	140904	4	12	21.0	117	23.6
BE(SIRE)	'Euro'	589	0	30	84.0	67.9	76.0	76.4	90.9	6.1	6262	49604	87509	664508	135584	4	12	19.6	120	22.9
CH(SIRE)	TTWA	2896	0	76	86.3	67.6	79.7	80.5	97.3	7.0	3373	21076	41971	501292	69213	5	30	38.1	192	35.8
CH(SIRE)	'Euro'	2896	0	107	85.6	67.7	78.5	78.9	97.3	6.9	1013	12049	29811	502635	60948	1	16	27.1	165	32.6
ES(SIRE)	TTWA	8098	10	477	90.3	68.1	82.6	82.8	97.3	5.9	3313	9289	31179	2217852	121791	1	8	17.0	216	25.8
ES(SIRE)	'Euro'	8099	9	798	89.5	66.9	79.5	80.1	97.3	6.3	1001	4572	18637	2220037	95058	1	5	10.1	177	17.0
FR(SIRE)	TTWA	36555	10	519	92.7	74.4	87.4	87.7	99.0	4.7	5338	24315	57368	4837335	223499	3	54	70.4	706	59.4
FR(SIRE)	'Euro'	36555	10	729	91.7	74.0	85.4	85.8	98.6	5.1	1713	13889	40842	4498780	178831	1	36	50.1	623	48.9
IE(SIRE)	TTWA	3448	10	38	90.2	69.4	82.3	82.1	94.3	6.8	3899	16551	30243	392551	63222	18	78	90.7	441	73.2
IE(SIRE)	'Euro'	3448	10	55	88.7	67.7	79.5	78.9	93.0	6.7	1210	10225	20895	393719	53513	9	46	62.7	447	64.8
SE(SIRE)	TTWA	289	0	124	88.8	69.0	84.2	84.0	96.2	6.6	3466	13950	31332	788168	78687	1	2	2.3	23	2.5
SE(SIRE)	'Euro'	289	0	166	87.6	67.2	81.7	81.4	96.2	7.2	1231	9645	23404	788168	68544	1	1	1.7	23	2.1

Notes
SD: standard deviation. *Assigned zones* refers to the number of LAU2 zones effectively assigned to a valid LMA (there is a discrepancy in the number of basic zones assigned in each case due to the characteristics of the different procedures). Global SC refers to global self-containment (the percentage of resident workers that occupy a job within the boundaries of the LMA where they reside over the total number of jobs in the country).

Appendix 2 to chapter 3 - Recommendations on the establishment of an EU-wide harmonised grid of comparable LMAs

In this Final Research Report (Section 3.5), a hypothetical set of 'Euro' criteria for LMA definitions was devised; with this set then used in the analyses of available national commuting datasets. The logic that drove the choice of criteria was two-fold:

A there should be target and minimum values for both self-containment and size;
B the values should derive from the lower ones in existing national LMA definition criteria.

Principle A reflects the decision to adapt the UK method of analysis, because it has pioneered this use of minima and targets to allow a restricted trade-off between size and self-containment, and this approach has not been questioned (probably due to the established transferability of this method). In any case, significant further experimentation with methods was beyond the scope of the brief additional research reported in this Appendix.

Principle B is probably well-founded in its emphasis upon existing national LMA definition criteria, but is more questionable in its prioritising of the lower values among the sets of criteria considered. To be very specific, this logic meant the hypothesised 'Euro' criteria were set so low that the number of separate LMAs they produce was very high: low criteria result in high numbers of qualifying LMAs.

To be of greatest practical policy value, a set of European LMAs will probably need to be 'populated' with relevant data on conditions and trends in each area, so their needs/potential can be compared. Much of the relevant data – such as from the LFS – would only be availability at NUTS3 level and this means that areas significantly smaller than NUTS3 would not currently be very useful on this basis. Table A8 compares the numbers of NUTS3 areas with the numbers of LMAs from the hypothetical 'Euro' criteria analyses (nb. "Euro(SIRE)" analyses were based on the datasets provided by SIRE – which had problems of data suppression in some cases – whereas the "Euro(web)" analyses used unsuppressed data obtained directly from the relevant NSI). With the one exception of Belgium – where the NUTS3 areas are notably small and thus highly prone to split realistic labour market areas in so urbanised and integrated a country – the number of 'Euro' LMAs is substantially larger than the number of NUTS3 regions in all the countries, which vary in nature from sparsely populated Sweden to the intensively developed UK.

Table A8 Comparison of the number of NUTS3 Regions and hypothetical 'Euro' LMAs

	NUTS3	Euro(SIRE)	Euro(web)
Austria	35	80	
Belgium	44	30	
Czech Republic	14	107	
Spain	59	798	892
France	100	729	
Ireland	8	55	
Sweden	21	166	166
UK	133		257

The way to define significantly fewer separate LMAs is to raise the criteria that they all must satisfy. Principle A (above) can still be observed by retaining the use of target and minimum parameters, because this enables a trade-off between size and self-containment that allows peripheral areas like (groups of) islands to remain separate LMAs even when the size of their employed population is low, due to the very high level of the self-containment of their commuting patterns. At the same time, across the vast majority of a country like the UK – little of which is far from a substantial urban area – all plausible LMA definitions will have large populations so the key issue there will be the minimum self-containment level that has been set.

Table A9 provides some results from sample variations of parameters in the LMA definition criteria. The first two rows – covering the numbers of NUTS3 Regions and of the hypothetical 'Euro' LMAs – are provided as a 'benchmark' by replicating data presented above (Table A8). The first additional set of results is based on analyses with criteria in which all four parameters have raised values: employed populations must have a minimum of 5,000 (but the target is 50,000), while a commuting self-containment minimum was set at 75% (but with a target value of 85%). The effect of these changes is to reduce to less than half the number of LMAs from the hypothetical 'Euro' set in both Spain the UK (nb. these are the only countries for which additional analyses have been carried out). The number of LMAs on this basis in the UK is found to be lower than the number of NUTS3 Regions, but the number in Spain is still very much higher than this 'benchmark' value there.

In fact it is a realistic outcome that sets of LMAs defined on a consistent basis in the two countries have a higher number of separable areas in Spain than the UK. While the number of NUTS3 Regions in a country primarily reflects its population size, hence there being more in the UK than in Spain, the number of LMAs should also reflect its territorial size and the extent to which there are rather thinly populated areas remote from the main urban areas. There are several of this latter area type in the UK, primarily in northern Scotland, but there are more such areas in numerous parts of Spain.

Table A9 explores the sensitivity of the results described so far to the minimum size requirement which is the most important parameter in determining how many LMAs are defined in these remote thinly populated areas. In the fourth and fifth rows are results from raising this minimum first from 5,000 to 10,000 and then to 20,000 (Table A9). The effect on LMA numbers in the UK is remarkably slight as a result of it including relatively few remote areas, but as expected the impact is very strong in the case of Spain. There is no simple way of determining the most appropriate set of parameters, with the key concern being the use value of the boundaries produced for the purpose which called for those definitions to be created. Thus it is a potentially relevant concern that the higher minimum size parameter set here (20,000) prevents some major island groups in the UK from remaining LMAs in their own right: this would not be a result that would be seen as useful for many purposes.

For the final set of results reported here the minimum employed population reverts down to 10,000 but the self-containment minimum is raised to 85% (and the target to 95%). Table A9 shows that the effect in Spain is to produce fewer LMAs: many of the 207 LMAs defined with the same population parameters but the lower self-containment minimum and target must have had self-containment values close to those requirements because when the requirements are raised only 114 Spanish LMAs meet these criteria. A similar rate of decline in separable LMA numbers also occurs in the UK. Here again it should be emphasised that the relevant form of evaluation is the use value of the areas defined by any particular analysis. In effect the 56 UK LMAs defined on this basis are really more like "city regions" than most labour market areas used in national policies. More localised definitions tend to be used for policies where it is hoped to target responses at areas of acute need, but at the same time there is a risk that if the areas are drawn too tightly then the higher level of commuting across those boundaries – which results from more narrowly defined areas – could mean that the benefits of policy action targeted at those areas are more likely to be gained by people of adjacent areas who commute in to take the newly created job opportunities.

Table A9 Comparison of the number of NUTS3 Regions and LMAs defined by alternative criteria

	UK (web)	ES (web)
[NUTS3]	[133]	[59]
['Euro' : minSIZE=1,000 targetSIZE=5,000 minS-C=66..6 targetS-C=75.0]	[257]	[892]
minSIZE=5,000 targetSIZE=50,000 minS-C=75.0 targetS-C=85.0	121	277
minSIZE=10,000 targetSIZE=50,000 minS-C=75.0 targetS-C=85.0	112	207
minSIZE=20,000 targetSIZE=50,000 minS-C=75.0 targetS-C=85.0	109	147
minSIZE=10,000 targetSIZE=50,000 minS-C=85.0 targetS-C=95.0	56	114

A final statistical point relates to the potential distortion to policy analyses resulting from analysing GDP values for areas with relatively low self-containments. This risk is related to the deviation of the measure known as Job Ratio (viz: no. jobs / no. employed people) from the ideal of 1.0 (viz: an area where the number of jobs equals the number of employed people). By defining the LMAs so that their self-containment minimum is relatively high, this risk is reduced. However this is in fact a rather 'blunt tool' to deal with problems arising from an unbalanced Job Ratio because some of the LMAs which have self-containments that are extremely high (eg. more than 90%) still have some of the most unbalanced Job Ratios. It would be possible to have Job Ratio balance as an explicit parameter in LMA definitional criteria, but this is a non-trivial extension to research conducted to date.

4. Annex I - Interim Research Report

TABLE OF CONTENTS

EXECUTIVE SUMMARY

This is the report on progress at an approximate mid-way point in a study to consider the potential for identifying a common European way of defining the boundaries of Local Labour Market Areas (LMAs).

The study has so far carried out four main activities.

- Outline the state-of-the-art in applied social science research in the definition of LMAs
- Compiled a preliminary inventory of necessary information and of information sources
- Compared national methods of defining LMAs, based on a survey of all Member States (MSs)
- Considered the implications in terms of establishing best practice in the definition of LMAs

The key purpose of this Interim Research Report is to inform stake-holders in all MSs of progress in the study and also to provide an early sight of emerging best practice in the definition of LMAs. It is hoped that as a result of this information sharing, MS stake-holders will respond to the invitation for feed-back on the report, and perhaps engage with a later stage of the study (as outlined at the end of the report).

The first substantive research activity was a comprehensive desk-based overview of relevant literature on appropriate methods for defining the boundaries of policy-relevant LMAs. The diversity of academic work which had to be reviewed hinted at the lack of a very strong convergence on agreed best practice. There were also new approaches being developed, as part of the emerging field of spatial economics and its analysis of large datasets.

All the same, there was no evidence any serious contradiction of the principles of LMA definitions that will be policy-relevant, as set out in *"Study on employment zones" (EUROSTAT Document E/LOC/20)*.
As a result, the state of the art review was able to build on these principles, combining them with other concerns evident arising from the newer and/or more academic literature. This led to enumerating a final total of 14 issues against which methods of LMA definitions can be compared.

In addition to the review of LMAs in the applied social sciences, this report also considers the information provided by national statistics institutes (NSIs) in reply to a tailor-designed questionnaire on LMAs, which was aimed to compile homogeneous information on this issue. The research team is extremely grateful for the information that was provided primarily by the respondents at the NSIs and, in some cases, by the academics that co-operated with them in those tasks. The answers received show that despite an obvious disparity in the national approaches to the issue of delineating LMAs, a vast majority of the MSs rely on the use of data on commuting flows from a variety of statistical sources. The questionnaire investigated the availability of such data in all EU MSs, including those countries where LMAs have not been defined until now. As a preliminary conclusion, the decennial Census of Population and Housing seems to be the

preferable source of commuting datasets. The responses to the questionnaire confirmed for the majority of countries that after the 2011 wave it will be possible to gather the commuting data that would permit the definition of LMAs in the diverse parts of the EU. This dataset will be available at LAU2 level for most MSs, and this could allow the characterisation of the sets of LMAs built over them through aggregation of data, something that the questionnaire confirmed as common practice in those countries where LMAs are in use.

The last section of this report compares existing national LMA definitions against the 14 relevant issues established in the course of the first research activity. This examination embraced both official and academic approaches included in the NSIs' responses to the questionnaire, and aimed at selecting the methods that could be then explored in the empirical analyses that will be conducted in the last stage of the project. The choice was guided first by recognising that a small number of issues are of primary concern. It was also argued that there were some rather different types of method, with the proposal emerging that it would be valuable to test one method that was based on the initial selection of centres, and another one that explores the whole dataset with an 'open system' approach that avoids imposing such a structure. The other factor taken into account was that methods whose results were adopted by the respective NSIs can be seen as thereby having had a positive endorsement. The combination of these factors led to the selection of the **UK** and **SE** methods as the recommended candidates' for the empirical research as respectively a more 'open system' method and one heavily based on a preliminary step of identifying core [urban] areas. There is also a case for a more 'academic' stochastic procedure (which performs simultaneously local and global optimisation of a within-region interaction index) as an appropriate comparator to the **SE** and **UK** methods already mentioned.

This forward look towards the empirical research activity later in the study was completed by some thoughts of quantitative indicators for the evaluation of the results of different methods applied to data for several MSs. A remaining task is to select MSs whose datasets it will be most valuable to analyse, and some factors relevant to this selection are suggested.

The remaining stages of the study will move on from the four essentially retrospective tasks described here to look forward at [a] the possible value added by having consistent EU-wide LMA definitions of LMAs, and also [b] the potential for identifying a single harmonised definition method that can be seen to delimit suitable LMA boundaries in diverse parts of the EU.

INTRODUCTION

Overview of the project

In 2007, when the implementation of the NUTS Regulation was reviewed by Commission Regulation (EC) No 105/2007, National Statistical Institutes (NSIs) proposed an investigation of alternative classifications to the administrative levels below NUTS for the EU management of territory. As a result, Eurostat committed itself to exploring functional regions and the possible application of that concept to the entire EU. The first step was then to investigate with the help of the research community the potential value-added and feasibility of, and best practice for, a consistent EU-wide definition of labour market areas.

The report that follows is the intermediate output from research activities undertaken by, and on behalf of, Eurostat in this direction. It is a result of the work of Eurostat Unit E4, coordinated by Mr. Oliver Heiden, and the external contractor DevStat – Servicios de Consultoría Estadística with its associated researchers, Prof. Mike Coombes from Newcastle University, and Prof. José Manuel Casado and Dr. Lucas Martínez from the University of Alicante.

Labour Market Areas (LMAs) represent a territorial breakdown which is seen as a valuable alternative to local and regional administrative areas for some statistical purposes, and for the design, implementation and monitoring of labour market and other public policies. These functional territories have gained importance for policy-makers, mostly during the last decade, because they provide a territorial grid with boundaries that do not derive from historical events or geographical factors, but they reflect the organisation of social and economic relations in each specific area.

Currently there is no common European definition of LMAs, although in several EU countries they have been defined and are in use. The purpose of the "Study on comparable Labour Market Areas" is to explore the possibility of a consistent statistical classification of the whole EU territory, defined on a functional basis. To be specific, the objectives are to: (1) outline the state-of-art of applied sciences in the field of LMAs; (2) compare the LMA concepts recognised and implemented in each Member State; (3) draw conclusions on relevant best practice; (4) explore the added value of a common definition for the entire EU; and (5) identify possible ways and means of harmonising LMA definitions across the EU.
What makes this study different from the previous ones[42] is the intention to cover all the territory of EU 27 Member States (MSs) while also empirically testing a proposed method for the delineation of LMAs so as to formulate proposals for a possible EU wide harmonised grid of comparable LMAs. This empirical research is to reflect the assessment of LMAs' potential applications in the policy fields of the EC, but will also rely on the involvement of MSs in the collection of comparable information and the evaluation of intermediate results of the study.

[42] The topic was previously investigated by other international organisations and DGs of the European Commission (EC). Two decades ago, Eurostat and Newcastle University carried out a study of principles upon which to base definitions of LMAs to be used in a policy context. In 2001, OECD undertook a cross-national survey in order to examine the relevance of functional delineation of regions on the basis of travel-to-work. OECD's study offers a clear view of different existing concepts of LMA in different states, although it does not include all EU countries. More recently, DG REGIO expressed a strong interest in LMA statistics and Eurostat carried out a survey of the LMA definitions in Member States.

Purpose of the Interim Research Report

This interim report details the progress of the work undertaken for the possible definition of a common EU way of defining LMAs. It is intended to provide the main stake-holders, Eurostat and NSIs, with a comparable view of the situation of LMAs within the EU.

The intention to undertake this research was communicated to NSIs at the Working Group on Regional Statistics and Rural Development (4-5 October 2011), when NSIs agreed to provide information for the study.

The report considers LMAs from both theoretical and practical perspectives by reviewing the state-of-art in applied social sciences, but also comparing existing national LMA definitions. It is intended to provide both experience and non-experienced users with the overall result of the research undertaken so far. In so doing, it aims to inform NSIs about the possibilities of a consistent LMA geography and to seek their views.

Structure of the Interim Research Report

Apart from this introductory section, the report includes the following chapters:

- **Chapter 1:** Overview of the state-of-art of applied sciences in the field of LMA

  This chapter identifies the key features of best practice emerging from the review of academic work on methods for defining LMAs and ends by proposing a set of principles to consider for the evaluation of LMA definition methods and their results. The 14 principles drafted here derive from the definitional criteria proposed in the earlier study undertaken by Eurostat, Coombes (1992), extending their scope so as to cover in more detail the technicalities of the methods of delineation and to enable not only the comparison but also the evaluation of existing LMA definition methods.

- **Chapter 2:** Overview of the information collected

  This chapter gives an overview of the information collected through the 2012 questionnaire on Labour Market Areas sent out to EU Member States in January - February 2012.

- **Chapter 3:** Cross-national evaluation of existing concepts of LMA

  This chapter analyses the existing national methods of LMAs definitions as implemented in Member States, and provides a detailed comparison of the methods using the principles identified in Chapter 1. The cross-national evaluation of existing concepts of LMA ends by drawing some conclusions in terms of identifiable best practices in the definition of LMAs.

- **Annexes**

  The report is accompanied by the detailed description of the national methods analysed and compared in Chapter 3 (for Belgium, Czech Republic, Estonia, Greece, Finland, France, Italy, Portugal, Sweden, Slovenia, Slovak Republic, and United Kingdom) and by the standard format of the 2012 questionnaire on LMAs.

CHAPTER 1: REVIEW OF THE STATE-OF-ART OF APPLIED SCIENCES IN THE FIELD OF LMAS

This study was motivated by the fact that there is no common European way of defining Local Labour Market Areas (LMAs). In this Activity there has been a review, based on desk study, of the present state of the art in the applied social sciences in methods of defining LMAs. The concept of the labour market area relates to the 'law of one price' in that it is an area within which demand and supply for labour meet and fix a price for labour. Although a few labour market segments have national or even international markets, for most forms of labour national markets are fragmented into different LMAs. This spatiality of labour markets is part of the functional reality within modern economies, and hence LMAs are a specific aspect of the functional regions that are increasingly recognised across Europe.

The first section of this chapter outlines the scope of the desk study reported here. The two main sections of this chapter then identify in the applied sciences *key features of best practice* emerging from the review of academic work on methods for defining LMAs. Finally there is a section that looks forward to the next Activities of this study, ending with suggested issues to consider in evaluating LMA definition methods and their results.

1.1 Policy Context

It has become increasingly common in European and other developed countries for LMAs to be defined for policy delivery or analysis, or the publication of data by national statistical institutes (Cattan 2001). This trend is notable because there is an inherent tendency to use local or regional administrative areas for official purposes, so considerable benefits must result from using LMAs for them to be used instead. The fundamental reasons lie in the key policy focus on sub-national contrasts in economic geography, with the recognition that addressing these contrasts coherently calls for analyses that are comparing *areas which represent labour markets*, because this is one of the key concepts in economic geography (Combes et al 2005).

Administrative boundaries do not very frequently match the functional realities of LMAs due to being the product of tradition and/or topographic factors that have become less relevant (eg. rivers that once were natural barriers but, when bridged, become the focus for economic development). Even when functional realities had informed the drawing of an administrative boundary (Andersen 2002), this boundary may have remained unchanged for so long a period that it no longer relates to changed economic patterns. Although some administrative areas may approximate LMAs, this is not consistently true either within countries or, more obviously, between them (Forstall et al 2009). The importance of this inconsistency stems from the fact that *policy analysis requires all the LMAs to be defined in a comparable way* so that data for the areas can be used with a minimum of anxiety about how far the way the boundaries were drawn affects comparisons between areas (eg. when identifying areas in most need of policy support). In short, the choice of areas in analyses raises the familiar risk of 'comparing apples with pears' but the general issue is in fact unavoidable: in spatial analyses it is termed the modifiable areal unit problem (Openshaw

& Taylor 1981) and is a key concern when analysing economic data for administrative areas (Mitchell & Watts 2010) .

European integration has led to the need for functionally coherent LMAs that are comparable not only within but also *between* EU member countries. This challenge calls for comparable data for comparable 'building block' areas to be analysed, ideally delineating LMAs in a single process across all countries. One advance would be an empirical evaluation of LMA definition methods by applying them to data from different countries: if such analyses prove possible they can be attempted later in this study. This chapter is a first step in that direction, because it aims to establish some best practice guidance that can inform the search for appropriate methods to devise 'European standard' LMA boundary definitions.

1.2 Scientific context

Scientific research into LMA definitions has not led to consensus on best practice (Schubert *et al* 1987), in part because of the need for an appropriate method for evaluating different methods. This need will be addressed in the last section of the present chapter, but the focus here is first on teasing out lessons of best practice emerging from the academic LMA definition literature (nb. a later Activity of the study will examine official definitions of LMAs in the countries of Europe).

Delineating sets of LMA boundaries is a specific form of partition (if done 'top down') or of grouping ('bottom up'), and hence is a type of taxonomic analysis. In principle, LMAs could be defined 'top down' but nearly all methods in scientific literatures are *based on grouping building-block areas ('bottom up')* by analysis of relevant data. The fact that flows are key defining features of functional regions has resulted in most *definition methods analysing patterns of commuting* because they are flows that are an aspect of the labour market through which the strength of the links between building-block areas can be measured. There are some alternative data sources relevant to the labour market which can show spatial patterns of workplace and home locations but the coverage of such data is limited and liable to bias: for example the movement of people into employment – 'hirings' – only covers new employees and so has a bias towards young people among others.

This study is concerned with methods to identify LMAs covering the whole European space and this can be seen as a form of taxonomy. Some taxonomic analysis methods such as cluster analysis are quite familiar but are unsuited for LMA definition because they do not readily handle the matrices on flows between areas that are understood to characterise LMAs. In addition, they do not readily handle issues of area topology and the particular *need to avoid non-contiguous LMAs* (LMAs comprising two or more bounded areas that are separated by other LMAs) as illustrated by results of Hensen & Cörvers (2003). In geographic research, cluster analyses readily produce what are termed *formal regions* (eg. groupings of textile towns), whilst the contiguous groupings which include LMAs are termed *functional regions* (Spence & Taylor 1970).

Before reviewing the literature specifically on LMA definitions it is important to recognise that other scientific literature can be at least indirectly relevant here. There is increasing interest in spatial issues within economics, aided by the diffusion of geographic information system techniques which make large spatially-referenced datasets more readily interrogated. In the growing field of spatial economics the

'discovery' of computer-assisted geographic analyses has led to exploratory work which is in effect looking for LMAs in data 'attribute space' (eg. Marcon & Puech 2003; Duranton & Overman 2005). These new approaches can be compared with those used when taxonomic definitions were first being computerised and methods were derived from general principles in mathematics and related fields (Sokal & Sneath 1963). For example, there was early work using graph theory Nystuen & Dacey (1961) which is now echoed by recent exploratory adaptations of social network analysis (Newman & Girvan 2004; Green 2007).

Many past studies defined LMAs with the argument that commuting patterns can be 'proxies' of the other forms of spatial interactions which would be part of a more ideal definition of functional regions (Sohn 2005). A question for this study is whether this logic might work in reverse: if there is no up-to-date commuting dataset available then might LMAs be defined with other interaction data? The dataset which is the nearest equivalent to commuting and is quite often available to the necessary level of reliability for all small building-block areas is migration. A relatively high proportion of the early functional region definitions relied on migration data (eg. Hemmasi 1980) but it has been less used recently and some work in England suggests that migration patterns in many regions differ markedly from those of commuting (Hincks & Wong 2010) so are not a plausible basis for defining LMAs.

Still at the experimental stage are studies attempting to define what are, in effect, functional regions based on analyses of mobile phone traffic (eg. Candia et al 2008) or even surname incidence data (Longley et al 2011). The former approach is analysing a form of interaction that may – but may not – mirror commuting patterns, whereas the latter uses non-flow data to surmise a pattern of interaction (viz. longer-term migration). The latter approach is slightly more familiar in modelling strategies used when there is no interaction dataset but some data on transport infrastructure or services. Hugo (2001) uses road networks to measure the relative ease by which more rural areas can access urban centres, rather as Green (1950) analysed bus service information to identify urban centres and their hinterlands, an approach which may still be relevant in countries where public transport is still key to mobility. In the absence of even these datasets it may be possible to predict patterns of commuting from data on jobs and employed residents: for example, Glover & Openshaw (1995) offer a version of the well established gravity model, while a simpler option assuming no "wasteful commuting" (Small & Song, 1992) has been illustrated by Coombes (2004).

With computational power continually growing there is also interest in the transferability of techniques such as location-allocation analysis to region definition (eg. Lolonis & Armstrong 1993), along with the automatic zoning program of Openshaw & Rao (1995). In recent years the challenge of distilling the patterns from huge datasets is an increasingly active research frontier due to the ever growing volumes of information available to scientific enquiry. Indeed new methods of taxonomy may emerge in computer science, for example, which are transferable to the definition of LMAs. One example of a more general-purpose computational technique with possible application to LMA delineation is simulated annealing (Kirkpatrick et al 1984), which can be set in the broader field of evolutionary computation (Fogel 2006). Exploratory research (Flórez et al 2008) suggests that such methods can be adapted to delineate LMAs by formulating objective functions to, for example, maximise cohesiveness in terms of commuting flows while requiring that minimum levels of self-containment and size are respected.

The potential scientific advantages of these approaches – such as their inherent rigour, potential self-optimisation and replicability – tends to be off-set by their rather abstract nature and risk of being seen as a 'black box' when policy-makers prefer readily understood definitions. This means that *transparency* in the definition process is important, so that is can be seen why any particular area has been given the boundary the analysis selected (nb. one problem with some computationally-intensive methods is that they are not deterministic, which means that different 'runs' of the same method on the same data can produce different results).

1.3 Geographic foundations

Defining functional regions such as LMAs builds on earlier scientific analyses in economic geography, with fundamental concepts dating back over a century. Christaller (1933) elaborated the *central place* model that saw whole territories orientated around urban centres, with southern Germany used as the possible exemplar. In the UK a more empirical tradition saw Mackinder (1902) identify economically integrated urbanised regions that were later termed *city regions* by Geddes (1915). Later technical change and growing prosperity made people more mobile, allowing more distant places to be linked with cities in particular, thus creating wider *urban systems* (see for example Pahl 1965).

In the closely-spaced urban systems found in much of northern Europe especially there are numerous examples of previously distinctive local economies being characterised as emerging *polycentric regions* (eg Burger et al 2011). In fact van der Laan & Schalke (2001) argued that LMA definitions would more fully reflect the complexities of modern patterns if they were *over-lapping* in many cases, but there are very few such definitions. In fact over-lapping LMAs would be outside the concerns of this study whose policy focus requires a set of LMAs covering the whole European space but, at the same time, *each building-block area should be in one, and only one, of the defined LMAs*.

Goodman (1970) provided foundations for subsequent research on defining LMAs by recognising the value of commuting flows to LMA definitions and identifying as the two essential requirements of LMAs:

(1) their boundaries are crossed by few journeys to work (ie. they are relatively self-contained)

(2) a relatively high level of intra-market movement results from the LMAs being as integrated as possible.

These two key factors have remained key foundations for LMA definitions subsequently. All the same, the lengthening and diffusion[43] of commuting flows makes meeting both requirements more difficult. Goodman (1970) warned against "the danger of seeking external perfection at the expense of losing the essentially *local* character of the market" – which implies priority of the second requirement over the first

[43] *The strong tendency for personal mobility to increase was enabled by several factors, most notably*

- *rising average earnings, making greater commuting costs more affordable*
- *increased car use, enabling more diffused and distended commuting patterns, and also*
- *decreasing real cost of car use.*

These processes have reinforced each other over recent decades, increasing the proportion of people who commute longer distances and thus increasing the integration of previously separate local areas.

– but the lengthening of average commuting trips in the 40 years since then has left some areas as not locally integrated at all (see for example, van Nuffel 2007).

Turning to existing LMA definitional practice, there is in fact no off-the-shelf measure of integration, whereas the self-containment measure is well established. As a result, best practice methods are characterised by the *analysis requiring a minimum level of commuting self-containment* for each LMA. Although there may well also be an aspiration to maximise integration, few if any definition methods set a specific level of integration as a constraint.

Smart (1974) set out to "produce a systematic definition of areas by which the main relationships between homes and workplaces could be indicated" and his pioneering efforts were followed by a novel computerised approach that brought scientific rigour to the definitions of TTWAs (Travel-to-Work Areas), the official UK LMAs (Coombes et al 1986). As often, this innovation relied on newly accessible relevant data being available for sufficiently small areas.

A key argument for analysing commuting data to define LMA boundaries is that the LMAs are where "jobs are sought and job decisions are made… in terms of ability to commute" (Wilcock & Sobel 1958). A focus on commuting may seem old-fashioned in a world where new means of communication allow many jobs to be done without the need for people to travel, but in fact *teleworking* displaced relatively few physical work trips so that "spatial patterns of commuting are more complex today than in previous decades, but no less important" (Arbuckle 1998).

In fact a secondary but crucial advantage of the usual dependence on commuting data in defining LMAs is that the 'friction of distance' that restricts people's patterns of movement causes most of the strongest interactions to be between nearby areas. This means that contiguous groupings of areas are *inherently* the most likely to be produced from commuting data: for example, less than 1% of building-block areas were allocated non-contiguously in an application of the unconstrained TTWA method to New Zealand (Papps & Newell 2002). By contrast, any methods that use an explicit contiguity constraint at every step of the analysis will see their results shaped by "irregular base areas" (Spence & Taylor 1970).

Contiguity constraints were once valued for vastly speeding up analyses – by greatly restricting the options considered as building-block areas are being grouped – but now vast numbers of permutations can be evaluated quickly. Thus little is gained from if a contiguity constraint is imposed throughout the analysis but, if one is imposed, it will reduce the options available to the analyses and this will inevitably risk creating some sub-optimal boundary definitions (Roca & Moix 2005). The clear implication is that *methods whose processes are not limited by contiguity constraints* are preferable in most cases.

Another distinction which can be drawn is between those methods that use one 'rule' from start to finish, and those which apply different rules at different stages of the process. The former type proceeds until all the LMAs satisfy a criterion that decides when the procedure stops (e.g. the population size of the smallest remaining region is large enough). Several such methods were developed in the 1970s at the time when computerised matrices first became available (eg. Slater & Winchester 1978), while some may still be sometimes used, as for example INTRAMAX which was created by Masser & Scheurwater (1980)

and used recently by Mitchell & Watts (2010). The new 'black box' methods mentioned earlier mostly involve multiple applications of a single rule.

Methods with numerous rules, by contrast, often identify LMAs in accordance with a geographical model (eg. a first rule uses a size criterion to find urban centres, then a later step analyses commuting patterns to group non-central areas as part of the LMAs around centres). As this implies, a single or multiple step method is usually adopted as a direct consequence of the *general* approach being taken towards the definition of LMAs, so past research does not readily indicate whether a single or a multiple rule method is the better practice in general. That said, if all 'other things are equal' then single step options gain from their *advantage of having fewer parameters to justify* and to then potentially have to put through sensitivity testing (cf. Papps & Newell 2002).

The single rule approach often repeatedly applies its single criterion within a hierarchical process and this can create sub-optimal results at larger scales. This is because the area groupings made at the early stages of the analysis inherently restrict the options available[44] at later stages. The solution is for methods to 'escape' from being hierarchical, although no 'universal' procedures exist to achieve this. The method to define TTWAs has since the 1980s used a procedure that aimed at self-optimisation (Casado-Díaz 2000): this procedure became the basis of the entire process in the latest version of this LMA definition method (Coombes 2010).

An important distinction that van der Laan & Schalke (2001) made among LMA definition methods was between *deductive* and *inductive* based approaches. In practice, deductive methods begin with the identification of urban centres around which the LMAs are constructed; by contrast inductive methods have no such pre-conceived structure to their definitions. Most deductive methods rely on the familiar notion of a single urban centre which attracts commuters from surrounding areas (eg. Arbuckle 1998), and this is readily understood by non-experts. However this model pre-dates the growth of polycentric urban systems, and decentralising employment to less urban centres. In fact many versions of the inductive approach exclude more remote areas from their results (eg. Cheshire & Gornostaeva 2002), an approach running counter to the all-inclusive objective of this study. The familiar deductive approach of metropolitan definitions has been seen as an 'urban bias' by those interested in rural development (Killian & Tolbert 1993).

The inductive approach, not surprisingly, has the converse strengths and weaknesses; it copes flexibly with diverse commuting patterns in different times and places, but this same flexibility means it does not have a simple form which is easily recognisable. The process is one of gradual integration, so that each step in the process only makes a small contribution to the final result. It reflects a less rigid conception of LMAs as clusters of commuter flows, within a wider 'space of flows' (cf. Castells 1989). In effect then, the choice between deductive and inductive

[44] *For example, a grouping of European countries might be expected to link Luxembourg with Belgium and then with the Netherlands in its early stages; in a hierarchical procedure, these early groupings would then prevent the later stages creating what may be the more optimal broader groupings in which, for example, the Netherlands could be linked with other northern countries which speak Germanic languages whilst Belgium and Luxembourg were grouped with France and more southern countries.*

approaches depends upon the evaluation criteria adopted. The search here is for LMA definition *methods to produce adequate results in many different conditions* across the European space and this leads to an emphasis on flexibility and so the inductive approach may be favoured. By contrast, the intuitive appeal of the deductive approach has the advantage that its urban centred model offers the policy relevant benefit of greater simplicity and transparency of method.

1.4 Moving forward

There is an emerging need for establishing systematic criteria allowing comparison among methods. One proposal is for a straightforward sensitivity analysis of methods by making small changes to some parameters in their definition procedures to assess the scale of the impact on the boundaries produced (see for example the tests conducted by Papps & Newell 2002). Such analyses are examples of what can be termed intrinsic testing: the test is of how well a method meets its intended objectives. One form of intrinsic test in the LMA context was provided by Feng (2009), with a membership function measuring how connected each building-block area is to others in the LMAs to which it was allocated. Casado-Díaz et al (2010) provide an example in which a more complete set of related indicators is tested.

Extrinsic testing strategies, on the other hand, assess how far LMAs are appropriate for spatial economic analyses because of their properties on variables *other than* those used in their definitions. Examples include Barkley et al (1995) who computed three indicators of spatial association to test the core-hinterland dependence in per capita income change, Cörvers et al (2009) who analyse variations in income levels, housing prices, and employment and unemployment rates, and others ranging from Baumann et al (1983) through to Maza & Villaverde (2011) who explore regional economic growth and compare the results from using functional rather than administrative areas.

It is noteworthy that these tests for internal homogeneity or homogeneous behaviour of the areas are not so relevant here because LMAs can be internally heterogeneous but still have the required intensity of flows between their constituent areas. If homogeneity in local attributes was more important to LMAs than cohesiveness in terms of flows then homogeneity would have been the more appropriate objective to use in the definitions. It should also be noted that these analyses are all very problem-dependent and likely to have results which are at least partly determined by the scale of the areas analysed.

These evaluation methods are approaches that can be used to compare different sets of LMA definitions covering the same territory (nb. it would be necessary to ensure that the sets of definitions are in key respects – for example, the number of LMAs that they divided that territory into – very similar, otherwise those more basic factors could dominate the results). The aspiration here is for a consistent set of LMA definitions covering the whole European territory but at present LMAs are either defined for one county only *or* they do cover many countries but only for selected parts (mainly around large cities).

It is a possible task for a later stage of the present study to create the situation needed to allow rigorous methods of evaluation to be applied. The intermediate step would be to apply selected methods to data for one or more countries: an earlier example of this strategy is reported in Eurostat & Coombes (1992).

This strategy requires the selection of LMA definition methods whose results will be evaluated, and this presupposes that the range of plausible methods has been evaluated to make that selection. The first step then has to be decide on the basis for this evaluation of methods; the best practice conclusions derived from the review of academic work here provides the necessary foundations (Table 1).

Table 1.1: List of the best practice conclusions drawn from the academic literature review

[A] *areas which represent labour markets*
[B] *policy analysis requires all the LMAs to be defined in a comparable way*
[C] *based on grouping building-block areas ('bottom up')*
[D] *definition methods analysing patterns of commuting*
[E] *need to avoid non-contiguous LMAs*
[F] *policy preference for readily understood definition processes*
[G] *analysis requiring a minimum level of commuting self-containment*
[H] *each building-block area should be in one, and only one, of the defined LMAs*
[I] *methods whose processes are not limited by contiguity constraints*
[J] *advantage of having fewer parameters to justify*
[K] *methods to produce adequate results in many different conditions*

To avoid the risk of 'reinventing the wheel' here it is very valuable to set the items in Table 1 against the "Principles" for LMA definition that were outlined in Eurostat and Coombes (1992), as shown in Table 2. There are two over-riding Principles that establish what makes a set of LMA boundaries fit for purpose: that they should be the product of a rigorous method of definition and reflect labour market geography. These two objectives are echoed in Table 1 by its first two findings on best practice (but note that **[A]** relates to Principle 2 while finding **[B]** relates to Principle 1). In addition finding **[D]** clearly also links with Principle 1 because its emphasis on commuting was derived from the concept of local labour markets.

Table 1.2: Principles to guide local labour market area definitions

Principle	Practice
OBJECTIVES	
1. Purpose	to be statistically-defined areas appropriate for policy
2. Relevance	each area to be an identifiable labour market
CONSTRAINTS	
3. Partition	every building block to be allocated to 1 and only 1 area
4. Contiguity	each LMA to be a single contiguous territory
CRITERIA *in descending priority*	
5. Autonomy	self-containment of flows to be maximised
6. Homogeneity	LMAs' size range to be minimised (eg. within fixed limits)
7. Coherence	boundaries to be reasonably recognisable
8. Conformity	alignment with administrative boundaries is preferable

Turning to the next Principles which establish what should be constraints in definition methods there are again parallels in Table 1: Principle 3 is directly echoed by finding **[H]** and similarly Principle 4 has its equivalent in finding **[E]**. The remaining Principles have a lower priority than the first four, which are the primary objectives and constraints. In practice, there is room for much debate over the extent to which any one of these should over-ride others or, as is perhaps more likely, which of them can be 'traded-off' against others. All the same, Table 2 does make clear that Principle 5 has a degree of precedence and it is echoed in Table 1 in the form of finding **[G]**. The following three Principles are less clearly reflected in the findings from the academic literature, although the policy-driven Principle 8 that administrative areas should be followed if there are no strong countervailing arguments is supported by a recent study (Cörvers et al 2009). The last of the Principles is echoed by finding **[K]**.

Looking back at Table 1 then reveals that there are four findings which are not obvious echoes of the Principles previously identified. These are all more specifically about the technicalities of methods, although finding **[F]** is in fact a warning that in the policy field there is preference for fewer technicalities to maximise the transparency of the way the results were produced from the data. There is little problem in this regard with finding **[C]** because building LMAs 'bottom up' is more intuitive, and can involve more simple methods, than a 'top down' method (of which there are few). Equally sympathetic with the policy aim for transparency is finding **[J]** because methods with few separate parameters will often be simpler. This leaves finding **[I]** as perhaps the 'odd one out' because it will call for greater computational intensity in order to achieve more optimal results.

The conclusion of this chapter involves deriving from this final section the issues by which to compare, and then also evaluate, potentially relevant existing LMA definition methods. Table 3 lists these issues (nb. it will also be necessary to consider issues such as the average size of LMAs the method defined).

Table 1.3: List of issues by which to compare/evaluate existing LMA definition methods

- does the method produce adequate results in different conditions (eg. metropolitan/peripheral)?
- is every building-block area in one, and only one, of the defined LMAs?
- are all the areas explicitly defined as labour markets?
- how consistently have the areas been defined so as to be comparable?
- are there any non-contiguous LMAs?
- how closely aligned are the LMAs to administrative areas (and was this an explicit constraint)?
- how readily understandable/transparent is the definition process?
- was the process a grouping of building-block areas or a subdivision of the whole territory?
- did the process analyse patterns of commuting and/or any other flow data?
- did the analysis explicitly require a minimum level of commuting self-containment?
- did the analysis explicitly require a minimum of population size or of any other dimension?
- did the definition processes have a contiguity constraint throughout?

- how many parameters are there which need to be justified?
- were parameter values set on a deductive basis or arrived at inductively (so readily modifiable)?

References

Andersen A (2002) "Are commuting areas relevant for the delimitation of administrative regions in Denmark?" Regional Studies **36** 833-844

Arbuckle D (1998) "Alternative approaches to defining metropolitan and nonmetropolitan areas" Federal Register **63** 70526-70561

Barkley D, Henry M, Bao S & Brooks K (1995) "How functional are economic areas? Tests for intra-regional spatial association using spatial data analysis" Papers in Regional Science **74** 297-316

Baumann J, Fischer M & Schubert U (1983) "A multiregional labour supply model for Austria: the effects of different regionalisations in multiregional labour market modelling" Papers in Regional Science **52** 53-83

Burger M, de Goei B, van der Laan L & Huisman F (2011) "Heterogeneous development of metropolitan spatial structure: evidence from commuting patterns in English and Welsh city region, 1981-2001" Cities 28, 160-170

Candia J, Gonzalez M, Wang P, Schoenharl T, Madey G. & Barabási A-L (2008) "Uncovering individual and collective human dynamics from mobile phone records" Journal of Physics A: Mathematical and Theoretical 41, 1-11

Casado-Díaz JM (2000) "Local labour market areas in Spain: a case study" Regional Studies **34** 843-856

Casado-Díaz JM, Martínez-Bernabeu L & Flórez F (2010) "Los mercados locales de trabajo españoles: una aplicación del nuevo procedimiento británico" JM Albertos & JM Feria (eds) La ciudad metropolitana en España: procesos urbanos en los inicios del siglo XXI Thomson-Civitas, Madrid

Castells M (1989) The informational city: information technology, economic restructuring and the urban-regional process Blackwell, Oxford

Cattan N (2001) Functional regions: a summary of definitions and usage in OECD countries DT/TDPC/TI(2001)6 OECD, Paris

Cheshire P & Gornostaeva G (2002) "Cities and regions: comparable measures require comparable territories" Cahiers de L'Aurif **135** 13-21

Christaller W (1933) Central places in central Germany [English translation] Prentice-Hall, Eaglewood Cliffs

Combes P-P, Duranton G & Overman H (2005) "Agglomeration and the adjustment of the spatial economy" Papers in Regional Science **84** 311-349

Coombes M (2004) "Multiple dimensions of settlement systems: coping with complexity" in A Champion and G Hugo (ed) New Forms of Urbanization: Beyond the Urban-Rural Dichotomy Ashgate, Aldershot

Coombes M (2010) "Defining labour market areas by analysing commuting data: innovative methods in the 2007 review of Travel-to-Work Areas" pp227-241 in J Stillwell, O Duke-Williams & A Dennett (eds.) Technologies for migration and population analysis: spatial interaction data applications IGI Global, Hershey

Coombes M, Green A & 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** 943-953

Cörvers F, Hensen M & Bongaerts D (2009) "Delimitation and coherence of functional and administrative regions" Regional Studies **34** 19-31

Duranton G & Overman H (2005) "Testing for localization using micro-geographic data" Review of Economic Studies **72** 1077-1106

EUROSTAT and Coombes M (1992) Study on employment zones (E/LOC/20) Office for Official Publications of the European Communities, Luxembourg

Forstall R, Greene R & Pick J (2009) "Which are the largest? Why lists of major urban areas vary so greatly" Tijdshcrift voor Economische en Sociale Geografie **100** 277-297

Feng Z (2009) "The fuzziness of Travel-to-Work Areas" Regional Studies **43** 707-720

Flórez-Revuelta F, Casado-Díaz J, Martínez-Bernabeu-Bernabeu L (2008) "An evolutionary approach to the delineation of functional areas based on travel-to-work flows" International Journal of Automation and Computing **5** 10-21

Fogel D (2006) Evolutionary computation: toward a new philosophy of machine intelligence (3rd edition) Wiley, Hoboken

Geddes P (1915) Cities in evolution: an introduction to the town planning movement and to the study of civics Williams and Norgate, London

Glover T and Openshaw S (1995) "The AGW spatial interaction modelling workstation" Available at http://wwwgeogleedsacuk/people/sopenshaw/w3agwhtml

Goodman J (1970) "The definition and analysis of local labour markets: some empirical problems" British Journal of Industrial Relations **8** 176-96

Green F (1950) "Urban hinterlands in England and Wales: an analysis of bus services" Geographical Journal **116** 64-89

Green N (2007) "Functional polycentricity: a formal definition in terms of social network analysis" Urban Studies **44**. 2077–2103

Hemmasi M (1980) "The identification of functional regions based on lifetime migration data: a case study of Iran" Economic Geography **56** 223-233

Hensen M and Cörvers, F (2003) "The regionalisation of labour markets by modelling commuting behaviour" Paper presented at the 43th European Regional Science Association Congress http://www.roa.unimaas.nl/seminars/pdf2005/M.Hensen.pdf

Hincks S & Wong C (2010) "The spatial interaction of housing and labour markets: commuting flow analysis of North West England" Urban Studies **47** 620-649

Hugo G (2001) "Addressing social and community planning issues with spatial information" Australian Geographer **32** 269-294

Killian M & Tolbert C (1993) "Mapping social and economic space: the delineation of local labor markets in the United States" in J Singlemann & F Deseran (eds) Inequality in local labor markets Westview Press, Boulder

Kirkpatrick S, Gelatt C & Vecchi M (1983) "Optimization by simulated annealing" Science **220-4598** 671-680

Laan L van der & Schalke R (2001) "Reality versus policy: the delineation and testing of local labour market and spatial policy areas" European Planning Studies **9** 201-221

Lolonis P & Armstrong M (1993) "Location-allocation models as decision aids in delineating administrative regions" Computers, Environment and Urban Systems **17** 153-174

Longley P, Cheshire J, Mateos P (2011) "Creating a regional geography of Britain through the spatial analysis of surnames" Geoforum **42** 506–516

Mackinder H (1902) Britain and the British seas Appleton, New York

Marcon E & Puech F (2003) "Evaluating the geographic concentration of industries using distance-based measures" Journal of Economic Geography **3** 409-428

Masser I & Scheurwater J (1980) "Functional regionalisation of spatial interaction data: an evaluation of some suggested strategies" Environment and Planning A **12** 1357-1382

Maza A & Villaverde J (2011) "EU regional convergence and policy: does the concept of region matter?" Journal of Policy Modeling **33** 889-900

Mitchell W & Watts M (2010) "Identifying functional regions in Australia using hierarchical aggregation techniques" Geographical Research **48** 24-41

Newman M & Girvan M (2004) "Finding and evaluating community structure in networks" Physical Review E **69** 026113:1-15

Nuffel N van (2007) "Determination of the number of significant flows in origin-destination specific analyses: the case of commuting in Flanders" Regional Studies **41** 509-524

Nystuen J & Dacey M (1961) "A graph theory interpretation of nodal regions" Papers of the Regional Science Association **7** 29-42

Openshaw S & Rao L (1995) "Algorithms for re-engineering 1991 Census geography" Environment & Planning A **27** 425-446

Openshaw S & Taylor P (1981) "The modifiable areal unit problem" in N Wrigley, R J Bennett *(eds)* Quantitative Geography Routledge, London

Pahl R (1965) Urbs in rure: the metropolitan fringe in Hertfordshire (Geographical Papers 2) LSE, London

Papps K & Newell J (2002) "Identifying functional labour market areas in New Zealand: a reconnaissance study using travel-to-work data" Discussion Paper 443 Institute for the Study of Labor, Bonn

Roca J & Moix M (2005) "The interaction value: its scope and limits as an instrument for delimiting urban systems" Regional Studies **39** 357-373

Schubert U, Gerking S, Isserman A & Taylor C (1987) "Regional labour market modelling: a state of the art review" in Fisher M and Nijkamp P (eds) Regional labour markets: analytical contributions and cross-national comparisons Elsevier, Amsterdam

Slater P & Winchester H (1978) "Clustering and scaling of transaction flow tables: a French inter-Departmental migration example" IEEE Transactions on Systems, Man and Cybernetics SMC **8** 635-640

Small K & Song S (1992) "'Wasteful' commuting: a resolution" Journal of Political Economy **100** 888-898

Smart M (1974) "Labour market areas: uses and definition" Progress in Planning **2** 238-353

Sohn J (2005) "Are commuting patterns a good indicator of urban spatial structure?" Journal of Transport Geography **13** 306-317

Sokal R & Sneath P (1963) Principles of numerical taxonomy Freeman, San Francisco

Spence N & Taylor P (1970) "Quantitative methods in regional taxonomy" Progress in Geography **2** 1-64

Wilcock R & Sobel I (1958) Small city job markets: the labor market behavior of firms and workers Institute of Labor and Industrial Relations, University of Illinois, Urbana

CHAPTER 2: OVERVIEW OF THE INFORMATION COLLECTED

In addition to the review of LMAs in the applied social sciences, this report also considers the information provided by NSIs in reply to a tailor-made questionnaire on LMAs (from now on called "the 2012 questionnaire" or Q-2012). The questionnaire was sent to NSIs in January – February 2012 and was aimed to compile the fundamental evidence needed for the comparison of LMA concepts recognised and implemented in each Member State. The questionnaire was customised for each of the EU MS: it included pre-filled questions, where the answers were retrieved from NSIs replies to a previous related questionnaire sent by Eurostat in 2007 or from other recent work carried out by Eurostat. Apart from the pre-filled answers some of the national questionnaires included methodological notes on the delineation method used in each specific country.

The standard template of the questionnaire included questions that were aimed at retrieving from all EU MS homogenous information covering four informational dimensions, specifically:

a. Characterisation of the existing delineation methods
b. NSIs' position towards an LMA geography and their expectations from this
c. Inputs for the delineation of LMAs
d. Possible characterisation of LMAs

According to the answers to the 2012 questionnaire, nine countries have official LMAs in use (Belgium, Germany, Estonia, Finland, France, Italy, Netherlands, Sweden and the United Kingdom), and seven NSIs reported on LMAs definitions which were not used officially (Cyprus, Czech Republic, Denmark, Greece, Portugal, Slovenia and Slovak Republic). While Chapter 3 of this report investigates the specific methodology (ie. characteristics of the existing delineation methods), this chapter, instead, gives an overview of the remaining information provided by MSs in reply to the 2012 questionnaire.
The overview information that follows covers all EU27 MS except for Malta. All the information was retrieved through the 2012 questionnaire, except in the case of Belgium, where the information is a result of the previous questionnaire sent in 2007 by Eurostat.

2.1 NSIs' position towards an LMA geography and their expectations from this

One of the goals of the dissemination of this Interim Research Report is to demonstrate the usefulness of an exercise of delineating LMAs, that is why it was considered relevant to gather information that could guide the process from the first steps. In this direction, the 2012 questionnaire included questions related to NSIs position towards an LMA geography and their expectations from this. This information unfolds in a mixture of qualitative and quantitative aspects.
An overview of the quantitative information that could reflect NSIs' position towards an LMA geography is presented in the table below. This type of information is mostly related to the issue of scale of LMAs, which is measured in terms of n° of LMAs, minimum/maximum/average area and/or population. The table below summarises this information for those countries where, according to the questionnaire, LMAs have

been defined either officially or as a result of academic exercises (in the latter case the table only includes the cases for which the information available was sufficiently detailed).

Table 2.1: Scale of existing LMAs

Country	Total population 2011	Area 2009 (km^2)	n° NUTS 3	n° LAU 1	n° LAU 2	National denomination	Produced by	Year when first produced	In official use	Basic building blocks	Current n° of LMA	Min. LMA population	Max. LMA population	Average LMA population	Min. LMA area (km2)	Max. LMA area (km2)	Average LMA area (km2)	Population density
BE*	10.951.266	30.666	44	-	589	bassins d'emploi	Stat. Office Wallonie	2000	yes	LAU 2	47	NA	NA	233.006	NA	NA	652	357
CZ	10.532.770	78.870	14	77	6.251	-	Charles University Prague	1961	no	LAU 2	184	10.304	1.442.991	57.243	101	2.719	429	134
DE*	81.751.602	357.114	412	1.351	12.066	Arbeitsmarktregionen	Statistisches Bundesamt	1972	yes	NUTS 3	258	60.560	3.442.675	316.867	75	6.999	1.384	229
EE*	1.340.194	43.452	5	15	226	Tööjõuareaal	University of Tartu	2001	no	LAU 2	15	10.168	522.147	89.346	1.023	4.807	2.897	31
EL	11.309.885	120.167	51	1.035	6.130	topikes agores ergasias	KEPE	2001	no	LAU 1	667	-	3.887.000	16.956	-	3.607	180	94
FR*	65.048.412	548.763	100	3.785	36.680	zones d'emploi	INSEE	1984	yes	LAU 2	321	10.044	5.851.493	202.643	119	40.945	1.710	119
IT*	60.626.442	301.392	110	-	8.094	Sistemi locali del lavoro	ISTAT	1981	yes	LAU 2	686	3.318	3.808.731	88.377	10	3.666	439	201
NL*	16.655.799	37.357	40	-	418	Regionale Platforms voor de Arbeidsmarkt	CBS	1991	yes	LAU 2	34	85.406	1.826.562	489.876	148	6.561	1.099	446
PT	10.636.979	88.840	30	308	4.260	-	INE	2007	no but planned	LAU 1	18	44.996	2.987.044	590.943	3.393	6.385	4.936	120
SK	5.435.273	49.013	8	79	2.928	zóny zamestnanosti	INFOSTAT	2002	no	LAU 2	51	10.392	737.484	106.574	292	3.581	961	111
FI*	5.375.276	335.765	20	70	336	Työssäkäyntialueet	Statistics Finland	1990	yes	LAU 2	58	4.150	1.414.693	92.677	NK	NK	NK	NK
SE*	9.415.570	449.159	21	-	290	Lokala arbetsmarknader	Statistics Sweden	1991	yes	LAU 2	76	2.878	2.435.363	123.889	883	27.410	5.910	21
UK*	62.435.709	244.436	139	380	10.310	Travel-To-Work-Areas	ONS & Newcastle Univ.	1998	yes	LSOA	243	7.907	8.952.972	256.937	53	5.174	1.006	255

***official LMAs**

Note : In Finland the municipalities with weak commuting figures were not allocated to LMAs.

Source: 2012 questionnaire on LMAs and Eurostat Chronos Database (own calculations)

Qualitative information on NSIs positions towards an LMA geography and their expectations from this (such as the purpose and the reasons which stood for the creation of LMAs, the main concerns of the exercise of defining a new statistical classification based on LMAs, and other issues about the statistical characteristics that should be met by LMAs if an EU-wide grid comparable areas would be proposed) was also collected. Although not summarised here, this information will be analysed and considered for the following activities of the project (ie. Elaboration of recommendations on the establishment of an EU-wide harmonised grid of comparable LMA).

2.2 Inputs for the delineation of LMAs and characterisation of LMAs

Despite an obvious disparity in the empirical and theoretical approaches to the issue of delineating LMAs, a vast majority of the MS rely on the use of data on commuting flows. The following table gives an overview of the sources of datasets used for delineating existing LMAs.

Table 2.2: Sources of datasets of existing LMAs

Country	Sources of datasets
BE*	Population Census 1991
CZ	Population Census 2001
DE*	Federal Employment Agency 1998 - 2011
EE*	Population Census 2001
EL	Population Census 2001
FR*	Population Census 1982,1984, 1999, 2006-2008
IT*	Population Census 1991, 2001
NL*	Population Census 2001, StatLine since 2006
PT	Population Census 1991, 2001
SI	SRDAP 2000-2010, Population Census 2002
SK	Population Census 2001
FI*	Register-based population statistics 1996-2011 discon.
SE*	Employment register 1993-2010
UK*	Population Census 1991,2001

***official LMAs**

The table above leads to the conclusion that a variety of statistical sources for information on commuting flows exist, although the most traditional and still the most widely-used are the decennial Census of Population and Housing (for Belgium, Czech Republic, Estonia, Greece, France, Italy, Portugal, Slovak Republic, United Kingdom).
In addition to identifying the sources of the commuting datasets used for delineating existing LMAs, through the 2012 questionnaire, the existence of such data in all EU MS, was also

investigated in order to accomplish the purpose of this exercise aimed at defining an EU-wide grid of comparable LMA. As a preliminary conclusion, the decennial Census of Population and Housing seems to be the preferable source of commuting datasets, especially after the 2011 Census, being this the first one legally regulated for all EU27 MS.

The "2011 European Census Programme" establishes the technical and legal framework required for census output harmonisation, leaving the Member States free to decide how to conduct the census, and to select the data sources, the methodology and technology applicable in each national context. The main requirements established at EU level refer to census topics, reference period, breakdowns, metadata, quality reporting and data transmission[45].
The breakdowns for geographical area in the 2011 round of censuses, as established by Commission Regulation (EC) N° 1201/2009, identify geographical areas from a high level of detail (LAU 2) to the national level.

Table 2.3: Breakdowns for geographical area in Population and Housing Census 2011

Geographical area	GEO.N.	GEO.L.	GEO.M.	GEO.H.
0. Total (in the territory of the Member State)	0.	0.	0.	0.
x. All NUTS 1 regions in the Member State		x.	x.	x.
x.x. All NUTS 2 regions in the Member State		x.x.	x.x.	x.x.
x.x.x. All NUTS 3 regions in the Member State			x.x.x.	x.x.x.
x.x.x.x. All LAU 2 regions in the Member State				x.x.x.x.

Source: Regulation (EC) N° 1201/2009
Note: The codes 'x.', 'x.x.' and 'x.x.x.' depend on the NUTS classification, the code 'x.x.x.x.' on the LAU classification, valid for the Member States on 1 January 2011. The annotation 'N' identifies the breakdown that refers to the national level.

Regulation (EC) N° 763/2008 of the European Parliament and of the Council establishes the topics to be covered by the round of 2011 censuses by geographical levels: for NUTS 3 and LAU 2, and for NUTS 1 and NUTS 2. While the "place of residence" is ordinarily included among population census topics at the highest level of detail (LAU 2), the Census Regulation includes the "location of place of work" as one of the obligatory topics to be covered, but at a lower level of detail (NUTS 2)[46]. However, this only applies to the data to be submitted to Eurostat, and not actually to the collection of data.

[45] EU legislation on the 2011 Population and Housing Censuses: Explanatory Notes, EUROSTAT, 2011 edition http://epp.eurostat.ec.europa.eu/portal/page/portal/product_details/publication?p_product_code=KS-RA-11-006
[46] Commission Regulation (EC) N° 1201/2009 on population and housing censuses as regards the technical specifications of the topics and of their breakdowns

The replies to the 2012 questionnaire showed that a majority of Member States collected through the 2011 Population Census the variable "location of place of work" at LAU 2 level. **Table 2.4** gives an overview of the availability of commuting datasets after 2011 Census and their source, as resulting from MSs' replies.

Table 2.4: Availability of commuting datasets in EU27 MS after 2011 Population Census

Country	Available	Source	Collection method	Level
BE	NK	NK	NK	NK
BG	YES	Population Census 2011	directly collected, exhaustive	LAU 1
CZ	YES	Population Census 2011	directly collected, exhaustive	LAU 2
DK	YES	Register-Based Labour Force statistics	register based	LAU 1
DE	YES	Federal Employment Agency	register based	LAU 2
EE	YES	Population Census 2011	e-census, directly collected, exhaustive	LAU 2
IE	YES	Population Census 2011	directly collected, exhaustive	LAU 2
EL	NK	NK	NK	NK
ES	YES	Population Census 2011	directly collected, 12 % sample	LAU 2
FR	YES	Population Census 2011	register based	LAU 2
IT	YES	Population Census 2011	directly collected, exhaustive	LAU 2
CY	YES	Population Census 2011	directly collected, exhaustive	LAU 2
LV	YES	Population Census 2011	e-census, directly collected, exhaustive	LAU 2
LT*	YES	Population Census 2011	e-census and directly collected, exhaustive	LAU 2
LU	YES	Population Census 2011	directly collected, e-census, exhaustive	LAU 2
HU	YES	Population Census 2011	e-census, partially register based, directly collected, exhaustive	LAU 2
MT	NK	NK	NK	NK
NL	YES	StatLine Database	register based combined with sample survey	LAU 2
AT	YES	Population Census 2011 & Register based	Register-based full enumeration	LAU 2
PL**	YES	Population Census 2011	register based, e-census and directly collected, exhaustive	LAU 1
PT	YES	Population Census 2011	e-census and directly collected, exhaustive	LAU 1
RO*	YES	Population Census 2011	e-census and directly collected, exhaustive	LAU 2
SI	YES	Population Census 2011 & SRDAP	directly collected, exhaustive & register-based	LAU 2
SK	YES	Population Census 2011	directly collected, exhaystive	LAU 2
FI	YES	Register-Based Population Statistics	register based	LAU 2
SE	YES	SCB Sweden database	register based labour statistics	LAU 2
UK	YES	Population Census 2011	directly collected, exhaustive	LSOA

Source: 2012 questionnaires on LMAs & Population Census 2011 official webpages

NK – not known

*according to the questions included in the 2011 Population Census questionnaire

** Works on the possibility of presenting data on commuting flows at LAU 2 in Poland are ongoing.

Finally, in what regards the availability of statistical variables that characterise LMAs, in the majority of the countries where LMAs have been defined, specific data for LMAs are not

collected. The information sources for the characterisation of LMAs derive from the statistics available at the basic building block level (LAU 1, LAU 2 or other level) and are constructed through aggregates. However, some exceptions still exist: for Germany, where four regional indicators (ie. unemployment rate, averaged over four years, annual infrastructure indicator, gross annual wages per employee, and employment forecast) are collected for the Joint action for improving regional economic structures, and for United Kingdom, where some specific statistics (unemployment claimant counts, job centre vacancies) are published.

CHAPTER 3: CROSS-NATIONAL EVALUATION OF EXISTING CONCEPTS OF LMAs

This chapter is considering the potential for identifying a common European way of defining Local Labour Market Areas (LMAs). An initial research reviewed the state of the art in the applied social sciences in methods of defining LMAs. Subsequent to that, information was sought from each member state (MS) on existing national definitions of LMAs: the research team is extremely grateful for the information that was provided, primarily by the national statistical institutes (NSIs). This report includes in its Annex extracts from the supplied information, converted into a standard format so that the different national methods can be directly compared. The next section of this chapter summarises the principles against which the national methods of LMA definitions will be compared. The followed section is the core of this report, which provides these comparisons of methods in detail. The chapter ends by drawing some conclusions in terms of identifiable best practices in the definition of LMAs.

3.1 Principles in the definition of LMAs

The concept of the labour market area is of an area within which demand and supply for labour meet and fix a price for labour. LMAs are a specific form of the functional regions which are increasingly recognised across Europe as the appropriate units for economic research and policy analyses which need comparable data for comparable areas. In most countries, the availability of data on labour market flows between areas is limited to commuting between home and workplace. The result is that in practice most official and academic approaches to defining LMAs focus on patterns in matrices of commuting flows between areas. Their objective is typically to draw boundaries which delimit LMAs that are highly self-contained and cohesive in terms of travel-to-work.

Looking back at Chapter 1 provides the basis for the assessment of the national methods for defining LMAs. That chapter drew on its review of best practice in the applied social sciences to set out the issues by which to compare and/or evaluate existing LMA definition methods in its Table 1.3:

1. does the method produce adequate results in different conditions (eg. metropolitan/peripheral)?
2. is every building-block area in one, and only one, of the defined LMAs?
3. are all the areas explicitly defined as labour markets?
4. how consistently have the areas been defined so as to be comparable?
5. are there any non-contiguous LMAs?
6. how closely aligned are the LMAs to administrative areas (and was this an explicit constraint)?
7. how readily understandable/transparent is the definition process?

8. was the process a grouping of building-block areas or a subdivision of the whole territory?
9. did the process analyse patterns of commuting and/or any other data?
10. did the analysis explicitly require a minimum level of commuting self-containment?
11. did the analysis explicitly require a minimum of population size or of any other dimension?
12. did the definition processes have a contiguity constraint throughout?
13. how many parameters are there which need to be justified?
14. were parameter values set on a deductive basis or arrived at inductively (so readily modifiable)?

Not all the issues listed above are equally relevant. The primary issues are those that should determine which methods to take forward to empirical analyses of the commuting datasets of different countries. These principal objectives, constraints and criteria are those that make it most likely that a method will produce appropriate LMA definitions for policy analysis across the whole of the EU:

- issue 2 (regions are exhaustive & non-overlapping)
- issue 4 (consistency: same method and parameters formally applied in a replicable process)
- issue 5 (regions are contiguous)
- issue 10 (minimum self-containment required).

By the same token, the other issues relate to more secondary concerns which a method may diverge from without contravening the basic objectives or constraints for this research:

- issue 1 (non-core-based)
- issue 3 (explicitly defined labour markets)
- issue 6 (respects administrative boundaries)
- issue 7 (clarity simplicity)
- issue 8 (aggregative)
- issue 9 (not based on other data apart from commuting)
- issue 11 (minimum size)
- issue 12 (not contiguity constrained)
- issue 13 (few parameters)
- issue 14 (analytically adjusted parameters)

The evidence base for assessing each national method in relation to each of the issues is rather varied. The principal source has been the set of responses to Q-2012 together with the further responses to the follow-up queries where necessary. In several of these responses there were links to further information and this has been drawn upon too as appropriate, along with previous knowledge of the researchers which is based in part on academic research. Despite this range of evidence, there are numerous cases of issues for which a robust assessment cannot be provided of numerous national methods: these cases are indicated by the essentially provisional

assessment of "?" (Tables 3.1 and 3.2). In such cases it may well be appropriate to update the information in Tables 3.1 and 3.2 if further information becomes available to the researchers.

The following notes summarise how the evidence has been used to reach a rather general assessment of each national method in relation to each issue. These assessments are reported in the form of the set of symbols used in Tables 1 and 2 where the 14 issues are considered in turn. The essential arguments supporting these assessments were outlined in the report of Activity 1: for example, the basic objective of this research – to seek a common European method of defining LMAs – was interpreted as meaning that any such method needs to produce adequate results in a range of different geographical conditions (issue 1 here). In this report, these rather general guide-lines have been made more specific so that the different methods can be more clearly contrasted. Thus for example this need to define adequate LMAs in all the very contrasting geographical circumstances across Europe is seen to 'count against' a method which has pre-suppose that all LMAs will take a particular geographical form. For the definition methods under review here, this leads to a negative assessment of method based on an initial step to identify urban cores, because this cannot be expected to be so appropriate in peripheral less urbanised regions.

1. Does the method produce adequate results in different conditions (eg. metropolitan/peripheral)?

All the methods analysed for this report have been either defined and/or utilised by NSIs or developed and tested in the academic sphere so their appropriateness to the territory they were designed for is not in question here. At the same time, to establish a cross-national definition of LMAs it is preferable that the method does not rely heavily on the definitions of foci or urban centres, otherwise it will be less likely that adequate results in a wide range of different geographical circumstances.

☑ = not core based
★ = cores are used but then reconsidered during the process
? = unclear evidence
☒ = core based

2. Is every building-block area in one, and only one, of the defined LMAs?

This issue too would be ideally resolved by a full set of maps, with the alternative source here being the information on methods provided in response to Q-2012 which is summarised in the Annex.

☑ = map evidence or the answer to Q-2012 supports the answer Yes
☒ = No

3. Are all the areas explicitly defined as labour markets?

For almost all countries the evidence is limited to the answer to Q-2012: in some cases the reported LMAs were defined more generally as functional or city/metropolitan regions and in these cases some scepticism about their validity as labour market areas is warranted.

☑ = Yes they are explicitly defined as LMAs
★ = whilst not explicitly defined as LMAs, they probably are reasonable approximations of LMAs

4. How consistently have the areas been defined so as to be comparable?

For this issue too the evidence on almost all countries is limited to the answer to Q-2012 and it should be acknowledged that there may be some cases where the definitions were reported to be entirely consistent when in practice some adjustments – which cannot be identified from the available evidence – had taken place to improve the results. This issue has been assessed in three47 separate items.

4a. Consistency in applying the same method to all the territory
4b. Consistency in applying the same parameters to all the territory
4c. Consistency in absence of 'manual' (non-formalised) adjustments

Each of the items has been assessed according to the following categories:
☑ = evidence of a high level of consistency
★ = probably consistent (but the evidence is not definitive)
? = there are certain grounds for doubt
☒ = inconsistency clear (eg. between regions)

5. Are there any non-contiguous LMAs?

The evidence on this issue would ideally have been a full set of maps of the LMAs produced by the method concerned but where they were unavailable the responses to Q-2012 were drawn upon.

☑ = map evidence or the answer to Q-2012 supports the answer No
☒ = Yes presumably (answer to Q-2012 says so)[48]

6. How closely aligned are the LMAs to administrative areas (and was this an explicit constraint)?

This issue too would be ideally resolved by a full set of maps, with the alternative source here being the information on methods provided in response to Q-2012 that is summarised in the Annex. As almost all countries use as 'building block areas' small areas from the administrative

[47] This division of the issue allows a more complete assessment of the characteristics of each method (nb. the same strategy has been applied to issues 7, 10 and 11).
[48] Even if contiguity was not a constraint upon the definition process (see issue 12), it is typically then enforced through a final manual adjustment: as a result it may be that if the answer to issue 5 is Yes a similar adjustment to the results could be applied subsequently.

hierarchy (most often LAU2), so the issue relates to alignment of the LMAs with higher order[49] administrative areas.

☑ = map evidence suggests close alignment
★ = non-trivial adjustments to increase alignment
? = uncertain
☒ = evidence suggests the answer is No

7. How readily understandable/transparent is the definition process?

The evidence on this issue is in the answers to Q-2012 and the assessment then relies on past experiences of policy-makers (mostly in Britain) viewing a certain level of complexity to be problematic. This issue is subdivided into two items, with the latter – "transparency" – giving a negative assessment to methods that rely upon informal judgements, because these mean their results may not be replicable.

7a. Clarity
☑ = clear and readily understandable
★ = understandable enough, based on the use of the results by the NSI
? = there is some lack of clarity
☒ = probably too complicated

7b. Transparency
☑ = complete
★ = non-formalised final step which makes minor changes based on local knowledge
? = uncertain
☒ = frequent reliance upon informal judgements

8. Was the process a grouping of building-block areas or a subdivision of the whole territory?

This issue recognises that although most of the familiar methods start with individual 'building blocks' (eg. LAU2 areas) which are then grouped by the algorithm, there have been some methods which begin with the whole territory and then sub-divide it into LMAs. Some research evidence suggests that the latter type is more likely to produce sub-optimal results, while their infrequent use also indicates that they are less intuitively understandable. The evidence for assessing methods on this issue is in the responses to Q-2012.

☑ = Yes the process was an aggregative grouping
☒ = No the process proceeded by dividing up the whole territory

9. Did the process analyse patterns of commuting and/or other related data?

This issue too was directly addressed in Q-2012 (and the relevant aspects of the responses have been summarised in the Annex). Given the focus here on finding a method potentially applicable

[49] It is worth noting that a close fit to administrative areas could be related to the nature of those areas in the particular country analysed; as a result, the most positive assessment of methods on this issue is here reserved for any method that includes increased alignment with administrative areas as a formal secondary criterion within the definition procedure.

to the whole EU there is a preference for a lower number of inputs because there are few relevant datasets which are available in every country.

☑ = no data other than commuting flows
★ = one additional dataset (eg. distance/time)
? = uncertain
☒ = more than one additional dataset

10. Did the analysis explicitly require a minimum level of commuting self-containment?

This issue also was directly addressed in Q-2012 along with the subsequent requests for clarifications, from where all the information necessary for the assessment has been extracted. The issue has been subdivided to allow a twofold assessment.

10a. Is reaching a certain degree of self-containment for all defined LMAs a criterion in the method?

☑ = Yes
★ = it is initially but it is not rigidly adhered to in the end
? = uncertain
☒ = No

10b Does functional dependence (eg. self-containment) act as the primary basis for grouping areas?

☑ = Yes
★ = no but it will be tend to be supported by the grouping based on the largest commuting flow
? = uncertain
☒ = No

11. Did the analysis explicitly include a minimum of population size or of any other dimension?

This is another issue directly addressed in Q-2012 but the responses have suggested that this issue should be subdivided into two items. A minimum population size can often be valued by policy-makers (for reasons such as a reduction on the volatility of statistical trends for the areas); at the same time, there are some policy-makers who value a restriction on the physical extent of areas.

11a. Minimum population

☑ = Yes a minimum population size criterion is used
★ = Yes, but it is not respected in final steps (eg. when assigning residual areas)
? = unproven
☒ = No there is no such explicit criterion

11b. Maximum area

☑ = Yes a maximum area is imposed, either directly or indirectly (eg. by time/distance commuted)

★ = Yes it is considered but it is not an absolute restriction (eg. when assigning residual areas)

? = uncertain
☒ = No there is no such explicit criterion

12. Did the definition processes have a contiguity constraint throughout?

This is another issue directly addressed in Q-2012: the responses are evaluated in line with the findings in the report on Activity 1 that a contiguity constraint leads to sub-optimal results in some cases, notably when the methods are very simple and consist on only a few deterministic steps.

☑ = No the process is not so constrained
★ = it is only imposed as a final step
? = uncertain
☒ = Yes the groupings are contiguity constrained during the whole process

13. How many parameters are there which need to be justified?

The evidence on this issue is in the answers to Q-2012 but in some cases this may have been less fully completed than in others so the assessment here may not be as accurate as for some other issues.

☑ = only one or two parameters were set
★ = three to five parameters are necessary
? = uncertain
☒ = more than five parameters

14. Were parameter values set on a deductive basis or arrived at inductively (so readily modifiable)?

The evidence on this issue in answers to Q-2012 requires a degree of interpretation, because for many countries the research has not had access to a description of the reasoning behind the set parameters, so the comparative assessments here may not be as accurate as for some other issues.

☑ = parameters based on analytical processes
★ = parameters very largely based on analytical process
? = uncertain
☒ = arbitrarily fixed (need sensitivity testing and local knowledge)

3.2 Cross-national evaluation of national methods of defining LMAs

According to the responses to Q-2012 the following classification of countries in relation to the definition of LMAs can be made.

A. Countries that have official LMAs in use: BE[50] DE EE FI FR IT NL SE UK.
A1. Countries that have an official analytical LMA delineation methodology: BE DE EE FI FR IT SE UK.
A2. Countries that have no analytical LMA delineation methodology: NL.

B. Countries whose response to Q-2012 reported[51] on LMAs definitions not used officially: CY[52] CZ DK[53] EL PT SI SK.
B1. Countries that are developing/considering an LMA delineation method, which is not official yet: PT.
B2. Countries where LMA definition research was done but is not officially used: CY CZ EL DK SI SK.

C. Countries reporting no officially or academically defined LMAs: AT BG ES HU IE[54] LT LU LV PL RO.
C1. Countries reporting that there is no official definition of LMAs, and not reporting any academic study: AT BG HU IE LT LU LV RO.
C2. Countries reporting that they are considering a LMA delineation method, but is not official yet: PL.

D. Countries that have never answered the questionnaires (1): MT.

For evaluation purposes the analytical methodologies (official or academic) reported in responses to the Q-2012 are assessed in the report in two Tables.

Table 3.1 covers both (A1) official methods in use: BE DE EE FI FR IT SE UK
and (B1) academic methods under official consideration: PT
Table 3.2 covers academic methods not currently under official consideration (B2): CZ EL SI SK

[50] The information about **BE** is based on the response to Q-2007 since the answer to Q-2012 was not received.
[51] It is important to note here that this classification is exclusively based on the information provided by the NSI (in fact in several other countries, like **ES**, some academic exercises have been conducted that were not reported by the respondent to Q-2012).
[52] The academic study that was mentioned by **CY** in their answer to Q-2012 refers to the same method described in **EL** applied over the **CY** dataset, therefore the assessment conducted here refer to both cases.
[53] None of the methods described in the answer to Q-2012 have been accepted by the NSI. Of them the one that would fit better this study consists in applying the official method **SE**; therefore no specific description is included here for **DK**.
[54] **IE** answered that NUTS 3 level regions are used for all purposes for labour market analysis, but it is likely that this also the case in the countries listed in the rest of the categories B and C.

It is not appropriate here to describe in detail the assessment, on each issue here, of all these methods. Tables 3.1 and 3.2 now summarise these assessment, and set out the basic grounds for selecting some methods to evaluate empirically, by using them to analyse commuting data in several different countries. Selecting the most suitable methods involves examining the evidence in Tables 3.1 and 3.2 so the approach taken here is to describe that evidence by proceeding through the 14 issues in a sequence which aims to progressively identify the methods that are less suitable[55] than others. To express this another way: some methods may be ideally suited to the specific territories they were designed for, but a method that can produce adequate results across the whole EU method needs to avoid being highly unsuitable for any country, rather than ideal for just one or even a few countries. One factor relevant for the selection of methods to evaluate further is that the method produced areas used by that NSI, because this implies a practical value rather than an 'optimality' that may be purely academic. The implication here is that there will be a preference here for the eight methods shown in Table 3.1 over those in Table 3.2.

The results in Table 3.1 indicate that at the outset it is possible to set aside issues 2, 3 and 8 since they do not strongly discriminate between the eight methods because they all have ? in these columns. The same is true of issue 14, although in this case it is because of the lack of any positive assessments. The remaining issues of primary concern are then 5 along with 4 and 10 (with the latter two issues now sub-divided into three and two separable issues respectively). The method with a negative assessment on issue 5 (contiguous final results) is that of **FI** but in fact it is possible for an 'extra' stage to be added to the method to resolve any problems of non-contiguity: this is the approach used by several methods (eg. **IT** and **UK**, leading to their ? assessments for issue 12). It is also true that while some methods have no non-contiguous areas in the results on their 'home' countries the same outcome may not occur if the method was applied in other countries (especially those with smaller LAU2 areas to analyse).

[55] It must of course be recognised that all the methods reviewed here could be reasonable options: they have been found to be appropriate sets of LMAs for the countries where they have been used by the NSI concerned and/or by academics.

Table 3.1: Summary evaluation of official national methods against the issues identified

MS: / issue:	BE[56]	DE[57]	EE	FI	FR	IT	SE	UK
1 (non-core-based)	☑	?	☒	☒	☑	★	☒	☑
2 (exhaustive & non-overlapping)	☑	☑	☑	☑	☑	☑	☑	☑
3 (explicitly defined labour markets)	☑	☑	☑	☑	☑	☑	☑	☑
4a (consistency – method)	☑	☑	☑	☑	★	☑	☑	☑
4b (consistency – parameters)	☑	☑	☑	☑	☒	☑	☑	☑
4c (consistency – analytical)	☒	☒	☑	☑	★	★	☑	★
5 (contiguous LMAs)	☑	?	☑	☒	☑	☑	☑	☑
6 (aligned to adm. boundaries)	☒	★	☒	☒	★	☒	☒	☒
7a (clarity/simplicity)	★	?	☑	☑	★	★	☑	☑
7b (transparency/reproducible)	☒	★	☑	☑	☑	★	☑	★
8 (aggregative)	☑	☑	☑	☑	☑	☑	☑	☑
9 (based on additional datasets)	☑	★	★	☑	☑	☑	☑	☑
10a (minimum self-containment)	☒	☑	☒	☒	★	☑	☒	☑
10b (self-containment guides the process)	☑	?	☒	☒	☑	☑	★	☑
11a (minimum population)	☒	☑	☒	☒	?	☑	☒	☑
11b (maximum area/distance/time)	☒	☑	☒	☒	☑	☒	☒	☒
12 (not contiguity constrained)	☑	☑	☒	☑	☒	☑	☑	☑
13 (few parameters)	☒	?	☑	☑	★	☒	☑	★
14 (analytically set parameters)	☒	?	☒	☒	☒	☒	☒	☒

Table 3.1 makes it clear that there are some uncertainties about the precise operation of the **DE** method, although other aspects, such as the information needs and the criteria all valid LMAs must fulfil are clear. One of these is the fact that the procedure includes a maximum commuting time within any valid LMA. The remaining uncertainties and the lack of readily usable information regarding this last point seriously hamper the possibility of considering this method for the empirical analyses later in this research study.

There is only one **?** marked in the **FR** column but issue 4b also has a negative assessment because there is use of local knowledge, and this could not be replicated across the whole of the EU.

On a similar basis the **BE** method can be discarded because of its ☒assessment on issue 4c (as well as the fact that the evidence here is limited to the Q-2007 response and so involves some uncertainty).

Of the remaining five methods, there are two (**IT** and **UK**) that are more 'open system' methods in that they are not heavily based on a preliminary identification of core areas. In fact the **IT**

[56] See Note 9.

[57] See the annex, where the DE method is reviewed.

method is basically the method applied in the UK several decades ago which has since been refined as a result of findings from applying it to several other countries around the world (Casado-Díaz et al 2010). As a result of this experience it is justifiable to select the **UK** method in preference to that of **IT** for the analyses of data from other countries in the EU.
Table 3.1 then has three remaining methods (**EE**, **FI**, and **SE**) and these would b potentially interesting contrasts to that of the **UK** because they all begin by identifying employment centres or foci. All the three methods lack an explicit minimum self-containment (giving a negative evaluation in issue 10a), in the **SE** method there is a clustering analysis that is less restricted to foci, producing a more positive assessment on the primary issue 10b. This suggests that the **SE** method is the most appropriate comparator to the **UK** method for the cross-national analysis in the next stage of this research study.

The case for choosing **SE** method as one to test empirically is also supported by the fact that it is, together with **UK** method, the one that has been applied more often beyond its own national boundaries. Examples of other applications of the **SE** method were cited in the responses to Q-1012 from fellow Scandinavian countries **DK**, **FI** and **NO** as well as **SI** (Drobne et al 2009).

Some of the responses to Q-2012 included reference to official methods of which the NSI was aware but whose resulting LMAs had not been adopted for any official purposes. Table 3.2 has evaluations of these methods on the same basis as that applied to the methods which produce official national sets of LMAs. As noted earlier, the lack of any adoption of the results of the method by the NSI concerned makes these methods less suitable candidates for the analyses of data on several countries later in this study.

Table 3.2: Summary evaluation of other national methods against the issues identified

MS: / issue:	PT	CZ	EL[58]	SI[59]	SK
1 (non-core-based)	☑	☒	☑	☑	☒
2 (exhaustive & non-overlapping)	☑	☑	☑	☑	☑
3 (explicitly defined labour markets)	☑	☑	☑	☑	☑
4a (consistency – method)	☑	☑	☑	☑	☑
4b (consistency – parameters)	?	☑	☑	☑	☑
4c (consistency – analytical)	?	☒	☑	☑	?
5 (contiguous LMAs)	☑	☑	☒	?	☑
6 (aligned to administrative boundaries)	☒	☒	☒	☒	☒
7a (clarity/simplicity)	★	★	☑	☑	?
7b (transparency/reproducible)	☑	★	☑	☑	?
8 (aggregative)	☑	☑	☑	☑	☑
9 (based on additional datasets)	☑	☑	☑	☑	☑
10a (minimum self-containment)	★	☒	☒	☒	☑
10b (SC guides the process)	☑	★	☑	☑	★
11a (minimum population)	☒	☑	☒	☒	☒
11b (maximum area/distance/time)	★	☒	☒	☒	☒
12 (non-contiguity-constrained)	☒	☑	☑	☑	☒
13 (few parameters)	☑	★	☑	☑	☒
14 (analytically set parameters)	☒	☒	☒	☒	☒

CZ and **SK** fall in the core-based category and can therefore be compared with the selected **SE** method. The latter has a better assessment for simplicity and clarity. Doubts about the **SK** method description (noted in the Annex) result in a better assessment for the **SE** method on primary issue 4c. On the other primary issues **SK** does have a better assessment than **SE** on issue 10a but this is probably due to the inherent size of LAU2 areas in **SE** which made a minimum size less necessary: the effect of adding this constraint to the **SE** method could be a valuable additional experiment to test in the empirical analyses.

Table 3.2 also has three more 'open system' methods in the academic procedures cited by **PT**, **EL** and **SI**. The larger number of uncertainties about the **PT** method may be due it being under consideration by the NSI at the present time; until such uncertainties are resolved, it cannot be considered a strong candidate for the empirical analyses later in this study. Both the other two methods (**EL** and **SI**) have at least one negative assessment in relation to the issues of primary concern here (2, 4, 5 and 10) and this then adds weight to the negative factor – in common with all the methods in Table 2 – that their results have not been adopted by their respective NSIs.

[58] See Note 11. This same procedure, originally proposed by Prodromídís (2010) has been tested in **CY** (and as so was mentioned in the response to Q-2012).

[59] Of the are six papers cited, the one evaluated here is Konjar et al (2010) because it is the most distinct and is an interpretation of commuting zones in the US

3.3 Towards an empirical test of 'best practice' in defining LMAs

While accepting that to some extent the assessments detailed above are provisional – because of the incompleteness of the available evidence – it is necessary to sketch the conclusions here on this basis. The immediate requirement is to propose the nature of the remaining research activity in this study.

The core activity will be the analysis of data for several countries by alternative methods to define LMAs. This basic outline raises the question of how the alternative methods will be evaluated. In the short term this can be simply outlined by identifying the need for a combination of simple and more sophisticated quantitative indicators such as:

- measures of the self-containment of the resulting LMAs (eg. how many are under a given threshold; how much variation across the different LMAs)
- measures of the cohesion of the resulting LMAs (nb. there is less consensus on how the integration of the constituent areas of LMAs are compared)
- more statistical indicators (eg. proportion of LAU2 areas not grouped 'optimally' in some sense).

Another element of the design of the final stage of the research is the choice of countries whose data will be analysed by selected LMA definition methods. There remains some uncertainty about which countries have all the necessary data available, so the recommendations at this stage can only be very general in nature. There is a strong case for choosing case studies which between them provide a good sample of the variation in territorial terms found across the EU: for example from heavily urbanised parts of the continent to more rural and perhaps peripheral circumstances (which might include island areas). On a more technical note, it may well be useful to include countries with a variety of LAU2 area sizes, because this factor can be influential in the likely effectiveness of methods. Another consideration may be to prioritise countries without existing official definitions, to maximise the new findings from the study.

The previous section of this chapter provided the basis for answering the next question in designing the next stage of this research study: which methods to apply to data for several countries. The choice was guided first by recognising that a small number of issues are of primary concern. It was also argued that there were some rather different types of method, with the proposal emerging that it would be valuable to test one method that was based on the initial selection of centres, and another one that explores the whole dataset with an 'open system' approach that avoids imposing such a structure. The other factor taken into account was that methods whose results were adopted by the respective NSIs can be seen as thereby having had a positive endorsement. The combination of these factors led to the selection of the **UK** and **SE** methods as the recommended candidates' for the empirical research.

The remaining question is whether there are alternative methods which were not cited in any of the responses to Q-2012 which should be evaluated in the remaining stage of this study. At the very least, such alternatives may provide a 'benchmark' against which to assess the established

national methods. From this more detached viewpoint, it can be suggested that all the methods cited in Tables 3.1 and 3.2 have at least one key feature in common: they are deterministic. Thus the case for assessing a very different method from the academic research sphere rests on the value of identifying a suitable method which does not share this characteristic.

Deterministic procedures are those that systematically produce the same output for given input data, together with a fixed set of parameter values. Contrasting with deterministic methods are stochastic procedures which introduce random variables to allow sub-optimal choices to be made in the short term, as a way of reaching the global optimum solution by the end of the procedure. (It is recognised that most sets of results will fall short of true 'optimality' but it is expected that adopting this strategy will produce results which are closer to optimality than those of a deterministic method.) Based on substantial and very recent research in this area (Watts, 2009) and related fields (Fortunato, 2010), it is proposed that the candidate stochastic method for evaluation in the final stage of this study is a grouping evolutionary algorithm (Flórez-Revuelta et al 1998, Martínez-Bernabeu et al 2012). Such a method conducts a stochastic search procedure to perform simultaneously local and global optimisation of a within-region interaction index (based on that used in the **UK** method in fact). It is therefore proposed that this method provides an appropriate comparator to the **SE** and **UK** methods that were identified above to represent centre-based and 'open system' approaches respectively.

References

Casado-Díaz, J.M.; Martínez-Bernabeu, L. and Flórez-Revuelta, F. (2010). "Los mercados locales de trabajo españoles. Una aplicación del nuevo procedimiento británico" ["Spanish local labour markets. An application of the new British procedure"], in La ciudad metropolitana en España: procesos urbanos en los inicios del siglo XXI (Albertos, J.M. and Feria, J.M., ed.). Madrid. Thomson-Civitas, 275-313.

Drobne, S., Konjar, M., and Lisec, A. (2009). "Delimitation of Functional Regions using Labour Market Approach". Proceedings of the 10th International Symposium on Operational Research SOR '09 in Slovenia, Nova Gorica, September 23-25, 2009.

Flórez-Revuelta, F., Casado-Díaz, J.M. and Martínez-Bernabeu, L. (2008). "An evolutionary approach to the delineation of functional areas based on travel-to-work flows". *International Journal of Automation and Computing*, vol. 5 (1), pp. 10–21.

Fortunato, S. (2010). "Community detection in graphs". Physics Reports, vol. 486, n. 3-5, pp. 75-174, Elsevier.

Martínez-Bernabeu, L., Flórez-Revuelta, F. and Casado-Díaz, J.M. (2012). "Grouping genetic operators for the delineation of functional areas based on spatial interaction". *Expert Systems with Applications*, vol. 39 (8), pp. 6754–6766.

Prodromídís, P. (2010). "Identifying spatial labor markets in Greece from the 2001 travel-to-work patterns". South-Eastern Europe Journal of Economics 1, pp. 111-128.

Watts, M. (2009). "Rules versus Hierarchy: An Application of Fuzzy Set Theory to the Assessment of Spatial Grouping Techniques" M. Kolehmainen et al. (Eds.): ICANNGA 2009, LNCS 5495, Springer-Verlag Berlin Heidelberg, pp. 517–526.

ANNEX 1: COUNTRY NOTES ON DELINEATION METHODOLOGY

Belgium's method for LMAs delineation [provisional description]

This description of the method used in Belgium is extracted from the description provided in answer to Q-2007 and from the description of the methodology attached to it.

Input data (bold) and indices:

- Commuting matrix T: $\mathbf{T_{ij}}$ is the flow (number) of workers that reside in municipality i and work in municipality j.
- R_i is the summation of flows with origin in municipality i (total number of workers residing in i – it includes T_{ii}). Similarly, R_A is the summation of flows with origin in any of the constituent municipalities of LMA A.
- SSA_A is the supply-side self-containment of A, equal to T_{AA}/R_A.
- SSD_{AB} is the supply-side dependence of A on B, equal to T_{AB}/R_A.

Parameters:

There are several values mentioned in the algorithm that could be considered parameters, namely the thresholds to classify municipalities into low (<30%), medium and high (>40%) self-containment and the ones to classify the dependencies into high (>5%) or low.

Algorithm:

1. Start considering each municipality as an LMA.

First phase:

2. Repeat:
   2.1. Find the pair of areas A and B that maximises SSD_{AB}.
   2.2. If A is a single municipality join A and B, otherwise start second phase.

Second phase:

3. For every remaining single-municipality area A with SSA_A<30% do:
   3.1. Find area B that maximises SSD_{AB} and area C that has the second highest SSD_{AC}.
   3.2. If B is a local hub (cluster of municipalities), merge A and B,
   3.3. else, if B is liable to become local hub ("would have been able to access the ranks of primary hubs, had they not found themselves 'encircled' by other hubs larger than them"), wait for step 4 to see if B becomes a hub,
   3.4. else, merge A and C **[even if C is not a primary hub nor liable to become one?]**.
   3.5. if maximum dependence of A is located abroad wait for similar cases **[there is only one case and it is not mentioned what to do then; see step 4.e]**.
4. For every remaining single-municipality area A with SSA_A>40% do:
   4.1. Find area B that maximises SSD_{AB}.
   4.2. For every area A that has SSD_{AB}>10% and SSD_{AX}<10% for any other area X, merge A and B **[B is a primary hub in all cases, it is unsure what would have**

been done otherwise].

4.3. For the remaining areas A, find the number of municipalities X that have SSD_{XA}>10%; "the higher the number of strongly dependent communes, the greater our chances of having a local hub" **[no threshold value is specified]**.

4.4. If SSD_{AB}<5% but A is not a local hub "it is necessary to weigh up in a specific fashion the hub (primary or secondary) to which its workers will tend to go and from which hub the persons performing a professional activity on its territory will tend to come" **[this step is not fully explained]**.

4.5. If the maximum dependence of A is directed abroad, it is necessary to look for a national hub that can serve as a national anchor point for it, or decide to set up a hub oriented towards another country (only in one case).

5. For every remaining single-municipality area A with 30%>=SSA_A>=40%, perform a case by case examination of the flows and attachment to other primary or local hubs or identification as a local hub following the previous criteria **[it is not explained how to apply those criteria]**.

Third phase:

6. Ad hoc comparison of the obtained delimitation with the one that would be obtained in the first phase if not stopped (full dendrogram) **[it is not specified how to perform this step]**.

Notes:

A. This method begins with the partial use of a regular open hierarchical clustering procedure that stops before merging two clusters (groups of communes), then the regionalisation is completed through a rule-based procedure.

B. Several questions about the methodology description in Q-2007 where addressed in Q-2012 but no answer has been received. Therefore, we cannot be sure of the interpretation in the previous algorithm. Moreover, the methodology was described as an ongoing work and without more recent input we cannot consider this method as definitive or official.

C. The third phase is not properly described in the methodology and it is assumed that it is a manual procedure. The same can be true for some parts of the second phase.

D. The self-containment levels of LMAs are not directly considered so a minimum level is not enforced, although the mergers are driven by dependency.

E. Contiguity is not considered in the algorithm, but no non-contiguous LMAs arise in the exercise.

Germany's method for LMAs delineation

The following description of the method used in Germany is based on the answer to Q-2012 and to the subsequent request for clarifications, part of which is literally quoted below:

> *To define the labour market areas, the indicator method was used. To apply the indicator method, the matrix of out-commuters and in-commuters is set up as a first step. For every Kreis (administrative district) of the area examined, the number of out-commuters into any other Kreis of the area examined is to be indicated. Based on that matrix, two matrices can be set up. The out-commuter matrix indicates for every Kreis the share of out-commuters to any other Kreis as a percentage of the persons in employment living in the Kreis concerned. The in-commuter matrix indicates the share of in-commuters from any other Kreis as a percentage of the persons employed in the Kreis concerned. Consequently, there are always four cells for any two Kreise, indicating the commuter links between the Kreise. In the indicator method, two Kreise form a labour market area if at least one of the four possible links exceeds the given critical value of commuter shares (e.g. 0.1 or 0.2).*
>
> *To define the LMAs as part of GRW (joint task of the Federation and the Länder to improve the regional economic structure), a proposal for definition was indicated for each of two given different critical values. The first proposal for definition groups all Kreise for which one of the commuter shares exceeds the value of 0.2. The second proposal for definition is less restrictive as the critical value is reduced to 0.1. Subsequent to that analysis, the resulting definition was checked for whether it meets the following constraints:*
>
> *- The definition of LMA borders must be identical to Kreis borders because, first, major statistical data are available only at Kreis level and, second, this enhances the political-administrative enforcement and implementation of GRW promotion activities. This also applies to the constraint that labour markets generally must not cross Land borders and that the entire territory must be covered without overlaps.*
>
> *- The self-containment of labour markets is covered by the fact that the area's self-supply in terms of jobs should be at least 65%. At the same time, this means that not more than 35% of the persons in employment living in the area commute out to other areas.*
>
> *- Another criterium ensuring the self-containment of the labour markets is that at least 65% of an area's persons in employment live in that area. Consequently, not more than 35% of an area's jobs may be filled by in-commuters from other areas.*
>
> *- To ensure a certain relevance of the labour market areas, the minimum number of inhabitants is set at 100,000 people.*
>
> *- Within a labour market area, the acceptable commuting time of 45 minutes per*

journey, that is, 90 minutes to work and back, must not be exceeded.

In cases where not all criteria can be met at the same time, a discretionary decision was taken regarding the definition.

This information describes in detail the constraints that a valid LMA should fulfill and the general criteria that guide the aggregation process, but the operation of the aggregation process itself is not fully described.

Input data (bold) and indices:

- Commuting matrix T: $\mathbf{T_{ij}}$ is the flow (number) of workers that reside in municipality i and work in municipality j.
- Distance matrix D: $\mathbf{D_{ij}}$ is the distance (in meters or minutes) between municipalities i and j.

1. R_i is the summation of flows with origin in municipality i (the total number of workers residing in i – it includes T_{ii}).
2. J_i is the summation of flows with destination in municipality i (the total number of jobs in i – it includes T_{ii}).
3. $SSD_{AB} = T_{AB}/R_A$ is the supply-side dependence of A on B (proportion of residents of A employed in B).
4. $DSD_{AB} = T_{AB}/J_A$ is the demand-side dependence of A on B (proportion of jobs in A hold by residents of B).
5. $SSA_A = T_{AA}/R_A$ is the supply-side self-containment of A.

Parameters:

1. minDependence: minimum supply- or demand-side dependence of a *kreiss* on another one to justify their merge, set to 10% or 20% (it is not stated which one is finally used, or the different uses of the LMAs resulting from both values).

2. minSC: minimum (supply- and demand-side) self-containment a LMA to be considerated as valid, set to 65%.

3. minSize: minimum population size (residents) of a valid LMA, set to 100000 (according to data some official LMAs have less population).

4. maxDist: maximum distance (in time or space) between two kreiss in a LMA to be considerated as valid, set to 45 minutes.

Notes:

- It is not possible to describe the algorithm itself using the information that has been

made available. More specifically, and among other aspects, no information is provided regarding the sequence in which the kreiss are considered. They could be for example simultaneously considered for potential mergers that would affect all pairs of kreiss reaching the minimum dependence without any recalculation of the dependences at that stage, or the process could be hierarchical, and based on the merger of the pair of municipalities that maximise the dependence indicator, and a subsequent recalculation of dependences between kreiss-kreiss or kreiss-group before considering the next pair of areas.

- The description of the method mention that a "discretionary decision" is taken in the cases were not all criteria can be met. As there is no explicit criteria to perform those decisions it can be considered as a manual final stage.

Czech Republic's method for LMAs' delineation

This description of the method used in Czech Republic is based on the answer to Q-2012 and the reply to a request for clarifications.

Input data (bold) and indices:

- Commuting matrix T: $\mathbf{T_{ij}}$ is the flow (number) of workers that reside in municipality i and work in municipality j.
- Municipalities' population (inhabitants) P: $\mathbf{P_i}$ is the number of residents (active or not) in municipality i. Similarly, P_A is the aggregated population of LMA A.
- N_A is the number of municipalities in LMA A.

Parameters:

- minPopLMA: minimum inhabitants in the LMA, 10000.
- minPopHinter: minimum inhabitants in the hinterland of an LMA: 4000.
- minMunLMA: minimum number of municipalities per LMA: 4.

Algorithm:

1. Identification of the dominating flow of each municipality i (Tij>=Tix for every municipality x).
2. Identification of work centres: municipalities that are the destination of the dominating flow of other municipalities.
3. Assignment of non-central municipalities to their dominating work centre (forming its hinterland).
4. Identification of non-valid LMAs: areas that do not count with (a) at least 10000 inhabitants (P_A>=minPopLMA), (b) at least three municipalities assigned to its work centre (N_A>=minMunLMA) and (c) at least 4000 inhabitants in its hinterland (P_A-Pworkcentre(A)>=minPopHint) are non-valid LMAs.
5. Each non-valid LMA is assigned to its own work centre's dominating work centre. In case of balance (when the number of commuters from the non-valid centre towards its diverse potential centres are of similar relevance), the decision is made considering also the flows from the non-valid centre towards its potential centres' hinterlands.
6. Application of local knowledge to resolve incoherence in the borders working over a map.

Notes:

A. The procedure could be described as a form of (hierarchical) clustering with a very simple linkage criterion: only the links between central elements in each cluster are considered.

B. All assignments in each phase are performed simultaneously.

C. The threshold values are not strict.

D. Contiguity is not included as a requisite, although it did not appear during the empirical

testing of the procedure.

E. According to the description, the procedure is not purely computerized and it is performed manually over spreadsheets and maps.

F. LMAs with very low self-containment levels could be created, as the aggregated flows are not considered. For the same reason, municipalities could be assigned to LMAs to which few workers commute (compared to the total number of outcommuters), despite having stronger aggregated dependence to other neighbouring LMAs (i.e. to the set of municipalities that conform them); phase 5 is intended to mitigate that, but this stage is not based on formal rules and cannot be codified.

G. Since the assignments are performed simultaneously, a rare case not considered in the procedure could cause ambiguity in phase 4: if there are two invalid LMAs A and B, with centres Ac and Bc, where dominating flow of Ac is towards Bc, and its second dominating flow is towards another work centre Cc, while the dominating flow of Bc is towards another work centre Dc. Would area A be assigned to Dc along with area B or would it be assigned to Cc?

H. The values of minimum number of inhabitants in core and hinterland and number of subordinated municipalities (the criteria of validity of a LMA) could and should be treated as parameters of the algorithm.

Estonia's method for the delineation of functional regions

The method used in Estonia is described in the partial answer to Q-2012 and a second piece of information by Dr. T. Tammaru.

Two concepts of core-based areas are considered in these documents:

I. Spheres of influence comprise county seats (cores) and the municipalities around them that send to the core more than 15% and 30% (two delimitations) of their workers. They do not allow overlapping (apparently this was possible in previous delimitations).

II. Labour-catchment areas comprise county seats and the municipalities that send at least 10% of their workers to the core or, if the municipality does send less than 10% to any core, to the core to which it sends more workers (regardless of any minimum threshold) or, if it is not contiguous, to the core of the municipalities that surround the one under consideration.

Input data (bold) and indices:

- Commuting matrix T: T_{ij} is the flow (number) of workers that reside in municipality i and work in municipality j.
- Contiguity matrix C: C_{ij} is the contiguity between municipalities i and j (1 contiguous, 0 non-contiguous). Similarly, C_{AB} is the contiguity between areas A and B (1 if any municipality in A is contiguous to any municipality in B).
- R_i is the summation of flows with origin in municipality i (total number of workers that reside in I, it includes T_{ii}).
- $SSD_{ic} = T_{ic}/R_i$ is the supply-side dependence of municipality i on core c.

Algorithm (labour-catchment areas):

1. Consider as core every county seat (previously they used jobratio>1).
2. For each free municipality identify its core, the one to which it sends more workers (maximise SSDic).
3. Attach each free municipality i to its core c if SSDic>10%.
4. For each remaining free municipality i (those where SSDic<10% for any core c) attach it to its core if it is contiguous, or attach it to the core of the surrounding municipalities otherwise.

Algorithm (spheres of influence):

1. Consider as core every county seat (in previous versions municipalities for which jobratio>1 were considered as cores).
2. For each free municipality identify its core, the one to which it sends more workers.
3. For each free municipality that sends more than 15% (30%) of its workers to its core, attach to it.

Notes:

A. It is a core-based clustering procedure. A core-identification step was previously used (municipalities with job ratio>1) but lately being one of the 15 county seats is the condition to qualify as a core.
B. Self-containment of the areas is not evaluated.
C. All municipalities must be directly connected through commuting flows to the core of the area they are assigned (no multilink is allowed).
D. Contiguity is enforced but the authors state it was correcting for it was rather necessary, and it is not clear if the algorithm itself considers the contiguity information in any way or if that restriction is ensured later, manually, while working with the plotted maps.
E. The spheres of influence regionalisation is **not exhaustive**: depending on the thresholds used, many municipalities can end up unassigned.
F. Although the maps attached to the answer to Q-2007 seemed to indicate that overlapping was permitted, the full answer to Q-2012 clarifies that overlapping is not allowed.

Greece's unofficial method for LMAs delineation

The description of this unofficial method used in an academic exercise in Greece is based on the response to Q-2012 which included a country note. The answer was prepared by Dr. Prodromos-Ioannis Prodromidis, the author of this academic study.

Input data (bold) and indices:

- Commuting matrix T: $\mathbf{T_{ij}}$ is the flow (number) of workers that reside in municipality i and work in municipality j. Similarly, T_{AB} is the summation of flows with origin in any municipality of area A and destination in any municipality of area B.
- R_i is the summation of flows with origin in municipality i (the total number of workers residing in i, it includes T_{ii}). Similarly, R_A is the summation of flows with origin in any area (group of municipalities) A.
- J_i is the summation of flows with destination in municipality i (the total number of jobs in i – it includes T_{ii}). Similarly, J_A is the summation of flows with destination in any area A.
- $SSD(A,B) = T_{AB}/R_A$ is the demand-side dependence of area A on area B.
- $DSD(A,B) = T_{BA}/J_A$ is the demand-side dependence of area A on area B.
- $MD(A) = \min(SSD(A,B),DSD(A,B))$ is the minimum dependence of area A on area B.

Parameters:

- minDep: minimum dependence of an area onto another in order to allow their aggregation, 15% (levels 10% and 20% were also tested).

Algorithm:

1. Start considering every territorial unit as a LMA.
2. Repeat:
   2.1. Find the pair of areas A and B that maximise MD_{AB}.
   2.2. If $MD_{AB} \geq$ minDep merge A and B (and recalculate flows between areas); otherwise terminate.

Notes:

A. The method can be described as an iterative hierarchical clustering procedure.

B. Contiguity is not imposed, but only three cases of non-contiguous LMAs appeared in their exercise.

C. Self-containment of the defined LMAs is not directly considered. Therefore, LMAs with very low self-containment levels could arise depending on the relation between the distribution of flows in the region and the parameter minDep (for example, in a region composed of territorial units with low autonomy levels and many relatively significant out-commuting flows it will be necessary a higher minDep to reach the same self-containment level than in the opposite scenario).

D. The answer to Q-2012 mentions another methodology recently applied in an academic exercise. The description of this method can be found in the academic paper Kallioras,

D. and Kandylis, Y. and Kromydakis, N. and Pantazis, P. (2011) "Definition of Local Labor Market Areas in Greece on the Basis of Travel-to-Work Flows", Univ. of Thessaly and National Centre for Social Research. The concept proposed there allows the overlapping of LMAs (one municipality can be attached to many poles).

Finland's method of LMAs' delineation

This description of the method used in Finland derives from the answer to Q-2012.

Input data (bold) and indices:

- Commuting matrix T: $\mathbf{T_{ij}}$ is the flow (number) of workers that reside in municipality i and work in municipality j.
- R_i is the summation of flows with origin in i.

Parameters:

- minAut: minimum supply-side (T_{ii}/R_i) self-containment, 75% (for identification of central municipalities).
- maxDep: maximum supply-side dependence (T_{ij}/R_i) from a single municipality, 10% (for identification of central municipalities).

Algorithm:

1. Identification of central municipalities: municipality i is central if $T_{ii}/R_i >$ minAut and $T_{ij}/R_i <$ maxDep for every municipality j.
2. While there are non-central municipalities unassigned to a central municipality:
   2.1. Find the non-central municipality i that maximises T_{ij}/R_i for every central municipality j.
   2.2. Assign i to j (that is, assign i to its dominating central municipality).

Notes:

A. It is a core-based hierarchical clustering procedure with a very simple linkage criterion: only the links between central elements of each cluster are considered.

B. Contiguity is not a restriction and non-contiguous LMAs are accepted (only a few cases in practice).

C. LMAs do not have to fulfil any criteria apart from having a central municipality and a hinterland (no minimum/maximum levels for self-containment, population or area).

D. LMAs with very low self-containment could be created, as the aggregated flows are not considered. For the same reason, municipalities could be assigned to LMAs to which few total workers commute, despite having stronger aggregated dependence to other neighbouring LMAs.

France's method for LMAs delineation

This description of the method used in France is based on the answer to Q-2012, the description of the method attached to the questionnaire, the answer to a request of clarifications, and another method's description as a diagram (it will be referred to as "Diagram" in this text). Some doubts about the method still remain (see numbered notes).

Input data (bold) and indices:

- Commuting matrix T: $\mathbf{T_{ij}}$ is the flow (number) of workers that reside in municipality i and work in municipality j. Similarly, T_{AB} is the aggregation of flows with origin in any of the constituent municipalities of area A and destination in a municipality of area B.
- Adjacency matrix C: $\mathbf{C_{ij}}$ is the contiguity between municipalities i and j (1 contiguous, 0 non-contiguous). Similarly, C_{AB} is the contiguity between areas A and B (1 if any municipality in A is contiguous to any municipality in B).
- Distance matrix D: $\mathbf{D_{ij}}$ is the distance (in meters or minutes) between municipalities i and j. Similarly, D_{AB} is the average distance between municipalities in areas A and B.

- R_i is the summation of flows with origin in municipality i (the total number of workers residing in i – it includes T_{ii}). Similarly, R_A is the summation of flows with origin in any of the constituent municipalities of LMA A.
- J_i is the summation of flows with destination in municipality i (the total number of jobs in i – it includes T_{ii}). Similarly, J_A is the summation of flows with destination in any municipality of LMA A.
- SSD_{AB} = supply-side dependence T_{AB}/R_A (proportion of residents of A employed in B).

Parameters: (see note A)

- minLink: minimum proportion of residents in the satellite area that work in the pole area needed in order to allow its aggregation (usually less than 1%).
- minSize: minimum population size (jobs) of a valid LMA (2).
- maxDist: maximum distance (in time or space) between two areas (measured as the average between municipalities in one area and municipalities in the other) to allow its aggregation [this restriction appears in the Diagram, but not in the texts that describe the method].

Algorithm:

1. Start considering every single municipality as a LMA.
2. Identify the pair of areas A and B with the highest SSDAB, where A is a non-pole area (**[J_A<minSize] (see note B)**) and B is any of its adjacent (C_{AB}=1), close enough (D_{AB}<maxDist), neighbouring **[pole] (see note C)** areas.
3. If $link_{AB}$>minLink, merge areas A and B (recalculation of T, C and D) and go to step 2. Otherwise finish.

Notes:

A. There is a general case algorithm, and major adaptations are made to fit diverse territorial realities (*grandes communes, territories proches des frontiers, zones se situant sur plusieurs regions, zones enclaves, poles isolés*) with the aim of for example avoiding too wide or too small LMAs. The associated criteria/values are not well specified; a general rule on how the decision of adjusting them is taken based on quantifiable territorial data is not made explicit.

B. There is a contradiction about the variable used to measure minSize in the different descriptions of the methodology:
   a. From description attached to the answer to Q-2012: "A chaque étape d'agrégation, on vérifie, en outre, la taille (en **nombre d'emplois**) de chaque unité (commune ou groupe de commune). Si celle-ci dépasse un seuil (appelé paramètre d'isolation et comptabilisé en termes d'**emplois**), elle devient pôle".
   b. From description in the Diagram: "Paramètre d'isolation: toute zone ayant une **population active au lieu de résidence** supérieur à ce seuil devient automatiquement pôle et ne pourra ainsi pas devenir satellite d'une autre zone".
   c. "**Population active au lieu de résidence**" would be R_A, while "**emplois"** is J_A.

C. The answers to the request of clarifications state that a pole area cannot be attracted by another pole area, but in the texts describing the procedure it is not specified that a non-pole area cannot attract another non-pole area when searching for the more intense link. However, at the end of the Diagram it is stated that both poles and satellite areas can "if possible" be outputs of the search of the highest link, so it is not clear whether this is a core-based procedure (only areas formed by municipalities identified as poles from the beginning can attract other municipalities) or not (aggregation of non-pole areas is possible and they can reach the status of pole area, and consequently become an independent LMA). The first interpretation is more likely after analysing all the information available. The notes in the Diagram about importance of a proper adjustment of the parameter minSize to avoid the creation of too small or too big areas reinforce that assumption.

D. Taking as correct the previously mentioned assumptions, this procedure could be described as a hierarchical core-based clustering procedure, with contiguity and minimum interaction restrictions for the aggregations and where the cores are determined by their size in jobs.

E. A final phase of manual adjustments to the delimitation can be performed after a round of consultations to the local authorities.

Italy's method for LMAs delineation

This description of the method used in Italy is extracted from the description provided in answer to Q-2007, confirmed in answer to Q-2012.

Input data (bold) and indices:

- Commuting matrix T: $\mathbf{T_{ij}}$ is the flow (number) of workers that reside in municipality i and work in municipality j.
- R_i is the summation of flows with origin in municipality i (the total number of workers residing in i, it includes T_{ii}). Similarly, R_A is the summation of flows with origin in any municipality of LMA A.
- J_i is the summation of flows with destination in municipality i (the total number of jobs in i – it includes T_{ii}). Similarly, J_A is the summation of flows with destination in any municipality of LMA A.
- $SSA_A = T_{AA}/R_A$, is the supply-side self-containment of area A.
- $DSA_A = T_{AA}/J_A$, is the demand-side self-containment of area A.
- MSC_A is the minimum self-containment of an area A, equal to $\min(SSA_A, DSA_A)$
- $SSD_{AB} = T_{AB}/R_A$ is the supply-side dependence of area A on area B.
- $\text{job ratio}_A = (JA\text{-}TAA)/(RA\text{-}TAA)$.
- $\text{interaction}_{iA} = T_{iA}^2/(R_i+J_A) + T_A^2/(R_A+J_i)$
- validity_A is the equation that measures how close is an area to be a valid LMA. It is calculated as $\min(MSCA/minSC, 1)*\min(J_A/minJob, 1)$.

Parameters:

- minSCcores: Minimum self-containment of the proto-cores, 50%.
- minSC: Minimum self-containment of the LMAs, 75%.
- minJob: Minimum size in number of jobs, 1000.
- minDep1: Minimum dependence of a municipality on a proto-core to consider its merger in step 2: 10%.
- minDep2: Minimum dependence of a proto-core on a municipality to consider its merger in step 2: 1%.
- minInteraction: Minimum interaction 0.002 (referred to as 0.2% in the source description).

Algorithm:

Start considering each municipality as an independent area.

1. Calculate *job ratio* and SSA for every municipality and consider as proto-cores the 20% of municipalities with higher *job ratio* and the 20% of municipalities with higher SSA.

2. For each proto-core A, in decreasing order of J_A-T_{AA}, do:
    2.3. While MSC_A<50% do:
        2.1.1 Find all the non-core municipalities i for which (SSD_{iA}>minDep1) and (SSD_{Ai}>minDep2) and ($interaction_{iA}$>minInteraction).
        2.1.2 Merge A and the municipality i from previous step (if any, otherwise end this loop) that maximises interactioniA and recalculate MSCA.
3. For every locality A (free municipality, core or multi-core), in decreasing order of $validity_A$, do:
    3.1 While $validity_A$<1 do:
        3.1.1 Find all the localities X for which SSD_{XA}>minDep1.
        3.1.2 Merge A and municipality X from previous step (if any, otherwise end this loop) that maximises $interaction_{XA}$.
4. For every municipality A where $validity_A$<1, in decreasing order of J_A-T_{AA}, do:
    4.1. Merge A with the proto-LMA X (locality where $validity_X$>=1, if any) that maximises $interaction_{AX}$.
5. For every proto-LMA A, in increasing order of MSC_A, do:
    5.1. If MSC_A <minSC, dismember A into its constituent municipalities and do:
        5.1.1. While there are any other free municipalities do:
            5.1.1.1. For each free municipality i, in order of J_i, assign i to the proto-LMA A that maximises $interaction_{iA}$.
6. (Optional) fine-tunning of the borders of the resulting LMAs.

Notes:

A. It is a core-based non-hierarchical agglomerative clustering procedure (also called rule-based) with an elaborated linkage criteria.

B. A minimum level of self-containment (minSC) is enforced for all the resulting LMAs.

C. Contiguity is not considered in the algorithm, but in the answer to the questionnaire it is stated that there is a treatment for non-contiguous areas, although it is not described and the **related question included in the country note of Q-2012 was not answered.** It could be assumed that non-contiguous LMAs are altered in the final (**manual?**) phase to produce fully contiguous LMAs that still meet the statistical criteria.

D. The algorithmic structure of this methodology is not fully represented in the source description. Although there are no doubts in the rest of the previous steps, the loop structure in step 5 of the previous algorithm is deducted, based on previous experience. The chosen structure is intended to cover the case where a dismembered municipality j, that only has interaction with other municipalities resulting from the dismembering of the former proto-LMA to which all of them belonged, is evaluated to find its (new) attracting proto-LMA before any of its functional neighbours have been assigned, therefore j cannot find an attracting LMA. As it is written, those residual municipalities are re-evaluated once the list of municipalities resulting from dismembering has been exhausted, and this sub-step is repeated until no residual municipalities remain.

Portugal's unofficial method for LMAs delineation

This method is used in an academic exercise in Portugal (by Mr. Pereira) and was presented to its NSI for its consideration (but still not official). Its description is included in the response to Q-2012 and in the answer to a request for clarifications.

Input data (bold) and indices:

- Commuting matrix T: $\mathbf{T_{ij}}$ is the flow (number) of workers that reside in municipality i and work in municipality j. Similarly, T_{AB} is the summation of flows with origin in any municipality of area (aggregation of municipalities) A and destination in any municipality of area B.
- Adjacency matrix C: $\mathbf{C_{ij}}$ describes the contiguity between municipalities i and j (1 contiguous, 0 non-contiguous). Similarly, C_{AB} is the contiguity between areas A and B (1 if any municipality in A is contiguous to any municipality in B, 0 otherwise).
- R_i (residents) is the summation of flows with origin in i (including T_{ii}). Similarly, R_A is the summation of flows with origin in any municipality of area A.
- J_i (jobs) is the summation of flows with destination in i (including T_{ii}). Similarly, J_A is the summation of flows with destination in any municipality of area A.
- S_i is the surface area of municipality i. Similarly, S_A is the summation of surfaces of all the municipalities that integrate area A.
- $interaction_{AB}$ is equal to the ratio of the (aggregated flow between areas A and B) and (the summation of both resident working populations), $interaction_{AB}=(T_{AB}+T_{BA})/(R_A+R_B)=interaction_{BA}$.
- SSA_A is the supply-side self-containment of area A, calculated as T_{AA}/R_A.
- DSA_A is the demand-side self-containment of area A, calculated as T_{AA}/J_A.

Parameters:

- minSC: minimum self-containment (supply- or demand-side) of the resulting LMA required to accept a merger: 85% (other values were also tested: 70%, 75% and 80%).
- maxSurface: maximum surface area of the resulting LMA required to accept a merger: 6,000km^2

Algorithm:

1. Start considering every single municipality as a LMA.
2. Repeat until no mergers are possible:
   2.1. Find the pair of contiguous LMAs A and B (so that $C_{AB}=1$) that maximise $interaction_{AB}$.
   2.2. Merge LMAs A and B if (a) the combined surface area is not over the established threshold ($S_A+S_B<maxSurface$) and (b) self-containment is over the established threshold ($max(SS_{AA}, DS_{AA})\geq minSC$).
3. Repeat until no isolated municipalities remain:
   3.1. Find the pair of isolated municipality i and LMA A, contiguous ($C_{iA}=1$)

that maximise interaction$_{iA}$.

3.2. Merge i and A.

Notes:

A. The method can be described as an open (not core-based) iterative hierarchical clustering procedure.

B. Contiguity is enforced during the whole process.

C. A minimum level of self-containment of the defined LMAs is directly considered a condition during the main aggregation phase, but it is not enforced in the last phase of the algorithm. Therefore, areas with self-containment lower than the parameter threshold could arise.

D. A maximum threshold of surface area is specified during the main aggregation phase, but it is not enforced in the last stage of the algorithm. Therefore, LMAs with greater surfaces could arise.

E. It seems controversial that the merger of two LMAs is not allowed if the joint self-containment does not surpass a minimum, regardless of their intensity of interaction and the self-containment of the separated areas. This would imply, in a territory composed by many municipalities with low self-containment levels, that the phase of aggregation of areas (step 2) would be ignored. Therefore all the clustering would be performed in the last phase, where the conditions of maximum surface area and minimum self-containment levels are not considered, so the resulting LMAs' characteristics could potentially be far from the ones envisaged.

F. A different version of this algorithm exists in a previous communication by the authors: "Para uma discussão de Regiões Urbanas Funcionais em Portugal". In that document the conditions are slightly different.

G. These methods are based on the previous work by Pereira (1997) "Bacias de Emprego em Portugal Continental", *Revista de Estatística* (INE), No 4, 1st semester.

Sweden's method of LMA delineation

This description of the method used in Sweden is extracted from the answer to Q-2012.

Input data (bold) and indices:

- Commuting matrix T: $\mathbf{T_{ij}}$ is the flow (number) of workers that reside in municipality i and work in municipality j.
- R_i is the summation of flows with origin in i.

Parameters:

- minAut: minimum supply-side self-containment (T_{ii}/R_i), 80% (for being considered as a central municipality).
- maxDep: maximum (supply-side) dependence in relation to a single municipality (T_{ij}/R_i), 7.5% (for being considered as a central municipality).

Algorithm:

1. *Identification of central municipalities*: municipality i is central if $T_{ii}/_{Ri}$>=minAut and T_{ij}/R_i<maxDep for every municipality j.
2. Two municipalities i and j that do not fulfill the conditions to be central but have their largest out-commuting flow directed to each other (that is, T_{ij}>=T_{ix} for every municipality x and T_{ji}>=T_{jy} for every municipality y) give place to a combined central locality.
3. Every non-central locality is assigned to the municipality that acts as the main destination of its largest out-commuting flow. Municipalities are then classified in the following classes:
    - Type-11: central localities.
    - Type-20: municipalities whose largest out-commuting flow is directed to a type-11 locality.
    - Type-30: municipalities whose largest out-commuting flow is directed to a type-20 locality.
    - Type-50: municipalities whose largest out-commuting flow is directed to a type-30 locality.
4. Every type-11 locality and its assigned municipalities form a LMA.

Notes:

A. It is a core-based hierarchical clustering procedure with a very simple linkage criterion: only the flows between central elements of each cluster are considered.

B. Although a contiguity restriction is not considered it does not appear in their exercise.

C. Each LMA that has type-30 (or type-50) municipalities or a type-11 combined locality is considered polycentric (the extra centres are the type-20 localities in the LMA with type-30 localities assigned to them).

D. LMAs do not have to fulfill any criteria apart from having a central municipality (no minimum/maximum self-containment, population or area levels).

E. LMAs characterized by very low self-containment levels could be created in the process, as the aggregated flows are not considered. For the same reason, municipalities could be assigned to LMAs towards which few (compared to the total number of commuters) workers commute, despite having stronger aggregated dependence to other neighbouring LMAs (if the main destination municipality for out-commuters from municipality A is in LMA X and the rest of commuters go to municipalities belonging to LMA Y, municipality A would be assigned to X although it is likely that the total number of commuters to Y exceeds the number of commuters to X).

Slovenia's unofficial method for LMAs' delineation (Drobne and others)

In Slovenia there is not an official delineation of functional areas or LMAs. However in the national answer to Q-2012 a set of recent academic studies that deal with these issues was attached. What follows is a short review of those papers, ordered by year of publication

S. Drobne, M. Konjar and A. Lisec (2009) Delimitation of Functional Regions using Labour Market Approach.

The so called Labour Market approach is explained in this paper. It is taken from Karlsson and Olsson (2006), who did not mention any other author for the LM approach but quote the Swedish Institute of Statistics SCB's method (SCB, 1992) as their inspiration (this is the same method reported by SCB in their answer to Q-2012). What follows is a brief description of the algorithm.

Algorithm:

1. Identifications of local centres (minimum supply-side self-containment and minimum number of jobs)
2. Assignment of non-centres to centres:
    - 2.1. non-centres which have their largest outgoing flow towards a centre are assigned to them
    - 2.2. non-centres which have their largest outgoing flow towards a non-centre already assign to a centre are assigned to that center
    - 2.3. pairs of non-centres whose largest outgoing flow are directed to each other are assigned to their (**combined?**) second largest flow.

Therefore, it is a core-based hierarchical clustering procedure that results in a map of non-overlapping LMAs that exhausts the whole territory.

M. Konjar, A. Lisec and S. Drobne (2010) Methods for delineation of functional regions using data on commuters.

Three approaches, so called, LM, CZ and the suggested Commuting Aggregation, are explained in this paper.

For LM approach check comments on previous paper.

The CZ (commuting zones) approach described in this work also cites Karlsson and Olsson (2006) as the source. It must be noted that the procedure described in Konjar et al. (2010) is not the same that is referred to in Karlsson and Olsson (2006), where it is not fully described since they forward to the source (Killian and Tolbert, 1993). The algorithm that Konjar et al. describe

as CZ does not exactly correspond to the original[60] commuting zones approach by Tolbert and Sizer (1987, 1996) and Sizer and Tolbert (1993). Instead, it is extremely similar to the already described LM approach: it uses the same core identification criteria, both are hierarchical clustering (the dismembering and reassignment of the constituent municipalities of a proto-LMA is not considered) and the measure of interaction that drives the aggregation process is calculated over non-aggregated flows (from the free municipality under consideration to the core of the candidate LMA, but not to the whole LMA). The only difference between the first and the second approaches described in this article (LM and so-called CZ) refers to the interaction index used to measure the links between two nodes: in the so-called CZ approach flows in both are considered ($T_{ij}+T_{ji}$) but only the working population of the smallest municipality is considered in the denominator ($\min(R_i,R_j)$). In contrast, the original CZ methodology from Tolbert and Sizer is an open system (not-core-based) and uses average linkage (instead of satellite-core linkage).

Finally, the so-called Commuting Aggregation approach proposed here as a new contribution is a variation of the previous one (called CZ in this article), but with an open (not core-based) system and considering aggregated flows. That is, it is more close to the original Tolbert and Sizer's original CZs approach than the previous method, although it does not use the same linkage criterion. It is also very similar to the academic exercises by Podromidis in Greece and Pereira in Portugal, with the difference in the interaction index employed (the former consider the sizes of both areas in the denominator).

S. Drobne and M. Bogataj (2012, accepted) A Method to define the number of functional regions: an application tu NUTS 2 and NUTS 3 levels in Slovenia .

They apply Intramax method ("Flowmap" software more exactly) to Slovenia, and analyse 29 alternative maps resulting from dividing the territory into (2-30) functional regions for each year and level of aggregation (NUTS 2 and 3). That analysis is then used to develop criteria that could guide the election of the number of functional regions, considering (a) the demographic criterion of the EU guidelines for the size of the region at NUTS 2 and NUTS 3 levels, and (b) the criterion on economic homogeneity of regions.

They use a weighted equation of the coefficient of variation of average monthly gross earnings per capita between functional regions and the coefficient of deviation of population in the region regarding the EU guidelines, measured for each (year/NUTS' level) regionalisation. According to their results more stress on homogeneity results in a greater number of (smaller) regions compared to the alternative considered: giving more importance to be in the centre of the NUTS population orientative thresholds.

[60] This approach is fully described in the summary of Karlsson and Olsson (2006).

The procedure used to produce the maps of LMAs is Intramax (Masser and Brown, 1975, 1977), and it is not described in this paper. It is a hierarchical aggregation procedure that focuses on the relative strength of interactions, once the effect of size variation in row and columns totals is removed, through the comparison of the observed flow and the expected value that is derived from the multiplication of the corresponding row and column totals, when the matrix is standardised to sum to unity.
Intramax's algorithm:

1. Transform the commuting matrix T into the standardised one S:

$$S_{ii} = \frac{T_{ii}}{\sum_j R_j}$$

2. Start considering every single municipality as an area.
3. Repeat:
   3.1. Find the pair of areas A and B that maximise $linkage_{AB}$.
   3.2. If $linkage_{AB} \geq minThreshold$ merge together A and B; otherwise terminate.

In this case, $linkage_{AB}$ is calculated as follows:

$$linkage_{AB} = \frac{S_{AB}}{\sum_P S_{PB} \sum_Q S_{AQ}} \square \frac{S_{BA}}{\sum_P S_{PA} \sum_Q S_{BQ}}$$

S. Drobne, A. Lisec, M. Konjar, A. Zavodnik Lamovšek and A. Pogačnik (2009) Functional vs. Administrative Regions – Case of Slovenia
This paper presents a discussion about three proposals of administrative division in 2, 6 and 8 provinces based on the LM approach.

S. Drobne, M. Konjar, A. Lisec, N. Pichler Milanović and A. Zavodnik Lamovšek (2010) Functional Regions Defined by Urban Centres of (Inter)National Importance – The Case of Slovenia
This paper presents a general discussion on urban and functional regions (using LM approach) in Slovenia.

Republic of Slovakia's method for LMAs delineation

This description of the method used in the Republic of Slovakia is extracted from the answer to Q-2012 and from the answer to a request for clarifications.
A number of ambiguities remain (in the answer to the request for clarification it is stated that "Mentioned material 'LMA' was elaborated ten years ago. The work on LMA was definitively finalized at that time. Person working on it does not work for our Office. We are not able to answer your questions".

Input data (bold) and indices:

- Commuting matrix T: $\mathbf{T_{ij}}$ is the flow (number) of workers that reside in municipality i and work in municipality j.
- Adjacency matrix C: $\mathbf{C_{ij}}$ is the contiguity between municipalities i and j (1 contiguous, 0 non-contiguous). Similarly, C_{AB} is the contiguity between areas A and B (1 if any municipality in A is contiguous to any municipality in B).
- R_i is the summation of flows with origin in municipality i (the total number of workers that reside in i – it includes T_{ii}). Similarly, R_A is the summation of flows with origin in any municipality of LMA A.
- J_i is the summation of flows with destination in municipality i (the total number of jobs in i – it includes T_{ii}). Similarly, J_A is the summation of flows with destination in any municipality of LMA A.
- $core_A$ is the core municipality of area A.
- $hinter_A$ is the combination of non-core (subordinated) municipalities of area A.
- A municipality i is subordinated to core municipality j if i sends to j more workers than to any other core.
- NSM_A is the number of subordinated municipalities (non-core) of area A.
- $SSA_A = T_{AA}/R_A$ is the supply-side self-containment of A.
- $jobratio_A = J_A/R_A$.
- $extwork_A = (J_A-T_{AA})/R_A$ is the proportion of jobs hold by residents from other areas.
- $closeness_A = W_A/(O_A+I_A)$
  where: W_A is the summation of flows with origin in any municipality of A and destination in its core; O_A is the summation of flows with origin in any municipality of A and destination in any core outside A; I_A is the summation of flows with origin in any municipality outside A and destination in the core of A.

Parameters:

- minSSAcore: minimum supply-side self-containment level that a municipality must meet to become core.
- minJR: minimum job ratio that a municipality must meet to become a core.
- minNSM: minimum number of subordinated municipalities that a municipality needs to become a core.
- minJF: minimum ratio of foreign job for a municipality to be a core.
- minSSA: minimum supply-side self-containment level for a valid LMA.
- minCloseness: minimum closeness for a valid LMA.

Algorithm:

1. Consider as core every municipality i that fulfills the following conditions:
   1.1. $jobratio_i$>minJR and (SSA_i>minSSA or $extwork_i$>minJF)
2. For each non-core municipality identify its core (the core that receives the largest outgoing flow from that municipality to any core).
3. While there are unassigned non-core municipalities, do:
   3.1. Assign all the non-core municipalities to their cores if the non-core municipality is contiguous to the core's area.
4. For each LMA A where NSM_A<minNSM, merge A with the other core that exerts more attraction over A. **[sic]**. Non-contiguities are solved individually by opinion of experts on the basis of cadastral maps **[sic].**
5. For each LMA A where $closeness_A$<minCloseness or SSA_A<minSSA do:
   5.1. Dismember A into its constituent municipalities and identify their (new) cores.
   5.2. While there are unassigned municipalities, assign all the non-core municipalities to their cores if adjacent to the core's area.

Notes:

A. The procedure can be described as non-hierarchical rule-based clustering, with some similitude with the one used in the TTWAs' revision based on the 1991 Census and the one currently in use in Italy.

B. A minimum level of self-containment is enforced for all the LMAs defined.

C. Contiguity is enforced during the whole process, but not necessarily in a systematic way (see notes below).

D. There are several issues that remain unclear, mainly:
   a. We assume that the assignments (steps 3.a and 5.b.1) are performed in parallel, since no order for a sequential assignment is specified (an explicit question on this issue was included in both Q-2012 and the request for clarifications).
   b. Step 4 of the previous algorithm (merger of LMAs with less than six municipalities) is only mentioned in answer to the last request for clarifications. In the initial description in answer to Q-2012, that condition (the number of subordinated municipalities of an LMA) was mentioned along with the conditions to identify cores, and no merger of whole LMAs (instead of dismemberment and reassignment of separated municipalities) was

mentioned. Moreover, in the answer to the request for clarifications it is stated that minNSM is not a condition for core identification.

c. Step 4 is not properly described and it cannot be reproduced without ambiguity: How is the attraction between a pair of LMAs measured? Alternative options are: from the core of the invalid LMA to the cores of contiguous LMAs, but also from the aggregated LMA to the neighbouring cores or aggregated LMAs. How is discontinuity treated?
d. The last two steps described in the answer to Q-2012 have been removed in the answer to the request for clarifications (it is stated there that "Points 4 and 5 were not used in the framework of regionalisation"). From those two steps, the first one seemed useless, as the criteria stated for it was the same as in the previous step, so it could be arguable that no unassigned municipalities would reach step 5. Step 5 is similar to what it is done in step 3 (dismemberment and reassignment). Maybe it was a misunderstanding and someone labelled as steps 4 and 5 what were extended explanations of step 3; but, as mentioned in the first note, the five-steps version of the algorithm is very similar to the five-step algorithm used in the TTWAs' revision after Census 1991 and the one used in Italy.

United Kingdom's method for LMAs (TTWAs) delineation

This description of the method used in United Kingdom is extracted from the description provided in answer to Q-2007, confirmed in answer to Q-2012 and from the description of the methodology available online.

Input data (bold) and indices:

- Commuting matrix T: $\mathbf{T_{ij}}$ is the flow (number) of workers that reside in municipality i and work in municipality j.
- R_i is the summation of flows with origin in municipality i (the total number of workers residing in i, it includes T_{ii}). Similarly, R_A is the summation of flows with origin in any of the constituent municipalities of LMA A.
- J_i is the summation of flows with destination in municipality i (the total number of jobs in i – it includes T_{ii}). Similarly, J_A is the summation of flows with destination in any of the constituent municipalities of LMA A.
- $SSA_A = T_{AA}/R_A$, is the supply-side self-containment of area A.
- $DSA_A = T_{AA}/J_A$, is the demand-side self-containment of area A.
- MSC_A is the minimum self-containment level of an area A, equal to $\min(T_{AA}/R_A, T_{AA}/J_A)$
- $interaction_{iA} = T_{iA}^2/(R_i * J_A) + T_{Ai}^2/(R_A * J_i)$
- $validity_A$ is the equation that measures how close is an area of being a valid LMA. It is calculated as follows:
  - $validity_A=1$ if $R_A \geq minSize1$ and $MSC_A \geq minSC1$ and $MSC_A \geq M*tam+N$
  - $validity_A=\min(R_A/minSize2, 1)*\min(MSC_A/minSC2, 1)$ otherwise,

  where $M=(minSC2-minSC1)/(minSize1-minSize2)$; and $N=(minSC1*minSize1 - minSC2*minSize2)/(minSize1-minSize2)$

Parameters:

- minSC1 and minSC2: relaxed and strict minimum (supply- and demand-side self-containments) of a valid LMA, minSC1 < minSC2.
- minSize1 and minSize2: relaxed and strict minimum population size (number of workers residing in i) of a valid LMA, minSize1 < minSize2.

Algorithm:

1. Start considering every municipality as a LMA and calculate its validity.
2. While there are invalid LMAs ($validity_A<1$) do:
    2.1. Find the LMA **A** with lowest validity index and dismember it into its constituent municipalities.
    2.2. While there are dismembered municipalities unassigned do:
        2.2.1. For each municipality **i** in **A** find the LMA $\mathbf{X_i}$ that maximises $interaction_{iX}$.
        2.2.2. Assign simultaneously every municipality i to its dominating LMA $\mathbf{X_i}$.
    2.3. Recalculate validity for every modified LMA.

Notes:

A. A minimum level of self-containment is enforced for all the resulting LMAs.

B. Contiguity is not considered in the algorithm but non-contiguous LMAs are altered in a final (manual) phase to produce fully contiguous LMAs that still meet the statistical criteria.

ANNEX 2: TEMPLATE QUESTIONNAIRE ON LMAs

The brief questionnaire that follows was an Action agreed by Member States at the Working Party on Regional Statistics and Rural Development (4-5 October 2011): please complete it on behalf of your Member State.

Introduction

Many countries now delineate labour market areas (LMAs). These sets of functional areas are seen as an alternative to local and regional administrative areas for some statistical purposes, and for the design, implementation and monitoring of labour market and other public policies. What makes LMAs attractive to policy-makers, economists and researchers is that they reflect the spatial pattern of economic and social relations in the area.

To summarise: LMAs group neighbouring 'building block' areas (eg. at LAU2 level) that are:

5. defined to be self-contained, so that little human interaction crosses their boundaries;
6. delineated so that commuting between homes and workplaces is mostly internalised;
7. defined for the whole country, which is covered by LMAs with no 'gaps' or overlaps.

Two decades ago Eurostat and Newcastle University carried out a study of principles upon which to base definitions of LMAs to be used in a policy context. Then ten years ago OECD undertook a review of the LMA definitions in its member countries. Five years ago DG REGIO expressed a strong interest in LMA statistics and Eurostat carried out a survey of the LMA definitions of Member States that same year (2007). Now this new questionnaire will provide evidence in a study to explore the feasibility of common definitions of LMAs for the entire EU.

This questionnaire is part of a larger study whose objectives are to:

(1) outline the state-of-art in applied research on LMAs;
(2) compare the LMA concepts recognised, and implemented, in each Member State;
(3) draw conclusions on relevant best practice;
(4) explore the added value of a common definition of LMA for the entire EU; and finally
(5) identify possible ways and means of harmonising LMA definitions across the EU.

The questionnaire aims at compile the fundamental evidence needed to meet objective (2). Member States are asked to kindly collaborate by completing this questionnaire by supplying the requested information on LMA geography in their country.

Guidance on completing the questionnaire

As mentioned above, Eurostat sent a related questionnaire in 2007 to Member States NSOs. This new and revised Questionnaire (hereafter Q-2012) derives from its predecessor (hereafter Q-2007), and its objectives are to:

- identify the progresses made at national level in the definition of LMAs (if any),
- gather more information on the availability of commuting data (used for the delineation of LMAs) at national level, and on NSOs positions towards an LMA geography across the entire EU territory.

The respondents are invited to kindly fill-in the questionnaire even if the concept of LMA is not defined in their country, nor they have plans to create LMAs.

For some questions, specific "notes" are formulated, which may also be interpreted as (partially) prefilled-in answers. These notes are based on the information already identified by Eurostat. Please consider these notes before answering the questions.
We hope that this approach will help you to complete the questionnaire quickly and without much difficulty, but if necessary you can contact the Eurostat officer responsible for the study:

Mr. Oliver HEIDEN
Eurostat, UNIT E4
Regional Statistics and Geographical Information
BECH A3/049
Oliver.Heiden@ec.europa.eu
Tel: (+352) 43 01 31 961
Fax: (+352) 43 01 34 029

THANK YOU VERY MUCH FOR YOUR COLLABORATION

2012 Questionnaire on Labour Market Areas in EU Member States

Country:

Name of respondent Q-2007:

Date of response Q-2007:

Name of respondent Q-2012:

Date of response Q-2012:

Section A - Existing delineation methods

A - Q.1 In your country, do you use the concept of Labour Market Areas, i.e. have LMAs been defined (either by your office or by another organisation in your country)?

A - Q.2 bis If the new answer is no, do you have any plans in the future to create LMAs, or are there existing boundaries, e.g. of administrative areas, that you consider adequately represent economic regions? If so, what are they?

A - Q.3 Which is the national concept (i.e. sistemi locale del lavoro (IT), zones d'emploi (FR), etc.) of these areas and what definition do you apply? If the definition is available on Internet, please send the link.

A - Q.4 There may be several alternative sets of LMAs definitions in your country; if so, please answer the questionnaire in relation to the set that is in the highest of the following list:

1. LMAs used for official statistics
2. LMAs used for some other government purpose (e.g. planning)
3. LMAs used only for academic or other purposes

Which of the above categories do the LMAs you describe in this questionnaire fall into?

A - Q.5 When were these LMAs defined?

General questions about the definitions of LMAs

A - Q.6 What main information sources were used in defining the boundaries of the LMAs (e.g. commuting or migration flows, or local knowledge)?

A - Q.7 Did the definition process consider the whole country, or only certain parts?

A - Q.8 Was there a single unified definition process, or were regions analysed separately (e.g. so that the boundaries of the regions could not be crossed by the LMAs)?

A - Q.9 Did the definition process end by allocating all the areas considered in LMAs, or could some areas be left unallocated (e.g. remote or island areas)?

A - Q.10 Please describe in detail the method you used to define the LMAs currently operational in your country, in a way that could allow it to be used to define LMAs in other countries.

Most probably the description of your methodology will be a longer text, to be attached separately to this questionnaire. We would prefer a text in English, but if it is only available in your national language, this is fine as well. We will translate it. If the text comes from a different institution, please organise the contact to that organisation, but send us the result of your research from your RESCO desk.

The following questions summarise issues which the description should cover; please provide summary answers here:

A - Q.11 What are the basic building block areas?

A - Q.12 Is there a criterion to decide which areas to group together (e.g. a minimum rate of commuting)?

A - Q.13 Is there a restriction on grouping non-contiguous zones?

A - Q.14 Are multiple "step" links allowed (e.g. will chains of cities, with strong commuting between each pair of cities, be grouped into a single LMA, or will the chain be broken)?

A - Q.15 Are all the initially defined groupings accepted as LMAs, or do they have to meet some other criteria (e.g. a minimum level of self-containment or population, or maximum size of area)?

A - Q.16 Is there a consideration of possible transnational LMAs?

A - Q.17 Please send us a detailed description of the boundaries of your LMAs, either through geo-reference layers or with the aid of a list of the administrative units contained in each LMA. Preferably use the LAU list that you have sent to Eurostat for the reference year concerned. A map of the LMAs in your country would also be very much appreciated.

[Please indicate the name of the file attached and the reference year]

A - Q.18 Please send us a file with the most recent area and population in each LMA.

[Please indicate the name of the file attached and the reference year]

Section C – Data for the delineation of LMAs

Please, note that with only a couple of exceptions, all EU-27 Member States collect data that may enable the delineation of LMAs at LAU2 or LAU 1 levels. This data mainly results from Population Censuses, but could be also retrieved from specific surveys, register based surveys, or other themes generally NSOs produce data for (e.g. Employment Statistics, Local or Regional Statistics, Territorial data, Labour migrations, etc.).

C - Q.19 Is a travel-to-work dataset available?

Please note that we specifically refer to data on commuting flows between geographical units in your country - number of persons that commute between each pair of units considered as origin and destinations of the flow. This information can be visualised as a matrix where all the units appear as rows and columns so that each cell Tij depicts the number of persons commuting from unit i to unit j.

C - Q.20 If such data on commuting is not available, please indicate it and state whether they will be available in the future (for example after the Census of Population 2011), and give details on this.

C - Q.21 If this type of data does not exist for your whole country, please indicate for which part of the country do they exist and proceed to the following questions.

C - Q.22 With regards to **the most recent data currently available**:

C - Q.22. a Which is the **lowest territorial level** for which commuting data are available:

☐ LAU 2
☐ LAU 1
☐ Other levels [Please, indicate nomenclature and number of units that constitute this level]

C - Q.22. b Please indicate the reference **year**:

C - Q.22. c Please, indicate the **source** (Census of Population, other, etc.)

C - Q.22. d **Organisation** in charge of producing the data:

C - Q.22. e Are data publicly **available**? If so indicate how these could be accessed:

C - Q.22. f When will an **update** of these data be available?

Indicate year, source and periodicity of the data (bear in mind that we always refer to data on commuting flows between all geographical units of a certain level in your country). Feel free to add comments.

C - Q.22. g Please specify any relevant **particularities** of the data

Do the data provide **disaggregated** information by gender / section of activity / other?

Do the data include information about **cross border incoming and outgoing commuting**? If so, which is the detail? Will it be possible to know the destination and origin of cross border

commuters by geographical unit (LAU2, LAU1, other) of origin and destination? Please add your comments on this issue.

Is/Are any specific sub-group(s) of workers (such as self-employed) excluded? If this is the case, please provide details on the percentage they mean of total employment.

Are the data based on direct information from all the population or extrapolated from samples? Please give details of the sampling strategy if this was the case.

Are data aggregated for more than one geographical unit due to confidentiality concerns caused by small population? If this is the case, how is this process guided?

Other comments (e.g. are the unemployed considered in any way?)

C - Q.22.h Please indicate for which of the **previous years** is it possible to obtain similar information and feel free to add your comments on this issue.

C - Q.23 With regards to the 2011 round of Population census and referring to your country, will it be possible to construct a matrix of commuting flows between geographical units of LAU 2 level? Which is the expected date of availability of such data?

C - Q.24 The study of which this questionnaire is part is researching commuting data for all EU Member States: can you facilitate our access to the most recent data on commuting flows between small areas in your country (ie. number of persons that commute between each pair of units considered as origin and destinations in a matrix)? If you are not the right contact point, please provide an internet link or email address where the data should be sought from.

Section D - Sources of information -possible characterisation of LMAs

D - Q.25 What socio-economic variables do you specifically collect for your LMAs?, where LMAs are sets of other basic geographical units (i.e. LAU 2). Which are the main indicators available for such units?

D - Q.26 How often are these variables / indicators updated?

D - Q.27 If you have published the definition and / or the variables / indicators for LMAs on the Internet, please send us the web link.

Section B - NSOs' positions towards an LMA geography and / or their expectations from this

B - Q.28 Do you find LMAs useful? Which purposes should they be fitted to?

B - Q.29 Which were the reasons for the creation of LMAs?

B - Q.30 Which are your main concerns regarding such an exercise (i.e. defining a new statistical classification)?

B - Q.31 Please share your opinion / comments about the statistical characteristics that should be met by LMAs if an EU-wide grid of comparable areas would be proposed (e.g. How a standard objective in terms of area / population, commuting self-containment or other variables should be established?, etc.)