Scales in Space

ANNE VAN DER VEEEN AND HENRIËTTE S. OTTER

University of Twente, Public and Social Administration, Enschede, The Netherlands

ABSTRACT

Economists have devoted more attention to the scale of time than to the scale of space. What has been done in the field of space is often general and abstract, not connected to an explicit observation set in time and space. Moreover, time scales and spatial scales are not tied, making the choice for a macro, meso or microeconomic theory a rather arbitrary process. We devote attention to the explanation of the phenomenon of emerging spatial structures. We will discuss the standard economic theories that describe the underlying processes and argue that by being more explicit about spatial scales explanatory power is added to current theoretical work.

Keywords: regional economics, scales and aggregation, emergent spatial structures.

1. INTRODUCTION

One of the underrated topics in economics is the issue of scale and aggregation. To be more precise, in regional economics spatial scale and spatial aggregation is a neglected item. This statement might sound a little bit strange in a world where transportation economics, regional economics and urban economics are well-established fields. However, it is our belief that in defining an observation set in order to understand the arrangement of spatial patterns and structures economists are poorly equipped. Economists have devoted more attention to the scale of time than to the scale of space. In addition, what has been done in the field of space is often general and abstract, not connected to an explicit observation set in time and space. Finally, it is our perception that in economics time scales and spatial scales are not tied, making the choice for a macro, meso or microeconomic theory a rather arbitrary process.

We cannot handle these critical remarks all at the same time so we will restrict ourselves in order to illustrate our point of view. In this article, we will devote attention to the explanation of the phenomenon of emerging [1] spatial structures [2]. We will discuss the standard theories that describe the underlying processes and argue that by being more explicit about spatial scales explanatory power is added to current theoretical work.

Given these introductory remarks on time, space and aggregation we will first pay attention to the choice of scales and aggregation levels in general. The issue of (spatial) aggregation as an almost insurmountable step will be discussed in some detail. Secondly, we devote a special section to ecology. We recently experienced that in ecology a discussion has taken place on exactly the same topic as we present here and we are convinced that by reviewing their findings on time and space, and especially their conclusions on aggregation, we can learn. Moreover, as an example, we will measure how location theory, as the heart of regional economic theory, is influenced by scaling. We evaluate how spatial resolution is handled in location theory and discuss how defining the problem in terms of spatial resolution might contribute to a better understanding of the phenomenon of emerging spatial patterns. Finally, we devote a section to the consequences for government in the design of spatial policy.

2. SCALES AND AGGREGATION

Models are abstract maps of empirical reality around us. Examples of these representations of reality are mental models, mathematical models, simulation models, physical scale models etc. A binding element in all is that we aim to frame theories and ideas to better understand the empirical chaos.

In every model, a choice has to be made on scales. Choosing a scale on which to project the objects and processes in a model refers to a quantitative and analytical dimension and to time and space [3, 4]. Concerning these
scales we may further discriminate between resolution and extent. For resolution in temporal and spatial scales we thus define a:

- time step (e.g., a day), and a
- spatial step (e.g., a grid of 100 by 100 meter),

For extent we can distinguish between:

- the extent of time (e.g., a year), and
- the spatial extent (e.g., a country).

As an example of the distinction above, abstract neoclassical models in economics have low temporal and spatial resolution. Moreover, they have relatively high extent in time and space. Large national-regional econometric models, on the other hand, may have a higher spatial resolution and consider a smaller extent in time.

Besides quantitative and analytical dimensions and time and space, there is another concept to introduce and that is level. It is defined as the unit of analysis along a scale [3]. Economists prefer to speak about aggregation level. Level follows from systematically making choices on time and spatial scale (and thus on resolution and extent) and on quantitative and analytical dimensions.

Before continuing with scales in economics, we make a remark on the aggregation process in economics. In economics, we do not have data at the coarse scale. Coarse data are aggregates used for macro or meso economic analyses. Two types of such economic aggregates can be distinguished: Aggregate quantities and aggregate agents [5]. Relationships between aggregate macroeconomic quantities can be derived from [6]:

1. a macro theory, e.g., the Harrod and Domar model,
2. a method based on analogies from micro behaviour, or
3. an aggregation of micro relations based on micro characteristics.

A macro theory under (1) always has more or less an ad hoc character. It is based on rigorous hypotheses on relations between aggregate variables and is not related to any micro behaviour. The analogy method under (2) is followed in consumption and production theory. Studies in this field start with an elaborated theory of individual behaviour, but they are also assumed to hold for per capita data of totals. However, as Van Daal and Merkies [6] note, “Usually any argument in defence of this jump in the train of thoughts is lacking.” More firmly, Malinvaud, in Harcourt [7], states the following about the microeconomic foundations of macroeconomics: “Aggregation was hardly ever justified, except in rather narrow cases, which were not often found in fact. Most of the times our macro economic theory therefore lacked the rigorous justification that we should like to find in micro-economic analysis.” The implication of these arguments thus is that forming an observation set in meso- and macroeconomics on basis of the analogy method (a representative agent) is a critical process.

For (3) a consistent aggregation procedure has to be followed. This procedure is related to what in natural sciences is called up-scaling and down-scaling [3]. However, as noted by Costanza et al. [4] such an aggregation procedure is far from trivial in complex, non-linear discontinuous systems. Indeed, Forni and Lippi [8] argue that macroeconomic modelling and testing would receive a new impetus if a better balance were reached between micro theory, aggregation theory, and empirical research on the distribution of the micro parameters over the population. Consequently, more importance would be given to heterogeneity on a micro level.

In spatial economics, there are even more perplexing aggregation problems. Whereas census data are collected for essentially non-modifiable entities (people, households) they are reported for arbitrary and modifiable areal units (enumeration districts, local authorities etc). This is the crux of the modifiable areal unit problem: there are a large number of different spatial objects that can be defined and few, if any, non-modifiable units [9].

The conclusion from the discussion in this section on scales and aggregation is that building an observation set in time and space on a certain aggregation level is far from a simple process. More strongly, by making mistakes or misjudgements in the design of our observation set we make misjudgements in the understanding of the processes we wish to describe.

Before continuing a discussion on building our observation set, we will review a recent dispute in the discipline of ecology on scales. Given the definition of ecology in the next section we see a certain analogy with spatial economics. There is an identical problem in identifying aggregation levels in relation to an observation set in time and space.

3. SPACE AND AGGREGATION IN ECOLOGY

Ecology attempts to explain the relationship between living organisms and their surroundings. Ecology is about the distribution and abundance of different types of organisms over the face of the earth, and about the physical, chemical but especially the biological features and interactions that determine these distributions and abundances [10]. In ecology there are supposed to exist several dependent (bio) diversities at different aggregation (organisational) levels. Processes can for instance take place in the biosphere, but also on ecosystem, community, population and individual species level. (Note the analogy with micro, meso and macroeconomics.)

In ecology, space and time are linked. Ecological processes that operate over large areas also tend to operate over long time scales. Modern ecology has focused mainly on those scales where local communities and short time periods are studied [11]. Thus, processes are simulated at
short time scales and treated entirely as recursive; conse-
quently high time resolution models are adopted [12]. Second-
ly, ecologists are interested in long time horizons and espe-
cially the long-term implication of human action [13].

Spatial dynamics are extremely important in ecology. Be-
des the physical flows of matter, the spatial arrangement of
habitats or land cover affect all ecological processes such as
species diversity, natural assimilative capacity and nutri-
ent cycling. The spatial pattern of habitats or land cover, the
landscape pattern, is thus linked with all ecological processes. Furthermore, the size and shape of the patterns themselves depend on the scale on which they are described.

These notions have led to the development of hierarchy
theory [14], which states that the variation that is observed in
ecosystems depends on the scale over which we measure it,
both in time and in space. Within such a hierarchy we
observe:

- Processes,
- Flows,
- Interactions, and
- Rates (which characterise the speed of change in the
  system).

Variables and processes on lower level in the hierarchy are
considered as noise, whereas variables on higher level act as
constraints. Rates appear to be a kind of distinctive variable
in relation to hierarchies. “High” levels show slow rates and
“low” levels show fast rates.

The notions on hierarchy and scale as presented above
have been shaken somewhat by authors who discuss the
relation between level and scale. An important first
observation by O’Neill and King [14] is that hierarchies
are less evident than they look because moving across scales
the dominant processes may suddenly change and relation-
ships may completely disappear. Moreover, within an
ecological observation set, processes may be located at
different levels by finding breaks or discontinuities in the
data. Otherwise stated, discontinuities in the ecological data
may suggest a change in level of organisation. The question
that is being raised is whether these levels of organisation, as
extracted from empirical data, are the same as adopted in
traditional biological literature: organism, population, land-
scape, ecosystem, etc.

Significantly, ecologists admit that they have confused
the words scale and level [14]. This implies, for instance,
that the application of the word scale in ‘landscape scale’ is
wrong. Landscape is a level of organisation. There is a
relation between scale and level, but changing the scale of
observation changes the observation set. Consequently, the
hierarchical organisation can change or disappear. Allen [15]
takes an even harder position: Landscape is a “type” as the
researcher constructs it and it is thus an organisation level
that is not scalar. Type-based levels of organisation contrast
with scale based levels, which are rooted in observations.

Higher levels of observation are materially larger, whereas
levels of organisation cannot be assigned to any particular
spatiotemporal size. Consequently, landscape is a model, a
choice in an analytical dimension.

A second observation by O’Neill and King [14] is that
hierarchies, as established by ecological theory, are rather
arbitrary. The authors do, however, like to keep the idea of
hierarchies, but these concepts should be sustained by
observations and should not be merely heuristic in the sense
of explaining very special problems.

From the above experiences in ecology, we firstly infer
that in ecology time and spatial scales are connected.
Secondly, we conclude that in ecology there is a relation
between scale and aggregation level, and that changing the
scale of observation also changes the observation set. More
strongly, even the hierarchical organisation can change or
disappear.

Having gone through the general discussion on scales and
through the particular application in ecology the question
may arise: “How about scales in economics”? Is there a kind
of hierarchy in economics sustained by observations? Or
is the distinction in micro, meso and macro a type based
level characterisation of organisation, which is a rather
arbitrary decision, made by the scientific economic
community?

4. SPACE AND AGGREGATION IN ECONOMICS

Economics is concerned with human behaviour. It studies
the allocation of scarce resources to different means.
Producers aim to maximise profits and thus minimise costs,
while choosing a certain technology, where labour and
capital are combined. Consumers aim to optimise their
utility given their income and the relative prices of different
goods.

Economics is thus concerned with choice and value.
Three main levels along the scale of analytical interest are
distinguished, the micro, meso and macro level. Each level
of aggregation has its own theoretical content. Microeco-
nomics studies consumer and producer behaviour, meso-
economics focuses on sectors, while macroeconomics
focuses on aggregates, aggregate behaviour and government
policy.

Analogous to ecology, the processes, interactions, flows
and rates in economics distinguish organisational levels.
Higher levels have slower rates (e.g., inflation), and different
levels show different interactions and processes. In Figure 1
we give an example of processes and interactions for three
levels of analytical interest. In the figure, we bring in a
traditional ‘natural’ economic order to resemble traditional
thinking in micro, meso and macroeconomics. The grey part
in the figure represents a dynamic area, where interactions,
flows and rates are relevant. Outside the grey part processes
and analytical concepts are not relevant. Above a certain
organisational level information acts as constraint and below a level information is supposed to be noise. For example, on the micro level price setting formation is given. On the meso level inflation will act as a constraint whereas individual producer maximisation is noise. Horizontally seen, for sectoral agents, behaviour of individual consumers and producers is noise, whereas the behaviour of aggregate agents is given. Note that in this reasoning, there is no explicit reference to space.

Given the processes and the analytical scale domain we distinguish, what are the time and space dimensions in economics? Moreover, what is the observation set? Here we notice an important difference between economics and natural science in general and ecology in particular. Economic theory is based on abstract social units. It is inter alia focussed on utility optimisation of households and price formation in markets. The consequence is that economic theory is not spatially explicit in terms of spatial resolution. Economists might research yearly changes in expenditure on housing for households in the Netherlands as influenced by changes in female labour force participation over a period of ten years. Or they might investigate changes in the quantity of steel sold by industry in the year 1999 in Portugal as a function of changes in GNP in Portugal. Or changes in Gross Regional Product in a time series for states in the USA. Or yearly changes in the demand for water in the UK because of privatisation.

To some degree, spatial extent and spatial resolution seem to coincide. Economic research on consumer and producer behaviour on the basis of individual data is not performed on or restricted to a local or regional level. In addition, sectoral observations can be collected at a local, regional and national level.

Does this imply that space does not matter in economics? No, space does matter; however, space is generally translated into transportation costs, and thus into prices, by the one-dimensional concept of distance. Thus, spatial differences come back in another fashion. But, note again that the resolution of space is not important.

Thus, we conclude that spatial resolution, as part of the concept of spatial scale is not taken into account in economics. Economic theory is on abstract social units. Concerning the related problem of aggregation, it is our observation that the ‘traditional’ division in micro, meso and macroeconomics does not have an explicit spatial connotation. As a corollary, the organisational division in economics in micro, meso and macro is a rather abstract distinction, a type based construct, as ecologists would call it.

4.1. Does Spatial Resolution Matter in Scientific Disciplines that Deal with Space?

Above we reached the conclusion that distance as a one-dimensional concept of space does matter, but we did not investigate two- and three-dimensional spatial issues in economics. Are there any applications in spatial sciences where spatial resolution is of importance? Of course there are; in agricultural economics crop results depend on technology as well as on soil conditions, climate and hydrology. In regional economics, inter alia, locational decisions made by households and by firms are spatially dependent. In land markets, land use and land cover change are at stake. Moreover, in (economic) geography, we are interested in differences between regions and countries and we try to understand the formation of patterns.

Yet, we are not impressed how spatial resolution is introduced in these disciplines. We will illustrate this statement by presenting standard theories on emerging spatial structures in regional economics and in geography. The evident example for emerging structures is that of location behaviour of firms and households in producing urban spatial patterns. We will evaluate how spatial resolution is handled in location theory and discuss how defining the problem in terms of spatial resolution might contribute to a better understanding of the phenomenon of emerging spatial patterns.

5. SPATIAL RESOLUTION AND EMERGING PATTERNS OF LOCATION BEHAVIOUR

In presenting theories on location behaviour, we will split between geography and regional economics as they approach location behaviour from different angles.

5.1. Geography

Geography focuses on where things are located and why. Location, maps and distribution help to answer the where question. The why question is addressed by researching the ability of people to adjust to their physical environment. Scale is of utmost importance in geography. Spatial scale (resolution and extent) is recognised in geography as the main mechanism whereby patterns can
be analysed and explained. Geographical Information Systems (GIS) is the essential tool to solve this why and where question [18]. It is our belief, however, that with help of GIS, geographers aim at merely a description of the adjustment process that comes with location behaviour. This means that the choice for a certain resolution is not decisive in explaining emergent patterns.

A good example is the work on land use dynamics. The high spatial resolution model of urban land-use dynamics developed by White and Engelen [19] aims to capture the spatial complexity of urban and regional areas, by making use of two basic techniques, cellular automata and GIS. Cellular automata use a set of transition rules that govern the local behaviour at each cell with respect to the cell’s neighbours and its own characteristics. It offers a means to study emergent global behaviour in systems where only local processes are understood.

By applying this technique of cellular automata, GIS is converted into a dynamic tool [20, 21]. The model of White and Engelen [19] for instance distinguishes between two levels, a macro and a micro level. The macro level includes a modelling framework, which integrates several component submodels representing the natural, social and economic subsystems. The micro level is developed on a cellular array in which the land use changes are calculated through transition rules [19, 22].

However, a drawback of cellular automata in general is the fact that the transition rules are not necessarily reproducible with an objective empirical methodology. The system performance depends highly on the skill of the modeller. Secondly, transition rules do not change during the course of a simulation and hence may be of limited importance because changes of landscape rarely are constant over time. Finally, a most important drawback is the difficulty of incorporating micro-economic behaviour. Geographers are relatively poor in formulating theories explaining behaviour in space. Indeed, Openshaw and Abrahart [23: p. 380] argue that human systems modelling is going to become an unavoidable area of considerable practical importance. People are too important to ignore. Currently, we have no good or even tolerably poor models of the behaviour of people.’

Our conclusion is that geographers, although they combine high spatial resolutions with GIS, do not succeed in explaining emergent location behaviour.¹

5.2. Regional Economics

In elucidating the role of spatial resolution in regional economics, we again discuss location theory and the appearance of spatial patterns and structures.

¹One of the reasons might be the intrinsic data problem geographers and economists have regarding the Modifiable Areal Unit Problem, as discussed above.

5.2.1. Location Theory and Spatial Patterns

In location theory, a distinction is made between location theories of the firm, location theories of households and the interaction between the two. In the literature on location of the firm, transportation costs (as an estimate of the notion of space and distance) are central to location choice. Here we may distinguish between models that assume a demand for goods and services continuously dispersed in space and models where demand is concentrated in one point [24]. The first type of models suggests Christaller [25] geographical patterns of firm location that are hierarchical ordered, whereas the second type presents structures that are dependent on the (point) location of markets and resources. Anas et al. [26] note that defining clusters in space is not so easy. The distinction between an organized system of subcenters and apparently unorganized urban sprawl depends very much on the spatial scale of observation. Here we find one of the very few remarks economists devote to the problem of spatial resolution.

For the explanation of agrarian land use the famous Von Thünen model [27] is important. The model has been criticised for the assumptions that production takes place around an isolated market and that soils are of constant fertility. Nevertheless his distance-cost relationship has become the basis of urban location theory. Some claim that Von Thünen’s approach has dominated the thinking about location exactly because of its simplicity and predictive ability [28].

In using an urban location model linked to Von Thünen’s theory, Alonso [29] developed a model that can be regarded as the basis for household location choice. Alonso’s approach is based on the principle that rents decrease outward from the centre of a city (lower revenue, higher operating costs and transportation costs). Rent gradients consist of a series of bid-rents, which compensate for falling revenue and higher operating costs. Different land uses have different rent gradients, the use with the highest gradient prevails. Competitive bidding (perfect information) determines patterns of rent and allocates specific sites between users to ensure that the highest and best use is obtained. Land is used in the most appropriate way and profit is maximised.

Criticisms to Alonso’s model are first of all that in reality information is incomplete; thus there is an imperfect market. He also fails to take into account the distinctive nature of buildings and their use, which are not easily changed (lock-in). Other points are the heterogeneity of property, public sector land and spillover effects of other uses.

The Alonso model and the literature based on it are characterised by other simplistic assumptions. Employment is centralised in the Central Business District (CBD), there is a dense radial road system and all households have the same taste [30]. Moreover, the model is static. Some of these assumptions have been removed [31–35], but the theories remain rather general and abstract.
Note that we did not refer to spatial resolution in interpreting Alonso type urban location models. By only applying distance as the one-dimensional concept of space, the location theories of Von Thünen and Alonso are not able to explain the complex spatial structures that we encounter. Anas et al. [26] discuss this problem by referring to alternative assumptions for the Pareto equilibrium of monocentric cities that make a uniform distribution unstable. Spatial inhomogeneities, internal scale economies, external scale economies and imperfect competition create polycentric agglomerations. In regional economic theory external economies of scale, agglomeration economies, or localisation economies are used as theoretical constructs explaining why firms locate in each other’s vicinity to arrive at increasing returns to scale [36, 37]. Business firms locate in each other’s vicinity in order to gain from the attractiveness of companies of the same type activity, but also to gain from the general atmosphere in such a region. These notions are regarded as a major contribution to economic theory [1]. However, the theoretical constructs remain more or less a black box failing to explain the occurrence of spatial structures and patterns on a high level of spatial resolution.

It is here that non-economic explanations have much more to offer in order to interpret the black box [2, 38]. Self-organising criticality and synergetics produce organised structures with which polycentric cities might be better explained. In these theories interactions between individual actors on a high level of spatial resolution give rise to meso and macro spatial structures. Anas et al. [26] therefore plea for an adaptation of standard economic theory. An explanation on a high level of spatial resolution is available where traditional economic theories seem to fail. Interaction between individual actors on high spatial resolution is, however, not in the heart of regional economic theory.

Before concluding that in traditional economics spatial scale is not taken into account, we would like to devote attention to a special branch in spatial economics that is related to special techniques caused by the features of space: spatial econometrics. In the same way as we discussed the combination of GIS and geography in the first section, it might be the case that the combination of spatial econometrics and spatial economics produces a powerful explanation for spatial behaviour.

5.2.2. Spatial Econometrics
Spatial econometrics is concerned with techniques that deal with the peculiarities caused by space [39]. It deals with spatial dependencies and with spatial heterogeneity. According to Anselin and Florax [39] spatial dependency is relevant in two cases:

- In case of a spatial structure underlying spatial correlation, where the main interest is the spatial interaction behind the variable of interest.
- Spatial dependence between ignored variables in the model as reflected in the error terms.

In neglecting these cases the estimation of an a priori specified model, based on observations for a finite set of spatial units, will cause a number of problems [40]:

- The modifiable areal unit problem [9], which concerns the aggregation of observations over space.
- Border or edge problems, pertaining to the problem that inferences are based on a finite set of observations whereas the spatial process extends to spatial units not represented in the data set.
- Specification of the spatial interaction structure, which is typically represented by a spatial weight matrix.
- Testing for spatial effects by means of spatial association or correlation.
- Estimation of spatial models for which adjusted estimators are needed.

The underpinning of spatial dependencies and heterogeneities in regional economics is based on the same ideas as we try to develop in this paper. Anselin and Florax [39: p. 5] state that there is a “renewed interest for the role of space and spatial interaction in social science theory. In mainstream economic theory this is reflected in the interest in the new economic geography.” It is our judgement, however, that spatial econometrics is mainly interested in the statistical and econometric problems of spatial dependencies and not so much in the extension of the theory of economic behaviour with a spatial context and component. It is our view that spatial complexity should acknowledge space as a context for decisions made by individual households and firms.

We conclude that the combination of spatial econometrics and spatial economics does not produce an additional explanation for spatial behaviour.

6. CONCLUSION
In discussing scaling and aggregation in (regional) economics, it is our first observation that the construction of an observation set may have strong limitations in relation to the spatial theoretical notions that are assumed. Secondly, the ‘traditional’ division between micro, meso and macroeconomics does not seem have an explicit spatial connotation.

Thirdly, in standard economic theory spatial extent and spatial resolution seem to coincide. Considering spatial resolution and human behaviour in regional economic theory, it seems as though there is a trade-off between two topics. Certain types of models are capable of capturing the spatial complexity of urban and regional areas, for instance, 

2But there are exceptions: Dubin [41] presents a wonderful paper of a logit model incorporating spatial dependencies on a GIS grid base!
by using cellular automata. These models have a high spatial resolution, but do not include choices made by individuals. On the other hand, current static and dynamic location models on the other hand do not guarantee a high spatial resolution. It is here that future researchers should concentrate their efforts.

REFERENCES