On the units of geographical economics

Joshua Olsen *

School of Geographical Sciences, University of Bristol, University Road, Bristol BS8 1SS, UK

Received 2 October 2000; in revised form 4 September 2001

Abstract

The new ‘geographical economics’ introduced by Paul Krugman and others has been dismissed by geographers for a number of reasons, ranging from (mis)treatment of externalities to reliance on formal modelling. These criticisms may be appropriate, but they miss some underlying unit assumptions that cause geographical economics to deal poorly with issues of scale and space. The theories produced by Krugman and others rely on the economic units of the firm, industry and economy for analysis. In contrast, geographers often use spatial units which not only reveal different processes, but can also delimit the economic units. Acknowledgement of the difference in the assumed units of analysis between geographical economics and economic geography is necessary before any useful discourse between the two fields can occur. © 2002 Elsevier Science Ltd. All rights reserved.

Keywords: Geographical economics; New economic geography; Paul Krugman; Units; Agglomeration

1. Introduction

In 1991, Paul Krugman published a short text and series of articles on what he termed ‘geographical economics’. In his own words, the point of these incursions was to prove first that ‘space matters’ to the study of economic activity (Krugman, 1991a, p. 8), and second that, ‘it is now time to attempt to incorporate the insights of the long but informal tradition in this area into formal models’ (Krugman, 1991b, p. 484). Since 1991, Krugman and his followers have elaborated upon his original model and the reasons that this new geographical economics is superior to other explanations of the spatial pattern of economic activity. The latest summary of Krugman’s approach appears in The Spatial Economy: Cities, Regions, and International Trade, which he co-authored with Masahisa Fujita and Anthony Venables (Fujita et al., 1999). This text not only links the current work concerning the location of economic activity to Krugman’s earlier work on ‘new trade theory’ (see also Venables, 1998), but it also represents a refinement of the field’s basic formal models (cf. Krugman, 1979, 1980, 1981, 1986, 1990).

Nevertheless, the general approach adopted in this most recent work is broadly similar to that of the last decade. First, geographical economics emphasises the importance of historical accident in determining the location of economic activity. This concept is represented in the models by a randomly uneven distribution of resources. Second, Krugman and his colleagues posit that cumulative causation and path dependence can result in an uneven distribution of activity among regions with approximately similar initial endowments. Essentially, successive firms entering a location make it a more attractive location for other firms. Third, the new geographical economics seeks to demonstrate how increasing returns, linkages and trade costs work to both maintain spatial order and create catastrophic changes once bifurcation points are reached. Fourth, these same factors, and especially resultant scale economies, are used to explain the evolution of a multi-city structure and regional specialisation. Lastly, throughout the development of all these ideas, the new geographical economics has remained inseparably wedded to the use of formal mathematical models to explain spatial and economic phenomena. The combination of this strict
formalism with the details outlined above, especially the awareness of increasing returns, is thought to make the field superior to previous attempts to address spatial and economic considerations made by Chisholm (1990), Christaller (1933), Henderson (1974, 1980, 1988), Isard (1956), Losch (1940), Pred (1966), and von Thunen (1826). Although only dismissively mentioned, it seems safe to assume that Krugman and his fellows also consider more recent economic geography to be inferior.

In a series of reviews and critical assessments, economic geographers have attacked the new geographical economics on a number of grounds (see Boddy, 1999; Dymski, 1996; Hoare, 1992; Johnston, 1992; Martin, 1999; Martin and Sunley, 1996). These critiques can be divided into three main categories. First, geographers have argued that Krugman and his followers ignore key externalities. For instance, Fujita et al. (1999) readily admit that of Marshall’s trinity of externalities, they only fully incorporate the concept of market size and market access effects, and only then by a narrow model of linkages represented by transportation costs. The externality of labour market pooling is mentioned, but not thoroughly addressed by the formal models, and the third Marshalian externality of knowledge spillovers is ignored almost completely. Geographers have objected to this omission, noting that knowledge transfers are crucial to understanding how conventions and innovations form and spread in the economy. The importance of knowledge spillovers is especially crucial for knowledge-based industries such as financial services (see: Leyshon and Thrift, 1997; Thrift, 1994) and high-technology businesses, such as those found in Silicon Valley and Boston’s Route 128 that Krugman highlights (see also Audretsch, 1998). Second, geographers have argued that the minimal treatment of externalities that geographical economics does offer deals poorly with the issue of spatial scale (see especially Boddy, 1999). Krugman and other geographical economists use the same model to explain agglomeration at the metropolitan level, regional level and international level, and they are never clear as to which mechanisms are more important at which spatial scales. Lastly, geographers have attacked the reliance of Krugman and his followers on formal mathematical modelling. This approach to understanding the social world relies on dubious assumptions and vast simplifications that the mainstream of economic geography has not used for several decades.

In general though, the criticisms levelled by geographers have addressed the new geographical economics on its own terms by noting the limitations of the specific model and formulaic approach. This strategy for explaining the faults of geographical economics has poignantly observed that Krugman’s theory deals poorly with scale and that his reliance on formal models is a poor method for understanding the complexity of space, but it has missed the underlying assumptions that lead to such dilemmas within geographical economics. In this paper, I argue that the failures of Krugman and others to effectively address the issues of scale and space stem from the traditional unit assumptions that underlie their work and all of economics. In the attempts of the new geographical economics to explore the economic there are three crucial units: the firm, the industry and the economy. These units are both a focus of analysis and an explanatory factor. In the context of this paper, they are entities whose actions are studied, but not for their own end. Rather, researchers use units to explain something greater. Here that greater concept is the economic - the world of trade for profit.

In exploring the role of these units, I first show how Krugman’s model is based on the assumption that the firm, industry and economy are the correct units for analysis. I then argue that it is this assumption which opens the analysis presented by Krugman and his followers to the scale problems and the mistreatment of space decried by geographers. Finally, in the last section I explore the different unit assumptions that can be seen as driving more accepted theories within economic geography, particularly those represented by Storper (1997). The disparity in unit assumptions between geographical economics and economic geography may prevent any future rapprochement of these two means of understanding economic activity in space.

2. The unit assumptions in Krugman’s model

Mainstream neo-classical economics has long focused its efforts on understanding the actions, interactions and existence of firms, industries and economies. In addition to this trinity, the rational independent actor or household might be considered a fourth unit traditionally used in economics. Krugman does occasionally refer to this unit, but with his emphasis on industrial location, he generally ignores the household in favour of the other three. The household is always in the background of his models, but never in the foreground of the accompanying explanation or analysis. It is a key factor in the models themselves in the form of labourers or ‘peasants’ that function both as an input factor for firms and also as a market for firm output, but it is never the focus of analysis itself. Geographical economics is not concerned with the location of people. Presumably, this would be dismissed as sociology, or at best demographics. Instead, it is concerned with firms and their location

---

2 As one of the anonymous reviewers noted, the firm, industry and economy are more exactly the three traditional units of mainstream industrial-organisation economics. This clarification is appropriate, since it is this brand of economics with which Krugman is concerned in his predominantly supply-side economics work.
decisions. When the equations are computed, the human actor makes a cameo appearance, but ultimately serves only as an index for firms.

Thus, it is the three units of the firm, industry and economy upon which Krugman and his economist colleagues rely. This is not to say that Krugman and his followers explicitly advocate the use of these units, but the units implicitly underlie all of their new work and treatment of others work. Even before the formal modelling in geographical economics begins, it is assumed that these units will be the basis for analysis, and the resultant model demonstrates the outcome of this initial bias. Indeed, although Krugman and others write a great deal about concepts like the region or the city, their work involves at best a conflation between economic and spatial units.

From the very first assumption presented in Krugman’s basic model, it is clear that the firm, industry and economy will form the basis for his analysis of economic activity in space. For example, instead of assuming a spatial unit for his domain of inquiry, Krugman (1991a, p. 15) assumes the existence of an economy. It is interactions within this simple economy that his model will attempt to explain. An economy need not be aspatial, but in this case the only spatial units in sight at this early stage are two nebulous ‘regions’ that turn out to be the mathematical equivalent of two points on a line (Krugman, 1993a). This ‘space’ plays no role itself, and only contains other economic units. Krugman’s simple economy essentially consists of two industries, agriculture and manufacturing, with the former serving as an analytical counterpart for the latter (Krugman, 1995a). Undifferentiated agriculture is practised in each of the regions by immobile ‘peasants’ (Krugman, 1991a). Manufacturing, on the other hand, is highly mobile and employs equally mobile workers which supposedly produce differentiated goods. As Fujita et al. (1999, p. 45) note: ‘Of course, the label “agriculture” need not always be interpreted literally; the sector’s defining characteristic is that it is the “residual”, perfectly competitive sector that is the counterpart to the action taking place in the increasing-returns, imperfectly competitive manufacturing sector’. In fact, Fujita et al. (1999) develop the conception of agriculture to a greater extent than any of the previous models, but as this statement shows it still serves as a counterpart to manufacturing in the simple economy that is being examined.

In discussing manufacturing, Krugman introduces the unit of industry. This unit operates on two levels in Krugman’s models. First, on a broad level he uses this unit as an example of the regional agglomerations of industrial activity that characterise Europe and the United States. This meaning of industry comes to the fore when he refers to the generalised activity of manufacturing. On a second level, he compares this concentration of manufacturing activity to clusters of specific industries such as high technology in Silicon Valley. In particular, this second level of meaning shows that Krugman sees the manufacturing component of his model as synonymous with the traditional unit of a single industry in economics. Krugman’s broad category of manufacturers may produce differentiated products, but in their generalised activity they are all essentially part of one unified sector.

The third unit upon which geographical economics relies is the firm. This is the decision-making unit that the models of Krugman and his followers are meant to track. This goal is made obvious in numerous statements, such as: ‘The intuition behind the approach in this paper is simple: firms that have an incentive to concentrate production at a limited number of locations prefer, other things equal, to choose locations with good access to markets’ (Krugman, 1993a, p. 131). Thus, in support of this intuition the second series of equations in the basic model for geographical economics are dedicated to describing the motion of firms. This emphasis on the firm outweighs other potential explanations for agglomeration like labour movement. For example, in the section of Krugman’s paper with the heading ‘Necessary Conditions for Manufacturing Concentration’, the discursive emphasis is all on firm strategy and only tangentially on worker preferences (cf. Krugman, 1991b). This emphasis on the firm continues in later publications, and is at the heart of the field’s acceptance of both internal and external scale economies. For instance, in a 1993 paper Krugman makes the bald statement: ‘The formation of cities is assumed to be driven by economies of scale at the level of the individual firm’ (Krugman, 1993b, p. 293). In fact, the focus must be on the firm before one can even grasp how Dixit and Stiglitz’s formalisation of monopolistic competition is incorporated into the model (cf. Dixit and Stiglitz, 1977). In this way, the firm compliments the two units discussed above as the key to Krugman’s understanding of the location of economic activity.

Krugman’s entire model is conceived around these three units of economic analysis. He sets out to study a hypothetical economy, in which there is a dominant manufacturing industry. Spatial agglomeration is then basically a function of decisions made by individual firms. These three units need not be aspatial, but traditional economics has always viewed them as such. Space, on the extremely rare occasions when it was addressed by economists, was incidental. Geographical economics claims to want to look at space, but it does so with the same units that economists have always used. For all their concern with space, Krugman and others are certainly not questioning this traditional unit assumption. This is not to say that they do not ever study spatial units, but in their main body of work on industrial location space is considered to be just the result of
actions taken by the economic units. It is the three units described above that geographical economists manipulate in their models, and the result is just new combinations of these units. At best, a certain conflation of space with economic units is apparent in Krugman’s writings, especially Geography and Trade (Krugman, 1991a). This text is explicitly broken up into a discussion of certain spatial scales, but each ‘scale’ is then conflated with certain traditional economic units. First, in his chapter on localisation in Geography and Trade, Krugman (1991a, ‘Chapter 2’) talks about the firm and the specific industry, with statements like, ‘Imagine for a moment that there is an industry that consists of just two firms’ (Krugman, 1991a, p. 38), or: ‘So, it is not surprising that localisation of industry, the emergence of sharply distinct industrial characters for particular cities or districts, became a striking character of the American economy’ (Krugman, 1991a, p. 52). Essentially, for Krugman’s geographical economics a city is nothing more than a collection of firms minimizing transport costs and maximizing scale economies. As a result of such conflation, one gets the impression that a cluster of specific firms, or at most a very specialised industry like rubber manufacture, is synonymous with a city. Furthermore, in explaining agglomeration, the models of geographical economics rely on manufacturing firms producing differentiated products. It is actually the uniqueness of the products produced by the firms that allows them to be spatially separated. As Krugman (1995a, p. 61) explains it, this firm-based conception justifies the assumption made by Weber (1909) that each good is produced at a single location. In this way, the unit of the firm is a necessary assumption to explaining the spatial dynamics of the model.

Elsewhere in Geography and Trade, Krugman (1991a) discusses the region with an emphasis on the broad meaning of industry as the unit of choice. Thus, a region is defined as the area where the hypothetical ‘manufacturing’ industry locates, or where industrial activity in general can be observed, as in the US rustbelt. Indeed, in ‘Chapter 11’ of The Spatial Economy Fujita et al. (1999) attempt to use their model to show how an urban hierarchy can emerge from a decentralised market process. They accomplish this task by showing that differences among industries in scale economies and transportation costs define a ranking of industries. This ranking is then used to generate a hierarchy of city types. As a result, the region containing this urban system is merely a set of industries with their simple differences in scale and trade factors.

Finally, geographical economics also widens its scope to address the national scale (for example Krugman, 1991a, ‘Chapter 3’). However, it is really the economy, consisting of both agricultural and manufacturing activities, that is the unit Krugman considers necessary for understanding industrial concentration at and beyond the scale of the nation. Thus there is nothing, ‘of inherent economic importance in drawing a line on the ground and calling the land on either side two different countries’ (Krugman, 1991a, pp. 71–72; see also Fujita et al., 1999, pp. 239–240). This is certainly true if the ‘ground’ being discussed is a straight line derived from a formula, but in the real world different nations have different cultures that cause their economies to operate in unique ways. Indeed, Krugman does not really consider nations to be spatial entities at all, just arbitrarily defined areas of economic policy (see Krugman, 1996).

In addition, the unit assumptions of geographical economics also explain Krugman’s treatment of previous attempts to theorise the spatial location of economic activity (cf. Krugman, 1995a). In particular, Krugman criticises the classical regional science and central-place theory works of Losch (1940) and Christaller (1933). In essence, these works are about settlement patterns and urban systems – concepts that Krugman finds difficult to immediately reconcile with the traditional units of economics. Nevertheless, these central-place theories are internally consistent. Krugman will admit as much, but he concludes that this earlier work must have been about geometry, not economics, because it was ‘exasperatingly blurry about who was making what decisions’ (Krugman, 1995a, p. 39). Krugman is less critical of the modern regional science of Isard (1956), who introduces the unit of the firm to the older German tradition (Krugman, 1995a, pp. 55–57). Indeed, Krugman admits that the work of Losch (1940) and Christaller (1933) makes intuitive sense, but only after he has switched their focus away from spatial units to a simple economy with two industries and firms producing differentiated products (Krugman, 1995a, pp. 59–64). This method of translation is also apparent in the more recent work in geographical economics: ‘Unfortunately, as soon as one begins to think hard about central-place theory one realises that it does not quite hang together as an economic model. In economic modelling we try to show how a phenomenon emerges from the interaction of decisions by individual families or firms’ (Fujita et al., 1999, p. 27). Thus, the treatment of classical regional science by geographical economists is one more indication of their reliance on a limited set of units.

In sum, in their discussion of earlier attempts to explain spatial phenomena, or in the new models that they propose, it is clear that while proponents of geographical economics do mention different spatial units in their analysis, they tend to either conflate them with, or

---

3 In other publications, Krugman conflates his economic units with spatial units in different ways. For example, in an essay on the difference between late 19th century Chicago and late 20th century Los Angeles, Krugman (1995b) uses each city as a metaphor for the national economy – a unit to which he can comfortably apply his type of analysis.
subsume them under, the comfortably-economic units of firm, industry and economy. The new geographical economics may be touted as ‘the fourth wave of the increasing-returns revolution in economics’ (Fujita et al., 1999, p. 3), but it remains conventional in its focus upon the three traditional units of mainstream economics.

3. The criticisms of Krugman’s model

Geographers have levelled a great range of criticisms at Krugman’s model (for a review, see Boddy, 1999), which are largely valid in their own terms. They fail, however, to address the implicit reliance of geographical economics on the units of analysis outlined above. This oversight appears in three principal categories of faults found with the geographical turn in economics: the intra-scale problem that arises from the failure of Krugman and his followers to account for certain externalities, the inter-scale problem caused by their spatially ambiguous treatment of the externalities that they do address, and the reliance of geographical economics on formal modelling.

First, as geographers have readily noted, even when only considering one scale of economic activity, there are still spatial factors at work that Krugman and his followers do not acknowledge. Specifically, geographical economics avoids the Marshallian externality derived from technology and knowledge spillovers in favour of ‘pecuniary’ ones derived from market-size effects. Krugman (1991a, p. 53) finds that the latter fit more easily into his increasing returns-based model, while the former is more suitable for sociological studies. Economic geographers have roundly criticised this avoidance of a factor that underpins much of their work. For example, Martin and Sunley (1996, p. 273) declare Krugman’s avoidance of externalities that he cannot model to be ‘too restrictive’, and instead argue that invisible spillovers ‘are an integral part of social conventions and meanings, and their formation should be an important area for regional research’.

Krugman’s critics would agree with him that technological and knowledge-based externalities do not fit into his model, but see this as a fundamental problem with his approach rather than something to be simply noted. They do not, however, explore the underlying reason for this mismatch. Fundamentally, the new geographical economics cannot handle this type of externality because it does not mesh with economic unit assumptions. Unlike the Marshallian externalities of labour pooling, which occurs at the level of industry, or intermediate inputs, which involve market interactions between firms, a knowledge spillover happens in an amorphous space between firms and industries. It cannot be understood with the firm as the basic unit, because it happens with only scant reference to a firm’s decisions to purchase an intermediate product or hire more workers. Nor can it be understood at the level of industry, because knowledge sharing often operates between industries, a fact noted by keen observers such as Jacobs (1969) and demonstrated empirically by other economists (Audretsch, 1998). The unit of the economy is also irrelevant since the localising powers of the technology externality would be lost at such a level. In short, the firm, industry and economy do not easily encompass social conventions and meanings. The avoidance by Krugman and his colleagues of non-pecuniary externalities and the criticism that he reaps as a result can thus be seen as stemming from his over-reliance on the three units.

Furthermore, geographers have observed that geographical economics avoids externalities associated with human agency. I have already noted that the household plays no significant role in geographical economics models, but there is also no room in the models for decisions of labour or business leaders that are not made completely rationally. For instance, the fact that distance may be a qualitative concept for most individuals (see Harries, 1997), and that this may affect labour movement, linkages and agglomeration does not enter into the equations. Beyond this, in the discussion of European integration and the agglomeration that should result, Krugman (1991a) does not acknowledge the fact that language and cultural barriers may prevent free labour movement even if it is technically allowed by government policies. Neither do Krugman and his followers admit the possibility that those in charge of ensuring economic rationality may not act in accordance with their best economic interests. Therefore, geographical economics does not consider firm leaders who may move businesses closer to their homes (Garreau, 1991, pp. 92–94; Whyte, 1988, pp. 287–290; for the same phenomenon in newly formed businesses see Cooper, 1998; Keeble and Gould, 1985), or decide to forego economies of scale in favour of a hedge against labour militancy (Harvey, 1996; Peet, 1984). Furthermore, along with Krugman’s disavowal of psychological externalities comes an ignorance of the various social externalities that may have shaped such mindsets. Thus, Dymski (1996, p. 447) can critique Krugman’s model as ‘ill-equipped to handling complexity in the sense of overlapping social relations and constraints’.

This criticism can also be explained as a unit problem. Human agency has no place in the new geographical economics because it has no place in the units that are at the base of the models promulgated by Krugman and his followers. This in part explains Krugman’s (1995a) frequent attacks on attempts to find alternatives to the rational, profit-maximizing firm such as those proposed by Galbraith (1987, 1992, 1993). While geographical economics wholly embraces imperfect competition in the
form of increasing returns and dis-equilibrium, the units to which these apply are still expected to operate rationally. For instance, firms are the decision making units in the model, not shareholders or CEOs, and it would be difficult to introduce non-rational firms without acknowledging other types of units, such as human actors. This mindset also leads to a reliance on other neo-classical economic units such as industries that can be modelled with simple factor assumptions and economies represented by lines or ‘racetracks’, neither of which are likely to display the vagaries of human agency.

Second, geographers have noted that the theories of Krugman and the other geographical economists encounter problems between scales. This reflects the fact that geographical economics is generally unclear about which factors operate at which geographical levels. As Boddy (1999) observes, Krugman seems to imply that increasing returns and transaction costs apply at the regional level, while the pecuniary externalities mentioned above work at the local metropolitan level. Similarly, Martin (1999, p. 78) attacks the cavalier attitude of the geographical economics school that the same model can be applied to such vastly different scales as intra-urban neighbourhoods and national industrial clustering. If anything, this problem appears more pronounced in Fujita et al. (1999), where international trade is added to the discussion of urban and regional phenomena.

Indeed, the new geographical economics is rather ambiguous when it comes to geographical scale, and once again this failing can be explained in terms of unit assumptions. As was shown above, Krugman and others often conflate the economic units of the firm, industry and economy with certain spatial units. By doing so, they can make it appear that their model applies to various patterns of agglomeration, but this effect is illusionary and falls apart when any attempt is made to move between spatial scales. In effect, an inter-scale movement requires a shift away from the conventional economic units. In other words, it is not that pecuniary externalities only work at the city level. It is, rather, that labour pooling and intermediate goods, as they function in the formulas, only apply to firms or a specific industry. Any shift upward to the region or nation will require a shift to a unit like ‘the economy’, where these externalities are meaningless. As already discussed, the so-called regions or nations in geographical economics consist of both agriculture and manufacturing, but the former, at least in the models, does not require intermediate goods, and the latter is the only piece that benefits from a concentrated group of workers. In essence, the model promoted by Krugman and others has inter-scale problems because it does not acknowledge spatial units as the basis of its analysis. Instead, it has its foundation in the three conventional units of economic analysis, which are largely scale-independent.

In fact, it is possible to go further than inter-scale or intra-scale problems and claim that the new geographical economics fails to adequately address space. In particular, this critique is directed at the use of formal modelling to explain complex spatial phenomena. Boddy (1999, p. 834) calls this ‘rigour mortis’ and notes: ‘From the perspective of economic geography… (Krugman’s models) do not appear to have made much of an inroad into the complexities of real-world processes’. Such complexities are what geographers believe is of spatial interest. As Martin (1999, p. 78) observes: ‘The spatial agglomeration models may well predict that, under specific assumptions, industrial localisation and specialisation will occur, but they are unable to tell us where it actually occurs, or why in particular places and not in others’ (original emphasis). Krugman’s writings in particular tend to acknowledge chance occurrence and QWERTY-type path dependence in nice vignettes, but the reality of the models is oblivious to local context. This is a failing that has only been mildly corrected in his recent writings (see Krugman, 1998a and Krugman, 1998b). In short, the basic proposition of geographical economics that a series of simple equations can explain spatial phenomena causes the theory to miss the full import of space and place to economic activity.

Even this broad critique can be explained by unit assumptions. If Krugman and his followers completely ignore the meaning of space in their over-reliance on formal models, they do so because they try to explain all processes in terms of the firm, industry or economy. In Development, Geography, and Economic Theory, Krugman (1995a) makes this intention clear by stressing that formal models are indispensable because they explain the microeconomic reasons for observed patterns. On the surface, the microeconomic reasons that Krugman refers to are increasing returns. However, one must remember that this mechanism is closely tied to the unit of the firm. For instance, in explaining the lack of space in formal economics by a lack of rigorous models, Krugman (Krugman, 1995a, p. 35) writes: ‘In spatial economics..., you really cannot get started at all without finding a way to deal with scale economies and oligopolistic firms’. Krugman may claim that his style of modelling is meant to explain basic processes, but the processes must in turn be attached to one of the basic units of economics. Thus, even seemingly rigorous models like those produced by central-place theorists and urban economists are dismissed by Krugman because they attempt to use the city, urban market area or agricultural hinterland as the units for analysis. This is far from stating a cause-and-effect relationship in which using the three units of mainstream economics results in a researcher reliance on mathematical modelling, but the two are closely intertwined. If the new geographical economics used space in formal models, then geographers could approach the field in a similar manner to
which they approach regional science. However, geographical economics relies on units that confuse the issue by seeming more appropriate to quantification. In fact, Fujita et al. (1999, p. 49) readily admit that qualities of space are deliberately manipulated to keep units such as ‘the economy’ in line with the model: ‘Depending on what we are trying to model, it is sometimes convenient to think of the economy as consisting of a finite set of locations (regions or countries), sometimes to think of it as spread across a continuous space’. If Krugman and the others who promote a geographical turn in economics miss the true importance of space, it is because space is not a unit that they will accept as the focus for analysis.

4. Economic geography and its units

Thus far, this paper has shown how geographical economics is based on certain unit assumptions. These assumptions underlie the models of Krugman and others, and can be seen as one important reason for geography’s critiques of the new geographical economics. However, the question of what other units might be used for understanding the location of economic activity has only been hinted at. In order to elaborate an alternative, it is necessary to examine some of the current work of economic geographers. In particular, I will concentrate this discussion on the recent work of Michael Storper, particularly _The Regional World_ (Storper, 1997). Although it would be naïve to declare Storper’s work to be the embodiment of economic geography, it does coherently bring together some of the dominant themes in the subject, including the institutional economics of Piore and Sabel (1984); (see also Amin and Thrift, 1994), the ‘California school’ of Scott (1986, 1988a,b, 1993a) and Storper and Scott (1989), and the concentration on high technology promulgated by Malecki (1984), Markusen et al. (1986) and Aydolet and Keeble (1988) amongst others (for strands brought together in Storper and Walker (1989), see Smith, 1992; Wood, 1993). In addition, one empirical study has already found Storper’s (1995, 1997) theory of ‘untraded interdependencies’ to be a better explanation for industrial agglomeration than the Krugman-inspired models of geographical economics (Pinch and Henry, 1999a; see also Henry et al., 1996).

It should first be noted that in terms of both problem and processes, the basic goal of economic geographers, as represented by Storper (1997) and others, does not differ greatly from geographical economics. In terms of the conceptual problem, just as Krugman (1991a, p. 1) sets out to explain ‘the location of production in space’, Storper’s efforts are dedicated toward constructing an analytical explanation for the locational characteristics of economic development. Furthermore, geographical economists and economic geographers are not as distant in their understanding of economic mechanisms as one might think. Both establish micro-foundations for a macroscopic world through an appreciation of historical accident and evolutionary economics. Specifically, it could be argued that it is really Krugman’s reliance on certain key processes such as increasing returns, imperfect competition and path dependence that form the foundation of his work. Geographical economists certainly think that it is an understanding of these processes that separates them from both their fellow economists and economic geographers. However, current geographers also readily embrace such concepts. In fact, Storper and Walker (1989) join Krugman in rejecting the neo-classical world of perfect competition, concluding that ‘the industrialisation process, with its increasing returns over time, drives location’ (Storper and Walker, 1989, p. 97). Admittedly, this finding is often couched in terms of the technological externalities that Krugman laboured to avoid, but as Storper (1997, p. 61) notes under the heading ‘Path Dependency in Technological Space’: ‘The starting point is to admit that many economic processes are amenable to increasing returns to scale’. In terms of mechanisms, this is the same position from which Krugman (1991a, p. 4) embarks: ‘Essentially, to say anything useful or interesting about the location of economic activity in space, it is necessary to get away from the constant-returns, perfect-competition approach that still dominates most economic analysis’. In addition, from these starting points, both groups acknowledge the role of circular causation in maintaining increasing returns. In sum, if Krugman’s group and economic geography are similar both in terms of their goals and understandings of economic processes, then it must be other assumptions that underlie their different approaches.

One explanation for the divergence between the two approaches can be found in their respective unit assumptions. Economic geographers certainly discuss firms, the industry and the economy. In drawing on institutional economics and placing organisations in a trinity with technology and territory, Storper (1997) is acknowledging the importance of such traditional economic units. His theories, however, do not rely on assuming these as basic units. Instead, a spatial unit is implicitly adopted as the proper domain for analysis in economic geography. This unit is usually the region or territory, although it may also be a city or political jurisdiction (see Scott, 1993b), and as a portion of space, it is understood as ‘a fundamental unit of social life in contemporary capitalism’ (Storper, 1997, p. 3). These spatial units are economic geography’s means of understanding both economic processes and the actions and interactions of the units of neo-classical economics that geographical economists like Krugman rely upon. Thus, Storper and Walker, 1989, pp. 138–139 write:
‘The territorial complex is an extensive work site that brings disparate production activities into advantageous relation with each other, at a larger scale and scope than the individual workplace, firm, or even, in many cases, the industry’. Similarly, for Storper (1997), the region itself is an active unit, not just the sum or conflation of other economic units as it is in geographical economics: ‘there is some complex dance of mutual constraint and coevolution, where both firms and territories are subject to temporal and spatial path dependencies that are intertwined with each other’ (Storper, 1997, p. 291). Thus, in economic geography the sites of economic activity are produced in and through development, not just selected after the fact by the rational path-dependent actions of firms and industries.

Perhaps afraid of the metaphors or positivism that might follow, some economic geographers might be tempted to deny the existence of something so simplistic as a unit, but is it really true that economic geography has no constituent part(s) upon which it focuses to explain the economic? A unit is something that a researcher thinks is worth watching. Its movements, its actions – however, conceived or contrived – are what is considered crucial to the understanding of a concept. I have shown above that the units of geographical economics are the firm, industry and economy. In the case of economic geography, the unit is space itself. Of course, different geographers treat the unit of space in different ways. In Storper’s case it is often the region or territory that is the crucial form of space. This concept is in line with the growing interest in the region as a key component of political economy, but in addition, there is room within Storper’s conception, and increasing room in geography in general, for spaces that are linked to social practices. These would include the spaces of networks and knowledge communities. In both cases, as long as space is considered crucial to understanding a larger concept, such as the economic, then it is effectively a unit. Granted, space is not as neat a unit as those used by economists. Krugman’s three units are largely aggregative: several firms make an industry; several industries make an economy. Space is less easily stacked and sorted, but in some form it still serves as the basis for what geographers study. In studying it, they usually name it, so that it becomes a designated (if not discrete) entity like a region or a network.

In fact, it is the assumption of space as a unit that allows economic geographers to explore issues alien to Krugman’s work, such as the mysteries surrounding Marshall’s third externality of knowledge spillovers. When space is the fundamental unit of analysis, research can go further to delve into the ‘untraded interdependencies’, including the conventions, informal rules, habits and institutional structures that explain uneven development (Storper, 1997, pp. 4–5). These elements combine to produce ‘worlds of production’ (Storper and Salais, 1997), which explain industry trajectories, but are still deeply rooted in specific places. In fact, these localised forms of knowledge are only relevant if the focus is not limited to individual firms or even industries, but instead acknowledges units like ‘the region’ as key elements in the ‘supply architecture’ for learning and innovation (Storper, 1997, p. 22). In contrast, in geographical economics, with its strict focus on firms, industries and the economy, nothing can really be ‘in the air’, as Marshall (1919) suggested. An ‘industrial atmosphere’ is a concept that resists being pinned to a traditional economic unit, but can easily be related to a territorial one. Such units of analysis allow geographers to recognise trends like flexible specialisation which threaten to change or even disintegrate the classical units of economics. In short, geography’s reliance on spatial units helps to explain both its divergences from economics and perhaps even its own unique insights into the landscape of economic activity.

A comparison of the approach of Fujita et al. (1999) and Storper (1997) serves to illustrate how analysis of spatial units reveals more than analysis of economic ones. Specifically, I will examine their approaches to understanding urban economies. The understanding of cities promoted by geographical economists was partially introduced above, with an emphasis on Krugman’s conflation of a city with a group of firms. Fujita et al. (1999) elaborate on Krugman’s original model, but the fundamental conclusion is the same: a city will form and function where firms choose to cluster. This analysis reveals that these clustering firms take advantage of the scale economies of their particular industry and the market potential they themselves create by drawing workers. The latter is somewhat city-specific, but in the models of Fujita et al. (1999) it is always balanced against the transportation cost experienced by individual firms. Therefore, a firm may choose to move or replicate itself elsewhere, even if the market potential of its original location has not changed. In the end, cities do not have much to do with the model, because firms and industries are solely controlled by the decisions of firms and industries.

Storper (1997) does not begin with firms, but with the city itself, and asks why the urban phenomenon is important to the economy. In doing so, he discovers more than a clustering of firms, and in fact, concludes: ‘‘The firm’’, which is the theoretical and accounting unit of world city theory, is certainly well wide of the mark, for it is only certain parts of firms, especially the world’s biggest firms, that are to be found in the world cities’ (Storper, 1997, pp. 244–245). Instead, Storper (1997) finds a local ‘socioeconomy’, that consists of local relations and conventions. These elements of the ‘socioeconomy’ can be tapped into by firms – indeed, that is one major reason firms locate in cities – but just as firms are not necessarily exclusive to cities, urban conventions
are not controlled by one firm or even industry. They are greater than those entities because they exist in the social and physical space of the city. By looking at the city, and not just firms, these forms of reflexivity become apparent. Furthermore, unlike scale economies and market potential-transportation cost ratios which are tied to firms and industries, the socioeconomic is tied to the space that is the city. As such, it cannot be easily moved, and serves to differentiate one city from another. In looking at the urban economy, Storper (1997) finds that not only is there more than market potential and scale economies at work, but that the additional elements that are specific to the city itself are crucial to understanding urban economies. Thus, Storper (1997) finds that in addition to the cost-based decisions noted by geographical economists, firms locate in cities for noncost-based reasons, including access to the production capacities of local action frameworks, specific contexts that help tailor good production for extra-city markets, the innovative advantage that comes from the spillovers of specific urban cultures, and the nature of local consumption habits. These many ‘worlds of purpose reflexive action’ uncovered by Storper’s (1997, p. 245) theory have the potential to explain the specific differences of cities and their economic activities, whereas geographical economic models can only explain the existence of an urban economy at one node versus agriculture at another. Fujita et al. (1999) start with firms and find that they have different markets and scale economies. The resultant cities are based on these firm and industry characteristics. Storper (1997) starts with cities and asks what makes them different. He finds that they contain different conventions. These conventions explain the developments of different urban economies along lines far more varied than those predicted by geographical economics.

Recently, Amin and Thrift (2000) have challenged geographers to question their relationship to mainstream economics. Their intervention does not mention the unit assumptions made by the discipline or Krugman’s genre of research, but they do make similar observations as those above about the benefits of economic geography. Specifically, they relate how the study of the economy by geographers has advanced an understanding of open systems, an appreciation of context, and the use of qualitative techniques. Furthermore, Amin and Thrift (2000) suggest that the findings of economic geography call into attention some of the traditional starting points for economics, in particular the theory of the firm. In short, they conclude that the approach of economists is not at all crucial to understanding the economic. Based on the analysis presented above, I agree, but stop short of their warning that economic geography should avoid any close relationship to economics. If part of the difference in approach and findings stems from a difference of unit assumptions, as this paper argues for the specific brand of geographical economics practised by Krugman and others, then some limited dialogue should still be possible. Economic geography can learn from geographical economics, but the geographers must bear in mind that they are learning primarily about how the three isolated units of mainstream economics behave. As already noted, Krugman uses those units in his models to build up agglomerations that he refers to as spaces. In contradistinction, geographers, as Amin and Thrift (2000) partially observe, show that space often determines the borders of the traditional economic units. Thus, the economic units can be made spatial, but when they are they suddenly seem much less important.

A further example might be helpful in illustrating the last point. Pinch and Henry (1999a) provide a ready one in their analysis of Britain’s Motor Sport Valley and the usefulness of Krugman’s theories to understanding its continued existence. First, they conclude that Krugman’s ideas about historical accident and path dependence are only partially helpful to understanding the agglomeration of motor sport related activity around Oxfordshire. Certainly, there were some ‘chance’ events that seem to have encouraged the development of a motor sport industry in this area, but Pinch and Henry (1999a; see also Henry and Pinch, 2000; Pinch and Henry, 1999b) argue that there were many factors affecting the generation of knowledge that were already present in England before the motor sport industry was fully established. Going back to the relationship between the focus of analysis in economic geography as compared to geographical analysis, British motor sport is an example of a space – in this case a pre-existing knowledge community – defining the formation of an industry. The motor sport industry moved in the direction of lightweight, mid-engine racing cars because there already existed a group of people clustered around the government-sponsored British aerospace industry. Motor sport evolved to benefit from this space. Looking just at the unit of industry, Krugman must rely on chance to start his model, but looking at space, much more than luck and fate becomes apparent.

Second, Pinch and Henry (1999a) observe that the continued existence of Motor Sport Valley is not as tied to scale externalities as a geographical economist like Krugman might predict. Instead, the authors find the ‘untraded interdependencies’ of Storper (1995) to be a much better explanation for the continued agglomeration of the industry. According to Pinch and Henry (1999a), the scale externalities that drive the models of geographical economists do exist in Motor Sport Valley, but they are hardly the only reason for the industry or its geographical distribution. Instead, the industry continues because of the knowledge that is contained in the region. Part of this knowledge is embodied by the workers themselves, a market-scale effect that
5. Conclusion

The arguments presented in this paper should cast doubt over any ideas for a substantial interaction between geographical economics and economic geography. Despite its promises, the new ‘geographical’ economics still appears fraught with geographical problems that prevent it from dealing adequately with the concepts of scale or space. This point has already been made by a variety of critics, many of who were dismissive of Krugman’s work when it first appeared (e.g. Hoare, 1992; Johnston, 1992). However, Krugman was persistent and recruited other economists to his cause, and economic geography seems increasingly pressured to respond.

Indeed, economic geography has been approaching its own subject in new ways, for instance by accepting the region as a dynamic entity that is more than the sum of its inputs and outputs (for a general summary see Thrift and Walling, 2000). It is therefore important that any response by economic geography to the new geographical economics be aware of the presuppositions that separate the two disciplines. This paper has argued that one reason for the failings of the models of the new geographically inclined economics to gain support in geographical circles can be found in unit assumptions, which in geographical economics take the traditional firm, industry and economy to be the proper arenas for analysis. These unit assumptions explain Krugman’s dismissal of classical regional science and central-place theory, and point to a misleading conflation with spatial units. These unit assumptions are also intricately interwoven with the criticisms found in geography journals of Krugman’s inter-scale, intra-scale and ‘rigour mortis’ problems. In contradistinction, economic geography, as represented by the work of Storper, relies on spatial units like the territory or region to organise its theories. This different focus helps explain the ultimate divergence between geographical economics and economic geography, as well as the valuable additional insights that the latter provides and that Amin and Thrift (2000) champion. Thus, this divergence is not just one of conventions of analysis or style of argument (Martin and Sunley, 1996), but of fundamental assumptions made before analysis even begins. In the end, while both the Krugman camp and its critics have tried to inform each other’s inquiries, there would seem to be clear limits as to how far this can go when each party builds their case on the basis of fundamentally different units of analysis.

The work of geographical economics is based on firms, industries and economies, and it tells us how these function within the realm of the economic. This is certainly important, but it is limited. The one thing upon which the geographically inclined economists and economic geographers seem to agree is the fact that space matters. Even within the realm of policy discourse, there is a similar realisation and a new found interest in how space – be it a network that promotes learning across time zones or a successful regional agglomeration – functions. If we agree on this, then would it not be better to focus our analysis on space as the important unit, rather than the units of economics that function within space?

Acknowledgements

This paper was substantially written while I was benefiting from a Fulbright grant that was partially sponsored by the British–American Chamber of Commerce Foundation. However, no funding organisation bears responsibility for the views expressed here. I am grateful to Martin Boddy, Leslie Hepple, Andrew Leyshon, and the three anonymous reviewers for their comments on previous drafts.

References

Cooper, S., 1998. Entrepreneurship and the location of high-technology small firms: some implications for regional development. In:
Oakey, R., During, W. (Eds.), High Technology Firms in the
Dixit, A.L., Stiglitz, J.E., 1977. Monopolistic competition and
optimum product diversity. American Economic Review 67,
297–308.
GeoForum 27, 439–452.
Cities, Regions, and International Trade. MIT Press, Cambridge,
MA.
Galbraith, J.K., 1980. Scale economies, product differentiation and
Krugman, P., 1981. Trade, accumulation and uneven development.
Krugman, P., 1986. Strategic Trade Policy and the New International
Economics. MIT Press, Cambridge, MA.
MA.
MA.
Krugman, P., 1993a. First nature, second nature, and metropolitan
Krugman, P., 1993b. On the number and location of cities. European
Economic Review 37, 293–298.
MIT Press, Cambridge, MA.
Krugman, P., 1995b. The localization of the world economy. New
Perspectives Quarterly Winter, 34–38.
MA.
Perspectives 12, 161–174.
Krugman, P., 1998b. What’s new about the new economic geography?
Leyshon, A., Thrift, N., 1997. Money/Space: Geographies of Mone-
Losch, A., 1940. The Economics of Location. Fischer, Jena, Germany
(English translation by Yale University Press, 1954, New Haven).
APA Journal 50, 262–266.
and Unwin, Boston, MA.
and its implications for regional development theory: a critical
Peet, R., 1984. Class struggle, the relocation of employment, and
Pinch, S., 2000. Book review: the spatial economy: cities, regions and
international trade. European Urban and Regional Studies 7,
287–288.
Pinch, S., Henry, N., 1999a. Paul Krugman’s geographical economics,
industrial clustering and the British motor sport industry. Regional
Studies 33, 815–827.
Pinch, S., Henry, N., 1999b. Discursive aspects of technological
innovation: the case of the British motor sport industry. Environment
and Planning A 31, 665–682.
New York.
Growth. MIT Press, Cambridge, MA.
Scott, A.J., 1986. High technology industry and territorial develop-
ment: the rise of the Orange County complex, 1955-1984. Urban
Geography 7, 3–45.
Scott, A.J., 1988b. Flexible production systems and regional develop-
ment: the rise of new industrial spaces in North America and
Western Europe. International Journal of Urban and Regional
Scott, A.J., 1993a. Technopolis: High Technology Industry and
Regional Development in Southern California. University of
California Press, Berkeley, CA.
Scott, A.J., 1993b. Interregional subcontracting patterns in the
aerospace industry: the Southern California nexus. Economic
Geography 69, 142–156.
The Capitalist Imperative: Territory, Technology and Industrial
Geography 69, 142–156.
Storper, M., 1995. The resurgence of regional economies, ten years
later: the region as a nexus of untraded interdependencies.
European Urban and Regional Studies 2, 191–221.
Storper, M., 1997. The Regional World: Territorial Development in a
Frameworks of the Economy. Harvard University Press, Cam-
bridge.
Storper, M., Scott, A.J., 1989. The geographical foundations and
social regulation of flexible production complexes. In: Wolch, J.,
Dear, M. (Eds.), The Power of Geography: How Territory Shapes


