

THE NEW ECONOMIC GEOGRAPHY, NOW MIDDLE-AGED

Paul Krugman

Prepared for presentation to the Association of American Geographers, April 16, 2010

It's almost exactly 20 years since I delivered a set of lectures in Leuven that became the monograph *Geography and Trade* (Krugman 1991a), which most people consider the beginning of the New Economic Geography. It has, from my point of view, been a great two decades – and not just because of that Swedish thingie.

What you have to understand is that in the late 1980s mainstream economists were almost literally oblivious to the fact that economies aren't dimensionless points in space – and to what the spatial dimension of the economy had to say about the nature of economic forces. You may find this implausible – how could economists fail to take into account facts of life that are part of everyone's daily experience? –but I can assure you that it was true. I recall a conversation at one conference on the “new growth theory” in which a fairly eminent economist challenged some of us, in belligerent tones, for any evidence that increasing returns and positive external economies actually play any important economic role. I think I replied “Cities” – to be greeted with a stare of incomprehension.

You might be tempted to say, so much the worse for mainstream economics – and that, as I understand it, has to a large extent been the response of “proper” economic geographers; indeed, over much of the past three decades the methodologies of geographers and economists have been steadily diverging. But my view, then and now, is that this isn't an adequate response. In a crude

sense, mainstream economics isn't going away: like it or not, the White House has a Council of Economic Advisers, not a Council of Geographical Advisers, the World Bank hires lots of economists and not many geographers, and so on. So if insights from geography are going to have the influence they should, there has to be some kind of rapprochement. More fundamentally, the economist's way of thinking may have blind spots, but it also has a great deal of power and depth; there should be a way to persuade economists to learn from geography without sacrificing the good in the field.

And that's what has been happening. Twenty years ago I hoped that economists could be induced to study an important but neglected aspect of the economy – and even a cursory glance at the frequency of papers on spatial and geographic issues in mainstream economics journals shows that this has happened. In 2006 the Federal Reserve held a major symposium on the new economic geography¹; in 2009 geography was the main concern of the World Bank's *World Development Report*.

I also hoped that the new approach would lend some guidance as national economies became more integrated, especially within Europe – and geographical economics has become a major concern in Brussels. And above all, I hoped that the new focus would induce economists to make use of the “intellectual and empirical laboratory” provided by intranational data – and that has taken place big time, with a flowering of studies using urban and regional data to address important economic questions.

¹ To which, curiously, I wasn't invited, perhaps because I started criticizing Alan Greenspan at an unfashionably early date.

Yet there have always been some questions about the justifications for this line of work. One critique has come from geographers proper, many of whom have argued both that there is little new in the new economic geography, and that its effort to reduce the complexities and richness of economic geography to stylized mathematical models makes it a fundamentally misguided enterprise. Another critique – one that has bothered me personally from the beginning – involves the sense that the new economic geography is far too concerned with old stuff, that in the interest of theoretical clarity it focuses on forces and processes that were important a century ago but much less relevant today.

In this paper I want to address the first critique briefly, not so much to argue that I'm right and you're wrong as to explain why the new economic geography is the way it is. Most of the paper will then be taken up with the second critique, which as I've already suggested is in part a self-critique. There's no question that in an effort to satisfy largely academic criteria – the desire to derive everything from first principles, the desire to pull analytical rabbits out of hats – the new economic geography, at least in its early incarnations, adopted an approach that in many ways seems more suited to the economy of 1900 than to that of 2010.

Or perhaps I should say more suited to the *advanced* economies of 1900. For that will be my theme at the end: while a focus on advanced economies might suggest that it's time to downplay the emphasis on tangible factors like transportation costs in favor of intangible factors like information spillover, the old new economic geography gains a new lease on life once you shift your focus to the developing countries that now account for most of the world's economic growth.

So: on to methodology.

Models and metaphors

Many economic geographers proper were furious at the rise of the new geographical economics. That was predictable: near the end of that 1990 monograph I foretold the reaction, and also explained why I was doing what I was doing:

“The geographers themselves probably won’t like this: the economics profession’s simultaneous love for rigor and contempt for realism will surely prove infuriating. I do not come here, however, to fight against the sociology of my profession, but to exploit it: by demonstrating that models of economic geography can be cute and fun, I hope to attract other people into tilling this nearly virgin soil.”

Actually, the reaction was even worse than I expected. As it happens, starting in the 1980s many geographers were moving even further from mainstream economics -- there was a widespread rejection not just of the assumptions of rational behavior and equilibrium, but of the whole notion of mathematical modeling and even the use of quantitative methods

As a result, there have been some detailed, impassioned critiques of the economists’ version of economic geography from the geographers. Notably, Martin (1999) has argued that “economic geography proper” involves “a firm commitment to studying real places (the recognition that

local specificity matters) and the role of historico-institutional factors in the development of those places.” And it involves a rejection of abstract models in favor of “discursive persuasion.”

So how do I respond to that critique? I have no problem with people investigating local specificity and engaging in discursive persuasion. But the new economic geography was designed to attract the attention of mainstream economists. And mainstream economics decided long ago that devising abstract models is an essential part of being a useful profession.

When and why was that decision taken? The answer, although not many people realize this, is that the Great Depression was the dividing line. Until the 1930s and to some extent into the 1940s, institutional economics, with a strong emphasis on “historico-institutional factors”, was a major force in American economics. But when the Depression struck, there was a desperate need for answers – and the answers wanted were to the question, “What do we do?”, not “How did we get here?”. Faced with that question, the institutional economists couldn’t deliver; all they could offer was, well, persuasive discourse on the complex historical roots of the problem.

The person who did deliver was John Maynard Keynes. Now, Keynes is a protean figure, whose writings can be read to provide support for many schools of thought. But *The General Theory of Employment, Interest and Money*, despite occasional historical asides, essentially presents an abstract, ahistorical model of the economy; at its core is a little two-equation equilibrium model of the level of employment. And here’s the thing: Keynesian economics, unlike institutional economics, was able to answer the question about what to do: it told you to boost demand with deficit spending.

But wait, you may say, what does Keynesianism -- which many mainstream economists now reject -- have to do with the field's overall mathematical-model-oriented ethos? The answer is that Keynesian economics was the "killer app" for the takeover of American economics by model-oriented thought. Paul Samuelson's textbook, which brought Keynesian economics to college teaching, also brought the modeling ethos and crowded out institutional approaches. And a key part of what mainstream economists want from their field is to be able to provide useful answers, the way Keynes did.

I don't necessarily mean policy guidance, although that's part of it. More broadly, what mainstream economists want is the ability to answer "what if" questions: if something were different, how would that change the economic outcomes? That's a kind of question that's almost by definition impossible to answer if your approach emphasizes the uniqueness of each individual case and the specifics of history. But it's very much the kind of question that the new economic geography was intended to answer. Indeed, what I saw as the big result in Krugman (1991b) was precisely the model's implication that the geographical structure of the economy depended on a few key parameters: transportation costs, economies of scale, and factor mobility.

To be honest, I don't understand the mindset of those who disdain the search for general conclusions about geography (or anything else) on *principle*. Surely the goal of all scholarship is ultimately to provide an understanding of the principles that govern the world, which necessarily means developing models on which one can perform thought experiments, asking "How would things be different if X happened instead of Y?" What I can appreciate is the argument that

attempts at general modeling are premature, that we need an accretion of detail before we can get to the abstract modeling stage. But I realize that this instrumental view of ground-level scholarship, the idea that it's all about paving a road to some higher abstraction, isn't something everyone shares. So be it.

Anyway, this difference in questions asked – and even in the questions researchers believe should be asked -- may serve as a partial excuse for the failure of most economists working on geographical issues to cite work by economic geographers proper. When a mainstream economist encounters a proper geography paper, he or she tends to skim through it looking for the results that can be brought to bear, directly or indirectly, on what-if questions – and all too often finds nothing along those lines. That doesn't mean that the paper is worthless, but it does mean that there isn't a lot of helpful overlap between the intellectual enterprises. That said, I and my colleagues definitely read too little by people from outside our tribe, and should look over the fence more often.

And with that, let me turn from broad methodological discussion to the choices made in developing the original version of new economic geography.

Core and periphery

There had, of course, been attempts at formal mathematical modeling of geographical issues before my 1991 core-periphery paper – and from mainstream economists as well as regional scientists. Notably, Henderson (1987) had developed a beautiful, though completely aspatial,

approach to modeling systems of cities, while Fujita (1988) had realized that the technical apparatus of monopolistic-competition trade theory could be used to provide a no-dormitive-properties explanation of urban agglomeration (I'll explain what that means in a minute).

But these efforts had not received broad attention. What I set out to provide was a loss-leader: a model that was not intended to be realistic, indeed was aggressively unrealistic, but would serve as a demonstration that economic models could yield new and interesting insights when applied to geographical questions. To that end, I tried to build a model with four main characteristics:

1. *Full microfoundations*: This was going to be a neoclassical model in the broad sense – not a model of perfect competition and efficient markets, but one in which economic outcomes can be represented as the equilibrium that results when maximizing individuals interact. You don't have to believe that this is a literal description of the economy – I certainly don't – to recognize that it's an approach that has shown a lot of power to make sense of the world, and my goal was to exploit that power, not challenge it.
2. *No dormitive properties*: The reference is to Moliere's doctor, who triumphantly explains that opium puts people to sleep because of its dormitive properties. The economics equivalent is the assertion that production clumps together because of agglomeration economies; I wanted to *derive* those agglomeration economies from something more fundamental, such as the interaction among economies of scale, transportation costs, and market size. This meant, among other things, downplaying invisible external economies like information spillovers.

3. *Distance between assumptions and conclusions*: This is a little hard to explain if you aren't in the modeling business, but there's a broad sense, which I share, that you're learning more from a model if the rabbit isn't stuffed too visibly into the hat just before the theorist pulls it out. That's why I didn't follow Fujita (1988) in assuming that there are nontraded inputs produced with economies of scale; even if that's reasonable in many cases, it would be more interesting to show agglomeration emerging without such a strong assumption.
4. *More than one possible outcome*: It seemed, from both an aesthetic and a historical point of view, that a model in which agglomeration didn't have to happen, in which results depend on parameters, would be more interesting than one in which emergence of an industrial core were preordained.

Oh, and one more thing: the model had to be tractable, something you could analyze with pencil and paper and understand what was going on.

What I came up with, working within these constraints, was the core-periphery model (Krugman 1991b). It was, as I said, aggressively unrealistic: only two industries and two locations, two factors of production – and worse, farmers themselves (as opposed to farmland) were assumed to be rooted in place. Manufacturing was assumed to be a Dixit-Stiglitz composite, a highly unrealistic but tractable way of getting increasing returns and imperfect competition into general equilibrium. And transportation costs were assumed to be “iceberg”, with products melting away en route, rather than including any sort of realistic transport sector.

Yet I found the model highly illuminating, and so apparently did many others. It used tools familiar to economists – by this time monopolistic-competition models were standard in the analysis of both international trade and economic growth – yet produced fairly exotic results. There was a sort of phase transition from dispersed activity to a core-periphery pattern when transport costs fell, economies of scale grew larger, or the share of manufacturing in the economy expanded. And there were multiple equilibria – the possibility that either region could end up as the manufacturing core after that phase transition.

Driving all this was the interaction among economies of scale, transport costs, and market size. Increasing returns at the plant level created an incentive for geographical concentration of the production of any given good; transport costs created an incentive to locate plants close to large markets (and large sources of goods from other plants); but the location decisions of producers themselves determined the location of large markets. Under the right circumstances, this could produce a circular causation in which concentrating production fed on itself. But that wasn't a necessary result, because the “centripetal” pull of market size was opposed by the “centrifugal” force of dispersed natural resources. All of this popped out of the math in a very gratifying way.

For all its unrealism, the model seemed to suggest an explanation for the 19th-century formation of real-world core-periphery patterns, notably the emergence of the U.S. manufacturing belt and Europe's “hot banana”. And it was, as I hoped it would be, a sort of door-opener for economists to other geographical issues, notably the formation of industrial clusters. Indeed, industry localization, rather than broad core-periphery patterns, has been the focus of the great bulk of empirical work since the new economic geography crystallized.

Now there are, obviously, factors that will loom much larger at the level of individual industrial clusters than at the level of broad core-periphery patterns. I found myself invoking the Marshallian trinity: specialized providers of industry inputs, thick markets for specialized labor skills, and information spillovers (I think, though I'm not sure, that I can claim credit for the new popularity of Marshall's wonderful line: "The mysteries of the trade become no mystery, but are, as it were, in the air.") I think it's fair to say, however, that my own work – and to some extent that of others – tended to favor the first two elements of the Marshallian trinity over the third. Why? It goes back to the no-dormitive-properties pledge: once you start invoking information spillovers to explain agglomeration, it's a bit too close to just assuming agglomeration.

But here's the thing: the no-dormitive-properties imperative led the new economic geography to emphasize tangible causes of economic concentration in a world that seems increasingly dominated by intangibles. How can you deemphasize technology and information spillovers in a world in which everyone's prime examples of localization are Silicon Valley and Wall Street? Ever since the beginnings of new economic geography, and up until very recently, I and others have had a slightly guilty sense that we were talking about the past, not the present, and much less the future.

A field behind the curve?

Rereading *Geography and Trade*, I realize that it has something of a retro – one might almost say steampunk – feel. I lovingly described the origins of the Dalton carpet cluster, which still exists

but is hardly the cutting edge of economic change. I delightedly described the Bulletin on localization of industry in the 1900 Census, with its descriptions and histories of the concentrations of underwear in Cohoes, costume jewelry in Providence, detachable collars and cuffs in Troy, gloves in Gloverville, pencils from Pennsylvania, and tents from Tennessee. (OK, I got a little carried away there – but up through Gloversville it's accurate).

The point, of course, is that in many ways these stories come from a world that no longer exists, in two senses. First, many of the industrial clusters of the past are pushing up the daisies, have gone to meet their maker, have become ex-clusters. Every regular train traveler in the Northeast Corridor knows the sign “Trenton Makes the World Takes”, which you see crossing the Delaware; but in fact Trenton's industrial complex, centered on ceramics, is long gone. The manufacturing belt itself has largely frayed away, with production shifting south, west, and overseas. Second, the old clusters have not been replaced with comparable new clusters. Yes, there's Silicon Valley – but it's very hard, looking at the U.S. economic landscape now, to identify industrial localizations in anything like the way that was easily possible a century ago. Actually, we can do better than that: we can quantify regional specialization. In *Geography and Trade* I proposed a simple measure of specialization between any pair of regions: the sum of the absolute differences in industry shares of employment between the two regions. This index is 0 if regions have the same industrial composition, 2 if they're totally disjoint. Sukkoo Kim (1998) has constructed this index at intervals going back to 1860, and taken averages to get an overall picture of U.S. regional specialization; the results are shown in Figure 1. The figure suggests that regional specialization peaked sometime around the Wilson administration, and that it has been downhill since the end of World War II.

To this quantitative evidence I would add a more subjective judgment: in 21st-century America, as compared with mid-20th-century America, there is much less sense that places are defined by what they do for a living. Compare Pittsburgh in 1950 with Atlanta today; one was a steel city, the other is a ... what? In general, I would doubt that many people – even residents – could identify the export base of most major metropolitan areas other than New York, where financial services are the obvious driver. To some extent this lack of obvious differentiation reflects a rising share of non-traded services in employment, so that most people in every metropolitan area are doing the same things – retail trade, local medical services, etc.. But it also presumably reflects a shift in the nature of local specialization. The word I guess I'd use for regional specialization in the contemporary United States (and, to a somewhat lesser extent, in Europe) is “subtle.” There is still extensive specialization – there must be, or there wouldn't be so much trade and travel between regions. But the specialization seems to involve relatively fine distinctions. There are medical equipment clusters in both Boston and Minneapolis; they're presumably not doing exactly the same things, but it would take close inspection to discern the differences.

And this subtlety of specialization probably also means that the subtler forms of agglomeration economies predominate. Specialized suppliers and thick labor markets for special skills are still factors; but information spillover, entrepreneurial chains of influence, and so on bulk larger. This doesn't pose any problem in principle for economic analysis, but it does make it hard to hold to the no-dormitive-properties principle. The new economic geography style, its focus on tangible forces, seems less and less applicable to the actual location patterns of advanced economies.

But what about the economic geography of less advanced nations?

NIEs and the NEG

In 1990, when I gave those initial lectures on geography, I focused entirely on advanced economies, and mainly the United States. That was largely a matter of convenience: the data were easy to find, as were the historical cases. But I could also defend it as reflecting the world's economic center of gravity. The new economic geography was mainly a story about manufacturing, and in 1991 high-income nations accounted for 84 percent of the world's manufacturing value-added.

Today that fraction is below 70 percent, with most of the *growth* in manufacturing taking place in emerging nations, above all China. And here's the thing: as I've said, new economic geography has a kind of steampunk feel, so that the stories it tells seem more suited to the U.S. economy of 1900 than that of 2010. Well, China is an economic powerhouse, but it's still quite poor; allowing for the vagaries of purchasing power parity estimates as well as GDP comparisons between very different eras, China today appears to have roughly the same level of per capita GDP as the United States at the beginning of the 20th century.

And guess what? Chinese economic geography is highly reminiscent of the economic geography of advanced nations circa 1900 – and it fits gratifyingly well into the new economic geography framework.

First of all, a dramatic core-periphery pattern has emerged, with massive migration from central and western China to the coastal region, and within that region to the manufacturing belt in the southeast (plus a secondary migration toward the Beijing area) (World Bank 2009). Admittedly, the picture is complicated by the fact that much of China's industrial production is aimed at the external world market, rather than the domestic market. Still, the differentiation of regions is a clear cousin to the emergence of the U.S. manufacturing belt in the 19th century.

And then there are the industrial localization stories. China is home to hundreds if not thousands of local industrial clusters, from the cigarette lighter town of Wenzhou, producing 95 percent of world output, to Yanbu, the underwear capital (the 21st century Cohoes). Case studies reveal the classic forces of localization at work; for example, the motorcycle cluster in Pengjiang consists of three cycle factories and about 30 factories producing accessories, and it's that local supply of specialized inputs that appears to be the main driving force behind agglomeration (Arvantis and Haijiang 2009).

Does the fact that these localized industries mainly produce for a global market require rethinking the logic of new economic geography? Yes and no. In a recent paper (Krugman 2009) I argued that to understand the welfare effects of Chinese-style industrial localization we have to think in terms of models in which external economies of scale at the national level are embedded in a wider trading system, which is not the way I was thinking in the original new economic geography work. But it's not a violent break -- and anyway, it basically involves integrating new economic geography with the new trade theory (or, rather, the old new trade theory -- it's in its fourth decade), which was where the new geography came from in the first place. Furthermore,

export-oriented localization, while not characteristic of late 19th-century America, was very much the norm in late 19th-century Britain.

In sum, it turns out that the concepts and approaches of the new economic geography aren't backward-looking after all. They're utterly relevant to understanding developments in the world's fastest-growing economies. Localization in America has become a subtle affair, but in China and other emerging economies, it's anything but subtle, and there's wide scope for the use of no-dormitive-properties models to make sense of what we see.

Oh, and to return briefly to the issue of general, abstract models versus historical specificity: surely the strong resemblance between China's industrial clusters today and the industrial clusters of the 19th-century – especially the export-oriented clusters of late-Victorian Britain – strikes a blow in favor of the argument for simple, common principles. The history behind the rise of, say, Dundee in the 19th-century – Juteopolis – and that of 21st-century Dongguan could hardly be more different. Yet they are recognizably the same kind of animal, which would seem on the face of it to suggest either that institutions either matter less than the most ardent advocates of an institutional approach claim, or that institutions tend to evolve to take advantage of the economic opportunities created by the invisible hand of market forces.

Final thoughts

A generation ago, mainstream economists hardly thought at all about the location of production within countries; they hardly looked at local and regional data for evidence on such matters as

the strength and nature of external economies. The new economic geography was conceived as an effort to change all that, bringing economists into an area the best way I knew how: by developing cute, nifty models.

That effort has succeeded. There are many ways in which the ongoing work in new economic geography can be criticized; one of these valid criticisms is the failure to pay sufficient attention to the work of more traditional economic geographers. But getting economists to think about location and spatial structure is nonetheless significant progress.

REFERENCES

Arvanitis, R. and Haixiong, H. (2009), "Research for policy development: Industrial clusters in South China" in Michael Graham and JeanWoo, *Fuelling Economic Growth: The Role of Public-Private Sector Research in Development*.

Brakman, S., Garretsen, H. and Marrewijk, C. (2009), *The New Introduction to Geographical Economics*.

FUJITA, M. (1988). A monopolistic competition model of spatial agglomeration: a differentiated product approach, *Region. Sci. Urban Econ.* **18**, 87-124.

Henderson, J. Vernon. 1987. „General equilibrium modeling of systems of cities.“ In Edwin S. Mills (ed.). *Handbook of Regional and Urban Economics*, 2. Amsterdam: Elsevier: 927-956.

Kim, Sukkoo (1998), "Economic Integration and Convergence: U.S. Regions, 1840-1990",
Journal of Economic History.

Krugman, Paul. 1991a. *Geography and Trade*. Cambridge: MIT Press.

Krugman, P.R. (1991b): "Increasing returns and economic geography," *Journal of Political Economy*, 99, 483-499.

Krugman, P. (2009): "Increasing returns in a comparative advantage world", forthcoming in
Deardorff festschrift.

Martin, Ron (1999), "The new „geographical turn“ in economics: some critical reflections”,
Cambridge Journal of Economics.

World Bank (2009), *World Development Report 2009*.

Figure 1: Regional specialization in manufacturing

