THE ECONOMICS OF CONTEXT, LOCATION AND TRADE: ANOTHER GREAT TRANSFORMATION?

Michael Storper
Professor, Sciences Po, LSE, UCLA
This version: March 21, 2008

Paper for delivery as the Roepke Memorial Lecture, Association of American Geographers, Boston

ABSTRACT

Economic geography has established the economic value of agglomeration, in both static (productivity) and dynamic (innovation) terms. Development economics suggests that localization of activity can be a barrier to convergence, in both static (home market effect) and dynamic (the spatial innovation hierarchy, where agglomerations are at the top) terms. Trade theory takes a broader view, suggesting that when de-agglomeration occurs, it brings welfare gains to the economy as a whole, even if there are high adjustment costs for the localities affected by employment and output losses. This paper argues that there are lacuna in both the case for de-agglomeration and that of agglomeration. In the former case, the Romer growth theory suggests that only with de-agglomeration are economy-wide growth effects of technologies likely to be achieved: Marshall-Arrow externalities are very different from Romer externalities. But even this “strong dynamics” case for the benefits of de-agglomeration does not exhaust the question of the welfare effects of the geographical spreading out of production. This is because at the heart of agglomeration are not just productivity and localized learning dynamics, but specific behavioral processes, which together constitute the “context” of decision-making and development. Trade theory has not established that all processes of de-agglomeration bring about a supply of contexts that is better for future economic development than the existing supply. Hence, there are gaps in the case for the welfare effects of de-agglomeration and trade.

JEL: D62, F12, F15, R12
O. A Story of Context

My franco-american godson has been passionate about food and cooking since he was about ten years old (he is now 22). In the last several years, in addition to completing undergraduate studies in both the USA and France, he has worked in high-level restaurants in Los Angeles, Chicago and New York, making friends with excellent restaurateurs in several cities. We’ve had an ongoing conversation, along with other members of our circle, about food quality, cooking, agriculture, and gardening. He keeps coming back to the conclusion that, though in the best American kitchens there is a great deal of innovation and excellence, it is nonetheless more difficult to get an excellent meal in even the most gastronomic cities in the US than it is in France. In order to get such a meal in the USA, you generally have to go all the way to the top of the food chain, to the “big deal” restaurants; by contrast, there are a lot of excellent but unassuming restaurants in France, and generally at more modest prices than in the US. This point does not challenge the statistics on starred restaurants in the Michelin guides of Paris versus New York, showing that the latter has many top kitchens; rather, it is that just below this very top level, there remains a huge difference between the US and France. A fair guess is that the same would be said of Japan versus the USA (Tokyo has double the number of starred restaurants than Paris).
For several years, I have pressed my godson on why he thinks this is the case. He has long thought about trying to open a simple, excellent restaurant in the USA, but increasingly thinks it would be almost impossible. In France, however, it remains possible, for several reasons. First, there is the supply chain for ingredients. This begins with the massive market at Rungis near Orly Airport, for which there is no equivalent in the USA. Perhaps more importantly, that market is complemented by numerous alternative ways to get very high quality ingredients, many of them just by going to the right vendors at the Paris markets. Though in some US regions, markets have been developing in recent years (northern California, the Hudson River Valley), the best ingredients are the province of only the very high-end restaurants, because relative prices for such ingredients are much higher in the US than in France. Supply chains for such ingredients continue to develop, but they are a far cry from those in France. This leads to the second difference. Because non-star chefs get hold of such ingredients relatively easily, there is an inherently lower barrier to entry to excellent cooking than they would in the USA. Thus, and this is the third difference, to use these ingredients profitably, American restaurants generally need compensate by operating at higher scale than their French counterparts.¹ They are also pushed to higher scale by the structure of real estate and location in America, where the “granularity” of urban structures is different, so “location”

¹ Thus, the Paris daily free newspaper, Métro notes, on December 4, 2007, that the prize for “Fooding” in 2007 goes to Pierre Jancou, with his restaurant Racines, which has only 36 seats, 10 of them outside. His secret: “I search for artisans for vegetables, meat and cheeses, and here we put forth the products, without chi-chi.”
and “size” become part of capturing a sufficiently big market to offset the higher fixed cost structure they face. Fourth, tends to be a different production process in the American kitchen. In the generally larger American restaurant, a greater proportion of the staff will be carrying out orders in a steeper hierarchy than in the excellent, but not big-deal restaurant in France.2

The point is that a relatively simple restaurant, that operates at modest scale, and offers excellent food, which is common in France, relies on easy access to external inputs, moderate scale of operations, and moderate hierarchy in the kitchen. These features enable it to keep its offerings focused and simple and hence control costs, and thus reproduce the model. All the incentives work in the opposite direction in the USA, leading to greater rarity of this type of restaurant, pushing producers and consumers toward the big, “star quality” restaurant model.

The fifth consequence is that the French restaurant production system has a much larger proportion of excellent, but not “celebrity” restaurants. This in turn leads to a recursive feedback effect in both countries. In France, there is a bigger, more open and competitive milieu for training cooking talent, and this is an ongoing “nursery” of talent for both the “normally good” and the higher-end restaurants. In the US, this type of training milieu is much thinner. The exception to the US pattern has been perfected by Italians in the US, and it is a

---

2 In the big deal restaurants in France, of course, there’s a very formal division of labor in the kitchen, very rigidly codified by education and practice.
telling exception. The average size of their successful restaurants are still bigger than their versions at home in Italy, but they are not huge affairs like the multi-starred restaurants. The Italian success comes from relying on the clan method: bringing in virtually the entire staff (often with the family at its core, plus investors from the home town) from Italy, sometimes (but more rarely) down to the waiters, but often including the architects and business managers.\(^3\) In essence, rather than drawing from a local context in the USA, they import it, which the French generally don’t do.

The sixth consequence has to do with the type of innovation carried out in the two places. According to my godson, while the top American restaurants innovate in the creation of new dishes, the innovations in fundamentals of technique, ingredients, and focused sauces and flavors, still mostly come from France, Italy and Spain. This is because the basic innovations are about artistry, and in this, the division of labor between conception and execution found in top American restaurants is a disadvantage. In France (and perhaps in the other countries), there’s more interaction between the merely “excellent” and “celebrity” levels of the food chain. Both countries innovate, but differently.

That such differences should persist is curious, given strong international knowledge spillovers in cooking and the restaurant industry. There has been the spread of a world-wide system and culture of gastronomic restaurants, involving the mobility of labor (celebrity chefs); knowledge (cooking schools, books,\(^3\)

---

\(^3\) Thanks to Monica Viarengo of Farina in San Francisco for explaining the Italian “method” to me.
magazines); capital (investors and owners); a sophisticated clientele that travels and eats and compares; and the creation of long-distance networks and collaboration, as well as a few truly multinational operators, with restaurants in more than one major city. In other words, the story does not seem to be one of knowledge so un-codified that it cannot travel. Increasingly, the knowledge does travel and there is more and more deliberate organization of the channels through which it is exchanged and mastered in far-flung contexts. Instead, it is about what happens to that knowledge when it arrives in different places: its use and further development and hence the qualities of the resulting product output. The causes cannot be simply assigned to “history, culture and ways of life,” either. The processes of learning, information exchange, and factor mobility should be able to alter history, culture and ways of life. Lots of people in France would like more imitation of things American there; lots of educated diners in the USA would like a Parisian dining scene. In this sense, their cultures have already changed.

This probably seems like a quaint story, but I believe it represents a more general reality of economic development: there remain important differences from place to place in how the details of economic systems work. Such differences, in the division of labor, supply chains, labor markets, and local “relational embeddedness” are major forces in determining why some regions and countries develop more and better than others, and to influence patterns of specialization, perhaps as much as the conventional sources of comparative
advantage. The examples are not limited to the “artistic” or design-based sectors: other examples include German versus American car engineering, or French service firms versus their American counterparts.

In this paper, we will theoretically re-interpret the restaurant industry example, with three principal goals. First, we will argue that the various differences in the two restaurant production systems described above can only be captured fully by introducing the theoretical notion of “context.” Second, we will ask whether new forms of geographically-distributed production systems represent a “great transformation” of the role for context in the economy. In this perspective, the Paris and Tokyo restaurant economies might be mere vestiges of the past, about to be replaced by a global model based on the New York system, which is duplicated in Chicago, Singapore, and Dubai? Third, we will place these ideas about context in the wider framework of trade theory, because this is where economic geography takes up the wider welfare effects of changing contexts. Let’s begin with the latter question, the welfare effects of the ongoing process of localization and de-localization in the economy.

I. Context and the gains from trade

The combined effects of the division of labor, specialization and gains from trade are widely agreed to be one of the two main forces behind world economic growth since its modern take-off around 1820, the other being
technological innovation (Mokyr, 1990; North, 2005). Contemporary debates about the geographical reshuffling of output and employment through outsourcing and off-shoring ask whether we have crossed some kind of new threshold in the world division of labor. Is the scale and speed of geographical fragmentation fundamentally altering the process of economic development, replacing local and national levels of interaction with a new type of geographically-distributed system?

Even though there is no definitive answer to this question, trade theory is confident about the issue of potential welfare effects of a great transformation: there may be new kinds of adjustment costs from recomposition of local and national economies, but there will be gains to fragmentation and trade for the world economy as a whole, and in the long-run for those local economies that successfully complete their “churn” and absorb its costs. This holds even when a particular country or region has benefited from a strong cluster: according to trade theory, by definition the long-distance relations that replace a local cluster will be more efficient and hence generate welfare gains. Development economics has identified some highly-contested exceptions to this assumption that the overall level of geographical concentration and diffusion is optimal: economies of scale; timing of entry; sequencing and terms of trade; infant industry. Most of it shows, however, that there is not enough diffusion or delocalization, not too much, i.e. less developed countries have insufficient opportunities to get into the world economy. Moreover, some critiques of the
ongoing transformation of localized into long-distance relations, i.e. of the
ongoing creation and destruction of local “contexts” tend to be naïve, nostalgic
or openly reactionary. These include various kinds of “neo-localism,” which are
often traditionalist and communitarian, as well as expressing wistfulness about
artisanal work versus large-scale divisions of labor.

An idea that has influenced much development policy consists in claiming
that the more an activity has a local supply chain, the more developmental
benefits for a locality can be captured over time through expansion of the
activity. This idea expresses a fundamental bias against long-distance linkages
and commodity chains that are highly fragmented over different territorial
jurisdictions. In simple accounting terms, it has some empirical validity, in the
short-run, in that the more localized the value chain, the more of it will be
captured locally for a given increment of output increase in the activity in
question. In more dynamic terms, however, there are problems with this way of
thinking (Puga and Venables, 1999). One is that there is reciprocity: so if
everyone adopts policies to capture these chains, it will negatively affect demand
for our exports, with possibly negative overall consequences on our output. A
second is that there are often political economy problems, i.e. signaling through
such policies to local firms that they can be lazy about quality and innovation
because they are protected. Thus, in the long-run, such a notion addresses itself
too simply to the static capture of output gains, and ignores the dynamic process
of creating local advantage. In the latter, it is not necessarily how much of an
input-output chain is concentrated locally that drives long-term growth, but rather whether certain local interactions are sufficient to re-create local comparative advantage, such that the ongoing re-composition of local output (sectoral and activity “succession”) increases local productivity levels enough to permit increasing local factor prices (and hence incomes). In other words, it is about local learning, innovation, and adjustment, not about mercantilist “capture” of supply chains. This is why many of the most successful economies in the world (national or regional) have been successful over time, in spite of the ongoing loss of certain activities via re-composition of their economies.

But there are doubts about this scenario as well. An economy might appear to do very well by shedding parts of input-output chains and re-composing itself, and there might be aggregate gains to trade. Nothing in trade theory establishes that even this is a long-term welfare maximization, however. For example, if the sectors in question \textit{would} have developed, in the future, in different ways if they were geographically fragmented from how they do when they are more geographically localized, \textit{and} if there are strong irreversibilities and sunk costs once geographical fragmentation gets under way, then the current decision to fragment them geographically leads to future “roads not taken.” This is a claim frequently made about lesser-developed economies as well as about some developed regions where, it is argued, a rich local feedback process within the supply chain encourages adaptive refinement of products and processes and generates “staying power” (Bardhan, 1971; Bruton, 1998). In this
case, retaining supply chains is a necessary precursor to learning and adaptation, which potentiate learning and adaptation later on. In trade theory, it is simply assumed that there cannot be foregone superior outcomes.

This is the main problem addressed in this paper. Both neo-localists and trade theorists have inadequate theory and evidence to deal with this issue. This is because there is a lacuna in economics and economic geography about the question of what determines how productive activities trace out their developmental pathways. This lacuna is rooted in the idea that organizational and geographical features of such systems, because they are driven by appropriate decision-making processes, are the right outcomes. These assumptions are highly questionable. The effects of decisions to fragment and relocate production are not just measured in terms of outputs and productivity levels - they also involve the creation, loss and change of contexts, which affect future development potential. Such change may or may not be optimal. Decisions about organization and geography of production may have unintended long-term welfare effects.

II. Why we need geographical fragmentation and trade: the geographical underpinnings of increasing returns

Before we consider a possible counter-narrative to the case for geographical fragmentation and globalization of production systems, let us
establish the strongest possible theory of why the economy must develop through a combination of organizational and geographical fragmentation of activity, and the ongoing recomposition of regional and national economies. The standard analysis of the welfare effects of geographical fragmentation concludes that that there are almost always aggregate gains to trade. There is a dynamic extension of it in political economy: that the “whip” of inter-place competition increases incentives to firms in different places to become more efficient over time (Wolf, 2001).\footnote{There are some areas where temporary local costs of openness are argued to outweigh the benefits, having to do with existence of economies of scale, infant industries, and local dynamic externalities, but rarely are they considered to outweigh the long-term aggregate benefits of trade to all parties concerned.} Trade is thus seen as the outcome of cost-minimizing decisions to geographically fragment production and markets, and this may in addition have \textit{dynamic incentive effects} on economic actors in different places.

There is another way to see the positive effects of trade and fragmentation, however, which suggests an even more powerful set of beneficial effects. This comes from growth theory rather than location theory, and centers on the geography of positive externalities and increasing returns. The Romer growth theory establishes economy-wide increasing returns as the principal source of long-run economic growth under resource constraints (Romer, 1990). Knowledge and technology are non-rival and only generally only weakly excludable over time; hence they can be infinitely re-used without loss and tend to spread their effects among communities of users (industries, for example) and geographical areas. The productivity of the R&D/innovation sector of the
economy is not subject to diminishing returns, and becomes a source of long-run unconstrained growth. This point is reconciled with standard theory because even though the different specific activities to which innovations are applied are perfectly competitive in the long-run, the economy as a whole is freed from diminishing returns (Romer, 1986, 1990, 1994). The link between these two seemingly contradictory propositions is that monopoly rents to innovation are bid away through diffusion and entry into each sector, but at the economy-wide level the recombination and re-use of technology creates increasing returns.

Applied to the geography of economic growth, the Romer theory is frequently allied to earlier contributions from Marshall and Arrow, respectively about technology spillovers at the regional scale (“the secrets of industry are in the air”) and “learning-by-doing.” Both capture key mechanisms of increasing returns later formalized by Romer: those of re-use, spillover and improvement. Many such Marshall-Arrow processes occur at definite territorial scales (local, regional, national innovation systems and spillovers). This leads many to claim that “Marshall-Arrow-Romer externalities” at the regional scale are the principal form of growth-enhancing increasing return in the economy as a whole.5

There is a key contradiction between this claim and a central point of the Romer theory, i.e. that increasing returns can be reconciled with allocative efficiency through perfect competition. If the principal source of increasing returns were a set of restricted local relationships, then in effect they would

---

5 The term “MAR externalities” shows 221,000 listings in Google on 1/25/2008.
merely generate localized technological or knowledge rents. We know that such
rents exist when knowledge is sufficiently complex or un-codified that some
kind of proximity (geographical, member of a network, often both) is necessary
to gain access and know how to use it properly. That such rents allow regional
wages and profits to rise above the economy-wide norm, at least for a time, and
hence contributes to (at least temporary) successes in regional development is
widely admitted among scholars of regional development and development
economics.⁶

But if this were the end of the story, then there would be no way to
account for long-term economic growth as a result of innovation: rather than
increasing returns at the economy-wide level, there would simply be long-term
accumulation of rents in certain lucky hands. Knowledge would be
geographically excludable; one of the defining pillars of the Romer theory would
be broken.

Moreover, though the spread of technology and knowledge are hardly
similar to the fluid process imagined by standard theories, there is still ample
evidence that much of the economic development process is driven by sharing
and diffusion of technologies (Mokyr, 1990). The reality is somewhere in-
between durable monopolization of knowledge and rent-seeking, and easy,

⁶ One major debate between more neoclassical models of economic development and more heterodox ones
is how long the rents can be enjoyed by the region or country before they are bid away due to rising costs,
technological change or some other force, and how such “short- to medium-term” rents relate to longer-
term economic development processes. Once again, strict neoclassical work emphasizes the return to
equilibrium, whereas work based on an evolutionary or “history matters” frame of reference emphasizes
circular and cumulative construction of advantage.
seamless sharing. There is sharing, but it is temporally- and geographically-uneven, with rents accruing to some firms and places for some periods of time, and then slowly breaking down, as the potentially non-rival and non-excludable character of technology progressively becomes real, allowing the economy-wide increasing return to become real in turn.\footnote{What of the scenario of localized learning, sharing and re-use of technologies, creating localized increasing returns? This is a well-documented phenomenon. But the circle of beneficiaries of such processes would be limited if the use of such technologies remained localized, if for no other reason than that producers would be strongly tempted to extract monopoly rents for the knowledge involved. That they do so is a claim of much of the regional economics literature, as it is one potential explanation for why certain regions can support very high labor and land prices. At some point in the development process, however, the benefits of such technologies seem to leak out to the wider economy, and this is likely to involve codification of the knowledge, delocalization, and the bidding away of monopoly rents through entry and wider application. \textit{Both} of these processes deepen and widen the basis for Romer growth effects.} In other words, even though increasing returns are the key source of economic growth, they are only realized through the geographical- and organizational process that leads from localization and monopoly rents to geographical (hence economic) diffusion of technology and the breakdown of those rents. This is a point that formed part of the intuition of early studies of the geography of innovation diffusion (Pred and Hagerstrand, 1967) but those scholars did not have the benefit of theoretical advances in growth theory or the economics of agglomeration. We now can see just how important their intuition really was to understanding how growth can occur, but also why it’s temporally and geographically not smooth. In the way we can formulate it today, there are no so-called “MAR externalities.” Rather, there are M-A sources of \textit{local} technological externalities and possible local monopoly rents; but the true R-sources of \textit{economy-wide} increasing returns are not essentially local.
M-A and R are linked through (a) the geography of where technologies and technological rents emerge; (b) sometimes, processes of localized learning and sharing, with both rent capture and rent-destruction effects; (c) a process of codification, so that such innovations can be more widely used; leading to (d) a geography of the destruction of technological rents that creates increasing returns in the entire economy.

A reformulation of this sequence looks as follows. In t1, innovations emerge in certain, specific places and organizational settings (firms). The geography of this innovation, involving M-A effects of proximity and localized learning (Henderson, 1999; Jaffe et al, 1993; Malmberg and Maskell, 2006). There are monopoly rents to such innovations, for a certain amount of time. These monopoly rents last as long as there are barriers to imitation, which include both knowledge barriers and trade/communication costs associated with using the technology – it is effectively, if not juridically, excludable. The places where there are agglomeration economies for such technology production and use capture rents.

In t2, these innovations can diffuse to a wider set of places, firms and uses. The knowledge is more amenable to imitation and copying because as it becomes more widely used, it tends to be codified, more people learn the codes for using it, and hence the trade costs associated with deploying it to other uses and places decrease. On the ground, there is geographical fragmentation of production and de-localization. The entry of other users will bid rents away. Schumpeterian
competition is replaced by standard (or at least semi-standard) market structures. The locational dynamics of this are described by the product cycle, but the latter does not have an explicit growth theory dimension.

Simultaneously, the Romer process of economy-wide increasing returns comes into being. This is because geographical diffusion allows more and more non-exclusive application of the knowledge, so that its potentially non-rival aspects now can become reality. The knowledge goes into many different applications, and back into recombination into further rounds of innovation, the drivers of long-run growth in the Romer model.

The process repeats itself over an unlimited number of cycles. The precise parameters for technology-creation, the trade costs and barriers to imitation, and Romer-like diffusion and re-application of technology will determine such things as: the spatial hierarchy of incomes at any given moment; the amount of time it takes for a shift away from the rent-earning (first mover) part of the innovation cycle to economy-wide increasing returns; and, the amount of increasing returns in the economy as a whole. Overlaid on all this is the geography of innovation: where additional innovation processes get started and new rounds of monopoly rents are earned.

This account is a more powerful case for the gains to geographical fragmentation and trade than the standard Ricardian/H-O accounts of allocative efficiencies through cost-minimizing locations, because it adds to the latter static analysis a dynamic process.
III. What is context?

In the account thus far, the local contexts of M-A processes need to give way to an economy-wide diffusion process if they are to generate long-run growth. Another way to state this is that local differences are good at certain moments in the economic process, but must ultimately be reduced at others (diffusion, end of monopoly rents), thus completing the development cycle from innovation-driven divergence to diffusion-led convergence and broad growth.

Anthropologists have long asked whether differences in economic organization between societies were simply reflections of great differences in resources (endowments), which then lead to different power structures or rules (incentives), or whether it is the actors from one place that have different processes of rationality, goal-seeking, learning and cognition (Sahlins, 1995). The anthropology of “primitive” economies uses their relative isolation as an experimental control for this question (“weak contamination” of practices and resources through contact), but has not arrived at consensus about it.

Contemporary institutional economics sees such divergence as empirically important, but theoretically places itself in the first camp: such differences are an aggregation phenomenon, not due to systematic at the level of individual actors (Acemoglu, Johnson, Robinson, 2004). Differences in endowments and initial conditions create different scarcities and collective action dilemmas, leading to
construction of different rules and market structures. From this, "history matters," and can lead to durably different and durable outcomes for economic organization and development. In spite of this, many analysts hold that the periods of relative institutional stability create particular "local" environments of cognition, learning and incentives, thus differentiating economic contexts from one another.

A third perspective simply acknowledges institutional differences as a fact. Economies are such complex aggregates of microeconomic phenomena that in any given place, even with common overall rules and laws, important differences of practice and history will emerge. These will affect choices and perceptions, and reproduce themselves as ‘varieties of capitalism’ (Hall and Soskice, 2001) ‘embedded firms and production systems,’ and ‘regional worlds’ (Storper, 1997; Storper and Salais, 1993). There can be ongoing differences in values and beliefs that influence political choices and hence institutional differences between places. In their different ways, all three of these perspectives insist that there are different “contexts” for economic activity, and these differences have something important to do with performance, both positive and negative.

Arrayed against all of the above is the notion that such conditions for institutional divergence and hence diversity of contexts, no longer exist in a world of intricate divisions of labor, trade and interpenetrating markets. In this view, there is no longer enough separation of contexts to allow such differences
to persist, except as residues of institutional rigidity (Friedman, 2005). For some, this leads to worries about loss of diversity of behaviors and material cultures, whereas – as noted above -- for the mainstream of economics it is a potential source of economic betterment as less efficient contexts are transformed through the “whip” of global integration and competition. This is another reason why geographical fragmentation, integration and diffusion – the destruction and transformation of local contexts – might be a key source of long-term economic development.

**Defining Context**

These different perspectives are nonetheless incomplete as inputs to a theoretical grasp of difference in economic environments. A step beyond them is to consider what psychologists who work on the economy call the “situation” of the actor. They provide us powerful analytics that enable us to see difference neither as a simple aggregation phenomenon, nor as simply “cultural.”

They start by challenging the notion that actors have universally powerful rational cognition and simple preferences (Ariely, 2008). Experimental evidence confirms that action is very strongly “localist” in the sense that choices are typically made with very imperfect information and on the basis of a set of limited criteria; hence, the “normal” world of economic decision-making is not one of ordered preferences, obtaining complete information and rationally weighing alternatives, but rather one of widespread and persistent cognitive
errors and biases. The unifying theme in behavioral economics is, thus, 
*situationalism* – the idea that decisions are always strongly conditioned by local 
influences, not perfect information and long-run optimization with it (Glaeser, 
2003, from Ross and Nisbet, 1991), and that this may have strong evolutionary 
roots (Heselton et al, 2005). There are five dimensions of the situation of the 
actor that can be drawn upon to define her context.

First, valuable specialized information is not uniformly available in this 
world. There are costs to obtaining it, and there are also barriers to access – one’s 
social and economic position define whether and under what conditions, and 
sometimes at what cost, one can get such information. This point is relevant to 
all behavior, but we can consider it in relationship to productive activity. The 
information one needs, but also to which one has access, is strongly defined by 
one’s place in the division of labor. Divisions of labor are dizzyingly complex in 
the 21st century, such that there is a great deal of opacity about what is going on 
outside of where we stand in a division of labor, or at the very least, outside of 
the “neighborhood” of roles surrounding our place in it. Most people don’t have 
an overall vision of the productive chain in which they work, not to mention the 
millions of other specialized systems that affect their lives as workers, consumers 
and citizens. This may be the beauty of the division of labor, i.e. that we don’t 
need to know these things in order to benefit from the productive power 
unleashed by such complex structures. But they also mean that what we know is
partial. Hence, the situation that we define for ourselves is a “local” – in the sense of partial -- one.

This issue is ignored in welfare analysis because it is assumed that if it’s worth getting information we currently lack, then we will search for it; this is the cognitive rationality assumption. But that assumption is challenged by prospect theory, which holds that people engage in very limited search behavior, and put enormous weight on limited reference points, even when these are quite arbitrary and ephemeral (Kahneman and Tversky, 1979). This is the second dimension of the actor’s situation.

A third dimension of situationalism is how actors form their goals. In the place of subjective preferences, on the one hand, or rationally-constructed ones on the other, an enormous body of evidence shows that goals are strongly influenced by comparison and emulation. Status comparison is a big element of this, but it affects virtually every dimension of preference formation. Since information is limited and prospecting is the norm, what we compare to and emulate is not some kind of exercise in global maximizing, but in many ways is highly dependent on where we are, socially, geographically, and in our position in the division of labor (Frank, 2001; Fine, 2006).

Fourth, most economic decisions are strongly present-oriented. Hyperbolic discounting, which places excessive weight on the present and not enough on even the medium-term future, is the norm, even in financial markets; and we use present-day cues, that stem mostly from ephemeral and situational
forces, rather than information over a relevant spatial and temporal span, to form our choices (Laibson, 1997). Though this may be tempered by structures of management, management studies show that it’s far from having been eliminated.

Fifth, the information we do access is generally processed in a relatively narrow way. Mental accounting shows that people mostly make decisions by ignoring events and consequences outside of a particular narrow domain (Thaler, 1985), and most decisions are made using rules of thumb that are far away from the processes that would be necessary to maximize. Even such “tricks” as “reframing” a situation can induce big changes in choice behavior, as can the way that local stimuli trigger emotions (Romer, 2000; Ariely, 2008). People are inherently conservative in their decisions, as shown by the endowment effect that they want to keep what they have over almost any other goal (Thaler, 1994). They are also averse to small-value risky gambles, and they are vengeful toward strangers (Fehr and Gachter, 2000). None of this means that all decisions are merely the unleashing of emotions, because economic actors may be able to learn to manage, interpret and manipulate their own emotions (Gul and Pesendorfer, 2001); but they cannot cut around them and construct a non-situational world for themselves. Intuitions are not inherently wrong, but they are strongly situational (Gilovich, Griffin and Kahneman, 2002).

---

8 See: [www.rulesofthumb.com](http://www.rulesofthumb.com)
Theories of corporate bureaucracy vacillate between the idea that organizational structure can correct this inherent tendency to be situational, and another, that such bureaucracies are themselves situations (Simon, 1986; Friedberg, 1993). Theories of entrepreneurship do the same: from those that argue that the entrepreneur does rational arbitrage of unexploited opportunities to those that hold that s/he leaps into the void with courage and flair. In the end, it would seem that the very structures set up to control our worst situationalist errors end up creating new situations. The interaction of such structures and the inherent dilemmas of being situational is the reality for economic actors. Action is unavoidably contextual.

We can now propose a definition of “context.” Its structural component is defined by the division of labor in which the actor finds herself, which has a decisive influence on the informational environment for the individual, hence her “input” structure of cues and reference points. In turn, the individual actor engages in her search behavior (prospecting) and goal formation (emulation/aspirations), and does so with certain time horizons (cues and discounting), leading to choice/evaluation behaviors (rules of thumb and framing).

This is not to argue that this geographical architecture of interactions fully determines what actors then do, but rather that it influences what they do by defining the situation in which they find themselves. Following Piore (1995), we can say that the meaning actors make of their situation depends on a process of
interpretation, itself in part dependent on the interactions that occur in the situation. Piore argues that time is a critical aspect of this – a string of interactions and interpretations, rather than a set of instantaneous and history-less discrete calculated choices. To that notion, we can add the importance of the geographical context of such interactions that lead to “meaning-making” and choices.

**Different types of context**

Though there are potentially innumerable forces that structure differences in context, two of them that have a particularly strong influence on the set of relationships between an economic actor and her wider environment: the division of labor and the geographical distance between activities. They influence the actor’s relationships because they have strong effects on the five aspects of situation identified above. The economy’s vast organizational and geographical complexity generates at least five dimensions that structure the actor’s situation: the immediate environment of her task in the detailed division of labor; her environment in the work unit; the environment between closely-related work units, possibly inside an organization (e.g. a firm); intended regular relationships to the environment (intra- or inter-organizational or market); and unintended or unplanned relations to the external world. A key consequence of the geography of the division of labor is to define how an activity is placed in its wider environment. Thus, some resulting environments are more specialized
and homogeneous, while other pieces of the division of labor are located in places with diverse, more heterogeneous local activity mixes. All these relationships are shaped by geographical location, which is in turn strongly related to productive organization, i.e. the division of labor and the trade costs between parts of it.

A way to begin to see this is to extend the standard logic of scale and scope to the structure of the actor’s situation. The more organizationally internalized the actor’s relationships, combined with the effects of the internal division of labor in an organization, the more an actor’s context is intra-organizational and, possibly, task-specialized. In turn, this will direct her communication within the organized chain and tend to simplify communications to the local, external environment. At the other extreme, shallow or artisanal divisions of labor, a great deal of externalization and narrow scope, less “purified” task definitions, will generally lead to more diverse, irregular, and uncertain external communications. Locational economics suggests the latter will tend to be located in more “diverse” and urban local economies, the former in more specialized and less urban ones, all other things being equal (Duranton and Puga, 2002).

This is just a simplified vision of two ideal-types at opposite ends of a spectrum. There is rarely a clean division between “locationally-fragmented, highly organized, specialized contexts” and “highly diverse, market-oriented, dense communication contexts,” but rather some fascinating mixes of them. For
example, in a local context such as a city, we can find firms involved in high
levels of internal communication, using algorithms and rules, and highly-organized professional supply-chain relationships, as well as local, spontaneous, diverse interactions. The actor’s context, in this case, is some mixture of the two, which is as yet rather poorly understood. Likewise, in the geographically-fragmented, highly organized supply chain environment, where information is strongly “stovepiped” to actors,9 there is still likely to be some leakage of unplanned information from the local or long-distance environment. The core issue is how the information and signals that are the key inputs into the actor’s situation are defined and channeled, and then what actors do in different types of situations: how they learn, what choices they make, whom and what they emulate.

Let us now think about some examples of this phenomenon. The apparel industry has a high level of product differentiation, both because of functional and fashion diversity, and because of steep quality ladders. Market structures, organization, and geography are correspondingly diverse within this industry, which exhibits everything from geographically-fragmented chains of mass production within large firms to highly-localized, specialized clothing “clusters” or districts. The processes of emulation, learning and innovation are very different in the many contexts defined by such a variety of organizational forms. Skills do not transfer readily between them. There is “spillover” from one to

9 Here I am adapting the idea of “global pipelines” advanced by Bathelt, Malmberg and Maskell, 2004.
another, in that mass production tries generally to knock-off and copy what comes from fast-fashion or high-fashion, which in turn learns from an alternative circuit of cities, the arts, and the “street.” So in this case, there is probably some kind of complementary relationship between two very different organizational and geographical contexts.

Another example comes from the literature on organizational diversity within capitalism (Sabel and Zeitlin, 1985; Hall and Soskice, 2001). German, Japanese and American carmakers have significant differences in the organization and geography of their core activities; so do Samsung and General Electric for consumer electronics. Unlike the clothing industry, the products here are roughly comparable, so we are confronted with alternative organizational and geographical arrangements for substitutable products (Krugman, 1991). This might be due to the increasing market interpenetration of these industries and the international technology spillovers it facilitates and requires.

This might tempt some to conclude that context, in the way that we are defining it, is unimportant, since apparently it leads to the same ending point. But upon closer examination, these industries struggle with dynamic change in different ways and with varying degrees of success. No close observer would claim that Detroit, with its far-flung system of suppliers and arms-length relations, does as well as its German and Japanese counterparts in getting out new models with constant engineering improvements (Goldberg, 1995). So, even in a world of enormous technology sharing and circulation, the actors in each
system are keyed into at least some differences in information, and have
different emulation dynamics, possibly discounting (time horizons), and choice
behaviors. If not, universally-available information, world market structures,
and international technological spillovers would push all three toward more
organizational and geographical and performance convergence than in fact we
observe. In this case, contexts are not simply complementary, around imperfectly
substitutable products, but they are competitors, different ways of doing
something. There are feedbacks from context to development and performance,
a recursive relationship (Essletzbichler and Rigby, 2004; Saxenian, 1994).

A third example is the extreme case of winner-take-all systems in the
world, such as the City of London, Silicon Valley, or Hollywood, or industrial
districts built around a combination of distinctive products and locally-
constructed techniques rooted in the local system as a whole (Becattini, 1993).
These highly-successful clusters serve the global market and are caught up in all
sorts of long-distance, formal organizational procedures, professional and
regulatory norms, and so on. Some kind of alchemy occurs in their core regions,
extended through the local and world-wide informal and semi-formal networks
at whose center they lie. As with any sector, they are subject to product cycle and
quality ladder dynamics that can modify them organizationally and
geo graphically. At the extreme, this might lead to the “end” of their existing
local contexts and their replacement by a different type of long-distance,
organizationally-formal, big-firm oriented (plus suppliers, etc) “context.” We
think of this as a natural and unproblematic dimension of economic
development, with benefits for outlying regions in particular, and for the
economy as a whole through technological maturity and Ricardian gains to
trade. But in the case of such radical restructuring and change in context, how
sure can we be about these presumed welfare effects? The Romer growth theory
was offered earlier as the strong case for such positive welfare effects, through
economy-wide increasing returns. But, viewed in the light of a theory of context
as the basis for discernably different economic behaviors, a different type of
question is asked: how does the change in context alter the emulation, choice,
and learning behavior of actors, and hence the dynamic development of the
sector in question. If Hollywood were a global production network with no local
core, would it be the same industry as with its present organization and
geography? A producer of different, but better outputs? A producer of worse,
less welfare-enhancing outputs? A fair guess is that the situations of its actors
would have changed, and with them, the processes of emulation, cognition,
learning and choice. It’s not just a question of choosing alternative techniques,
then, but that such a change in techniques has unplanned effects on the
productive environment and hence on the industry’s dynamics – what its people
know, what they want to do, and what they ultimately do. The world movie
industry would not just be a change in technique, but a change in what it
produces (Martin and Sunley, 2006; Boschma and Lambooy, 1999; Jaffe et al,
1993).
Distributed context: a great transformation?

Finally, let us think about the contemporary process of globalization in the age of digital technology. Major changes are occurring in the feasibility and costs of managing long-distance relationships in the economy, and hence in the possibilities for fragmenting production, organizationally and geographically. Are we at the brink of a “great transformation” in local versus long-distance relations (Polanyi, 1944)? Such a transformation toward distributed contexts would be four-fold. First, previously-existing borders between using intra-organizational hierarchies and external markets to coordinate activity would be moved by the ways new technologies allow coordination and monitoring to occur. In production, the big vertically-integrated producers who were once necessary to coordinate complex transactions and eliminate bilateral hold-up risks in production, appear to be increasingly giving way to networked “heterogeneous” production. The notion of a great transformation assumes the “completion” of this process. Second, the role of geographical distance would be profoundly altered, with a lesser role for proximity, even for the most complex intra- and inter-organizational relations. In the past, integration had very big distance costs. The same forces that reduce the need for integration also reduce the need for proximity: much more efficient communication technologies that permit complex monitoring and contracting at great distance (Baldwin, 2006),
and permit the fragmentation of production systems at the fine-grained level of
tasks, rather than subsystems (Grossman and Rossi-Hansberg, 2006).

Third, and underlying the first two, the boundary between formal and
informal processes of coordination, contracting and monitoring would be
modified in favor of new forms. It would be possible to “quasi-formalize,” thus
combining ease of deal-making with long distance and flexibility – the “best of
both worlds” – and greatly reducing the relevance of the current trade-offs.

Fourth, information would be obtained via “global pipelines” (Bathelt,
Malmberg and Maskell, 2004). Such distributed contexts would be more
organizational than geographical, and the organizational basis would be
radically different from the past, owing to the possibility of managing complex
interactions via the radically increased intermediation of information
technologies. The five dimensions of context would be radically altered for the
actors in distributed contexts.

What we have just laid out is speculative. Much of the literature counters
this vision, or at least parts of it, by asserting the ongoing (even heightened)
complexity of deal-making, increasing and more volatile output environments,
thus requiring face-to-face contact, and social networks to carry out production
(Storper and Venables, 2004; Olson and Olson, 2000). Communities of practice
would remain central in this vision of things, and would become more important
in the core “nodes” of geographically-extended production systems in this
alternative view. But at the very least, it is worth considering this stylized scenario of what a great transformation might look like.

IV. Some economic aspects of context

We argue above that contexts have strongly recursive relationships to individual action. This means that outcomes are contextual, especially the development of knowledge through different use of know-how, learning, experimentation and choice. In the restaurant example, the products of the two systems are highly imperfect substitutes –with emphasis on creation of amazing concoctions in the USA, on basic flavors in France. If this is the case, then contexts create different pathways than can be achieved using an alternative organizational (hence contextual) form of production organization. It follows that the ongoing rearrangement of the organization and geography of production is much more than the interplay of trade costs and comparative advantage in seeking the most efficient organizational and solution to a given production (allocation) problem. This is because, by changing the availability of contexts, it would also be rearranging the possibility set for future development of the economy. Properly defined, the question of context cuts right to the heart of whether and under what circumstances processes of globalization (and economic development more generally) are truly welfare-maximizing or not.
Antecedents

There is considerable existing reflection on some of the dimensions of whether contexts are welfare-maximizing, though it uses different vocabularies from ours. For example, in economic geography, agglomerations are said to be sources of technological spillovers and to facilitate innovation, and in some cases to shape it “locally.” In development economics, there is a notion that a country or region needs to develop an “ecosystem,” consisting of a complex geometry of parts of the division of labor and supporting activities (notably in R&D, training, and so on), if it wants to move up the hierarchy of functions in the world division of labor. Behind both of these currents is a debate about whether an economy does better being specialized (Marshall) or diversified (Jane Jacobs), i.e. what kind of spillovers and interactions are maximized (Duranton and Puga, 2002).

Another idea is that economies move up the developmental hierarchy when they are both able to absorb widely-available knowledge, often said to be “codified” so that it travels well, as well as create their own innovations that cannot be easily copied (tacit knowledge). These notions are central to the geography of growth explored above in section II.

All of these contributions are of great interest, but notice that they are about specific places, and do not tackle wider questions about resource allocation and welfare effects of an increasingly inter-related global economic space, involving specialization and competition, and ongoing re-organization of productive activities between places. As noted, in mainstream trade theory, any
such questions disappear axiomatically. Trade and monitoring/contracting costs change and permit new forms of fragmentation and patterns of location, then they will lead to long-term gains to trade.

But if context is more than just a residual effect of transportation and communication costs, then even an efficient process of fragmenting and redistributing a production chain does not exclude the possibility that there will be other changes in the way the production system performs, in the areas of emulation, cognition, learning and choice (Arthur, 1994). Moreover, as the restaurant example suggests, the resulting products, even in the present, can be imperfect substitutes, so that there is not just a change in production organization, but in the envelope of outputs in some cases. In other words, the decision to restructure a production chain involves not just discrete choice processes today, but “roads not taken” (Arthur, 1989, 1994; Hodgson, 1998; Essletzbichler and Rigby, 2004; Dosi, 1998). We can compare the road taken to the existing road, but what we cannot measure is the possible qualities and quantities of future outputs of contexts that are being eliminated or restructured. Moreover, the creation or elimination of contexts is largely an unintended outcome of decisions made about the organization of production and location, with these dynamic effects not figured into the decision-making process. They are externalities.

Globalization and the Diversity of Contexts
Competition among contexts is viewed as the healthy basis of the contemporary globalization process. The digital revolution seems to be pushing this to new heights, as the possibilities for fragmentation and locational choice are vastly augmented by it. At the same time, new “local” economies are able to use long-distance knowledge spillovers and inputs from far away to learn new production tasks and insert themselves in the global economy of competition among products, effectively introducing another layer to the competition among contexts.

Is this competition among contexts neutral to the supply of products, today and in the future? Hence, is it neutral to the developmental and learning dynamics of both organizational communities (in a technological area, let’s say) and territorial economies? These are the questions that need to be tackled by economic geography and trade economics, but are currently out of their range, because context is not considered.

Consider an extension of our earlier speculation about a great transformation. The emergence of very large-scale distributed contexts might be reducing certain dimensions of competition among territorially-anchored contexts, in that the major feature of the new context for economic action becomes the spatially-extended network. In this case, there might be enhanced “within technique” competition among places, but within a world economy generating significant institutional convergence between production systems at a global level (single best practices), such competition is more about how places
deliver up resources to the global network context, hence reducing the importance of local or national extra-organizational context in generating a variety of techniques and products (Gertler, 2001). The information used by actors would be “stovepiped” via the global division of labor. This information would increasingly take precedence over information coming from outside such a network (e.g. from a local or national context of information exchange) in the key processes of emulation, cognition and learning. Certain formal models of trade with heterogeneous firms predict precisely this reduction in variety (Baldwin and Forslid, 2006).

This brings us back to issues of long-term economic development. One key perspective on why the modern economic take-off occurred in Western Europe argues that it was the combination of a balance between inter-relationship – the dense trading and political relationships that emerged during the mercantile period – with separation and competition – the proliferation of different technological systems and their products – that promoted competition and overcame the possibility of progress-killing local monopolies (North, 2005). Standard theories of globalization argue that a similar process is occurring today, at the global rather than continental scale. But if global distributed contexts are emerging and reducing such local and national environments to a subservient role of delivering up resources to global systems, with a strong degree of organizational convergence, then this assumption would be questionable.
If the process of globalization leads to elimination of some contexts in favor of others, mainstream economics would say that such a loss of diversity leads to a better output of goods and services, and that it is therefore justified in welfare terms. But now we have seen that this conclusion may or may not be merited: though there are clearly many circumstances where eliminating or changing contexts is justified in efficiency terms, there is the problem of imperfectly substitutable products, on one hand, and dynamics on the other. Mainstream economic geography and international economics has had virtually nothing to say about these issues.

V. Conclusion

The argument of this paper consists of four main points. First, the local or national differences in the ways production systems, labor markets, supply chains and relationships come together are, as in our restaurant example, just the entry point into their most interesting possible dimension, which is their dynamic properties – what people learn, develop, and innovate in them. This point is not new in the empirical sense: the vast literature on national and regional systems of innovation and on the role of geographical proximity in innovation suggests that certain kinds of innovations depend on proximity, while others can be carried out at long distances. We made a related, but different point: the qualities and types of product and innovations fundamentally
depend on their context. The question is not just about the line between local and non-local innovations, but about the contexts for all innovations.

A second point of this paper has been to ask what we mean by “context.” Most economists reject out of hand the possibility of significant differences in expectations and choices from one place to another, while geographers and sociologists tend to assume the importance of such differences without giving them precise analytical definition. Thus, most economists see local or national economies as outcomes of the forces that spatially distribute economic activity (trade costs plus comparative advantage): the residual effect is a local or national economic environment. Geographers and sociologists look for the “bottom up” sources of economic differentiation and dynamism, and sometimes consider how the local and the long-distance interact, but this leads to a curious (again largely implicit) assumption that territorial economies are shaped by “local” rationalities, and overlain on top of them are more abstract economic systems that are somehow not contextual, expressing instead some kind of systemic rationality. Further, the ongoing process of fragmentation of production (divisions of labor) and reintegration through trade is one of replacing contextual economics with non-contextual economics, with the implicit assumption that the latter is better than the former. Situationalism, based on strong results from observation and experimental evidence, shows that all economic action is contextual. In some cases, the context is principally structured by formal, long-distance relationships, while at the other extreme, it reflects highly localized,
informal and unplanned interactions (and then there is everything in-between). This leads us to the third point, which rejects the common opposition of “local contexts” to globally-abstract systems of production; instead, we have applied the analytical concept of context both to local economies and to new forms of geographically-extended and fragmented production, and ask how each actor-environment shapes economic learning and development.

Finally, though much research questions how the process of globalization affects economies, there is virtually no inquiry into the question we have asked in this paper, about whether the assumption that such organization and geographical changes may be presumed always to generate big gains is justified. Looked at from a situationalist and dynamic perspective, there may be reason to start asking such questions. In addition to this overall welfare issue, we may soon need to understand how radically geographically-distributed contexts function as economic systems, and as systems of context for economic actors and human creativity.

REFERENCES


Rules of thumb: www.rulesofthumb.com


