Abstract
Spatial heterogeneity at the city level is crucial for explaining local inequality. This heterogeneity will continue from a cultural perspective, an age profile perspective, and a productivity perspective. The process of catching-up to the technological leader is influenced by geographic proximity and local input-output linkages, particularly to upstream industries. Heterogeneity of urbanization patterns and spatial linkages affect FDI flows. The relative power of the economic agglomerating and spreading forces are not scale-neutral but heterogeneous, more specifically: spatial linkages are found to be more important at higher levels of aggregation.
countries’ on behalf of its 50th anniversary, 30-31 October, 2008, which provided the basis for this special issue of the Journal of Regional Science. This issue of the Journal of Regional Science provides a selection of the articles presented on that occasion.
1. Introduction
The Nobel Prize in Economics 2008 was awarded to Paul Krugman for combining two fields in economics (trade theory and economic geography). According to the noble prize background report: ¹

“Traditionally, trade theory and economic geography evolved as separate subfields of economics. More recently, however, they have converged [to] become more and more united through new theoretical insights, which emphasize that the same basic forces simultaneously determine specialization across countries for a given international distribution of factors of production (trade theory) and the long-run location of those factors across countries (economic geography).”

It was the second time that these two fields were jointly mentioned by the Nobel prize committee as the 1977 press release for Bertil Ohlin (joint with James Meade), stated that “Ohlin … demonstrated similarities and differences between interregional (intra-national) and international trade, and the connection between international trade and the location of industries.”

Evidently, the fields of international trade, focusing on interaction at the national level, and regional science, focusing on interaction at the regional or even city level, from time to time evolve separately and jointly. In the past two decades they have become more integrated than before. This paper briefly reviews recent contributions on the various economic forces active at different spatial scales fuelled by the observation that the majority of population and economic activity is now concentrated at the city level.

2. Urbanization and unbalanced growth
Currently, more than half of the world population lives in cities, as announced with some aplomb by the United Nations Population Fund (2007): “In 2008, the world reaches an invisible but momentous milestone: For the first time in history, more than half its human population, 3.3 billion people, will be living in urban areas.” Also according to the World Bank (2009a, p. 12) “The average size of the world’s largest 100 cities has grown to almost 10 times their size in 1900 .. , and almost two-thirds of these cities are in developing countries.”

¹ See the scientific background report by the committee “Trade and Geography-Economies of Scale, Differentiated Products and Transport Costs” which can be found at the homepage of the Nobel Prize at: http://nobelprize.org/nobel_prizes/economics/laureates/2008/index.html.
Figure 1 Urbanization and income per capita, 2004

Data: own calculations based on World Development Indicators online; regression line depicted.

Figure 1 illustrates why people move to the city by depicting the positive correlation between the degree of urbanization (per cent of total population) and the GDP per capita level. Empirical and theoretical work focuses on analyzing the underlying forces for this correlation, based on scale economies, knowledge spillovers, market access, proximity to specialized inputs, specialization, thick input markets, etc. ensuring not only that productivity is higher in urban centers but also that firms and consumers want to be located there for various other reasons (see The World Bank, 2009a and 2009b, for recent overviews).
Changing focus from the cross-section perspective of Figure 1 to the time dimension, Figure 2 illustrates the dynamics of economic development in conjunction with the associated urbanization process for the case of China. After Mao’s death (1976) and the reinstatement of Deng Xiaoping (1977), Economic Reform policies were adopted (December 1978) to expand rural incentives, encourage enterprise autonomy, reduce central planning, open up to international trade flows, establish foreign direct investment, and pass new legal codes. Since then, China’s PPP income levels in the period 1978-2006 have risen from 22 per cent to 83 per cent of the world average, while its degree of urbanization has risen from 18.7 per cent to 41.3 per cent.

Other countries have similar experiences as China. Consequently, a recent comprehensive study of the World Bank (2009a, pp. 5-6) concludes: “World Development Report 2009 has a different message: economic growth is seldom balanced. Efforts to spread it prematurely will

---

2 Note that the figure measures economic progress in China relative to the outside world by expressing China’s GDP per capita as a percentage of world average GDP per capita. Correcting for global developments (van Marrewijk, 2004) is crucial for understanding the links between urbanization and production (see Annez and Buckley, 2009, Figure 1.3 when this correction is not made).

3 China might not be using its full urbanization potential as migration to urban centers is limited by the government, with the consequence that cities might be ‘too small’ (see Au and Henderson, 2006)
jeopardize progress. Two centuries of economic development show that spatial disparities in income and production are inevitable.” This raises the question what forces are active on what spatial scale of analysis, and – as figure 1 illustrates – understanding these agglomeration patterns is of crucial importance to understand economic development and prosperity.

3. Economic forces at different spatial scales

The data strongly suggest a trend towards more agglomeration of economic activity within nations. This raises the question of what economic forces are active to stimulate agglomeration. Cities are attractive because they offer benefits compared to the countryside. Cities create markets that benefit not only consumers but also firms. A large variety of products, services and amenities offered in cities make living in a city attractive. Proximity gives firms and consumers an opportunity to organize daily interaction in an efficient way. City centers might be the location of specialized shops and services, which differ from the shopping availabilities for the daily needs in the outskirts of cities. For firms, cities provide markets for specialized products and services. These factors make cities attractive for both firms and consumers, and attracting new firms and consumers further stimulates growth. The firms located in cities, however, do not only serve the local market but also ship products to surrounding locations or across international borders. The theory of comparative advantage argues forcefully that specialization patterns differ between countries. These nation-wide characteristics of countries, however, cannot always be traced back to characteristics of a particular city or region. The comparative advantage of the flower sector in the Netherlands, for example, is the result of technological spill-overs of the agricultural sector, but also the access to a large airport, proximity to the (warm) North Atlantic Current (ensuring a mild climate), a specialized labor force, and institutions that favor the agricultural sector in general, making the flower sector relatively efficient. Comparative advantage is the collective result of all these country-wide characteristics, cities included. Interactions like these might explain why more urbanization within countries can be consistent with convergence of income per capita between countries (Puga, 2002).
Thisse (2009) points out that location theory can in principle be divided into three subfields: spatial competition theory, urban economics, and economic geography. The differences between these fields also stress the different forces that are active on different spatial scales. Spatial economics often deals with the location of firms on ‘main street’, urban economics focuses on cities. Economic geography is somewhat silent on the exact scale for which it is relevant. The latter has been applied to all spatial scales ranging from cities to nations Brakman, Garretsen, Van Marrewijk, (2009a). Combes, et al. (2008), however, argue that economic geography is most likely to be more relevant at larger spatial scales, such as countries. The economic forces that operate at different spatial scales are illustrated in Figure 3.

*Figure 3 Intra-regional, inter-regional and international economic forces*

Intra-regional connections: solid lines; inter-regional connections: dashed lines; international connections: dotted lines. The figure illustrates these connections only for cities 1 and 7.

The agglomerating and spreading forces tend to be different on each of these spatial scales. Within cities consumers and firms have to be near each other, because they interact on a daily basis. City specific agglomeration and spreading forces determine the ultimate sizes of cities. High densities produce the Marshallian externalities that favor the formation of cities. With high commuting and communication costs, that become larger with distance, cities tend in general to be smaller as economic interaction becomes more difficult. Larger cities might arise when communication and commuting costs decline and these large cities may in turn be surrounded by smaller residential areas (that might eventually become cities themselves). However, high densities also produce congestion.
costs that act as a spreading force. For regions, the mobility of firms and workers determines spatial patterns of economic activity. Differences in regional factor endowments become important and determine why specific regions are attractive for specific firms or workers, which in turn determines the specific functions of cities in that particular region. The quality of the infrastructure between regions becomes important on this scale and also institutions that foster or hinder mobility of production factors.

The interactions between nations – the field of international trade theory – are largely determined by transportation and communication costs in the broadest sense. As a stylized fact, production factors are often assumed to be mobile within countries, but not across borders. Empirical evidence confirms this notion. As is shown time and again, the level of transportation costs is a crucial variable for understanding international trade patterns (Brakman, Garretsen, and Van Marrewijk, 2009a). Countries that have a central position in a network of trading nations benefit the most from international trade and are able to make efficient use of the international division of labor. As transportation and communication costs become smaller, fragmenting the production process also becomes possible. Different stages of the production process can relocate to other countries. In general, this fragmentation process strengthens the existing pattern of comparative advantage (Neary, 2007).

4. Heterogeneity at different spatial scales
Different economic forces are active at different spatial scales. In this sense, the various forces that determine the spatial organization of economic activity at a particular spatial scale resemble the weak and strong forces of nature. These forces range from gravity, a relatively weak force (but noticeable on a large scale and responsible for phenomena like galaxies) to the forces that are active on an atomic scale, which are extremely strong, but have a small range. This special issue of the Journal of Regional Science highlights the
forces that are active on different economic scales, ranging from forces that work at the city scale to those that are relevant at the international level.\textsuperscript{4}

First, Glaeser, Resseger, and Tobio (2009) stress the importance of local inequality (e.g. income- and skills distributions) at the city level rather than the national level: crime rates are higher in more unequal cities and more people say they are unhappy. Consequently, once they control for the initial distribution of skills, local inequality and the growth rates of city-level income and population are negatively associated. These issues should be studied in the urban context taking the high mobility of workers between cities into consideration, which contrasts with the low mobility across nations. Hysteresis plays a role in determining skill inequality, through historical schooling patterns and immigration, which in turn is able to explain about one third of the variation in city-level inequality. Similarly, the historical context of local industrial patterns also partially explains the differences in returns to skills.

The next two papers focus on the interaction between different groups within and between cities. Ottaviano and Prarolo (2009) analyze the cultural economic forces for two mobile groups (‘diaspora’), each with a localized externality that enhances the productivity of individuals within the group in the same city. On the other hand, such interaction dilutes cultural identities and lowers the consumption of culture-specific activities. In general, the two groups segregate in different cities when communication at a distance is difficult and integrate in multicultural cities when such communication is easy. In their model it is possible that the two groups segregate in different cities while multicultural integration Pareto dominates this equilibrium. Arguably, the falling communication costs at a distance should further stimulate the emergence of multicultural cities.

The third paper by Gaigné and Thisse (2009) similarly focuses on the forces within and between cities, this time from an economic perspective of an aging population. Like the Ottaviano and Prarolo paper, there are two groups of individuals, namely workers (who

\textsuperscript{4} All articles have gone through the regular refereeing and revise-and-resubmit procedures.
produce and consume) and retirees (who only consume). There are two sectors, namely local services and manufacturing, such that the location decision of workers is determined by wages and the cost-of-living, while the location decision of retirees is only determined by the cost-of-living. They argue that the future of working cities is still bright and the emergence of rentier cities is unlikely because the costly supply of business to consumer services prevents complete segregation. The emerging multicultural city is thus going to be populated with workers and retirees at the same time.

The next two papers focus on productivity and spillovers. Griffith, Redding, and Simpson (2009) analyze the impact of geographic proximity on technological catch-up using a comprehensive British data set. Other things equal, establishments that are further behind the industry productivity frontier experience faster rates of productivity growth. This catch-up process accounts for about 9 per cent of annual growth for the median establishment. Moreover, they show that being geographically close to frontier firms leads to faster catch-up. Nonetheless, since the productivity gap with the regional frontier tends to be smaller than with the national frontier, the overall contribution of catch-up to the national frontier is larger than to the regional frontier (in their calculations a factor of three to six times more important).

The fifth paper by Lopez and Suedekum (2009) analyzes productivity and spillovers from a different perspective using data from Chilean plants. Their focus is on intra- and inter-industry agglomeration spillovers (within the same region). In the initial set-up they find, like previous studies, evidence for agglomeration spillovers within the same industry but not between different industries. This picture changes when they combine their regional information with the economy-wide input-output table, which they use as a proxy for vertical inter-industry linkages. Taking this additional information into consideration they show that there are important productivity spillovers from plants in upstream industries, but not from plants in downstream industries. These patterns suggest that the technological spillovers more or less follow the production processes downstream. Like the Griffith, Redding, and Simpson paper, this work therefore provides an important rationale for regional clustering of specific production processes.
The final two papers move to the national level, combined with an analysis at a lower level of aggregation, namely cities for the Poelhekke and van der Ploeg (2009) paper and regions for the Brakman, Garretsen, and van Marrewijk (2009b) paper. The sixth paper by Poelhekke and van der Ploeg focuses on the role of urban concentrations for Foreign Direct Investment (FDI) using data from American multinationals. In a novel setting, and controlling for the usual determinants of FDI, they show that urban concentration matters. In particular, a nation with several medium-sized cities attracts more FDI, but too much concentration in the shape of primacy hurts FDI. This suggests that both accessibility of market potential through urban connections and the number of potential investment locations are important channels for FDI to play its role as a means to spread technology and catch-up with the economic leaders. As in the other papers of this special issue, the spatial context is important. Taking spatial lags into consideration, they find that high FDI flows to surrounding countries boosts national FDI, while high surrounding market potential deflects FDI. Neighbouring attractive locations thus complement each other.

The final paper by Brakman, Garretsen, and van Marrewijk analyzes the relative importance of the economic forces illustrated in Figure 3 and the associated dominant theoretical explanations, namely urban economics and new economic geography, at two different spatial scales within and between European nations and regions. A main difference between the two theoretical approaches is that new economic geography stresses the role of spatial linkages whereas urban economics does not do so. They find not only that spatial linkages are more important at the country level than at the regional level, but also that the relative importance of the spatial economic forces varies over time.

5. Conclusion: heterogeneity

The main arguments and conclusions in the papers of this issue of the Journal of Regional Science can be characterized by spatial heterogeneity, which is in line with the focus of the most recent World Development Report (2009). Within-city heterogeneity is crucial for explaining local inequality (Glaeser, Resseger, and Tobio). This is likely to continue, both from a cultural perspective (Ottaviano and Prarolo) and an age perspective
Heterogeneity will also continue from a productivity perspective, not only because the catch-up process to the technological leader responds to geographic proximity (Griffith, Redding, and Simpson), but also because regional technological spillovers are based on input-output linkages to upstream industries (Lopez and Suedekum). These forces strengthen regional heterogeneity in the shape of clusters for specific processes. At the national level, urban heterogeneity (in the shape of different urbanization patterns) and the spatial distribution of FDI flows and market potential are all shown to lead to heterogenous FDI patterns (Poelhekke and van der Ploeg). Finally, the relative power of the various economic agglomerating and spreading forces are not scale-neutral but heterogeneous, more specifically: spatial linkages are found to be more important at higher levels of aggregation (Brakman, Garretsen, and van Marrewijk).

References


Thisse, J. (2009), “Toward a unified theory of economic geography and urban economics,” mimeo, CORE.