# Integration as a spatial institution: Implications for agglomeration and growth

by Tim Deeken and Ingrid Ott

# Integration as a spatial institution: Implications for agglomeration and growth[★]

Tim Deeken

Karlsruhe Institute of Technology (KIT) [*]

Ingrid Ott

Karlsruhe Institute of Technology (KIT) [❋]

and Kiel Institute for the World Economy (IfW)

July 7, 2014

**Abstract**

In this article we analyze the interdependent issues of urbanization, growth, and globalization by presenting key empirical facts and relevant underlying economic theories on each. We look more closely, but without providing a detailed formal analysis, at a model by Baldwin and Forslid (2000) that combines a seminal model from the endogenous growth literature (Romer 1990) with one from the new economic geography literature (Krugman 1991). In the analysis the significance of a sophisticated consideration of the concept of integration is pointed out. We investigate the issue of scale, scale economies, and density and the important role integration plays in these considerations as well. We especially argue that future research should more precisely focus on integration as a dynamic concept that does not only affect agglomeration and growth, but which is itself the endogenous outcome of various interdependencies and which complements the institutional settings of the territories that are linked to each other.

Key words: integration; spatial institution; agglomeration; growth

JEL: O4, R5

---

# 1 Introduction

The 2009 World Development Report by the World Bank opens with the statement "Production concentrates in big cities, leading provinces, and wealthy nations. Half the world's production fits onto 1.5 percent of its land" (World Bank 2009, xiii). This immense concentration of economic activity has its counterpart in the fact that urban areas currently account for more than 50% of the global population (United Nations, 2012). Economic activity and people are thus unlikely to be randomly distributed across space. Together these facts indicate the existence of benefits from concentration. These benefits are well known and date back to Marshall (1890) and his description of external economies. More specifically, these can be broken down to market size effects operating through forward and backward linkages, thick labor markets, and pure external economies like, for example, knowledge spillovers (Krugman 1998). However, if only these agglomeration forces were at work, the implication was that the whole world would end up in one gigantic agglomeration. In effect, these forces that are conducive to agglomeration are balanced by a variety of dispersion forces like the existence of land rents or pure external diseconomies (e.g. pollution, crime). The resulting economic landscape is the outcome of the tension between these opposing forces.[1]

In addition to this concentration of economic activity, recent economic history over approximately the past 70 years is furthermore characterized by the ongoing process of globalization in its various forms as well as by an unprecedented increase in per capita income levels. Agglomeration and growth are connected via integration (globalization) to each other. In this paper we present key empirical facts on each aspect and comment on some important models in the theoretical literature concerning these. We look in more detail, but without delving deep into the formal analysis, at a model by Baldwin and Forslid (2000) that combines an endogenous growth model of the Romer (1990) variety with a new economic geography model along the lines of Krugman (1991). This combined model incorporates all three characteristics of interest and also allows for a more sophisticated consideration of the impact of integration on agglomeration than is the case in traditional new economic geography models. The reason is that it not only considers the costs of trading goods, but also accounts for the costs of trading information, a lowering of which tends to weaken the agglomeration forces. A detailed understanding of the concept of economic integration is particularly relevant, as it is the "way to get both the immediate benefits of the concentration of production and the long-term benefits of a convergence in living standards" (World Bank 2009, 1).

The issue of size and scale economies, and the role integration plays herein is impor-

---

[1]Expressed differently, these forces are second-nature determinants for location decisions in contrast to first-nature or exogenous determinants like natural resources, climate or natural harbors.

tant in these considerations. For instance, the implication of many endogenous growth models that larger economies exhibit higher growth rates is not necessarily born out empirically (Jones 1995b). Another aspect is that scale economies might not be relevant at all levels of aggregation (i.e. city, region, nation). We especially argue that future research should more precisely focus on integration as a dynamic concept that does not only affect agglomeration and growth, but which is itself the endogenous outcome of various interdependencies and which complements the institutional settings of the territories that are linked to each other.

The paper is organized as follows. Section 2 presents stylized facts and theoretical issues on agglomeration, growth, and globalization. In Section 3 a model by Baldwin and Forslid (2000) is described that links these aspects. Section 4 deals with the issue of scale, size, and density while Section 5 derives policy implications from the insights of the previously presented arguments. Section 6 offers some concluding remarks and perspectives for future research.

# 2 Building blocks

## 2.1 Spatial concentration

The earliest urban structures date back to the time of the Neolithic Revolution (Bairoch 1988). Prior to the Industrial Revolution, however, "the urban way of life had for thousands of years been the exception, it now became the rule" (Bairoch 1988, 213). This tendency towards increasing urbanization continues today and reflects an ongoing global pattern as can be seen in Figure 1, which illustrates a positive trend that is expected to last during the next several decades.[2] Though the trend is global, there are marked differences between the various 'global regions'. Western Europe, the United States and Australia have crossed the threshold of more than half of the respective population living in urban areas before 1950 and now have urbanization rates between 75% and 90%. Africa and Asia on the other hand had urbanization rates of ca. 15% in

---

[2]The coarse aggregation scheme is chosen to ensure comparability of urbanization rates and GDP per capita throughout the subsequent section. Since there is no perfect overlap in data availability as regards urbanization rates (Figure 1) and growth (Figure 2), we adopt the notion of 'global regions' from the United Nations World Urbanization Prospect and apply it to the New Maddison Project Database (Bolt and van Zanden 2013). Details on this can be found in Appendix A. In Figure 1, the concept of 'urban population' is based on the definitions of the respective national statistical agencies and thus may vary across regions. For an alternative agglomeration measure that aims at enabling cross-country comparability see Uchida and Nelson (2010). Note also that the population forecasts (from 2012 onwards (United Nations 2012)) depend upon national census data that are also only comparable with restrictions.

1950 and are expected to cross the value of 50% in 2035 and 2020, respectively. Latin America has followed yet another path. It reached an urbanization rate of 50% in the early 1960s, had a higher rate than Western Europe in 2000, and currently has a rate of nearly 80%.
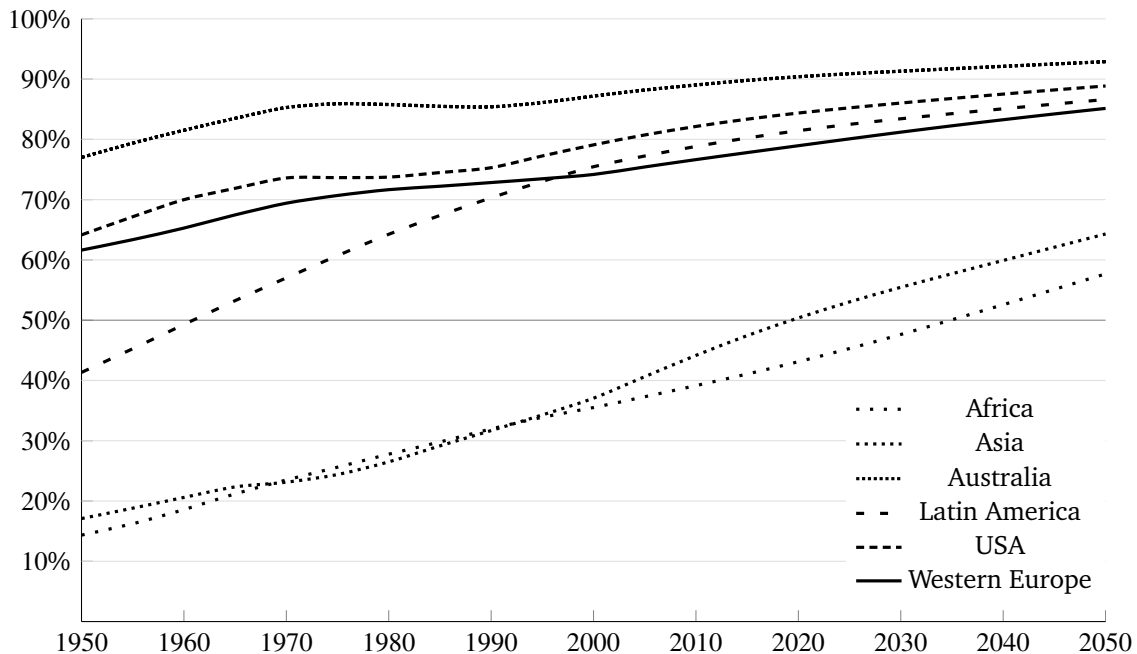


**Figure 1:** Urban population by major geographical area.

*Note:* The data are from United Nations (2012). See Appendix A for a list of geographical entities included in the aggregates.

These numbers for the urbanization rate provide little information however on whether people are living in large cities of several million people or if they are living in comparatively small cities of a few hundred thousand inhabitants. Table 1 remedies this situation to some extent by distinguishing the absolute number of urban residents according to five city-size classes.[3] What can then be inferred from combining the data in Figure 1 and Table 1 is that not only are more and more people living in cities, but they are also increasingly living in large cities. In 1970, for instance, approximately 14%

---

[3]Note that, since the entries in Table 1 are absolute numbers of urban residents, the data offers only limited information about the number of cities in each group. The 630 million people, who are forecasted to live in cities of over 10 million residents in 2025 might, for instance, be distributed more or less evenly over some 60 cities or on the other hand be concentrated in a few gigantic cities of 50 million residents and many comparatively small ones with "only" approximately 10 million residents. Any other combination that distributes 630 million residents over a number of cities with a minimum size of 10 million is also possible. More details on the city-size distribution can be found in (Gabaix and Ioannides 2004).

of the world's population lived in cities of more than 500,000 inhabitants, whereas the corresponding figure was approximately 26% in 2011 and will rise to approximately 33% in 2025.[4]

The share of the urban population living in cities with less than 500,000 inhabitants is forecasted to fall by 20 percentage points from 62% in 1970 to 42% in 2025, but the share of urban residents in cities with more than 10 million inhabitants is expected to increase from 3% to 14% over the same period. Combining these aspects with the empirical evidence on city-size distributions implies the emergence and even intensification of core-periphery structures at various levels of spatial scale.

**Table 1:** Evolution and forecast of total population in millions according to city-size classes.

|      | < 0.5m | 0.5-1m | 1-5m | 5-10m | >10m |
|------|--------|--------|------|-------|------|
| 1970 | 833    | 128    | 244  | 109   | 39   |
| 1990 | 1333   | 206    | 456  | 142   | 145  |
| 2011 | 1849   | 365    | 776  | 283   | 359  |
| 2025 | 1966   | 516    | 1129 | 402   | 630  |

*Note:* The data are from United Nations (2012).

In other words, the trend of ongoing concentration may be observed at different levels of spatial aggregation. Brakman et al. (2009, 13) neatly summarize this phenomenon by stating: "It appears that the highly uneven distribution of economic activity across space has a fractal dimension – that is, it repeats itself at different levels of aggregation." Put differently, ongoing concentration might be observed across several spatial scales whereupon from a global perspective the resulting core-periphery structure remains unchanged.

These facts on urbanization are strong indicators for the existence of 'local' scale effects. The resulting spatial pattern is the outcome of the location decisions of firms and households, and the underlying economic reasoning may be summarized as follows: Individuals are indifferent as regards relocation if benefits and costs are equalized. In other words, for any degree of aggregation a spatial equilibrium is reached whenever forces attracting people (so-called agglomeration forces) and those pushing off people (dispersion forces) are balanced.[5] From a slightly different perspective, a location may be attractive for people or firms due to characteristics that are external to these actors.

---

[4]The shares are calculated from the data in United Nations (2012, Figure II and Table A.5).

[5]This argument holds with the caveat that in formal theoretical models it is only valid for interior equilibria, but not for equilibria in which complete agglomeration occurs.

The associated (external or) agglomeration economies that act as attractors for firms and individuals to a certain place date back to the seminal work of Marshall (1890), who identified three main sources of spatial concentration processes:[6] (i) *Market-size effects:* Local concentration leads to large local markets and vice versa. Via 'forward' or 'cost linkages' a local concentration benefits both consumers, who profit from more varieties and lower prices, as well as firms, since the local production of intermediate goods reduces the costs of downstream producers. On the other hand, firms gain from producing in a large market with good access to customers ('backward' or 'demand linkages'). (ii) *Thick labor markets* ease the matching problem between supply and demand of (today frequently specialized) labor. (iii) *(Pure) external economies* that are the more distinct the more densely populated a region is, since proximity allows for more frequent interaction and increased information spillovers. Infrastructure fixed costs are also spread over more heads. These forces mostly cover *specialization advantages* as the considered actors can benefit from scale economies and the related cost decreases, which usually are meant to apply within single sectors.[7]

Equally important for the attractiveness of a certain location, though not specifically mentioned in the context of Marshall's triad, are: (iv) *Selection effects* that occur in highly competitive markets where only efficiently working firms are able to survive. This effect is reinforced, since efficiency acts as an attraction point for internationally mobile capital and/or frequently highly qualified labor. (v) *Diversity:* Especially in light of the aforementioned increase of large urban agglomerations (Table 1), it is quite reasonable that the economic centers are not only characterized by specialization in a single but in several different fields at the same time. Groups of interconnected companies together with the supporting institutions, which are all located close to each other, are frequently said to form a cluster (Porter 1990). In particular, if several clusters are co-located, it is not just specialization, but in contrast the opportunity to interact with various – also heterogeneous – actors that attracts new firms and people. In this connection, Jacobs (1969) was the first to point to the positive effect of the co-location of diverse actors.[8] The resulting productivity and consumption economies are associated

---

[6]In this context, Alfred Marshall shaped the notion of the 'industrial districts' – a formulation that must be understood in light of the 19th century's economic conditions. The basic mechanisms, however, are still valid today only that aside from the industrial sector an important share of value creation is realized within the service sector. Marshall's concept has been picked up by Arrow (1962) and Romer (1986), who introduced it in theoretical models. As a consequence, agglomeration economies that are related to spillovers and (industry) specialization are mostly denoted as Marshall-Arrow-Romer (MAR) externalities; see also Krugman (1998) for a compact overview.

[7]Considering current production conditions, this includes the service sector and the joint use of research infrastructure.

[8]The associated external effects are nowadays called Jacobs or urbanization externalities (thus being distinguished from the previously mentioned MAR or specialization/localization externalities).

with the diversity and intensity of economic activity. In addition, large agglomerations also are more efficient in solving the matching problem that may arise in the factor markets thus linking the reasoning to the aforementioned thick labor markets and adapting them to a more dynamic environment that also might account for structural change. It is therefore reasonable to assume that as a consequence of 'density', the corresponding region's productivity increases and thereby drives local productivity above average.

But if only agglomeration forces were at work, the spatial equilibrium would result in a unique agglomeration of economic activity thus contradicting the empirical facts summarized in Figure 1 and Table 1. In fact, there exist opposing dispersion forces that are mainly based on the following sources: (i) *Immobile factors* such as land, natural resources, but also workers, imply that industries have to some extent go to where factors and their owners are located. At the same time, institutional arrangements may also foster or hamper the mobility of factors. Two arbitrary descriptive examples are the four freedoms within the EU or the necessity of work permits for non-residents in the US. (ii) *Land rents* increase as a result of concentration both for producing firms, but also for private individuals, who seek to minimize commuting costs. In the extreme, increasing land rents may even lead to the relocation of economic activity not only from the core to peripheral regions, but also as regards the location of different operating areas of firms within urbanized areas (functional specialization, Duranton and Puga 2000). (iii) *(Pure) external diseconomies of scale* such as congestion, pollution, and crime also hamper concentration.

As argued before, a spatial equilibrium is reached when both, agglomeration and dispersion forces, are balanced. Ongoing urbanization in this context implies the emergence and reinforcement of core-periphery structures in economies with increasing population size. In Section 4 we relate the identified core-periphery structures more precisely to scale effects and size in growing economies.

## 2.2   Growth and prosperity

Aside from the spatial concentration of population and production at certain locations, another major trend since the era of industrialization is ongoing economic growth as measured by gross domestic product (GDP) per capita. Figure 2 depicts GDP per capita based on 1990 international dollars[9] for the period 1850-2010 and strikingly highlights the world-wide growth story for the same major global regions as in the last

---

[9]In this unit 1 international dollar has the same purchasing power as $1 US had in 1990, and the GDP per capita values in other currencies are adjusted by purchasing power parities. Keeping in mind the difficulties of estimating these time series (see e.g. Bolt and van Zanden (2013)), it is nonetheless to a certain degree possible to compare the values across time and countries.

Subsection 2.1.[10] Around the world, centuries of Malthusian stagnation in which per capita income was near the subsistence level precede the Industrial Revolution until economies entered a period of sustained economic growth.[11] The positive trend might already be discerned in the 19th century. However, the effect of exponential growth becomes distinctly visible after World War II, though there are clear differences between the various illustrated global regions.
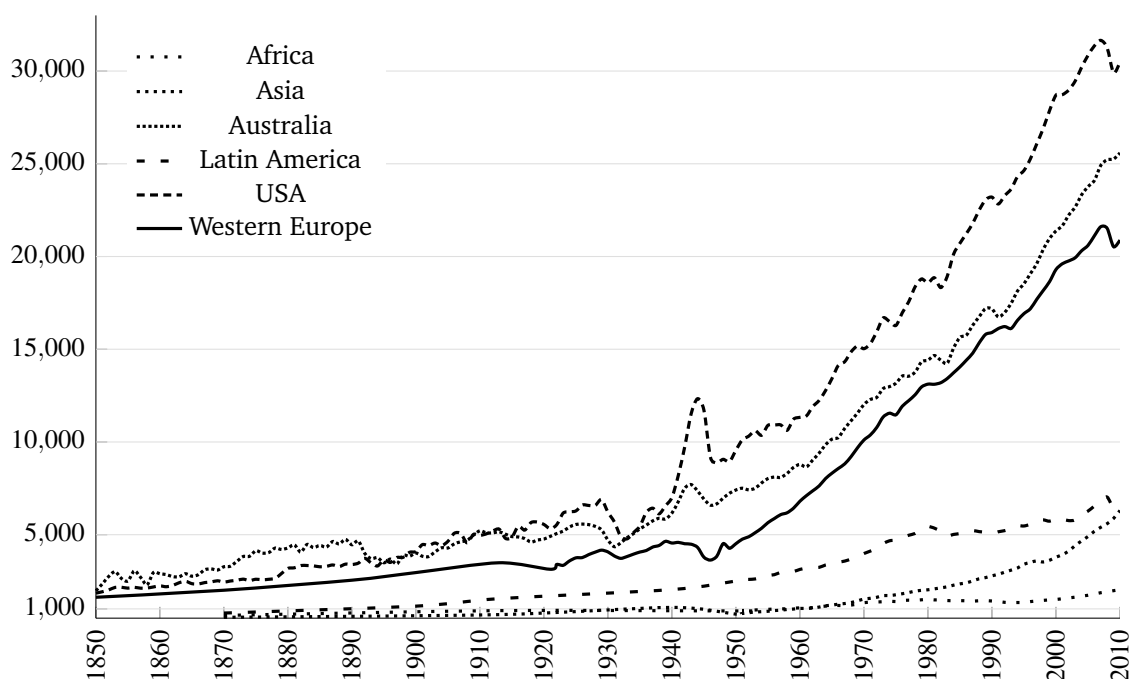


**Figure 2:** GDP per capita in 1990 international dollars.

*Note:* The data are from the New Maddison Project Database (Bolt and van Zanden 2013). See Appendix A for a list of geographical entities included in the aggregates.

The dominating positive trend for the global regions tends to hide drastic differences in income at the less aggregated level of individual countries. At the top, the USA has a GDP per capita of $30,491 in 2010 while the poorest country in the Maddison sample in that year is the Democratic Republic of the Congo with a GDP per capita of $260. Another noteworthy characteristic of the data is the possibility that a laggard overtakes the leader in the GDP per capita ranking ('leapfrogging'). This is exemplified by the USA becoming richer than the previous leader Australia in 1899.[12]

---

[10]The New Maddison Project Database provides information on the economic performance of countries and geographic regions from the year 1 onwards (at least for a subset of countries). The selected regions were chosen to allow for an immediate comparison to Figure 1.

[11]See, for instance, Galor (2005, 2011) and the references therein for empirical evidence and a theoretical model explaining this transition to sustained growth.

[12]For additional empirical evidence on the changing leaders and laggards for the period 1-2003,

The pattern of ongoing growth in Figure 2 has been accompanied by additional economic regularities. Kaldor (1957, 591) stated for capitalist economies that there are "remarkable historical constancies revealed by recent empirical investigations." His observations led to the nowadays well recognized 'stylized facts of growth' (Kaldor 1961).[13] At the same time, these observations set the frame for a research agenda to develop a consistent, comprehensive theory, which led to the *neoclassical growth model* (Solow 1956, Swan 1956). Based on the empirical observation of ongoing economic growth – often interpreted as increasing prosperity that in turn allows for a life above subsistence level – economic theory in general tries to gain a deeper understanding of the driving forces of this process. The major goal is to identify the underlying determinants, to comprehend their interaction, and, given any indications for forces hampering this process (e.g. market failures), to derive appropriate policy recommendations in order to maximize overall welfare. Growth thereby refers to the evolution of the GDP per capita, mostly analyzed at the country level, which results from the accumulation and use of more and/or better inputs in the production process. Across time and in spite of its rich explanatory power of the stylized facts detailed in Footnote 13, the neoclassical growth model has been criticized with respect to several dimensions (e.g. Jones and Romer 2010). It only covers one state variable, namely physical capital. The lacking microfoundations imply an exogenously assumed savings rate, since the resulting growth rate is not derived from individual optimization behavior. The model is also not able to explain the variation of growth rates between countries. In the model's equilibrium, growth per capita comes to a standstill with the consequence that the empirically observed positive growth rates can only be explained by exogenous technological progress. Due to its specification, technological change naturally remains a black box so that ultimately no clear-cut policy recommendation can be derived from the model in this respect.

Solow's basic growth model set the ground for several extensions, among them the inclusion of human capital or productive governmental activity. Further efforts led to the emergence of *endogenous growth theory* in the mid 1980s, which also provided insights into the utility maximization–growth nexus (i.e. the derived aggregate growth rate is based on individual optimization decisions (microfoundations)) thereby especially addressing the role of technological progress, human capital accumulation, or the role of institutions. Recent discussions of growth theorists distinguish 'proximate' from 'funda-

---

compare e.g. Figure 9 in Brakman and van Marrewijk (2008).

[13]In detail, the Kaldor facts are the following: (i) Labor productivity has grown at a steady rate, (ii) capital per worker has continually increased, (iii) the real interest rate (return on capital) has been stable, (iv) the capital-output ratio has been constant, (v) the shares of capital and labor in national income have been stable, and (vi) among the fastest growing countries the growth rate has varied in the range of 2.5%.

mental' causes of growth. The latter cover conditions such as luck, geography, culture, and institutions while the former refer to production inputs such as physical and human capital as well as to their overall productivity, which is enhanced by technological progress.[14]

These new theoretical frameworks have now themselves undergone an empirical assessment. Again the goal has been to review whether or not they are suited to contribute to a better understanding of the empirical facts and thus to link theory and empirical research within a consistent framework. Jones and Romer (2010) recently revisited the aforementioned Kaldor facts and updated them thereby also accounting for structural change of the last five decades.[15] Summarizing these findings leads to the recognition that in order to match the requirements of today's empirical results and to develop a consistent theory, aside from physical capital more state variables – namely human capital, ideas, population – and also institutions as reflecting the "rules of the game" (North 1990, 3) need to be considered. In the context of growing economies, the important role of institutions is emphasized in the seminal work of Acemoglu et al. (2001) and highlighted as well by, for instance, Rodrik et al. (2004).[16]

Altogether, within growth theory usually the platform of analysis is aggregate economies (countries or continents) where spatial components are not explicitly considered.[17] In case they are, the notion of 'geography' refers to natural conditions, which include a country's endowment with natural resources or its climatic conditions (e.g. Gallup et al. 1999, Dell et al. 2012). Put differently, as regards the spatial dimension, both the models and the stressed empirical studies refer to what sometimes is called 'first-nature geography', but do not make explicit man-made 'second-nature geography' conditions in a territorial sense.[18] The latter might, however, to some extent be understood as or be linked to institutional settings. Especially in Section 4 we argue that the degree of

---

[14]Excellent overviews on issues related to economic growth are provided by Acemoglu (2009) – with a special emphasis on the distinction between fundamental and proximate causes of growth – or more broadly by Barro and Sala-i-Martin (2004).

[15]In detail, the Jones-Romer facts identify the following empirical regularities: (i) Increases in the extent of the market, (ii) accelerating growth, (iii) variation in modern growth rates, (iv) large income and total factor productivity differences, (v) increases in human capital per worker, (vi) long-run stability of relative wages.

[16]Note that in many discussions and theoretical models, the concept of institutions is often connected to the presence of property rights. Throughout this paper, however, we assume a broader interpretation and embed institutions as regards their impact on the degree of integration – as later e.g. in Section 3 where institutions are seen as being related to the variable 'freeness of trade' and the environments that allow for knowledge spillovers.

[17]A winged word in this context is that 'the world is flat', which refers to the title of a book by Friedman (2005).

[18]An exception is Bosker and Garretsen (2009), who demonstrate that second-nature (or relative) geography interpreted as the institutional quality of neighboring countries has an impact on a country's

integration might be interpreted as reflecting an explicit spatial dimension of institutions. In addition, Jones and Romer's (2010) stylized fact of 'increases in the extent of the market' implicitly incorporates a spatial dimension, as the observed increase is the outcome of both growth (at any place) and access to foreign markets due to globalization which extends the relevant market for any economy involved in (international) trade. To conclude: Until today, spatial aspects are, if at all, only implicitly or indirectly addressed in most models of endogenous growth.

## 2.3 Globalization

A glimpse at the discussion so far suggests that "agglomeration can be considered the territorial counterpart of economic growth" (Fujita and Thisse 2002, 389). However, this viewpoint is too simplistic and without any explanatory power as regards the underlying interdependencies that drive this result given that countries or regions are not isolated actors, but are increasingly embedded in international value-creation processes in which due to globalization the places of production and consumption frequently are not located within the same national borders. Globalization in this context covers several dimensions, namely mobility of goods, people, ideas, and capital.[19] The exchange of these items is the immediate outcome of increased integration.[20] Important grounding in this context is given by rules that facilitate international/cross-border economic activity like the reduction of trade barriers, reduced transportation costs or the recognition of foreign degrees. The key driver that fosters the joint emergence of growth and agglomeration, however, is the increase in the international trade of goods (compare Baldwin et al. 2001). Concerning mobility of people, migration mainly occurs within, but also between countries (World Bank 2009, 147), and global migration is increasing at least in absolute numbers. The world-wide stock of migrants has risen to 165 million in 2000 from a starting value of 92 million in 1960, although the migrant's share of the world population has fallen from 3.05% to 2.71% over this period (Özden et al. 2011, 15). For migration flows, analysis by Abel and Sander (2014) suggests that for the three 5-year periods from 1995 to 2010 a relatively stable share of approximately 0.6% of the world population has migrated internationally. With respect to the mobility of ideas, Jones and Romer (2010, 229) note that, for instance, the change in the share of patents granted to non-US entities by the U.S. Patent and Trademark Office can be

---

economic development.

[19]The latter, for instance, via foreign direct investment (FDI). The concept may be understood even more broadly as in the KOF Index of Globalization, which includes, amongst other things, information on cultural proximity (see `http://globalization.kof.ethz.ch/`).

[20]See Meissner (2014) for additional details on the concepts of globalization and integration.

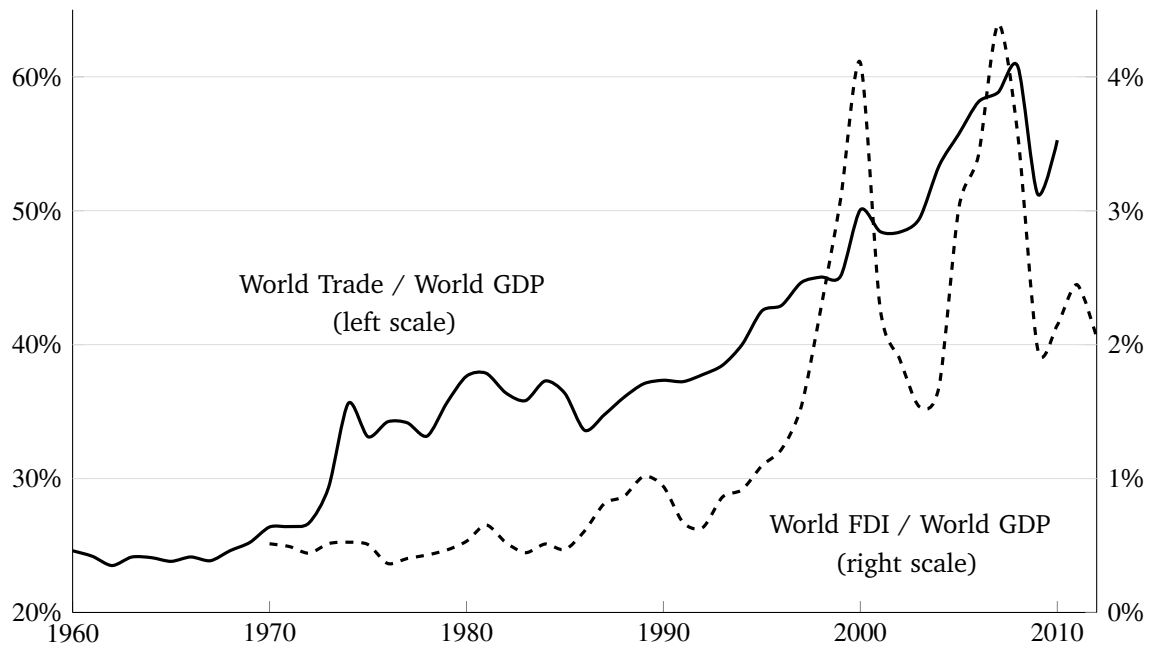interpreted as an indicator for the international flow of ideas. This share has increased from 18% in 1963 to 52% in 2012.[21]



**Figure 3:** Globalization.

*Note:* World FDI over World GDP is series BX.KLT.DINV.WD.GD.ZS from the World Bank's *World Development Indicators* (World Bank 2013), and the variable World trade / World GDP is calculated from Heston et al. (2012).

Figure 3 illustrates the phenomenon of globalization along the two dimensions 'trade' (mobility of goods) and 'FDI' (mobility of capital), which both have significantly increased during the last several decades although not steadily so. The sharp drop around the year 2000 coincides with the bursting of the dot.com bubble, and the onset of the recent financial crisis has led to a setback in the trend towards increasing globalization. These effects are particularly stark for the series of the ratio of World FDI to World GDP.

The increase in international trade over the past 50 years shown in Figure 3 overlaps with a period of ongoing trade liberalization. One measure for liberalizing trade are preferential trading agreements, and the cumulative number of these agreements in force has increased from a value in the single digits in the early 1950s to nearly 70

---

[21]Compare the column 'Total Patent Grants, Foreign Origin Percent Share' under `http://www.uspto.gov/web/offices/ac/ido/oeip/taf/us_stat.htm`. The exchange of ideas is also immensely fostered by the fall in communication costs and the accompanying spread of the internet to which approximately 35.6% of the global population had access in 2012, whereas the corresponding shares were 0.05% in 1990 and 6.7% in 2000 (World Bank 2013; series IT.NET.USER.P2).

in 1990 before reaching almost 300 in 2010 (World Trade Organization 2011, 54-55).[22] On a closer look, the increase in world trade over time hides an important aspect concerning the evolution of the composition of world trade. The share of *intra-industry* trade as opposed to *inter-industry* trade has increased as well. Brülhart (2009, 426) notes that the share of global *intra-industry* trade rose from roughly a quarter in 1962 to over 50% in 2006.[23]

A noteworthy aspect that accompanies the increasing share of *intra-industry* trade underlying this development is the composition of this form of trade in horizontal and vertical versions (see e.g. Thom and McDowell 1999 for this distinction). *Horizontal* intra-industry trade refers to trade in differentiated products at the same processing stage in a sector and is closely related to the love-of-variety effect that is, for example, present in the model discussed in the subsequent section.[24] *Vertical* intra-industry trade on the other hand takes place within a given sector, but at different stages of processing and is linked to comparative advantage and its resulting specialization. Distinguishing between final and intermediate goods instead, Brülhart (2009) notes that trade in final products has increased globally from a starting value slightly above 10% in 1962 to roughly a third in 2006, and the corresponding series for intermediate goods has followed the one for final goods closely until the mid 1970s, after which the shares were consistently higher and reached about 40% in 2006.[25]

On a theoretical level, neoclassical international trade theory in the form of the Ricardian and Heckscher-Ohlin models relies on the presence of comparative advantage due to different technologies or relative factor abundances to explain *inter-industry* trade between countries. However, neither is all trade between countries of the *inter-industry* variety nor is all trade between countries that differ in income, size, and relative factor endowments. Germany, for instance, exports cars to France and vice versa. In order to explain this *intra-industry* trade that takes place even in the absence of comparative advantage various assumptions of the neoclassical trade theory need to be relaxed. This is done in the field of new trade theory, which, for instance, no longer assumes that firms produce with a constant returns to scale technology and operate in an environ-

---

[22]This is not to say that trade liberalization is the only or most important factor in the increase in world trade. See Baier and Bergstrand (2001) for an analysis of the relative importance of trade liberalization, transport costs, and other factors in this development.

[23]The underlying measure for *intra-industry* trade is the Grubel-Lloyd index at the 3-digit Standard International Trade Classification (SITC) level. At the more detailed 5-digit SITC level *intra-industry* trade has increased from ca. 10% to about 30% over the period.

[24]This effect is however not constrained to final goods, but is also relevant for intermediate goods or services in firms' production processes (Hewings and Oosterhaven 2014, 912-913).

[25]Data in Brülhart (2009) distinguishes only between trade in intermediate and final goods, but not between *horizontal* and *vertical* intra-industry trade. The shares are for the 5-digit SITC level.

ment of perfect competition. Instead, increasing returns to scale at the firm level and imperfect competition are introduced.[26] The second change concerns the assumption of a homogeneous product in the neoclassical trade models, which is replaced in new trade theory by heterogeneous goods which are assumed to be imperfect substitutes for the consumers who exhibit love-of-variety preferences that are captured via a constant elasticity of substitution function.[27] The most widely adopted model of monopolistic competition with these characteristics is the one introduced by Dixit and Stiglitz (1977). Notice that these assumptions also mirror the transition from neoclassical to endogenous growth theory discussed in Subsection 2.2.

Increased freeness of trade also enhances international competition between firms and industries and induces possible relocations of economic activity allowing for production at the most productive places. The consequences on the spatial distribution thereby involve two dimensions. Concerning *inter-industry* trade, industries that exhibit a comparative advantage will grow, whereas the disadvantaged industries will shrink. As regards *intra-industry* trade, liberalization increases international competition, and some firms are not able to cope with these conditions. Empirical studies have shown that the corresponding reallocation of firms is more pronounced as regards *intra-industry* than *inter-industry* trade (Brakman et al. 2009).

Analogous to the reasoning in Section 2.1, the resulting spatial equilibrium is the outcome of the interaction between local increasing returns to scale and trade costs in which agglomeration and dispersion forces offset each other. Reduced trade costs (which enhance the freeness of trade) are frequently interpreted as being the outcome of increased economic integration between formerly more or less autarkic economies. The driving force in all these models is a reduction of trade costs or put differently, increased integration. Trade costs are mostly argued to capture a reduction of trade barriers such as duties or other non-tariff barriers. Major components are also transportation costs that have undergone a significant decline throughout the last several decades.[28] This implicitly incorporates aspects of technological change, which allows e.g. for an increase in shipping capacities, but also reduced communication costs that are due to improvements of information and communications technology. Nevertheless, within the considered models and empirical studies the degree of integration is assumed to be exogenous. In Section 4 we argue that integration itself undergoes an

---

[26]As Davis (1995) notes though, it is not necessary to assume increasing returns to scale at the firm level, since *intra-industry* trade can be accounted for without this assumption based on comparative advantage.

[27]These preferences imply that consumers receive more utility from consuming e.g. one unit each of seven different varieties instead of seven units of one particular variety.

[28]The decline in costs is clearly present for air shipping, but less pronounced for ocean shipping as Hummels (2007) notes.

evolution together with the processes of growth and agglomeration. In order to capture the implications, integration should be related to the institutional view detailed throughout Subsection 2.2. A more pointed view would be that integration is a fundamental cause for the spatial shape of the economic landscape and has explicitly to be seen as a dynamic institution with a strong spatial dimension.

# 3   Model

Section 2 described a variety of stylized facts for three defining characteristics of 'modern' economic history: Growth, agglomeration, and integration. We detailed these several lines of argumentation in the former part of the paper. Due to the dynamic perspective and the various interacting effects, their final impact on the spatial distribution of economic activity in growing economies is far from being trivial. To provide some guiding to disentangle the main lines of reasoning within a consistent model, this section presents, without going deeply into its formal structure, a model by Baldwin and Forslid (2000) that highlights the connections between these three characteristics and thus attempts to account for the joint endogeneity of the location of industrial activity and long-run economic growth.

Baldwin and Forslid's starting point is the two region core-periphery model developed by Krugman (1991), which takes care of agglomeration and integration. In this model there exists a traditional sector with constant returns to scale and perfect competition and an industrial sector with increasing returns to scale at the firm level, which is characterized by monopolistic competition. Goods from both sectors are traded, although trade costs occur only in the manufacturing sector. Whereas the global labor supply and the labor supply in the traditional sector are fixed, labor in the industrial sector is mobile, and the interregional labor distribution in this sector is determined endogenously, since migration depends on differences in the real wages in the two regions.[29] This model setup produces circular (cumulative) causation, which can be broken down to three forces that ultimately depend on the degree of integration as measured by transportation/trade costs. Two of these, the backward and forward linkages, are agglomeration forces and have already been described in Subsection 2.1 under the label market-size effects. The third force is the competition or market-crowding effect, which works against agglomeration, as firms prefer to locate away from their competitors. Whether or not a process of cumulative causation will be set in motion, depends on the relative strength of these forces. If, for instance, agglomeration forces are stronger

---

[29]In a slight modification to the Krugman (1991) model, Baldwin and Forslid (2000) allow for forward-looking behavior in the migration decision instead of static expectations.

than the dispersion force, then a shock to the system has the result that all industrial activity will locate in a single region, whereas the immobile factors of the traditional sectors determine the economic power of the periphery.

Baldwin and Forslid (2000) combine this framework with the endogenous growth model by Romer (1990) thereby incorporating the third characteristic. Both models are built upon the Dixit-Stiglitz approach to modeling monopolistic competition (Dixit and Stiglitz 1977) mentioned in Subsection 2.3. This identical understructure greatly facilitates the integration of the two separate models. The new aspect in Baldwin and Forslid's model is to not only consider the implications of changes in the trading costs for goods as in standard new economic geography models, but also the implications of changes in the trading costs for ideas. New ideas or knowledge are the driving force of economic growth in the model, and they ultimately show up in the form of an increasing number of varieties produced in the monopolistically competitive manufacturing sector. Producing a new variety requires a fixed cost of one unit of capital $K$ in the model in addition to a variable cost for labor. Capital in this model is viewed as "new knowledge embedded in a manufacturing facility that is immobile across regions" (Baldwin and Forslid 2000, 310). Production of capital requires only labor as an input and takes place in the model's third sector, the investment good or innovation sector, which is characterized by perfect competition. The crucial feature of this sector is the presence of knowledge spillovers or technological externalities in the sense that the unit labor requirement falls with an increasing level of production in the investment good sector. The specific distribution of manufacturing activity over the two regions moreover has a bearing on the extent of knowledge spillovers. More precisely, knowledge accumulated in a given region, e.g. the north, is more beneficial to firms in the north than knowledge accumulated in the south. This specification receives empirical support by, for instance, the work by Eaton and Kortum (1996, 276), who demonstrate for the OECD that even though there is substantial diffusion of technology between countries, large impediments to its diffusions exist that "are sufficient to generate large differences in productivity across countries." Put differently, location matters for firm productivity.

Baldwin and Forslid formally model production in the innovation sector via the following production function:

$$Q_K(t) = \frac{L_I}{a_I(t)}; \qquad a_I(t) = \frac{1}{K(t-1) + \lambda K^*(t-1)}; \qquad 0 \leq \lambda \leq 1$$

in which the variable $Q_K(t)$ denotes the flow of new capital at time $t$, and employment in the investment sector is denoted by $L_I$. $K$ and $K^*$ are capital in the north and south, respectively, and the parameter $\lambda$ signifies the degree of knowledge spillovers. The

one-period lag, $t-1$, for the capital in the two regions in the expression for the unit labor requirement in the investment goods sector, $a_I(t)$, indicates that it takes a certain amount of time until knowledge produced in one region becomes available in the other region.[30]

A complete analytical derivation of the results in Baldwin and Forslid (2000) is outside the scope of this article. Suffice it to say at this point that the model's dynamics can be completely described by a system of three difference equations in the variables for labor in the north, $L(t)$, the north's share of global capital, $K(t)/(K(t)+K^*(t)) \equiv \theta_K(t)$, and the shadow value of migration, $W(t)$. As in the standard core-periphery model, three stable long-run equilibria exist:[31] One in which manufacturing activity is spread symmetrically between the north and the south, and in the other two equilibria manufacturing activity is either completely agglomerated in the north or the south. The stability of these equilibria is verified via an analysis of the eigenvalues of the Jacobian matrix of the system of difference equations linearized around the steady state under consideration.

Despite the fact that the model is analytically rather complex, it has the neat feature that the stability properties of the respective equilibria depend only on two parameters. One is the already mentioned $\lambda$, denoting the degree of knowledge spillovers. The higher its value, the less localized are the technological externalities. Hence, this parameter can be interpreted as representing the cost of trading information. The second parameter, $\phi$, is an index capturing the notion of freeness of trade and can be shown to vary between 0 and 1.[32] This specification makes it possible to summarize information on the stability of the equilibria in the comparatively simple diagram shown in Figure 4, which is divided into three sections.

Analysis of Figure 4 shows that the results from the standard core-periphery model of the Krugman variety carry over to the growth-augmented model. A fall in transport costs (equivalently a higher value for $\phi$) has the implication that agglomeration becomes the only stable equilibrium.[33] This process is illustrated by the horizontal line with the three arrows in Figure 4. For a high enough level of knowledge spillovers, the economy moves from a situation in which only the symmetric equilibrium is stable (Section *I*) through a situation in which both agglomeration and the symmetric equilibrium are stable (Section *II*) to one in which only agglomeration is a stable equilibrium (Sec-

---

[30]Baldwin and Forslid (2000, 313) take one period to last approximately 10 years.

[31]Additional unstable interior equilibria exist as well.

[32]For infinitely high trade costs the index is zero, whereas in the absence of trade costs it equals one.

[33]Note that Figure 4 only establishes that agglomeration is a stable equilibrium, but not which region is or becomes the core. To determine this, the initial conditions and the specific shock that disturbs an unstable equilibrium need to be analyzed.
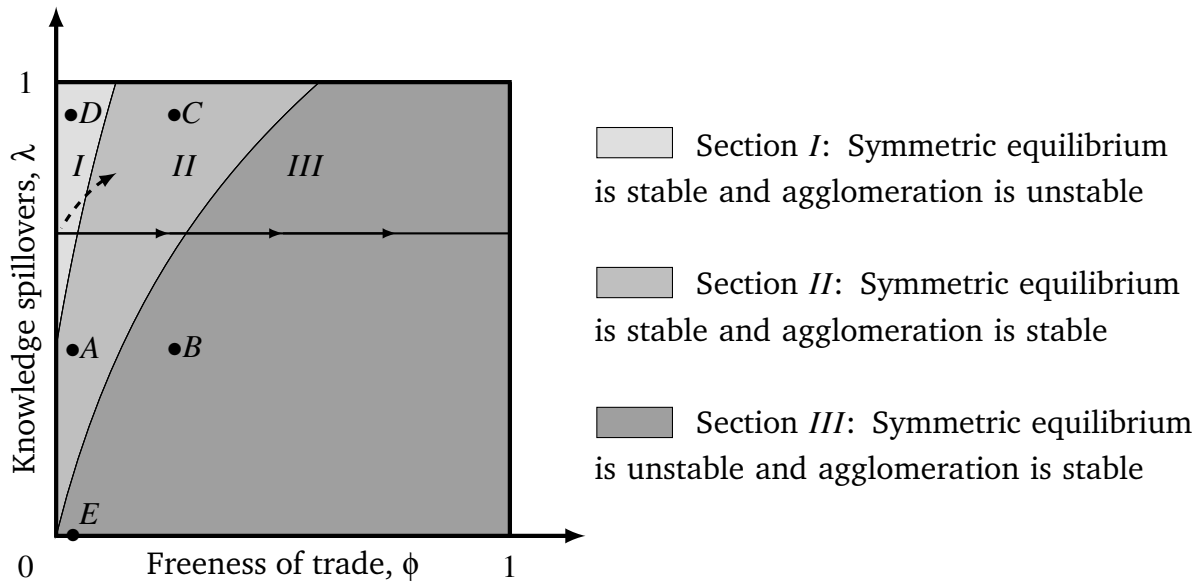
**Figure 4:** Stability of the equilibria in dependence on λ and φ.

tion *III*). This destabilizing aspect (in the sense of a tendency towards agglomeration) of closer integration is linked to the presence of an additional agglomeration force due to the introduction of endogenous growth into the model. However, the presence of knowledge spillovers influences the strength of this third circular causation chain and Baldwin and Forslid (2000) indicate that the strength of this effect can be counteracted by the model's second policy parameter, λ. Hence, combining a lowering of the costs of trading goods with a lowering of the costs of trading ideas opens up the possibility that the symmetric equilibrium remains stable for a wider range of values for φ. This possibility is shown by the dashed arrow moving from Section *I* to Section *II*.

It is also possible that a policy that lowers the costs of trading information sufficiently, leads to a spreading out of industrial activity. In Figure 4 this is captured by a 'world' economy starting in point *A* in a core-periphery equilibrium, which then becomes more integrated through higher knowledge spillovers and moves to point *D* in which only spreading is a stable equilibrium.

Therefore, integration needs to be viewed as a more complex process than in standard new economic geography models and embrace aside from mobility of goods and people also mobility of ideas and capital. Ever closer integration through a given reduction in the costs of trading goods does not necessarily lead to complete agglomeration. This process can be counteracted (compare a movement along the dashed arrow or from point *A* to *C* with a policy that moved the world economy from *A* to *B*) and possibly even reversed[34] through adequate policies that lead to a fall in the costs of trading

---

[34]Consider a situation in which both regions are living in autarky, introduce some free trade which

information. Note that although integration is now more broadly specified than in the standard new economic geography model, it is still specified as an exogenous and static concept.

# 4   Size vs. scale: When, where, and why does it matter?

The considerations so far have mainly been focussed at a highly aggregate level, be it the arbitrarily chosen global regions highlighted within the empirical presentations in Figures 1 and 2, the world-wide view in Figure 3 or the economy-wide perspective assumed within the discussed theories in Subsections 2.2 and 2.3 as well as within Section 3 and Figure 4. Besides, it is nearby to assume that the aggregation level also determines the size of the considered economy. Benefits of large economies arise e.g. as size co-determines the financial, political, business and cultural environments of people and firms. One example is the provision of public goods like national defense, public security, or the judicial system, for which the costs in larger nations can be spread over more taxpayers. Further benefits of large countries are insurance against asymmetric regional shocks[35] or the possibility of nation-wide redistribution schemes that affect the after-tax income distribution in a way not feasible if the territories were independent entities (Alesina and Spolaore 2003, 3-4). Counteracting these positive effects of size are more heterogeneous preferences in larger countries, leading to a higher probability that some subset of regions (or individuals) in a nation is not in agreement with the policies of the central government and thus poses a danger to its stability (Alesina and Spolaore 2003, 4-5). Ethnic, linguistic, and religious heterogeneity play a role in this context. More prosaically, administrative costs may rise with ever larger size and issues of congestion, crime, and pollution may become a problem. It is thus nearby to assume that there is some endogenously resulting 'optimal' size of economic spaces, which differs according to their respective characteristics and evolves as production conditions, dynamic environments, or institutional settings (especially of interest are the rules underlying international trade or knowledge diffusion) change.

That size per se is not necessarily beneficial as regards prosperity is confirmed if one takes a more precise look at empirical regularities. Less aggregated data suggests e.g.

---

moves the world economy to point $E$ where the core-periphery equilibrium is stable and then drastically reduce the cost of trading information so that the world economy ends up in point $D$ where only spreading is a stable equilibrium.

[35]Just to mention one example: The lack of such an insurance in the Eurozone during the recent economic crisis has inhibited the recovery dramatically. Adversely affected US states, for instance, received federal transfers from their 'stronger' counterparts, whereas such fiscal transfers (e.g. from Germany to Spain) were not possible within the monetary union of the Eurozone as it is not a fiscal union.

that out of the twenty countries with the largest population only four (USA, Russia, Japan, and Germany) belong to the group of high-income economies according to the World Bank's classification.[36] Contrariwise, the mentioned high-income group includes countries with a relatively small population like Singapore, Luxembourg or Iceland. A similar result holds for selected US states. Out of the top ten according to population, only New York is also in the top ten with respect to per capita income. On the other hand, small states like New Hampshire are relatively rich.[37] Prosperous regions at different sizes have at various times included, for example, agglomerations in north-western Europe for the continental scale, the Ruhr district in Germany for the country scale, or the city state of Hamburg in northern Germany on the local scale. Even at the city level, more and less prosperous districts might be observed.[38] From a dynamic perspective (i.e. focussing on the growth rate and not on the level of GDP per capita) OECD data for the second half of the 20th century at the country level also suggests that there is no conclusive evidence of a unique and positive relationship between GDP growth and measures of scale (compare e.g. Jones (1995a,b) or Backus et al. (1992)).[39]

The discussed ambiguity of the empirical studies is also mirrored within growth theory, which considers both the level but especially the growth rate of GDP per capita. An important class of growth models exhibits scale effects in the sense that variations in the size or scale of the economy permanently alter the long-run equilibrium growth rate per capita.[40] Within these models the relationship between scale and growth is unique – thereby contradicting the previously discussed empirical ambiguity – and depends upon both the existence and the nature of production externalities (e.g. via the provision of productive public inputs or various sorts of spillovers). This class of models, however, has strong formal requirements leading to restrictive knife-edge assumptions that have to be fulfilled in order to allow for long-run equilibrium growth.[41] Both the theoretical and empirical limitations of the mentioned models have led to the formulation of so-

---

[36]An overview of the classification scheme can be found here: `http://data.worldbank.org/about/country-classifications/country-and-lending-groups`.

[37]See the 'State and County QuickFacts' dataset available under: `http://quickfacts.census.gov/qfd/download_data.html`.

[38]Compare e.g. for Karlsruhe the two districts of almost identical size 'Südweststadt' and 'Oststadt' with the former being relatively prosperous and the latter being inhabited by poorer individuals.

[39]Over an extraordinary long time frame though, beginning in 1,000,000 B.C., Kremer (1993) finds support for the hypothesis of a positive link between population growth and economic growth. Notice that this long-run time scale includes the transition of various organizational forms of economic activity from hunter-gatherer over subsistence to nowadays industrialized economies.

[40]Examples are the models of Romer (1990) or Barro (1990); compare e.g. Turnovsky (2000, chap 14) for a comprehensive overview.

[41]These assumptions include constant returns to scale to rivalrous factors to allow for competitive factor markets. See Solow (1994) and Dalgaard and Kreiner (2003) for more on this.

called non-scale growth models, i.e. models that exhibit equilibrium growth rates that are not subject to scale effects in spite of endogenously accumulated production factors and the existence of externalities. However, those models suffer from other limitations, e.g. they exhibit special stability characteristics together with transitional dynamics and hence are again only suited to explain details of the complex growth story, but not to resolve the aforediscussed tensions.[42] One might summarize that there is no clear-cut evidence on the relationship between size and prosperity neither as regards theory nor the empirical analyses. Thus it is also not clear how integration that alters the size of a considered economy affects wealth. This is a strong indicator for the fact that not all relevant aspects have been addressed yet to understand the relationship between agglomeration, growth, and integration and how their respective interdependencies and feedbacks shape the economic landscape.

One starting point for further thoughts consists in following the line of reasoning of Alesina and Spolaore (2003, 82), who argue that "whether country size matters for economic prosperity depends on a country's degree of integration with the rest of the world." This is a quite plausible argument, since, for instance, small countries like Luxembourg or Switzerland (0.5m and 7.5m inhabitants, respectively) dispose of intensive trade relations and are well embedded in international value-creation networks. At the same time, land-locked countries of larger size as e.g. Uzbekistan (27.5m inhabitants), but with little access to world markets, are less prosperous.[43]

Though the argument of Alesina and Spolaore again starts at a national level, it is nearby to relax this perspective and to analyze the role of integration for prosperity and agglomeration also at a less aggregated level. It is conceivable that size may not matter for productivity at an aggregate level, but be quite relevant at a regional scale thereby relying on externalities associated with proximity-productivity linkages. It furthermore clarifies that the notion of scale (in the sense of size) has to be distinguished from scale effects, and one has to be precise by using the different wording.

Scale may thus not simply be used as a synonym for population size. From the perspective of the production conditions, scale effects are especially linked to market forms[44] or various sorts of externalities as determinants of ongoing growth.[45] From a firm

---

[42]See Turnovsky (2000), Eicher and Turnovsky (1999) and Eicher and Turnovsky (2000) for more details on scale and non-scale models of growth.

[43]In 2010 GDP per capita in Switzerland was more than four times the respective value for Uzbekistan (Bolt and van Zanden 2013). The population numbers are also for 2010 and are taken from United Nations (2012, Table A.5).

[44]An example is the necessity of constant returns to scale of the private factor inputs so that competitive factor remunerations are guaranteed.

[45]A detailed overview on different scale effects is provided by the World Bank (2009, 128) in Table 4.1 entitled 'A dozen economies of scale'.

perspective, usually internal and external economies of scale are distinguished. The optimal firm size is reached whenever firms fully exploit existing internal economies of scale. In addition, there exist external economies of scale, and both scale effects in conjunction determine whether or not a firm decides to relocate. In this context, in particular firms that dispose of production conditions characterized by increasing returns to scale and the associated spillovers are important drivers for regional development. Nailing these thoughts down to a spatial component, it seems quite plausible to assume that e.g. knowledge flows are easier realized in places where more people are concentrated. Economies of scale are then rather linked to the idea of density than to mere size. It is thus important through which channel integration acts: Does it enhance the mere number of actors within the economy (scale) or does it affect their way of interaction (scale effects)?

In what follows we argue that integration can and should be understood in a much broader sense than just a reduction of trade costs and also cover, as argued within the model presented in Section 3, the impact of knowledge spillovers. It affects the environment of the firms in the sense that it has the power to transform size into density the latter being quite well an agglomeration force. In doing so, integration is especially apt to activate latent agglomeration and dispersion forces thereby shaping the economic landscape at various levels of aggregation. Besides, integration may be deployed more effectively as an agglomeration force if it can build upon a solid institutional base. So far, the role of institutions has been discussed as a fundamental cause of economic growth both within and across economies. Our approach exactly starts at this point and argues that the design of integration pins down the impact of institutions to a spatial dimension and may thereby alter the effective economic scale. This might be justified as follows: Integration affects both the size of the relevant market and the effectiveness of local increasing returns to scale. It thereby has the power to act as a major agglomeration force and thus has an impact on the spatial structure of economic activity at different levels of aggregation. Being more precise, integration may alter various returns to scale relevant for an individual firm. It acts via two separate channels. On the one hand, as it enhances the mobility of goods, people, capital, and ideas, integration, in the sense of higher freeness of trade also increases the size of the relevant market. On the other hand, it also affects the firms' environmental conditions, especially with respect to density as integration also impacts on knowledge spillovers. Formerly latent economies of scale might become active at a firm level, if, as a consequence of integration, a region becomes more dense. In sum, integration is not just an additional argument accompanying the agglomeration-growth nexus, but across time it is *the* key driver of spatial concentration in growing economies. It especially becomes powerful not only as an enabler but also as a magnifier of scale effects.

We conclude that although the empirical regularities discussed in Section 2 and the presented theory in Section 3 are strong indicators for a co-evolution of urbanization and growth, spatial scaling (in the sense of zooming in from continental over national and regional to even city level) together with a differentiated analysis of the impact of integration both requires and allows for a more sophisticated view.

One caveat concerns the missing linkage between scale and institutions and how these are related to space; another concerns the dynamics and endogeneity of institutions. There is a large discussion on the role of national and regional innovation systems that deals with institutional and organizational dimensions that already address the evolution of institutions within dynamic economies. The analysis, however, is carried out mostly within isolated economies (compare e.g. Cooke et al. (1997)). Analogous reasonings as regards the institutional embedding of integration and its feedback on diverse levels of aggregation are still far from being understood in depth. What most theories do so far is an analysis of how static and exogenously given levels of integration affect the dynamic concepts of agglomeration and growth. However, the corresponding feedback link from agglomeration and growth on the evolution of integration and its interpretation as an institution that is shaped by those economies that are linked via integration is missing. But these considerations are mandatory to understand integration as a dynamic and endogenous concept that is linked to space. In what follows we go even one step further and argue that if we interpret integration as a dynamic spatial institution, the concept and its design become a fundamental cause not only for growth but also for agglomeration on which economic policy has a large impact. The corresponding policy implications therefore have to accommodate these various interdependencies.

# 5   Policy implications

The previous sections have presented a multitude of empirical facts on the issues of growth, urbanization, and globalization and explored the connections between these also on a theoretical level before the matters of size and scale and the role integration plays herein have been discussed more deeply. Implications for policy have so far been delegated to the background.

Which policy implications can be derived from the Baldwin and Forslid (2000) model? Naturally, these depend on the particular objectives set by policy makers and how they relate these objectives to overall welfare. For instance, a preference for a symmetric outcome[46] can be achieved by adjusting the policy variables for the costs of trading

---

[46]Perhaps inspired by the Basic Law for the Federal Republic of Germany, which tasks the Federation

goods and information adequately. However, as long as one neglects arguments of redistribution, such an objective contrasts with the goal of maximizing overall real income. The technological externalities in the innovation sector imply that agglomeration of economic activity is growth enhancing in the sense that more varieties are produced if all manufacturing activity is located in a single region (unless $\lambda = 1$). Workers in every region and sector gain from this increased number of varieties via the love-of-variety effect. This dynamic gain needs to be seen alongside the static welfare loss for consumers in the periphery's traditional sector, which arises since these workers have to import all manufactured varieties, leading to a higher price index in the peripheral region (for $\phi < 1$). Baldwin and Forslid (2000) demonstrate that the dynamic gains can only mitigate, but not compensate the welfare losses for the workers in the periphery's traditional sector. Hence, starting from a symmetric equilibrium no clear-cut policy implications exist, which lead to a Pareto improvement. The presence of the externality moreover implies that, as in standard expanding-variety growth models, the decentralized growth rate in the Baldwin and Forslid (2000) model is not Pareto optimal. Combining completely free trade (i.e. $\phi = 1$) with, for instance, a subsidy to production in the innovation sector would be an example for a policy that leads to the socially optimal growth rate and, assuming the starting point is the symmetric equilibrium, brings about a Pareto improvement for all workers (Baldwin and Forslid 2000, 323). As described above, building upon an institutional foundation, policies for closer integration become more important in the process of development. Along the lines of the model in Section 3, these fall into two categories: Integrating policies that lead to lower costs of trading goods and integrating policies that lower the costs of trading information. That tighter integration brings immense benefits is highlighted by e.g. Eaton and Kortum (1996, Table 5), who note that for all OECD countries except the US the majority of the contribution to productivity growth comes from abroad.

Taking a step back from the model, it is evident that economic activity is spatially concentrated.[47] This need not imply differences in living standards between regions in which economic activity is concentrated and regions in which it is not even though it can – the relationship resembles an inverted U over the course of economic development (World Bank 2009, 74) and can be interpreted as a variant of a 'Kuznets curve'.[48] In reaction to this situation many policy makers pursue the goal of convergence of liv-

---

with the "establishment of equivalent living conditions throughout the federal territory" (see `http://www.gesetze-im-internet.de/englisch_gg/`).

[47]For just an additional illustration of the global situation see for example: `http://gecon.yale.edu/large-pixeled-contour-globe`.

[48]The relationship considered here is not quite the same as the one for the original Kuznets curve, which considered interpersonal income inequality instead of spatial inequality, though the underlying reasoning is the same for both curves (World Bank 2009, 293).

ing standards as measured by GDP per capita. The European Union, for instance, tries to promote this goal by allocating a significant share of its funds for cohesion policy for the period 2014-2020 to economically lagging regions.[49] It is important to keep in mind here that pursuing economic convergence of regions does not imply striving for an equal spatial distribution of economic activity, as this would prevent regions from benefiting from agglomeration economies with possible negative repercussions on aggregate economic performance.[50]

According to the World Bank (2009, e.g. 41), economic integration is expected to resolve the tension between the fact of concentration and the objective of convergence. However, the utilized concept of integration is fuzzy, and the corresponding policy implications need to be differentiated according to the scale at which the policies are applied as well as to a country's level of development thereby taking an implicit dynamic perspective, which also addresses the evolution of integration and thus represents a dynamic concept. Such an approach allows taking into consideration the relative importance of the various agglomeration economies over the course of development and furthermore means that policies that aim at integration within or between cities differ from policies intended to integrate regions or nations.

Consider, for example, the policies the World Bank (2009, 229) advocates for an urbanization strategy that explicitly aims at increasing density in order to foster benefits from the proximity-productivity linkage. The report categorizes territories according to their level of urbanization. Given *incipient urbanization*, characterized by a share of the urban population of less than 25%, density is comparatively low so that policy makers' objective should be to enable the aforementioned agglomeration economies. In addition, at this stage of development there is room for individual plants and firms to more fully exploit internal returns to scale. However, the choice of policy instrument is important. Policy makers, for instance, face a risk of favoring one place or industry over another by spatially targeted interventions, whereas markets may assess the situation in a different light. This may lead to inefficient economic structures, which may persist over time. Hence, a strategy avoiding such an outcome would place emphasis on establishing adequate institutions like secure property rights in land markets and provision of basic social services for health and education without applying a spatial focus. This aspect remains important in areas with *intermediate urbanization* shares of about 50%.

---

[49]Defined for this purpose as regions with a GDP per capita less than 75% of the EU-27 average. See the European Union's Regional Policy website for more detailed information: `http://ec.europa.eu/regional_policy/index_en.cfm`. At the national level in Germany, the federal financial equalization system (Länderfinanzausgleich) indirectly pursues a similar objective by utilizing a large budget to redistribute money from prosperous to weaker federal states.

[50]The example of the industrial policy in the former Soviet Union has demonstrated the problems with such an approach (World Bank 2009, 256).

Firms located in these areas are embedded in various emerging networks where co-operation arises both within and between different places. Through the co-location of firms in the same or closely related sectors these firms benefit from localization (MAR) externalities, which lead to increasing returns to scale within the considered region. The promotion of these agglomeration economies should be high on the policy makers' agenda in these regions, but not the only item on it. Investments in infrastructure are important as well in order to ease congestion and to better integrate people and places. These policies continue to be relevant for areas of *advanced urbanization* where around 75% of the population live in urban areas. Benefits in these areas arise mostly from urbanization economies dating back to Jacobs (1969) and which point to diversity and intense economic activity. 'Livability' is the watchword from the World Bank (2009, 201) for these areas, and this could be reflected in polices that try to reduce crime and pollution or provide amenities.

Policies should thus reflect that as an economy evolves, it passes from states with constant returns to scale through states characterized by specialization/localization economies to states of urbanization economies. Throughout this process of transformation, frequently internal economies of scale turn into external economies, which then simultaneously act as agglomeration forces for those being located in close proximity. This argument is especially intuitive for growing agglomerations where development prepares the ground for the gradual evolution of clusters, which on their own are specialized, but which – given sufficient size – due to their interaction finally also allow for diversification. Integration plays a major role herein.

# 6   Conclusion

Agglomeration, economic growth, and integration are three main aspects that characterize recent world economic history. This article has highlighted the relevant empirical facts on these and summarized how they relate to each other by presenting a model by Baldwin and Forslid (2000) that combines two seminal models from the endogenous growth (Romer 1990) and new economic geography literature (Krugman 1991). A non-formal analysis of the combined model has shown that lowering the costs of information tends to weaken the agglomeration force that is set in motion by a lowering of trade costs. Despite these important insights, the treatment of integration in this model is still on a very rough level.

The issue of scale in its various guises and how it is affected via integration has received particular attention. It has been demonstrated that the presence of agglomeration economies possibly depends on the level of aggregation (continent, nation, region,

city, district). Analyzing this issue in more depth and investigating the presence of possible threshold effects seems to be worthwhile. Major weaknesses are that the degree of integration is exogenously given in most analyses and thus does not allow neither for economic development nor structural change or spatial scaling. Especially in dynamic economies where the economic landscape is shaped by the interactions of growth and agglomeration such a perspective of integration is much too simplistic. Integration in a comprehensive sense goes far beyond being the mere enabler and magnifier of mobility of various items. It also implies changes in the organization of economic processes at a spatial scale as it changes the way in which different agents in a system relate to each other. The design of integration thus complements respective institutional settings of interacting territories. Due to the dynamic environments, also integration itself evolves and becomes a dynamic and endogenous concept. Future research thus requires a far more precise view on the endogenous determination and evolution of integration and how this interacts within supra-national, national and regional institutional settings thereby shaping the economic landscape. Integration should be clearer understood as a process that itself undergoes dynamic changes, too. As the economy evolves, also the design of the prevailing integration mechanism has to be continuously adjusted. In doing so, the respective level of agglomeration at which integration becomes active has to be considered as well as the dynamic characteristics of the investigated economic spaces.

# A   List of geographic entities

A list of countries and entities sorted into the various regions in Figures 1 and 2 according to the classification in the New Maddison Project Database is given below.

**Table A.1:** Countries aggregated into major geographical regions according to the classification in the New Maddison Project Database.

| | |
|---|---|
| **Africa:** | Algeria, Angola, Benin, Botswana, Burkina Faso, Burundi, Cameroon, Cape Verde, Central African Republic, Chad, Comoro Islands, Côte d'Ivoire, Democratic Republic of the Congo, Djibouti, Egypt, Equatorial Guinea, Eritrea, Ethiopia, Gabon, Gambia, Ghana, Guinea, Guinea Bissau, Kenya, Lesotho, Liberia, Libya, Madagascar, Malawi, Mali, Mauritania, Mauritius, Mayotte, Morocco, Mozambique, Namibia, Niger, Nigeria, Republic of the Congo, Rwanda, Saint Helena, Sao Tomé & Principe, Senegal, Seychelles, Sierra Leone, Somalia, South Africa, Sudan, Swaziland, Tanzania, Togo, Tunisia, Uganda, Western Sahara, Zambia, Zimbabwe |
| **Asia:** | Afghanistan, American Samoa, Bahrain, Bangladesh, Bhutan, Brunei, Burma, Cambodia, China, Cook Islands, East Timor, Fiji, French Polynesia, Guam, Hong Kong, India, Indonesia, Iran, Iraq, Israel, Japan, Jordan, Kiribati, Kuwait, Laos, Lebanon, Macao, Malaysia, Maldives, Marshall Islands, Micronesia, Mongolia, Nauru, Nepal, New Caledonia, North Korea, Northern Mariana Islands, Oman, Pakistan, Palau, Papua New Guinea, Philippines, Qatar, South Korea, Samoa, Saudi Arabia, Singapore, Solomon Islands, Sri Lanka, Syria, Taiwan, Thailand, Tonga, Turkey, Tuvalu, United Arab Emirates, Vanuatu, Vietnam, Wallis and Fortuna, West Bank and Gaza, Yemen |
| **Latin America:** | Anguilla, Antigua and Barbuda, Argentina, Aruba, Bahamas, Barbados, Belize, Bermuda, Bolivia, Brazil, British Virgin Islands, Cayman Islands, Chile, Colombia, Costa Rica, Cuba, Dominica, Dominican Republic, Ecuador, El Salvador, Grenada, Guatemala, Guyana, Haiti, Honduras, Jamaica, Mexico, Montserrat, Netherlands Antilles, Nicaragua, Panama, Paraguay, Peru, Puerto Rico, St. Kitts and Nevis, St. Lucia, St. Pierre and Miquelon, St. Vincent and the Grenadines, Suriname, Trinidad and Tobago, Turks and Caicos Islands, United States Virgin Islands, Uruguay, Venezuela |
| **Western Europe:** | Andorra, Austria, Belgium, Cyprus, Denmark, Faeroe Islands, Finland, France, Germany, Gibraltar, Greece, Greenland, Guernsey, Iceland, Ireland, Isle of Man, Italy, Jersey, Liechtenstein, Luxembourg, Malta, Monaco, Netherlands, Norway, Portugal, San Marino, Spain, Sweden, Switzerland, United Kingdom |

# References

Abel, G. J. and Sander, N.: 2014, Quantifying global international migration flows, *Science* **343**(6178), 1520–1522.

Acemoglu, D.: 2009, *Introduction to Modern Economic Growth*, Princeton University Press.

Acemoglu, D., Johnson, S. and Robinson, J. A.: 2001, The colonial origins of comparative development: An empirical investigation, *American Economic Review* **91**, 1369–1401.

Alesina, A. and Spolaore, E.: 2003, *The Size of Nations*, MIT Press, Cambridge MA.

Arrow, K. J.: 1962, The economic implications of learning by doing, *The Review of Economic Studies* **29**, 155–173.

Backus, D. K., Kehoe, P. J. and Kehoe, T. J.: 1992, In search of scale effects in trade and growth, *Journal of Economic Theory* **58**, 377–409.

Baier, S. L. and Bergstrand, J. H.: 2001, The growth of world trade: tariffs, transport costs, and income similarity, *Journal of International Economics* **53**, 1–27.

Bairoch, P.: 1988, *Cities and Economic Development - From the Dawn of History to the Present*, Mansell Publishing Limited.

Baldwin, R. E. and Forslid, R.: 2000, The core-periphery model and endogenous growth: Stabilizing and destabilizing integration, *Economica* **67**, 307–324.

Baldwin, R. E., Martin, P. and Ottaviano, Gianmarco, I. P.: 2001, Global income divergence, trade, and industrialization: The geography of growth take-offs, *Journal of Economic Growth* **6**, 5–37.

Barro, R. J.: 1990, Government spending in a simple model of endogenous growth, *Journal of Political Economy* **98**, 103–125.

Barro, R. J. and Sala-i-Martin, X.: 2004, *Economic Growth*, 2 edn, MIT Press, Cambridge, MA.

Bolt, J. and van Zanden, J. L.: 2013, The first update of the Maddison Project; re-estimating growth before 1820, *Maddison Project Working Paper 4*.

Bosker, M. and Garretsen, H.: 2009, Economic development and the geography of institutions, *Journal of Economic Geography* **9**, 295–328.

Brakman, S., Garretsen, H. and van Marrewijk, C.: 2009, *The New Introduction to Geographical Economics*, 2 edn, Cambridge University Press, Cambridge UK.

Brakman, S. and van Marrewijk, C.: 2008, It's a big world after all: on the economic impact of location and distance, *Cambridge Journal of Regions, Economy and Society* **1**, 411–437.

Brülhart, M.: 2009, An account of global intra-industry trade, 1962-2006, *The World Economy* **32**(3), 401–459.

Cooke, P., Uranga, M. G. and Etxebarria, G.: 1997, Regional innovation systems: Institutional and organisational dimensions, *Research Policy* **26**, 475–491.

Dalgaard, C.-J. and Kreiner, C. T.: 2003, Endogenous growth: A knife edge or the razor's edge?, *Scandinavian Journal of Economics* **105**(1), 73–85.

Davis, D. R.: 1995, Intra-industry trade: A Heckscher-Ohlin-Ricardo approach, *Journal of International Economics* **39**, 201–226.

Dell, M., Jones, B. F. and Olken, B. A.: 2012, Temperature shocks and economic growth: Evidence from the last half century, *American Economic Journal: Macroeconomics* **4**(3), 66–95.

Dixit, A. K. and Stiglitz, J.: 1977, Monopolistic competition and optimum product diversity, *American Economic Review* **67**, 297–308.

Duranton, G. and Puga, D.: 2000, Diversity and specialisation in cities: Why, where and when does it matter?, *Urban Studies* **37**(3), 533–555.

Eaton, J. and Kortum, S.: 1996, Trade in ideas: Patenting and productivity in the OECD, *Journal of International Economics* **40**, 251–278.

Eicher, T. S. and Turnovsky, S. J.: 1999, Non-scale models of economic growth, *The Economic Journal* **109**, 394–415.

Eicher, T. and Turnovsky, S.: 2000, Scale, congestion and growth, *Economica* **67**, 325–346.

Friedman, T. L.: 2005, *The World is Flat - A Brief History of the Twenty-First Century*, Farrar, Straus and Giroux.

Fujita, M. and Thisse, J.-F.: 2002, *Economics of Agglomeration: Cities, Regions, and International Trade*, MIT Press.

Gabaix, X. and Ioannides, Y. M.: 2004, The evolution of city size distributions, *in* J. Henderson and J. Thisse (eds), *Handbook of Regional and Urban Economics*, Vol. 4, Elsevier B.V.

Gallup, J. L., Sachs, J. D. and Mellinger, A. D.: 1999, Geography and economic development, *International Regional Science Review* **22**(2), 179–232.

Galor, O.: 2005, From stagnation to growth: Unified growth theory, *in* P. Aghion and S. N. Durlauf (eds), *Handbook of Economic Growth*, Vol. 1A, Elsevier B.V., chapter 4.

Galor, O.: 2011, *Unified Growth Theory*, Princeton University Press.

Heston, A., Summers, R. and Aten, B.: 2012, Penn World Table Version 7.1, Center for International Comparison of Production, Income and Prices at the University of Pennsylvania.

Hewings, G. J. D. and Oosterhaven, J.: 2014, Interregional trade models, *in* M. M. Fischer and P. Nijkamp (eds), *Handbook of Regional Science*, Springer-Verlag Berlin Heidelberg, chapter 46, pp. 903–925.

Hummels, D.: 2007, Transportation costs and international trade in the second era of globalization, *Journal of Economic Perspectives* **21**(3), 131–154.

Jacobs, J.: 1969, *The Economy of Cities*, Random House, New York.

Jones, C. I.: 1995a, R&D-based models of economic growth, *The Journal of Political Economy* **103**(4), 759–784.

Jones, C. I.: 1995b, Time series tests of endogenous growth models, *The Quarterly Journal of Economics* **110**(2), 495–525.

Jones, C. I. and Romer, P. M.: 2010, The new Kaldor facts: Ideas, institutions, population, and human capital, *American Economic Journal: Macroeconomics* **2**(1), 224–245.

Kaldor, N.: 1957, A model of economic growth, *Economic Journal* **67**(268), 591–624.

Kaldor, N.: 1961, Capital accumulation and economic growth, *in* F. A. Lutz and D. C. Hague (eds), *The Theory of Capital*, St. Martin's Press, New York, pp. 177–222.

Kremer, M.: 1993, Population growth and technological change: One million B.C. to 1990, *The Quarterly Journal of Economics* **108**(3), 681–716.

Krugman, P.: 1991, Increasing returns and economic geography, *Journal of Political Economy* **99**, 483–499.

Krugman, P. E.: 1998, What's new about the new economic geogrpahy?, *Oxford Review of Economic Policy* **14**(2), 7–17.

Marshall, A.: 1890, *Principles of Economics*, Macmillan, London.

Meissner, C. M.: 2014, Growth from globalization? A view from the very long run, *in* P. Aghion and S. N. Durlauf (eds), *Handbook of Economic Growth*, Vol. 2B, Elsevier B.V., chapter 8, pp. 1033–1069.

North, D.: 1990, *Institutions, Institutional Change and Economic Performance*, Cambridge University Press, Cambridge.

Özden, Ç., Parsons, C. R., Schiff, M. and Walmsley, T. L.: 2011, Where on Earth is everybody? The evolution of global bilateral migration 1960-2000, *The World Bank Economic Review* **25**(2), 12–56.

Porter, M. E.: 1990, *The Competitive Advantage of Nations*, Macmillan, New York.

Rodrik, D., Subramanian, A. and Trebbi, F.: 2004, Institutions rule: The primacy of institutions over geography and integration in economic development, *Journal of Economic Growth* **9**, 131–165.

Romer, P. M.: 1986, Increasing returns and long-run growth, *Journal of Political Economy* **94**, 1002–1037.

Romer, P. M.: 1990, Endogenous technological change, *The Journal of Political Economy* **98**(5), S71–S102.

Solow, R. M.: 1956, A contribution to the theory of economic growth, *Quarterly Journal of Economics* **70**(1), 65–94.

Solow, R. M.: 1994, Perspectives on growth theory, *Journal of Economic Perspectives* **8**(1), 45–54.

Swan, T. W.: 1956, Economic growth and capital accumulation, *The Economic Record* **32**(2), 334–361.

Thom, R. and McDowell, M.: 1999, Measuring marginal intra-industry trade, *Weltwirtschaftliches Archiv* **135**(1), 48–61.

Turnovsky, S. J.: 2000, *Methods of Macroeconomic Dynamics*, 2 edn, MIT Press, Cambridge/Mass.

Uchida, H. and Nelson, A.: 2010, Agglomeration index: Towards a new measure of urban concentration, *Working Paper 2010/29*, World Institute for Development Economic Research.

United Nations, Department of Economic and Social Affairs, Population Division: 2012, *World Urbanization Prospects, the 2011 Revision*, Final Report with Annex Tables, New York.

World Bank: 2009, *World Develoment Report 2009: Reshaping Economic Geography*, World Bank.

World Bank: 2013, *World Development Indicators*, World Bank.

World Trade Organization: 2011, *World Trade Report 2011 – The WTO and preferential trade agreements: From co-existence to coherence*, World Trade Organization.

# Working Paper Series in Economics