THE ROLE OF OPENNESS IN ECONOMIC GROWTH OF REGIONS

THE PERSPECTIVES OF ECONOMIC GROWTH AND NEW ECONOMIC GEOGRAPHY THEORY AND EMPIRICS

TOMASZ BRODZICKI

Institute for Development
Working Papers no. 001/2018
The role of openness in economic growth of regions
The perspectives of economic growth and new economic geography theory and empirics

February 2018
Tomasz Brodzicki

Abstract

The bond between economic growth and openness of regions is a complex and to some extent ambiguous both from theoretical and empirical perspectives. The relationship becomes even more blurred if we take into account spatial nonlinearity into account. In the present paper we critically review the theoretical and empirical literature in the two strands – economic growth and new economic geography. It seems that the analysis of the impact of openness on economic growth of regions should take into account the various potential channels, try to prove the direction of causality and acknowledge the presence of spatial interlinkages between regions. This is in particular important for mid-size and large-size economies with a significant number of regions and large regional heterogeneity frequently due to more deep-rooted factors. It implies the use of spatial econometric techniques.

JEL Classification: \( F21, F43, O47, R11, R12 \)

Keywords: economic growth, economic development, openness, trade, FDI, regional economics

The study has been supported by a grant from the National Science Centre [grant no. 2015/19/B/HS4/01704] ‘Regional Exporting Activity. Assessment of Determinants in Light of Contemporary Foreign Trade Theory for Poland and Spain’.

---

\(^a\) Tomasz Brodzicki, PhD
Economics of European Integration, Faculty of Economics
University of Gdańsk, Sopot, Poland;
e-mail. t.brodzicki@ug.edu.pl
ul. Armii Krajowej 119/121, 81-824 Sopot, Poland

and
Institute for Development, Sopot, Poland
e-mail: t.brodzicki@instytut-rozwoju.org
ul. Czyżewskiego 6/1, 81-706 Sopot, Poland
**Introduction**

The extent of economic openness is not easy to define and many definitions coexist. An open economy could be defined as “a market-economy mostly free from trade barriers where exports and imports form a large percentage of the GDP. No economy is totally open or closed in terms of trade restrictions, and all governments have varying degrees of control over movements of capital and labour. The degree of openness of an economy determines a government's freedom to pursue economic policies of its choice and the susceptibility of the country to international economic cycles.”¹

The above definition stresses the extent of liberalization of flows of both goods as well as of flows of major production factors, such as labour and capital. Nonetheless, the most widely utilized matric of openness - the openness index is calculated as a ratio of country's/region’s total trade to its GDP. According to the World Bank’s national accounts database, the global openness index in 2015 was 58.0 percent and it reached its maximum level in 2008 (60.34 percent). In comparison, it was only 24.21 percent in 1960, 38.78 in 1980 and 51.37 percent in 2000. With exception of the recent years, in relation to the global financial crisis, the overall level of openness has significantly and visibly increased. At the same time, the extent of openness varies greatly between countries, and to an even larger extent, between regions within individual countries.

As an illustration of the above, in the present study, we are dealing with NUTS2 regions of Poland and Spain, observed over the period 2005 to 2014. Over the period the openness ratio increased in most of the regions (on average by 9 percent). The openness ratio dropped only in the case of Mazowieckie, Illes Balears, Canarias and Comunidad de Madrid (please refer to Figure 1). On the other extreme, the highest increases have been reported in Andalucía, Łódzkie, Dolnośląskie and Opolskie (by more than 15 percent), Lubuskie by approx. 25 percent, Región de Murcia 33 percent and Pomorskie by 34.3 percent. The lowest openness ratio in 2014 was observed in Canarias (island region of Spain) and the highest for Pomorskie – seaside region of Poland with two significant seaports located in Gdańsk and Gdynia with the biggest container terminal in the Baltic Sea Region located in Gdańsk (Deepwater Container Terminal).

¹ [http://www.businessdictionary.com/definition/open-economy.html](http://www.businessdictionary.com/definition/open-economy.html)
The role of openness in economic growth theory

The nexus between openness and economic growth is one of the most important areas of academic analysis both on the theoretical and empirical front. The issue has been and still is addressed by various fields of enquiry such as classical and new growth theory, international economics or economic geography and finally the so-called new economic geography (NEG).

In the neoclassical growth theory (Solow 1956, 1957; Swan, 1956) openness does not matter in the long-run as growth is independent of economic policy. It can only lead to the so-called level effect, altering just the level of the real GDP per capita in the steady-state and not the growth rate. In the short-run, capital deepening is the major source of growth, as income per capita is directly proportional to the level of capital per capita. The level of real GDP per capita in the steady state is a positive function of the rate of saving (and thus investment rate) and a negative function of the population growth rate and depreciation of capital.
Technological progress of exogenous character affects the level positively. The only factor affecting the long-run growth rate is the rate of exogenous technological progress. Namely, it is equal to the rate of technological progress. In this setting, the impact of an increase in the extent of openness, due for instance, to changes in trade policy on economic growth is only temporary.

In an augmented model of Mankiw, Romer & Weil (1992) human capital accumulation and the human capital endowment are taken into account, additionally. The model is able to better explain the observed variation in the level of economic development. It thus not, however, modify the prior conclusions. The augmented neoclassical model by Brodzicki (2015) takes further the impact of infrastructure into account. In the model, in accordance with Mincerian tradition, the average level of education may be specified as a function of average years of schooling and average years of experience (Bils & Klenow, 2000).

The emergence of the endogenous growth and new trade theories (Lucas 1988, Romer 1986, 1990, Aghion & Howitt, 1992, 1998) has led to the reopening of the debate on the role of trade, and in more general, the degree of openness in determining economic growth in the medium and long-term. The models of the first and second generation endogenized the rate of growth of technology either by allowing for the impact of human capital or introducing a separate R&D sector, purposefully producing knowledge in the form of patent. It is worth to point out, however, that even in a semi-endogenous model of Ben-David & Romer (2002), openness to trade through its impact on the process of accumulation of knowledge and technology transfer (knowledge diffusion) leads to endogenization of the economic growth process.

The new growth theory models of Rivera-Batiz & Romer (1991) or Grossman & Helpman (1991, 1992) lead to different policy conclusions. A policy shift leading to a greater extent of openness (such as trade liberalization), could lead to a permanent effect. The long-run growth rate could be affected but not only positively, an adverse impact is also possible. In brief, the balance of costs and benefits of greater openness (liberalization) depends on the nature and the exact product structure of trade – in other words, it's not merely related to the intensity of trade but also by its composition.

Greater openness to trade affects the rate of accumulation of knowledge mostly through imports. They work as a channel allowing absorption of more advanced knowledge positively
affecting overall efficiency and thus the growth rate. Rivera-Batiz & Romer (1991) show however that whether the effect is positive or adverse, depends on the distance of economy from the global technology frontier and the nature of diffusion of knowledge (perfect versus imperfect). Imperfect knowledge flows coupled with openness can actually harm underdeveloped states or regions.

Barro & Sala-i-Martin (2004) or Aghion & Howitt (2009) emphasize the role of technology diffusion in both absolute and conditional convergence. It is mostly due to the fact that imitation and implementation of innovation are cheaper than the initial innovation itself. Imitation and adaptation still entail substantial costs which can be however lowered through more intense trade (mostly imports) or superior human capital and skills base. Nelson & Phelps (1966) stress that followers tend to grow faster, the greater is the initial gap to the leader. The gap diminishes over time and thus the followers’ growth rates tend to decrease alongside. In the steady-state, the leader and the follower grow at the same rate. Nelson & Phelps (1966) argued that education could positively affect the speed of adoption of new technologies. They distinguished the theoretical level of knowledge from the prevailing (existing) level of technology. Benhabib & Spiegel (1994) extended the original model by adding an extra innovation term, which controls for the impact of own capacity to develop knowledge on top of the ability to absorb external knowledge. Abreu, de Groot & Florax (2004) further accounted for potential spatial dependencies between bordering economies. Ciolek & Brodzicki (2017) in their empirical model extend it even further by including two potential channels for technology diffusion through imports and FDI inflows. The channels find support in many empirical studies. In their seminal study, Coe & Helpman (1995) studied the impact of trade on technology diffusion and found that international R&D spillovers were related to imports and in particular to the composition of imports. Furthermore, a strong correlation between R&D embodied in (bilateral) trade flows and TFP growth was identified. Coe, Helpman & Hoffmaister (2009) endorsed the impact of domestic and foreign R&D capital stocks on the TFP even after controlling for human capital. They extended the analysis by the inclusion of institutional variables, allowing for parameter heterogeneity based on institutional characteristics. The results suggested that institutional differences were significant determinants of TFP, and they had an impact on the degree of R&D spillovers. In the context of the second channel considered, Hejazi & Safarian (1999)
identified significant R&D spillovers through FDI from largest industrial countries to smaller OECD Member States. Xu (2000), in contrast, found that technology transfer of US multinationals contributed to productivity growth but only in the group of developed economies.

According to Keller (2004), for most countries, foreign sources of technology are crucial (90 percent or more) in productivity growth. They are at the same time more valuable for small and relatively poorer countries. It could be related to the significance of the variation in domestic R&D investments. There is no indication that the process of international diffusion (and thus learning) is inevitable, simple or automatic. Imports are the major channel of international technology diffusion with no indication of learning-by-exporting effects. FDI effects are also present. However, the impact is highly asymmetric. Keller (2004) points that technological knowledge spillovers appear to be resulting from a deliberate commitment to learning and matching international performance standards through ongoing interactions with foreigners. At the same time, local efforts seem to be necessary for successful technology adoption.

It could, therefore, mean that technology diffusion is not only spatially bounded but also not-universal. The follower has to have a minimum level of endowments in order to be able to absorb technology. Technological change can thus be skill-based, which could lead to technology-skill mismatch and thus non-convergence in TFP levels (Acemoglu & Zilibotti, 2001).

Eaton & Kortum (2001) constructed a model of innovation, growth, and trade with technology spillovers, which pointed to convergence in income levels. The benefits of the larger market can be exploited by an innovator through exports. However, the innovator, in an open economy, has to compete not only with domestic rivals but also with imported technologies. These are two offsetting forces. If they offset completely, only static gains from trade arise with no dynamic gains through technology accumulation. Decreased barriers to trade stimulate research activity characterized by the presence of scale effects. Real wages depend on the productivity of workers but also the size of the population. In the extreme case of autarky, relative real wages in the model depend on relative labour forces weighted by research productivity. Decreasing barriers to trade, however, benefit smaller economies to a large extent. On the other extreme, with zero gravity (a costless trade),
relative real wages depend solely on relative research productivity with the size of the economy playing no significant role.

In an extension, Eaton & Kortum (2002), construct a Ricardian model accounting for realistic geographic features, where bilateral trade is a function of absolute advantages, trade-promoting comparative advantages and trade-resisting geographic barriers. Gravity is thus fully integrated and accounted for.

Howitt (2000) in a multi-country endogenous Schumpeterian growth model, shows that due to technology diffusion, only R&D-performing countries grow in the long run and converge to similar growth paths while non-R&D-performing countries stagnate. In this framework, an increase in the investment rate or the R&D-subsidy rate in any R&D-producing country can increase the overall growth rate. In an extension, Aghion, Howitt & Mayer-Foulkes (2005) attribute the emergence of the so-called convergence clubs in income, to similarities in the R&D potential and the process of knowledge diffusion. In their stylized model, countries sort themselves into three groups: members of the highest group converge to a steady state where they perform leading-edge R&D (at the global technology frontier), while the intermediate group converges to a steady-state where they only implement technologies developed elsewhere. High and intermediate group countries share the same growth rate in the long run as a result of technology diffusion, nonetheless, inequality between them in terms of development levels, increases. Economies of the lowest group grow at a slower rate and are unable to converge due to their inability to absorb knowledge from the global technology frontier (GTF). In this set-up, the initial distance to the technological frontier matters and economies initially lagging by a significant distance can be entrapped in the low-income level.

It is also worth addressing the direction of causality between openness and economic growth. If openness affects growth than we deal with the so-called export-led growth (ELG) process through the channels described above. The term was introduced by Balassa (1978) and was later investigated among others by Marin (1992). Some theoretical work was conducted on the related concept of learning-by-exporting (Krugman, 1980, Grossman & Helpman, 1991 or more recently at firm level e.g. Melitz, 2003 or Helpman, Melitz, & Yeaple, 2003. Marin (1992) states that: “outward-looking” regime favours the productivity performance of developed market economies as well as that of developing countries” after analysis of four case countries. 

---

2 Marin (1992) states that: „outward-looking” regime favours the productivity performance of developed market economies as well as that of developing countries“ after analysis of four case countires.”
The initial evidence was weak and the general agreement on the existence of ELG is sometimes challenged (e.g. Dreger & Herzer, 2013).

On the other hand, the causality could be just the opposite – going from the growth to greater openness. Higher productivity in the larger domestic market (the so-called home marker effect, HME) could lead to greater international competitiveness and a subsequent increase in the regional exports. At the same time, demand for imports increases in the size of the regional economy. Thus a bidirectional relationship is likely to exist if the processes described above hold simultaneously (Liu et al. 1997).

It is worth stressing that in the new approach, Rodrik (2002) perceives openness or as he puts it the extent of integration as one of three fundamental deep determinants of economic growth alongside the quality of institutions and geographical conditions. Openness is treated here as a semi-endogenous factor shaped by purely exogenous geographical conditions (in particular of the first nature of geography) as well as institutional factors. Nonetheless, it, directly and indirectly, affects the shallow determinants of growth – related to the endowment of basic factors of production and the process of their accumulation as well as the overall productivity of the economic system.

Summing up the initial literature review, it is evident at least from a theoretical standpoint, that openness affects growth through a number of channels.

- First of all, it leads to reallocation of factors of production to more productive sectors and thus to specialization in accordance with the comparative or competitive advantage thus resources are allocated efficiently.
- Secondly, it leads to increased diffusion and accelerated absorption of knowledge and technology (technology transfer) in particular through imports (Coe & Helpman, 1995) or inflow of FDI (Branstetter, 2006).

3 The deep determinants have been analyzed for Poland in Brodzicki (2014). The results obtained lead to the following conclusions with respect to the impact of openness on economic growth of Polish voivodeships: „voivodeships with a higher level of economic openness, stronger export potential and higher attractiveness for foreign investors attain on average higher economic growth rates. These regions are also characterized by higher exposure level to external macroeconomic shocks“.
Thirdly, it stimulates the rate of innovation as it is frequently associated with an increase in the expenditures on research and development.

Fourthly, it allows better utilization of scale economies and agglomeration externalities as a result of greater specialization. At the same time, it leads to enhanced accumulation of factors of production.

Fifthly, it stimulates competition in national and international markets thus forcing companies to be more innovative.

**Critical review of the empirical growth literature**

In the empirical literature, two strands of studies clearly dominate – the macro approach with mostly cross-sectional or panel analysis of global or more homogeneous groups of countries and the micro approach – analysis for individual countries based on sectoral or firm-level data.

In both strands of literature, various variables are utilized as proxies for the overall extent of openness. These include: openness ratio (trade/GDP), openness dummy – open/closed economy, rate of exports, share of imports in GDP, real foreign exchange distortions, average level of tariffs in general or the average level of US tariffs (Romalis 2006), dummies for preferential trade agreements (PTAs) and regional trade agreements (RTAs), the outward orientation index of the World Bank, black market premium on foreign exchange, exports distortion index as well as various geographic variables (landlockedness, common border), etc. It is worth mentioning that even the openness dummy could, in fact, have a rather complex structure. For instance in Sachs & Warner (1995) an economy is referred to as an open economy only if it fulfills simultaneously the following set of conditions:

- average tariff rate on capital and intermediate goods is below of 40 percent;
- non-tariff barriers are assigned to less than 40 percent of import of capital and intermediate goods;
- black market premium does not exceed 20 percent of official foreign exchange;
- the economy is a non-socialist country;
- state monopoly in exports of key branches is not present.
Nonetheless, the openness ratio defined as the ratio of total trade to GDP seems to be the most popular.

In his famous cross-sectional regression analysis, Barro (1991) identified a positive and statistically significant impact of the level of openness on economic growth in a cross-section of countries. Dollar (1992) noting a potential bias utilized an index of exchange rate disturbances, finding it to adversely affect economic growth. The result was further confirmed by Easterly et al. (1993) and Lee (1993) using similar approaches.

Sachs & Warner (1995) utilized a dichotomous index of openness conditional on meeting 5 criteria, finding openness to matter for growth in a cross-section of countries. The index was also utilized by Gallup et al. (1998), which led to analogous results even if deep-rooted geographical factors were taken into account. In his seminal study, Vamvakidis (1999) identified a positive and statistically significant effect of multilateral economic integration.

Wacziarg & Welch (2003) found the earlier studies applying the Sachs & Warner (1995) index to be sensitive to the time period considered.

Edwards (1998) in his seminal study analysed the impact of 9 different indices of openness/disturbances in the exchange rate on productivity as measured by the total factor productivity and thus indirectly on the real GDP per capita in a large cross-section of 93 countries. The impact was identified to be positive however its magnitude was found to be less significant in comparison to the traditional determinants of economic growth such as the initial level of GDP per capita (in line with the absolute/conditional convergence literature) or the initial level of human capital endowment (as postulated by the augmented Solow-Swan model of Mankiw, Romer & Weil (1992) and some of the models of the new growth theory strand (Lucas, 1988, Lucas, 1990). Human capital endowment affects directly the rate of domestic innovation (Romer, 1990) as well as the speed of adoption of technology from abroad or knowledge diffusion (Nelson & Phelps, 1966).

Due to a potential endogeneity, the instrumental variables approach is frequently utilized. For instance, Frankel & Romer (1996, 1999), propose an instrumental variable based on geographical factors that determine to a large extent the intensity of bilateral trade while being exogenous with respect to the level of income. The impact of openness proved surprisingly to be insignificant in two large cross-sectional datasets of countries considered.
Irwin & Tervio (2002) reiterated the test by Frankel & Romer (1999) in a slightly modified manner for a panel of countries. The results pointed to a positive relationship between the intensity of trade and the level of GDP per capita. Later on, Romalis (2006) found similar results using instrumental variables (IV) approach in a large panel of countries (135) observed over a period of 40 years (1960-2000).

In his comprehensive analysis, Vamvakidis (2002) tested six different measures of openness for an elongated period (1920-1999) finding that the positive relationship between openness and growth existed only after the 1970s. It could be related to a noticeable increase in the extent of openness with the beginning of the next phase of globalization.

In contrast to aforementioned empirical studies, Wacziarg & Welch (2003) utilized a different approach to analyzing the effects of cases of significant trade-policy liberalizations and finding that they were on average followed by an increase in the investment rate of 1.5 to 2 percent, and in the share of trade in GDP by 5 percent (increase in openness) while the ex-post growth rate was higher than ex-ante growth rate by a mean of 1.5 percent.

Using the extreme bounds analysis approach (EBA) Levine & Renelt (1992) found the index of openness to be one of the variables indirectly affecting the growth rate in a cross-section of countries through the impact on the process of accumulation of capital (investment rates). In the same study, they rejected the existence of the direct linkage.

In contrast, Doppelhoffer, Sala-i-Martin & Miller (2000) using the Bayesian Averaging of Classical Estimates (BACE) approach for a balanced panel of 88 countries and 68 significant determinants of economic growth found the time since the liberalization of an economy to positively affect economic growth. It could mean that the benefits of liberalization or openness accumulate in time and thus short run and long-run impact could significantly differ. In the same study, the impact of the overall openness level was found to matter significantly less.

The quick review of the most important empirical studies on the openness-growth nexus leads to the conclusion that the results obtained to a large extent depend on: dataset – scope and period (duration), methodological approach adopted (cross-sectional vs. panel), allowing for linear or non-linear impact, the choice of the openness proxy, taking care or not of potential endogeneity, the choice of the correct instrumental variable, taking care of
outliers and other standard econometric problems (sample selection bias, heteroscedasticity, potential co-linearity).

**Empirical studies at the regional level**

In comparison to the vast empirical literature on the nexus between openness and economic growth, the analyses conducted at the regional level are surprisingly rather scarce.

Soukiazis & Antunes (2011) conducted the analysis on the role of openness, export shares or trade balances on the regional growth in Portugal at the NUTS3 level over the period 1996–2005 further conditioning for the role of the human capital. The dynamic panel model estimated with the use of GMM proves that factors associated with external trade, including openness, human capital endowment and sectoral labour shares (in particular in the industrial sector) are important determinants of regional growth as well the conditional convergence processes. Furthermore, the interactions between the key variables also play an important role and to a large extent explain different performances between regions of the Portuguese Littoral and Interior.

Boschma & Iammarino (2009) analyzed the impact of trade linkages for Italian provinces at NUTS 3 levels and three-digit sectors over the period 1995–2003 looking for the impact of the so-called related variety. The hypothesis was positively verified - related variety contributed to regional economic growth. Well-endowed in complementary sectors Italian provinces showed superior performance. Boschma & Iammarino (2009) postulate furthermore that openness alone (more trade or more knowledge flows) does not affect growth, but this is due to the presence of related extra-regional knowledge which leads to intersectoral learning across regions.

In recent years a number of studies have been performed on Asian economies and their regions. Sun et al. (1999) show in a study of Chinese regions at manufacturing industries level that openness to trade (trade orientation and FDI) has a positive effect on technical efficiency.

Leong (2013) analysing the impact of special economic zones, as cases of liberalization on regional economic growth in China and India found that both FDI and exports positively affect growth. The presence of SEZs increases regional growth, however, an increase in the
number of SEZs has a negligible effect on growth. Leong (2013) identifies greater openness (wider liberalization) to be a precondition of further growth. Wei et al. (2009) in a panel of Chinese regions over the entire period 1979-2003 proved that FDI inflows were one of the forces behind the observed regional discrepancies in growth. The authors claim however that FDI cannot be blamed for the extent of regional inequality as it was due to the uneven distribution of FDI and not the FDI itself.

Anwar & Nguyen (2010) using simultaneous equations model in a panel of 61 Vietnamese provinces from 1996–2005, found a mutually reinforcing two-way process between FDI and regional economic growth. The benefits of FDI inflow could be further strengthened by more investments in education and training, development of the financial market and reducing technology gap between foreign and local firms.

Shafiullah et al. (2017) analyzed the export-led growth (ELG) hypothesis in Australia taken as a whole and its regions at the sectoral level over the period 1990 to 2013 using quarterly data. The authors state in particular that mining and fuels sector’s exports played a crucial role in driving economic growth of Australia and in three of its regions in the long run - namely New South Wales, Queensland, and Western Australia. Shafiullah et al. (2017) identify each Australian region’s experience with ELG to be region-specific. Traces of ELG in the short run were identified for South Australia, Tasmania and Northern Territory. It, in turn, seems to depend on the composition of exports (product or sectoral-structure).

According to Kanbur, Ravi & Venables (2005), rising spatial disparities in regional development in many developing states are mostly due to uneven impact of increased trade openness and globalization. It leads to efficiency gains mostly due to concertation of economic activity in major cities and coastal districts adversely affecting inland regions. In a study on Latin America, Serra et al. (2006) argue that regional disparities modestly increased, at least temporarily, in the wake of trade liberalization. It was especially marked for Mexico.

Redding (2012, 2016) used a version of the quantitative spatial model in order to investigate the effects of a fall in trade costs between the US and Canada, leaving internal trade costs unchanged and allowing for heterogeneous worker preferences across locations. The analysis is conducted at the regional level. Redding states that given greater trade intensity with US states, Central Canada would gain more than Western Canada, under population immobility. But in the case of the mobile population across regions, the improved market
access of Central Canada would cause it to gain population, while Western Canada would see a decline in population. The reallocation of the population would continue until all Canadian regions gain equally from the fall in trade costs, in the absence of costs to mobility. In analyzing counterfactuals, Redding (2016) states the welfare gains from trade (liberalization) depend on changes in both domestic trade shares and reallocations of the population across locations. Furthermore, factor mobility introduces quantitatively relevant differences in the counterfactual predictions of the constant and increasing returns to scale models. Therefore, models excluding spatial interlinkages and factor mobility at the regional level of analysis can lead to falsified results. It is not the case for studies performed at the country-level, where the estimates for welfare are pretty correct.

When analysing the nexus between openness and economic growth at regional level we have to note the direct or indirect impact of other accompanying variables or processes.

For instance, Sachs et al. (2002) studying $\sigma$-convergence, and $\beta$-convergence show that more than 80 percent of the cross-state variation in growth rates among Indian states can be explained solely by an urbanization variable. Agglomeration factors are also strongly postulated by NEG theories.

The role of human capital accumulation is clear on the theoretical and empirical ground, however, the scope of the definition of human capital differs. For example, Boschma & Fritsch (2009) point similarly to Florida (2004) to an important contribution of the so-called creative class for regional growth in seven European countries. They are however not able to determine whether human capital as measured by the creative occupation, outperforms standard indicators based on formal education and whether formal education has a stronger impact. The creative class endowment is positively affected by the regional climate of tolerance and openness as well as regional job opportunities.

Other factors could matter as well such as the economic structure or the size and the share of an industrial sector. For instance, the study by Hansen & Zhang (1996) points to the key role of the industrial sector in explaining the regional variation in growth among Chinese province. The result supports the Kaldorian approach to regional economic growth with cumulative causation between trade liberalization, the rise in export demand, the growth of industrial sector (industrialization) and its impact on overall productivity and thus increases in international competitiveness.
One of the issues that cannot be overlooked is the path-dependency in regional development. For instance, Felice & Vecchi (2015) indicate that the regional North-South variation in Italy was already present at the moment of the unification and then subsequently increased further. The explanation of the present variation involves endogenous factors - natural resources, human capital endowment, and social capital.

Several studies have been also performed for Poland. Brodzicki (2015) attempted to identify shallow determinants of growth of Polish regions and in addition the sign and magnitude of macroeconomic education and infrastructure externalities. Brodzicki proposed an augmented neoclassical growth model incorporating a Mincerian approach to human capital accumulation, further assuming a direct impact of infrastructure on the overall productivity. The estimated panel model accounting for fixed region-specific effects was robust and explained approx. 90 percent of the observed variation in GDP per capita. The return to the accumulation of human capital through education and experience for Polish regions was found to be statistically significant, robust and positive. The macroeconomic infrastructure externality proved to be in turn positive however overall insignificant with the impact of quality of railway.

In a recent article Ciołek & Brodzicki (2017) analyzed the determinants of spatial variation and spatial spillovers of TFP at the level of LADs. They utilized and tested an extended empirical version of the aforementioned Nelson-Phelps (1966) model accounting for potential spatial interactions among regions at a high level of spatial disaggregation. It required the use of dedicated spatial econometric methods. The authors stress that the TFP assumes the highest values in the metropolitan centres and spreads out on their nearest surroundings with the maximum value for the capital region – Warsaw. The secondary local hills in TFP are located in cities or towns with county rights. The range of TFP spillover is found to be of roughly 175–200 km and is nonlinearly decreasing from the local productivity hills. Furthermore, the rate of growth of TFP shows spatial autocorrelation and was found to depend positively on the rate of increase in human capital endowment and on the gap from the leader under certain assumptions. The result is in line with the postulated of Nelson and Phelps. Furthermore, Ciołek & Brodzicki (2017) accounted for two channels of productivity diffusion related to international trade and factor flows namely through imports and through the inflow of foreign direct investments. The authors found no evidence of the
channel through imports (trade relations). However, the FDI channel was found to be robust and strong. It points to the need of broadening of the concept of openness to account for FDI inflows and outflows as well.

For comparative purposes, it is worth to mention results of similar studies. Bottazzi & Peri (2003) using R&D and patent data for European regions over the period 1977–1995 found knowledge spillovers to be relatively weak and localized within a distance of 300 km. Bronzini & Piselli (2009) analyzed relationship between TFP, R&D, human capital and public infrastructure on a panel of Italian regions over the period 1980–2001. The results indicated the existence of a long-run equilibrium between TFP levels and the three types of capital with human capital affecting the TFP the most. Moreno, Paci & Usai (2005) investigated the spatial distribution of innovative activity and the role of technological spillovers in the process of knowledge creation across 138 regions of 17 countries in Europe over the period of 1978–1997 at sectoral level (3 digit ISIC sectors). The authors identified a strong initial central-periphery pattern of distribution of innovation activity with concentrations in Northern and Central regions with a tendency to decline (diffuse). They identified furthermore a robust and positive spatial autocorrelation in the innovative activity. External effects were also identified pointing to the role of technology diffusion within a distance of roughly 250 and 500 kilometres. Sterlacchini (2008) adopted Fagerberg’s technology-gap model of economic growth (Fagerberg, 1988) and examined the relationship between the economic growth of 12 European regions over the period 1995–2002 and their knowledge and human capital endowments. Strelacchini took into account foreign and domestic knowledge, the ability to utilize both sources and the distance from the technology frontier. He controlled for potential agglomeration effects by including the log of population density. The GERD and the share of the population with tertiary education were found to be the most important determinants of growth in incomes per capita.

**Empirical studies on the role of openness to FDI in the economic growth**

One of the important aspects of openness is the extent of openness to the inflow as well as an outflow of foreign direct investments. The theoretical channels have been discussed above.
In a seminal paper Borensztein, De Gregorio & Lee (1998) in a cross-country regression framework analyzed the channels of impact of FDI on growth, utilizing data on FDI flows from industrial countries to developing countries over two decades. The results obtained supported the notion that FDI is an important vehicle for international technology diffusion affecting the growth rate directly a bit stronger than indirectly through the impact on domestic investment. The authors noted, however, that the impact is conditional on the existence of a minimum endowment of human capital or in other words it is conditional on a sufficient absorptive capability for advanced technologies. Growth process and knowledge acquisition could thus be therefore said to be skilled-biased.

Balasubramanyam et al. (1996) stress the significance of trade openness in the case of developing countries in order to acquire beneficial effects of FDI inflow. In particular, export promotion policy proves to be more conducive to the nexus than the import substitution policy. In turn, Blomstrom et al. (1994) argued that FDI had a positive growth effect only when a country was sufficiently prosperous in terms of the level of development. Alfaro et al. (2004) analyzed the linkage between FDI and growth in a cross-country data set over the period 1975-1995 controlling for the level of development of the financial market. The role of FDI was identified as ambiguous as such, however, countries with more developed financial markets benefited significantly from FDI. The studies point to positive, however, the conditional impact of FDI on economic growth. The analysis by Lee & Chang (2009) using panel cointegration and panel error correction models for a panel of 37 countries over the period 1970-2002 proved the result showing a fairly strong long-run relationship. The short-run relationship was, however, identified as weak.

Carkovic & Levine (2002) using a novel statistical method, postulated in contrast, that FDI inflows did not exert an independent influence on economic growth. The authors control for potential biases in the estimation process due to endogeneity, the omission of country-specific effects and of the initial level of income. In the words of the authors “, while sound economic policies may spur both growth and FDI, the results are inconsistent with the view the FDI exerts a positive impact on growth that is independent of other growth determinants”. Also, De Mello (1999) in the non-OECD sample found no causation from FDI to growth based on fixed effects regressions with country-specific intercepts, and even a negative short-run impact of FDI on GDP using the mean group estimator.
De Mello (1997) stressed that the ultimate impact of FDI on output growth in the recipient economy depended on the scope for efficiency spillovers to domestic firms, by which FDI led to increasing returns in domestic production, and increased in the value-added content of the FDI-related production. Tang et al. (2008) point to complementarity between FDI and domestic investments in China. No crowding out of domestic investment has been observed.

In a recent paper, Iamsiraro & Ulubaşoğlu (2015) scrutinized 108 published studies (meta-analysis) and tested the relationship in a global sample of 140 countries in the period 1970 to 2009, proving the clearly positive impact of FDI on growth both in developed as well as developing countries. Iamsiraro & Ulubaşoğlu (2015) stress the role of regional variation and contemporaneous FDI rather than past FDI, for growth. Furthermore, trade openness and financial development rather than schooling (human capital) were identified as appropriate absorptive capacity indicators for positive growth (please compare to results of previous studies).

Nair-Reichert & Weinhold (2001) emphasize that FDI on average had a positive impact on growth, however, the result proved to be highly heterogeneous across the group of developing countries. Furthermore, there was some evidence that the efficacy of FDI was higher in more open economies.

Similarly to trade openness we have to address the issue of causality between FDI and economic growth. Asheghian (2004) stressed that the US economic growth was mostly driven by TFP growth and the growth in domestic investment as well as FDI. Furthermore, the relationship between FDI and economic as well as TFP growth (and thus indirectly affecting growth) is uni-directional and going from FDI. Similarly, Hansen & Rand (2006) analyzed the causality between FDI and growth in a sample of 31 developing countries covering three continents over the time period 1970–2000. They identify a strong causal link from FDI to GDP in the short and in the long run. And interestingly state the long-term impact is independent of the level of development of the recipient.

There are also studies pointing to the bi-directional relationship, for instance, Choe (2003), Jayachandran & Seilan (2010) for India or Malaysia and Thailand (Chowdhury & Mavrotas, 2006). The bi-directional nature of the relationship is further postulated by Iamsiraroj (2016) using a simultaneous system of equations approach of 124 cross-country data for the period 1971–2010. Furthermore, trade openness alongside labour force and economic freedom are
identified as key determinants of FDI, which in turn accelerates economic growth process. Basu, Chakraborty & Reagle (2003) stressed the existence of a cointegrating relationship between FDI and economic growth using a panel of 23 countries. Interestingly trade openness seems is stressed as a key determinant of the beneficial impact of FDI on growth. The authors namely identified two-way causality both in the short and the long run only in open economies, whereas unidirectional long-run causality from growth to FDI in closed economies.

Fidrmuc & Martin (2011) analyzed the role of FDI, trade and growth in eleven CESEE countries at the country level over the period 1995-2009. They tested for the hypotheses of ELG and FDI-led growth in the CESEE region. The authors found the stock of FDI to be positively related to industrial production and economic growth. In nearly all CESEE countries, exports and FDI had a significant impact on industrial growth performance with exports playing a stronger role. The relationship between trade, FDI and growth was identified as complex, with both variables identified as endogenous. Output growth for Poland, Romania, Slovakia and Slovenia was found to profit from exports and FDI. The long-run industrial production was found at the same time not to be affected in Bulgaria, Croatia, Estonia, Hungary and Lithuania, and as a result, the export-led/FDI-led growth hypothesis was rejected in these particular cases.

Summing up the empirical literature conducted at the level of countries brings rather mixed results on the nexus between FDI and growth in terms of existence, accompanying determinants and direction of causality.

Chen & Fleisher (1996) identified the role of FDI in the process of conditional convergence of production per capita in Chinese provinces over the period of 1978 – 1993 alongside physical investment share, employment growth, human-capital investment and coastal location of a region. The authors utilized an augmented Solow growth model framework. We can conclude that FDI could have a potential impact on the observed regional variation in the levels of productivity and thus levels of development.

Sjöholm (1999) analyzed productivity growth in Indonesian manufacturing companies across regions. He found that regional characteristics at the district level, rather than at the province level, seemed to explain productivity growth. This spatial scale seems to be more conducive to inter-industry knowledge spillovers. These, in turn, are positively affected by
FDI as “domestic establishments benefit from a regional presence of foreign establishments in neighbouring industries”. Sjöholm (1999) identified in addition intra-industry spillovers from FDI at the national level as domestic establishments in industries with a large foreign presence have shown high productivity growth.

Buckley et al. (2002) analyzed the impact of FDI inflow into 29 Chinese provinces over the period 1989 – 1999, trying to identify the impact of host country conditions on the nexus between FDI and growth. They proved to matter both at the national and regional level. FDI concentrated in economically stronger regions and the benefits related to FDI were a function of the strength of local competition.

Lessmann (2013) studied the impact of FDI inflow on regional variation in China, and a wider sample of 55 countries over the period of thirty years, 1980-2009. Economic theory says that the level of economic development could have an impact on the FDI–regional inequality nexus. FDI could actually increase the extent of regional inequalities as they are usually spatially concentrated (FDI agglomeration effect). Lessmann (2013) found that FDI inflows increased regional inequality in low and middle-income countries, while there were no negative consequences for redistribution observed in the high-income economies. Furthermore, he stated that the higher mobility of individuals in developed countries, as well as government policies (linked to better institutional quality), were likely to at least partially mitigate the adverse impact of FDI on regional inequality.

Su & Liu (2016) in their econometric analysis a panel of Chinese cities over the period 1991–2010, identified the positive role of FDI on the per capita GDP growth rate. The effect was intensified by the human capital endowment of the city which authors interpret as human capital contributing to growth is to serve as a facilitator for technology transfers stemming from FDI.

The review of theoretical and empirical literature proves that the nexus between openness and growth is complex and still is being investigated both at the country as well as regional levels. The different channels are being gradually verified with more and more elaborated methods and using better datasets. Nonetheless, some questions related to the direction of the causalities still remain open and many contributing factors are still being identified. It is
important to note that when we deal with subnational units such as regions, the spatial aspects of the analyzed phenomena and in particular potential spatial interactions between regions should be considered. As the main growth theory is aspatial it requires the review of theoretical and empirical studies in the economic geography literature.

The perspective of new economic geography on the nexus

A further insight into the analyzed nexus can be brought by the new economic geography literature. The seminal paper by Krugman (1991a) is generally regarded as the foundation of the new economic geography. It is an extension of the standard new trade theory model (Krugman 1980) allowing for interregional mobility of factors of production. Breinlich et al. (2013) stress the fact that NEG theory is based on trade theory, and thus the relationship between external trade, internal economic geography, and regional disparities, is at its core. Fujita et al. (1999) suggest that openness could work to disperse manufacturing industry as a whole, but also lead to the spatial clustering of specific industries. External trade thus affects spatial patterns of activity by changing market access considerations (Hanson, 1996).

Incorporation of space into the theoretical framework and at the same endogenization of location choices requires moving beyond the neoclassical paradigm. It is crucial to reject the assumptions of the zero transport costs and allow for the presence of the increasing economies of scale (both internal and external to a firm).

The NEG brought the issue of space into the mainstream economic theory. From a modelling point of view, NEG theories can be considered as an extension of new trade models allowing for simultaneous flows of goods as well as factors of production (capital and/or labour) and knowledge flows.

Fujita & Thisse (2002) state that there are three important assumptions in spatial modelling within the NEG:

1. space is heterogeneous, which leads to the presence of comparative advantages in technology, natural resources, facilities or the presence of transport hubs or markets;
2. externalities exist in both production and consumption – agglomeration forces arise from the bottom-up as a result of the non-market interactions between market agents - enterprises and households; classic Marshallian effects (Marshall 1879,
arise and specialization related to the snowball effect in the event of the concentration - we deal with external economies of scale and scope;

3. markets are imperfectly competitive thus firms are endowed with some degree of market power – shaping the price above marginal costs as a consequence of the utilization of internal economies of scale; two possibilities exist:
   a. monopolistic competition – many companies, lack of strategic interaction, the product is diversified;
   b. partial oligopolistic competition (few actors with similar market power, strong strategic interactions, the existence of the Nash equilibrium – game theory).

Combes et al. (2008) show that the choice of the above modelling strategy has significant implications for the results obtained. Approaches 1, 2 and 3a from the macro perspective bypass the role of individual companies. Approach 3b takes into account strategic interactions between agents. The first approach generates a solution that is socially efficient in the sense of Pareto, and the remaining models lead to socially inefficient solutions. Combes et al. (2008) stats that the classic NEG model chooses the third approach, where location decisions become endogenous at the expense of assuming the homogeneity of space (thus omitting the first nature of geography).

The basic NEG model of Krugman (1991a, 1991b) with a typical core-periphery structure (CP) is based on the concept of the monopolistic competition of Dixit & Stiglitz (1977)\(^4\). NEG models focus on the relationships between three factors determining location decisions of market agents: agglomeration benefits, non-zero costs of transport and interregional migration (Fujita et al. 1999; Fujita & Thisse 2002; Baldwin et al. 2003). The cost of transport (resistance to overcome space) is primarily a function of the distance between trading parties, but may also be a function of institutional barriers.

A key feature of the NEG models is endogenization of the location decisions, and thus endogenization of location and distribution of economic activity in space (Brülhart, 2001; Brakman et al. 2004). Endogeneity of the location decisions means that market players and households consciously decide on the selection of their location in order to maximize profits

\(^4\) The framework is hence frequently referred in the literature as DSK or Dixit-Stiglitz-Krugman framework.
(firms) or their total utility (consumers) based on the information (including prices) generated by the market system.

The market structure in the NEG models gradually evolves in the spatial domain under the influence of such factors as:

- the size of the economy of individual regions,
- the costs of transport,
- the scope of internal economies of scale associated with the size of the production in individual production plants,
- the scope of the external economies of scale related to the degree of concentration of economic activities in the different sectors – the classic Marshallian externalities,
- the presence of linkages (ascending and descending or forward/backward linkages) in value-added chains,
- the effect of demand have the foggiest idea about the internal market – the home market effect (HME),
- the intensity of competition in a market is a function of the quantity of firms and size-distribution (market size and economies of scale, competitive effect) as well as the interaction between them (e.g. the existence of restrictive collusion or exclusionary free competition).

The degree of spatial concentration or dispersion of economic activity depends on the balance between centripetal (pro-agglomerative) and centrifugal (pro-dispersal) forces. It is worth to note, that various NEG models differ in the direction of effect and the significance of particular forces. For instance, external economies of scale are a natural agglomeration force leading both to the creation of within-sector concentrations of firms (industrial districts or clusters) or multi-sectoral agglomerations of firms or households (cities).

In the standard NEG model with monopolistic competition, we deal with a differentiated product, homogeneous companies specialize in the production of its individual varieties, firms pose a certain degree of market power and the nature of competitive rivalry is thus imperfect. At the same time, the number of enterprises in the market is large enough to get rid of the problem of potential strategic interactions characteristic for oligopolistic models. The framework is quite often criticized for oversimplification and some authors try to take
account of the strategic interactions in the location decisions of agents, for example using the Cournot oligopoly model (see e.g. Combes & La Fourcade, 2011).

In NEG models we typically deal with multiple equilibria of unstable or stable character (as evidenced by the tomahawk diagrams). Several corner solutions are possible:

- the total concentration of production activities in one of the two regions concerned (catastrophic agglomeration),
- total dispersion of production activities (the symmetric layout, equal sharing of potential between territorial units),
- asymmetric equilibria of typically unstable character.

Furthermore, NEG models distinguish between short- and long-term equilibria. The long-term allows for structural changes in the economy and thus adjustment in the location of economic agents.

A level of mutual openness plays an important role in regional development processes (in other words, the degree of integration of the regional economy), which is by definition linked to the level of transport costs. These are typically modelled as the so-called ice-berg transport costs attributed to Samuleson (1954) in which transport costs linearly relate to distance, and they work by extracting from the arriving volume. Transport costs reflect the specific resistance of space, leading to concentration or dispersion of production in space. This relationship between the level of transport costs and the degree of concentration of business is non-monotonic and nonlinear having the shape of an inverted U (see, for example, Ottaviano 2008). High transport costs lead to dispersion of economic activity. As transport costs fall to medium levels, we observed gradual concentration. If they fall further to low levels, dispersion of economic activity emerges once again.

The lowering of the cost of the exchange is the result of among others:

- technological progress through the use of more efficient means of transport, better organization of logistic processes, etc.
- utilization of economies of scale in transport that leads to decrease in the unit costs,
- improvement in the transport infrastructure leading to an increase in accessibility of the regions,
and finally reduction in the level of protection of the relevant market, and thus increase the level of mutual openness of the economies in the region.

In this framework, one can explore the impact of integration (asymmetric reduction in trade costs) on the location of economic activity. For example, Hanson (1994) examined the effects of fundamental change of the trade policy regime in Mexico from a protectionist to liberal in 1984, on the location of the processing industry of the country. In a relatively short span of time from the moment of liberalization, it led to a significant contraction of the prior concentration of manufacturing industry around Mexico City (the so-called Mexico City manufacturing belt) which emergence was related to the implementation of the costly policy of import substitution. Liberalization resulted in the relocation of production plants to the areas directly on the border with the United States.

The results of empirical studies indicate that the integration processes within the framework of the EU also affect the location of business activity – leading to an increase of the degree of spatial concentration of production activities (see e.g. Amiti, 1999; Brülhart, 2001; Midelfart-Knarvik et al., 2002). The flow between the core and peripheral regions of the EU depends on the specific circumstances of each sector (including the intensity of economies of scale, dependence on transport costs, and the intensity of the backward and forward linkages) and we observe heterogeneity in this area. Furthermore, each new enlargement leads to an adjustment in the location decisions. Brülhart (2001) is of the opinion that at an aggregated level, the characteristic core-peripheries system of the European Union seems to gradually fade away. Midelfart-Knarvik & Overman (2002) state that the equalization-oriented structural policy of the EU has a significant impact on the location of economic activity, not necessarily leading to efficiency of the block taken as a whole.

Models of NEG, aiming to resolve the problem of the analytical burden, by standard cover two regions and two sectors (2 x 2). In reality, both the number of regions and sectors is significantly higher with multi-region and multisector framework. Behrens et al. (2004) point out that in the multi-regional system, even the static location of companies is determined by mutually reinforcing spatial effects (accessibility) and aspatial effects (attraction), which affects the overall distribution of demand in all the regions concerned (third country effect). The economic potential of a specific region is a function of not only the size of the region but
also its relative accessibility within the framework of the multiregion system (hence the popularity of the market potential approach).

However, even the most advanced models of spatial equilibrium for France (by Combes & La Foucard 2001 or Combes & La Foucard 2011), taking into account the problem of strategic interaction (Cournot competition) in the location decisions of operators, or Teixeira (2006) for Portugal, have a static character and thus ignore the problem of growth dynamics in the short medium and long period. Models NEG trying to explain the evolution of location of activities in space, simply overlook the problem of the economic growth process. Combes & La Foucard (2011) state that they cannot tell whether France is close to a short or long run spatial equilibrium. They actually note that from the point of view of regional perspective, France is gradually moving from the system with one core (îles de France – Paris) to a system with two cores – the other one emerging in the south-eastern part of the country.

These complex scenarios including endogenous location choices and endogenous growth are only possible in models combing static NEG with postulates or mechanisms characteristic for the previously described new growth theory, the so-called dynamic NEG models, these included Martin & Ottaviano (1996) or Baldwin & Forslid (1998, 2000).

Dynamic NEG models with endogenous growth assign a major role to externalities from human capital accumulation or the wider concepts of knowledge. The externalities effects are decreasing in distance, which means they are localized (Hanson, 2000). The same factors that determine the location of economic activity in the context of dynamic models of NEG are the same ones which are responsible for endogenization of economic growth.

Baldwin et al. (2003) distinguish basically three types of dynamic NEG models. These are:

- core-periphery models (CP),
- footloose entrepreneur models (FE),
- footloose capital models (FC).

The standard NEG model in the DSK formula includes the following structure and set of assumptions:

- two regions - developed North (N) and the underdeveloped South (S);
- two sectors of the economy – agriculture (A) and the manufacturing industry (M);
- two factors of production – physical capital (K) and labour (L);
one of the factors of production is mobile, the second is immobile;
A sector – Walrasian agricultural sector with perfect competition producing a homogeneous product with constant economies of scale;
M sector – manufacturing sector producing n varieties of a differentiated product in the presence of increasing economies of scale within monopolistically competitive market;
each of the varieties of the differentiated good is produced by another company which within this variety possess market power – a monopolist rent, and thus sets prices above of marginal costs;
firms are homogeneous – a representative firm model;
preferences of firms and households are homogeneous;
regions share the level of technological sophistication.

A typical mainstream NEG model a la Krugman (1991a, 1991b) is characterized by the following features:

circular causality on the demand side (offset production moves) and on the supply side (production costs, offset production affects the level of costs);
hysteresis in the location of economic activity, that is, the present distribution in space depends on the preceding situation (path-dependency);
the occurrence of multiple sable or unstable equilibria;
the presence of symmetric (N = S, dispersed equilibrium) as well as severely unbalanced equilibria (the tomahawk diagram);
the Bell chart depicting the evolution of regional development;
the presence of catastrophic agglomeration due to the homogeneity of preferences;
the presence of the home market effect (HME);
the existence of the so-called spatial structure of wages (SWS) reflecting the variation in productivity;
Non-linearity (non-monotonic character) of many associations, such as the impact of transport costs on the level of the spatial concentration of economic activities.

In dynamic models, in contrast to static approaches, processes have a temporal dimension, which creates a fundamental difference compared to the static models. In terms of the
medium-term as economic policy changes result is a one-time adjustment of capital, that is the flow and allocation of capital between regions, although the rate of accumulation is unchanged – we thus deal with the level effect. In the long run – economic policy changes can affect the pace of capital accumulation – and thus we deal with the classic rate of a growth effect.

Effects for the overall level of prosperity in the case of models with a fixed level of capital (K) are as follows:

- the effect of prices on the border – the decrease in import prices rises prices in exports which causes an increase in the prosperity of the region N according to the classic definition of the terms of trade effect;
- the effect of location on cost of living – if the number of varieties of differentiated good n is constant, the dislocation of the production of one variety of the differentiated good from the region S to N causes an increase in the prosperity of the region; the significance of the effect increases with the increase in transportation costs;
- the effect of migration – depends on the type of model, in CP and FE models an inflow of mobile capital K raises real income; in the FC model, it has no impact on the level of prosperity.

**Dynamic NEG models**

In the case of dynamic models with capital accumulation the capital accumulation occurs until the steady state level is reached, for which the value of an additional unit of capital aligns with the cost of its production. In the steady-state further accumulation of capital bring zero effects for the general welfare.

Baldwin et al. (2002) distinguish two NEG models with an accumulation of capital:

- the constructed capital model (Baldwin, 1992).
- localized (LS) or global (GS) spillover models.

In the CC model with the accumulation of capital and each new unit of capital is associated with the emergence of a new variety of a differentiated good. The increase in varieties is the manifestation of technological progress, by analogy with the horizontal differentiation models of endogenous growth literature. The process of accumulation of capital thus
translates into an increase in a number of available varieties. It causes a progressive decline in the general level of the level, and therefore a gradual increase in the real product and real wages.

LS and GS models also allow for endogenization economic growth. The key is to the adoption of a broad definition of capital – physical capital, human capital and knowledge considered together, bypasses the problem of diminishing returns through the incorporation of learning curve effects in the production process. Thus the cost of broadly defined capital decreases in time. Such capital may spill into adjacent areas (imperfect diffusion, LS – localized spillovers) or globally (perfect diffusion, GS – globalized spillovers).

Knowledge diffuses primarily within individual economic sectors (intra-sectoral spillovers, the so-called MAR externalities) and to a more limited extent between sectors (inter-sectoral spillovers, or the Jacobian externalities). The localized spillovers are the main forces responsible for the spatial concentration of sectors or overall spatial agglomeration of economic activity and population, explaining the phenomena of industrial districts, clusters, and the emergence and development of cities and metropolitan centres.

In the LS models, we are dealing with a perfect diffusion of knowledge between enterprises in a given region (for example, within a single cluster) and hindered interregional knowledge diffusion. The scale of localized diffusion location choice has an impact on the long-term growth rate.

In the GS models, knowledge defuses between firms from different regions. In this setting, capital accumulation can lead to catastrophic agglomeration. Spatial considerations do not affect the long-term growth rate, however, they have an impact on the level of development (through transitional effects).

In the LS and GS models, private atomistic innovators overlook the presence of externalities in their activities and thus the impact of their activities on the evolution of the general price levels. At the same time, the learning effect occurs in the innovation sector. The apparent market failure leads to the socially suboptimal growth rate in the Pareto sense and thus creates an opportunity for a potentially favourable public intervention (e.g. R&D subsidy). The dynamic NEG models with knowledge spillovers include models of Martin & Ottaviano (1999), Baldwin & Forslid (1999), Baldwin et al. (2001) or Baldwin & Martin (2004).
Endogenization of growth in the NEG model requires similarly to models of the new growth theories, the extension of the concept of capital and taking into account externalities in its accumulation, or endogenization of the technological progress by the introduction to the structure of the model of an R&D sector responsible for the creation of new knowledge and innovation. In the process of knowledge generation, knowledge diffuses to the neighbouring regions.

From the theoretical point of view, diffusion of knowledge can be perfect (global and immediate diffusion) or imperfect (spatially restricted or localized). Numerous empirical results confirm that the diffusion of knowledge is imperfect and is strongly localized in spite of the progress in the field of information technology (see e.g. Jacobs, 1969; Ciccone & Hall 1996; Coe & Helpman, 1994; Coe et al. 1997; Jaffe et al. 1993; Zucker et al. 1998; Eaton & Kortum, 1999; Keller, 2002; Keller, 2004; Thompson & Fox-Kean, 2005). The recent evidence for Poland also points to the localized character of knowledge diffusion (Ciołek & Brodzicki, 2016; Ciołek & Brodzicki, 2017). The localized knowledge spillovers are in brief due to the tacit character knowledge which diffusion requires direct face to face interaction.

Dynamic models of NEG are far from ideal. They are foremost, not able to fully capture the complex nature of space or the interactions between actors in the overlapping (Zaucha, 2008). To some extent, they represent a compromise between the desire to capture as much of the spectrum of conditions and factors and the objective limitations on the modelling side.

NEG models cannot be considered as an exhaustive coverage of spatial issues in the analysis of economic growth. They are excessively one-sided, and respond to the question on the role of space in the process of economic growth, refer only to the cost of tackling space. In this situation, the spatial concentration is the factor, which promotes development. The alternative approach is the postulate of space without the cost of its overcoming. The NEG, however, omits the costs of "density", which may negatively impact the growth rate of GDP through congestion (Zaucha, 2007; 2008).

Fujita & Krugman (2004) stated that future models must have even richer microeconomic foundations, rely to a greater extent on the results obtained through empirical analyses and
relate directly to the consequences for the well-being of individual societies, and therefore, generate better recommendations for economic policy.

**Firms heterogeneity and the new new economic geography**

The newest theoretical models a la Melitz (2003) move away from the traditional assumption of firm homogeneity (representative company models) and take into account the actual heterogeneity of firms in terms of e.g. productivity, size or scope of activities. These models, taking into account the stochastic distribution of productivity, lead to important and new theoretical postulates, for example on the cause of the occurrence of the exporters, competitiveness, internationalization and innovation, but at the same time lose the simplicity and thus transparency characteristic for former models and thus lead to unambiguous economic policy recommendations.

Ottaviano (2010) notes that the future NEG models should account for both macro-heterogeneity across locations and micro-heterogeneity across firms within sectors and across people with various preferences. Taking the firm heterogeneity into consideration leads to the emergence of the next generation of models new new economic geography (NNEG) models. He states: “*still based on the pillars of scale economies and imperfect competition but with a stronger emphasis on how individual heterogeneity across people and firms maps into aggregate behaviour*”

For instance, Baldwin & Okubo (2005) built a model of the NNEG-type integrating a heterogeneous firm's Melitz-style model of monopolistic competition with a simple NEG model obtaining a number of interesting results. First of all, only most productive firms can benefit from reallocation to larger regions. The selection effect exists decreasing the extent of traditional agglomeration economies. Furthermore, a spatial sorting effect arises that induces the highest productivity firms to move to the core and the lowest productivity firms to the periphery. Furthermore, the HME effect is weaker due to firm heterogeneity. According to Ottaviano (2010), in the model firm heterogeneity acts as an additional centrifugal force – the greater, the larger are the trade costs and the larger the substitutability between firms’ products.
Forslid & Okubo (2010) extended the Baldwin & Okubo (2006) model by introducing different capital intensities among firms (that can move between regions) and sectors. More productive firms were assumed to be more capital intensive. As a result, the model postulates sorting to the large regions from both ends of the actual productivity distribution (Pareto or log-normal distribution of TFP is observed in the actual firm-level data). Specifically, firms with high capital intensity and high productivity as well as firms with very low productivity and low capital intensity tend to relocate to the core (core region premium). Using Japanese data authors provide some evidence for the predicted two-sided sorting and in particular in high capital intensity sectors. The early NNEG models include as well Baldwin & Okubo (2009), Nocke (2006) or Okubo (2009).

Some critical remarks on NEG

Brülhart (2011) in conclusions of his survey of implications of trade liberalization for intra-national geographies of individual economies (intra-national regional inequality) states that the results of both urban systems and NEG models are inconclusive. Everything depends on the modelling strategies and choices adopted. Empirical results are rather inconclusive as well, however, a majority of cross-country studies finds no significant effect of openness on urban concentration or regional inequality. Whether trade liberalization (or greater openness to trade) raises or lowers regional inequality depends on country-specific geography which is to a large extent of exogenous character. The greatest benefits are likely to emerge in the case of regions, ceteris paribus, with inherently less costly access to foreign markets - border or port regions.

Rivas (2007) postulates that an ability of a given region to capture the benefits related to greater trade openness depends primarily on critical endowments and thus the degree to which trade is likely to reduce regional inequality in a given country is mediated by the geographic distribution of its endowments. He tested the hypothesis for Mexico on a sub-national dataset from 1940. The results indicate that liberalization benefits to a greater extent region with lower levels of education, thereby tending to reduce regional inequality. However, it also benefits more regions with higher levels of income and infrastructure, thereby tending to increase regional inequality. The second effect is stronger, thus trade openness increases regional inequality.
Krugman (2011) states that recent developments in the regional evolution of China are in line with the postulates of the standard core-periphery model that predicts increasing regional specialization as a result of economic integration (liberalization).

Redding (2010) points to a major weakness of NEG theory in the empirical domain. Earlier Krugman (1998) noting the theoretical contribution of NEG by incorporating space into mainstream economics, fully inspecting the impact of increasing returns and identifying many non-linear aspects, stated that the weakness lied in lack of convincing empirical verifications. Nonetheless, some of the analyses for instance by Davis & Weinstein (1996, 1998, 2002, 2008) were very encouraging, however, pointed to some important problems such as the assumption of homogeneity of preference (not in line with the rebuilding of Nagasaki or Hiroshima after WWII). On the other hand, Davis & Weinstein (2003) proved the existence of the home market effects for a broad segment of OECD manufacturing industry.

It is worth pointing out that economic geographers question frequently the basic assumptions, the structure and logic of NEG models. Martin (1999) considers it just as a reworking of traditional location theory and the regional science using recent developments in formal (mathematical) mainstream economics. Therefore they also question the policy relevance of this kind of modelling (e.g. Martin & Sunley, 2010).

Krugman (2011) on the other hand thinks that the possibilities of convergence between economic geography and NEG are rather dim by stating “Although both economists and geographers study these spatial processes, no fruitful exchange between the two is expected because of the use of different methodologies”.

Garretsen & Martin (2010) point to two important weaknesses of NEG models that is the oversimplified treatment of geography (pre-given, fixed and highly idealized abstract geometric space) and history (logical time and not real history). One of the first steps in this process is made in an empirical paper by Bosker et al. (2010) where the authors move from unidimensional NEG model and propose a strategy combining estimation and simulation accounting for heterogeneous and complex geographical structures. Combination of the estimation of structural NEG parameters with a simulation of the underlying multidimensional NEG model more accurately links the empirical results to the theory.
Storper (2011) postulates that only big changes in the openness to trade (for instance adjustments in the tariff levels or establishment or abolishment of trade barriers or trade liberalization schemes) considered as major shocks to economic system may result in significant structural adjustments in firms decision resulting in their reorganization and potentially re-location as long as the shocks are big enough to overcome existing sunk costs and/or agglomeration economies and then readjustments in the location of labor.

Yu, Zhao & Ming (2006) analyzed the causes of industry agglomeration in China at the provincial level of spatial disaggregation over the period 1987—2001 using the NEG modelling framework. The authors conclude that liberalization led to industry agglomeration. Furthermore, the factors such as market size, the level of urbanization or investments in the infrastructure promoted industrial agglomeration. Coastal Chinese regions enjoy clear geographical advantages promoting industrial growth.

Conclusions

Summing up, the bond between economic growth and openness of regions is a complex and to some extent ambiguous. The relationship becomes even more blurred if we take into account spatial nonlinearity into account. Nonetheless, overlooking the spatial interactions at the regional level of analysis, and thus of NEG or NNEG postulates, could lead to falsified conclusions or policy recommendations. It is not the case if the analysis is conducted at the national level.

Therefore the analysis of the impact of openness on economic growth of regions should take into account the various channels discussed, try to prove the direction of causality and acknowledge the presence of spatial interlinkages between regions. This is in particular important for mid-size and large-size economies with a significant number of regions and large regional heterogeneity frequently due to more deep-rooted factors.
References


Carkovic, M. V., & Levine, R. (2002). Does foreign direct investment accelerate economic growth?.


