Author:
Miroslav N. Jovanović
University of Geneva, Global Studies Institute, Dušan Sidjanski Centre of Excellence in European Studies, Switzerland

SPATIAL LOCATION OF FIRMS AND INDUSTRIES: AN OVERVIEW OF THEORY AND PRACTICE V2.0*

ABSTRACT

Where economic activity will locate and stay in the future is one of the most important and challenging questions in economics. Even though advances in technology have reduced the cost of transport and communication, which has curtailed the 'distance penalty' for business operations, local proximity of firms that produce similar, competing and/or related products still matters, especially at times with proliferating economic sanctions and trade wars. This reinforces the absolute, rather than relative, advantages of many small areas. The location of firms depends not only on costs of production and marketing, but also on economies of scale, activity-specific backward and forward linkages (indivisible production), accumulated knowledge, path dependence, innovation, the existence of sophisticated customers (markets) and on unpredictable chance events and historical accidents. 'Global' competitiveness often depends on highly concentrated 'local' knowledge; capabilities and common tacit codes of behaviour which can be found in spatially concentrated firms (clusters). Widespread global value chains are ageing and are being redressed because of uncertainties related to supply routes, 'near-shoring' and economic sanctions. Although there are certain principles and lessons learned for the location of firms and economic policy, the issue is still complex, evolving and subject to further theoretical and empirical analysis.

Keywords: Location; Geography; Cluster; History; Accident; Chance; Multiple Equilibria; Lock-in; Path Dependence; Links; Firms; Multinationals; Economies of Scale; Global Value Chains; Sanctions; Trade War
JEL Classification: F1; F23; L52; R10

---

* Over two decades elapsed since the publication of the first version of this article. Judging by the continuous downloads (https://papers.ssrn.com/sol3/papers.cfm?abstract_id=451800) of the article from the Social Sciences Research Network (SSRN) site and a revealed interest in it, I was encouraged by professor Amedeo Amato to update, extend and prepare the present version 2.0. I dedicate this article to the memory of this great expert and personality.
I have benefited from intellectual capital and discussions with many friends, colleagues and students, but I owe a special debt of gratitude to my teacher Richard G. Lipsey for the shaping of my view on how the economy exists and operates. I am also indebted to referees, Giovanni B. Pittaluga and Angela Procopio for their full and continuous support. This article was edited by Charles Toby Pearce. The views expressed are my own and do not necessarily reflect the position of the institutions for which I work. I am solely responsible for any potential inaccuracies stated herein.

RIASSUNTO

Localizzazione geografica delle imprese e delle industrie: un'analisi teorica e della pratica 2.0

Uno dei più importanti e stimolanti quesiti dell'economia è capire dove le attività economiche si stabiliranno in futuro. Anche se il progresso della tecnologia ha ridotto i costi dei trasporti e delle comunicazioni, diminuendo così la penalizzazione portata dalla distanza nelle operazioni commerciali, la prossimità locale delle aziende che producono merci simili, concorrenziali e/o correlate è ancora un fattore da considerare, specialmente in tempi in cui le sanzioni economiche e le guerre commerciali proliferano. Ciò rafforza i vantaggi assoluti, più che quelli relativi, di molte piccole aree. La localizzazione delle aziende dipende non solo dai costi di produzione e marketing, ma anche da economie di scala, da integrazioni lineari in attività specifiche (produzione indivisibile), dal saper fare, da dipendenze dal percorso, dall'innovazione, dall'esistenza di clienti sofisticati (i mercati) e da eventi casuali imprevedibili e da circostanze storiche. La competitività 'globale' spesso dipende dal sapere 'locale' che è altamente concentrato; da capacità e codici comportamentali tacitamente riconosciuti che si riscontrano nei distretti imprenditoriali. In tutto il mondo le catene globali del valore stanno diventando obsolete e vengono ripensate a causa dell'incertezza legata ai percorsi delle forniture, agli aiuti ed alle sanzioni economiche. Benché ci siano alcuni principi e lezioni apprese dalla localizzazione delle aziende e dalla politica economica, il problema è ancora complesso, in evoluzione e soggetto a ulteriori analisi teoriche ed empiriche.

Anyone who hears and obeys these teachings of mine
is like a wise person who built a house on solid rock.
Rain poured down, rivers flooded, and winds beat against that house.
But it did not fall, because it was built on solid rock.
Matthew 7:24–5

1. INTRODUCTION

Many changes and developments have taken place since the publication of version 1.0 of the present article in *this Journal* in 2003. These include the spread and, later, shortening of global value chains (GVCs); the accelerated death of multilateralism; deglobalisation; the coronavirus pandemic (COVID-19); wars; blockades of transport corridors; the spread of economic sanctions and retaliations; enormous subsidies; a policy shift towards 'the green economy'; Brexit; turmoil

in the North American trade deal; mass immigration of low-skilled labour; the gig economy; increasing energy costs; as well as the United States' (US) continuous and aggressive trade war against everyone[1]. This begs for an update and revision, so here it comes with a reliance on version 1.0 and Jovanović (2009, 2019, 2020, and 2022).

Where economic activity will locate in the future is one of the most important and challenging questions in economics. Progress in technology and moves towards and away from a liberal economic policy create new challenges for theorists, policymakers and business executives. As a number of economic activities became 'footloose' and highly mobile, one of the most demanding and intricate questions in such a situation is where firms would locate, re-locate or stay. The purpose of this article is to provide a survey of the theoretical foundations of spatial location of firms and industries. It deals with the rationale for the location of business in a certain geographical location, as well as with various considerations and conditions that influence the issue of how firms and industries fit, come and stay in a certain space.

The structure of this article – which has the flavour of evolutionary economics – is as follows: this introduction leads towards section 2 which considers the theories of location of national firms. Multiple equilibria are the subject matter of section 3. Section 4 explains location possibilities and outcomes. Reasons for the spatial clustering of firms and the formation of cities are presented in sections 5 and 6, respectively. International firms are tackled in section 7. Tax incentives are the subject matter of section 8. Section 9 explains how the historical lock-in effect and expectations shape the geographical location of firms. GVCs are covered in section 10 before the impact of economic sanctions on location matters are discussed in section 11. The conclusion (section 12) is that although there are certain principles for the spatial location of firms, the issue is still tricky and open to further theoretical and empirical analysis.

Spatial economics (geography of production, i.e. the spatial relation among economic units) has been separately analysed as a sub-branch of a number of subjects within economics. It has, however, been treated with different intensity within those research disciplines. Contributions are scattered widely across fields which include microeconomics; planning; development; economic geography; regional science; urban economics; location theory; industrial organisation;

[1] The US trade war against everyone, especially China and Europe, forces trading nations to search for partners elsewhere. The locus of international trade shifts towards China, Asia, the Middle East and Europe, away from the US.

international trade; foreign direct investment (FDI); transport economics; business economics; innovation studies; public finance; price theory; imperfect competition; labour economics; environment; and resource economics. The common research denominator in all these fields was the spatial dimension, which was seen as an opportunity, a medium for interactions, and a limitation. As a multidisciplinary research field, spatial economics and its evolutionary research strand has integrated these previously separate research areas.

The problem of the spatial location of firms does not exist in a situation with homogeneous space and perfect markets. The problem exists only in a situation with market imperfections. Without market imperfections such as transport cost (so that location matters), the production location decision becomes irrelevant. This is to say that without market failures, firms can split into units of any size and operate in all locations without any cost disadvantage. Empirical research found

"that the degree of production indeterminacy is greatest when trade barriers and trade costs are relatively low" (Bernstein and Weinstein, 2002, p. 73).

In addition,

"The Spatial Impossibility Theorem implies that when space is homogeneous and transportation is costly, the only possible competitive equilibrium is the so-called backyard capitalism in which every location operates as an autarky" (Fujita, 2005, p. xvii).

While considering feasible locations for its business in a situation with market imperfections a firm may, according to the neoclassical theory, either:

- consider *cost minimisation* (prefer the spatial location with the lowest-cost production function, and ignore the demand side); or
- emphasise *profit maximisation* (demand/revenue, and neglect everything else).

In practice, neither of these choices on their own necessarily offers the optimal location in terms of profit. Therefore, when considering possible locations for its business, a firm studies both options simultaneously. The firm attempts to maximise profit (in the long term), taking into account the operating 'cost penalty' linked with alternative geographical locations.

Apart from these basic, static, 'mechanical', 'clean' and direct neoclassical elements that are

strongly entrenched in theory, it is essential to employ at least two other approaches in the consideration of spatial location of firms:

- The *behavioural approach*. This is employed because of the internal dynamics of a firm. The real situation is charged with imperfections, risks, ambiguity, diversity of beliefs and strong endemic uncertainties. Therefore, profit optimisation or maximisation may not be the ultimate choice and goal for the firm. In a situation with bounded rationality and imperfect information, firms may choose to have a 'profit-satisfactory' behaviour as their business goal.
- The *institution-environment-related approach*. Firms make their decisions actively in a dynamic and risky environment in which there are multiple choices and limitations. Firms have to obey laws; fulfil obligations (taxes) and negotiate subsidies with the government; negotiate contracts with suppliers, clients and unions; organise and control production, distribution and service; consider new products, technologies and markets; decide about marketing and prices; and the like. In these circumstances, small and medium-sized enterprises (SMEs) often have relatively little advantage, while large firms fare much better.

In the distant past, the resource endowment of a specific location, together with the available technology, was the principal determinant of location for the manufacturing industry. In the new situation, in which most determinants for the spatial location of firms and industries are mobile and man-made, the location determinants for (footloose, fragmentable, mobile and internationally connected) industries and firms include the following five sets of considerations:

- *Costs and prices* (*C+P*): availability, substitutability, quality and prices of inputs (raw materials, energy, labour); cost of market access (trade costs); economies of scale; utility costs; infrastructure (physical, social and cultural); transport cost of inputs and output; earlier sunk costs in other locations (not yet depreciated); availability of investment funds; and costs, including sunk costs, of the project.
- *Demand* (*D*): its real and potential size, its growth and consumer preferences.
- *Organisation and technology* (*O+T*): input-output production links; externalities; competition and market structure; location decisions of other competing or supplying

firms; technology and the speed and direction of its change; and local research and development (R&D) resources and capabilities.

- *Policy-related factors* (*P*): incentives; taxes; subsidies; public procurement; rates of exchange; permissions; national defence; environmental standards; and mandatory or voluntary unionisation of labour.
- *Social factors* (*S*): spread of literacy and bureaucratisation of the society; system of education of management and training of labour; brain drain; income distribution[2]; general quality of life (soft location factor); and retirement patterns[3]. One should also include here the whims of investors.

$$L = f(C+P, D, O+T, P, S) \quad (1)$$

The decision about the location (L) of a firm or an industry in space is also linked with sunk costs and usually with the long term. The decision is in the function of these five composite variables (equation 1). Making a decision about the location of a firm or an industry is a complex, uncertain and risky task. It comes as no surprise that the national or regional production geography does not always change very quickly. Hence, according to the 'spatial indifference assumption', firms locate in spots where they think they may maximise profit. This is workable and tractable if one assumes a perfect divisibility of the production process and flexibility of factors. If, however, there exist economies of scale, then the idea of perfect divisibility of production must be relaxed.

There is a developed business in selling advice, including advice on location, because there are clients for such 'products'.

"Friedrich Hayek once said that he knew few people who had made money from acting on economic forecasts, but a good many who had made it from selling them. It is difficult for those in any profession to

[2] Income distribution also affects the location of production. If, for example, income is distributed to a small segment of the population (landowners or owners of capital) who spend it on imports of luxury goods and services, there may not be a spread of development of industries in their country. If, however, income is distributed in favour of a large number of families that demand domestically produced goods and services, this may have a positive impact on the location of firms and industries closer to the domestic market.

[3] Ageing population and retirement patterns influence the location of business and trade. The old demand health-related goods and services. For example, some retiring British and Germans move to Spain to enjoy a milder climate and a lower cost of living, as well as their pensions and life savings. As such, they create new demand in Spain for the local supply of services such as financial advice, health care, catering, cleaning . . .
In Japan, for instance, owing to the birth rate falls, greater longevity and an ageing population, the '*biggest paper company is to stop making nappies for children to concentrate on adult incontinence pads*' (Lloyd Parry, 2024).

stand out against the spirit of the times"[4].

If a system is unstable, then almost any small event (the butterfly effect - a chance) may trigger a 'train crash'. There is always some unwanted 'train' somewhere that should not be 'there'. Predicting triggers is hard and uncertain, especially when the system as a whole is unstable. What to do? Well, analysts and policymakers need to detect warning hints and sources of instability, rather than focus on triggers. To detect triggers may be too late to avoid a 'train crash'.

'Stuff happens', but does not happen rarely. Hence, the bell curve distribution of events needs to be flattened in reality. 'Black swans' may not be all that rare, for instance, in capital markets and elsewhere. What matters is not necessarily the 'black swan', but rather the response to it. Barbarians (black swans) did not destroy the Roman Empire, but rather it was ruined by its internal weaknesses and sicknesses that wrecked its social cohesion. They were corruption, anarchy, internal intrigues, criminality, violence, avoidance of military service; that is, the same phenomena that torment the contemporary Western society. The last emperors feared the attack by the barbarians; hence they erected grandiose walls around Rome. In vain. The walls are still there as witnesses of such illusions, because Rome was its own greatest enemy (Kalajić, 2015, pp. 203, 238).

Over-regulation, internal disputes and disagreements represent the existential dangers for the European Union (EU). The top EU statesmen have publicly expressed this danger a number of times[5]. When the US president Joe Biden gave excessively generous subsidies of over $1,200 billion to domestic firms (and foreign ones that wish to locate to the US) for 'green industries' announced in 2021 and implemented a year later, the EU was terrified because of the possible exodus of industries to the US. When the president Donald Trump undid that in 2025 and gave priority to fossil fuels and artificial intelligence, the EU panicked again. Has the EU lost its compass? To avoid the EU red tape, firms are either holding back investments or leaving for the US and China. The EU is losing its location-related attractiveness and competitiveness on a daily

[4] Brittan (2007).
[5] Matteo Renzi, Prime Minister of Italy (EurActiv, 2016).
Jean-Claud Junker, President of the European Commission (Brunsden, 2016).
Emanuel Macron, President of France (Abboud *et al.*, 2024; Le Monde, 25 April 2024; Waterfield, 2024).
Christine Lagarde, President of the European Central Bank (Fleming *et al.*, 2025).
Friedrich Merz, Germany's Chancellor-to-be (Connolly, 2025).

basis[6]. The way out of the existential crisis may be found in the simplification and avoidance of burdensome legislation.

The prevailing liberal (non-interventionist) attitude in the field of regional policy in the market economies had certain validity until the economic crisis of the 1930s. With free trade and perfect allocation and use of resources, intervention in regional policy is not necessary because markets clear and bring the economy to equilibrium[7]. Such a 'classical' or convergence approach was not validated by experience. In spite of such expectations, even the *laissez-faire* governments may consider it beneficial to intervene as the adjustment process may take too long to be politically acceptable.

Governments intervene in regional matters (location of economic activity) mainly for the following three reasons:

- *Equity*: this is a strong social motive based on public pressure on the government to try to achieve and keep a 'proper' balance, as well as an 'orderly' geographical distribution of national wealth among different regions.
- *Efficiency*: this is a desire to employ, sustain and increase national economic potential and capabilities.
- *Strategic behaviour*: this gives public authorities a chance to shape comparative advantages and influence the output potential and capabilities of the country.

The 'new' geography of production (spatial economics) emphasises growing structural disparities, rather than the equalisation among regions and countries. It introduces market imperfections and cumulative causation[8] effects into the old model. With the new assumptions, it indicates that integration and trade liberalisation may increase the geographical concentration of industries in the core countries and regions, rather than decrease it. Imperfect markets with

[6] "*A survey by the German Chamber of Commerce and Industry last September found that 43 per cent of large industrial companies were planning to relocate their operations outside of Germany, with the US being the top destination. On top of cheap energy with gas prices one-sixth the prices in Europe, lucrative subsidies in the Inflation Reduction Act for decarbonisation technologies are attracting European companies*" (Chu *et al.*, 2024).

[7] The tacit assumption is that there are no transport and adjustment costs.

[8] Cumulative causation mixes causes and effects of an event. They are combined in a chain reaction that is increasingly circular, snowballing, herding or perpetually accumulative. This type of self-reinforcement has different labels in economics which include: economies of scale, path dependency, virtuous and vicious circles, as well as threshold effects. The sources of this process are large sunk costs, learning and coordination effects.

economies of scale, entry and exit barriers[9], constrained mobility of factors or externalities may increase the attractiveness of already advanced areas. In so doing, it may reinforce their absolute advantage and it may widen regional gaps, rather than reduce them. However, *structural* disparity among the regions does not necessarily mean there is an *income* disparity between them. If there are two regions and if each region specialises in a different output activity, then economic structure among them would change, even though incomes in both regions may still remain comparable.

Intervention in the form of a regional policy is not a simple task as there are constant changes in technology, taste, competition, environment and demography over time. In spite of intervention in regional affairs over the past decades, a trend appeared during the 1990s in many EU member countries towards abandoning or easing regional policies. However, at the EU level, intervention in regional affairs remained fervent, but with questionable results.

The inclusion of transport costs and market imperfections (economies of scale and other externalities) expanded the classical theoretical concepts and moved them closer to reality. Given that spatial concentration is the most striking feature of the geography of production, there exists clear evidence of some kind of increasing returns to scale (Krugman, 1992, p. 5). New technologies in selected industries may overcome some of the obstacles to the spread of production, but not much. Hence, the pattern of regional specialisation and trade can be arbitrary and potential gains from specialisation and trade are likely to be ambiguous[10].

Even in a perfectly homogeneous world, production activities will tend to cluster, at least because of economies of scale. In these circumstances, government intervention (industrial, competition, trade and regional policies) is generally directed towards an increase in national or regional production potential, competences and capabilities through a (supposedly) superior allocation and use of resources from the national standpoint[11]. The objective of such intervention is to influence the spatial distribution of economic activity, as well as to create and redistribute wealth

[9] Entry and exit barriers include: huge sunk costs (in project investment and advertising), economies of scale, product differentiation (for example, market demands a specific brand name type of a drink), access to distribution channels, R&D, regulation (product quality), marketing, restriction of access to complementary assets (such as computer reservation system for tickets), reactions of competitors such as predatory pricing, exclusionary pricing, as well as trade, competition and industrial policies.

[10] The interested reader is invited to consult Brülhart (1998) for a brief survey of theoretical strands.

[11] A strategic industrial (and trade) policy is based on a number of assumptions that include non-retaliation by foreign partners and next to perfect information and forecast.

in order to ease and, eventually, solve the 'regional problem'. If policies mattered, then one would find an abrupt change in production activity when crossing the (national) frontiers at which policies change.

2. THEORIES OF LOCATION

The study of the location of production (spatial economics; economic geography) has a long history. In spite of the general interest, there is no wide consensus in the literature about the factors that influence the location of new businesses. The problem was noted, for example, by Ohlin who wrote that

"the theory of international trade is nothing but *internationale Standortslehre*" (teaching about international location) (Ohlin, 1933, p. 589).

In the same vein, Isard wrote that

"location cannot be explained without at the same time accounting for trade and trade cannot be explained without the simultaneous determination of locations ... trade and location are the two sides of the same coin" (Isard, 1956, p. 207).

A country's or a region's level of income and its structure of production (size, scope and variety of firms and industries) depend on at least two factors. The first includes a country's endowment of primary and derived stock of factors such as land, labour, capital, infrastructure, technology, as well as production, organisation, entrepreneurial and control skills and competences. The other element is how the country or region interacts with the external environment; in particular, the extent, level and form of its access to the international market for goods, services, capital and knowledge.

The classical English-speaking economists of the 19th century did not include institutions in their analysis. They also considered distance (space) only as a transaction-related issue. Elsewhere, in Germany and Sweden, researchers took space and geography into their studies. Heterogeneous space is important both for the location of economic activity and for trade. As such, they planted the academic seed for the theory of organisation and for the study of imperfect competition. The Swedish school (Hägerstrand, 1952, pp. 13, 18) emphasised local knowledge, accumulated human capital and face-to-face networking as some of the most important origins of agglomeration and

clustering. Myrdal (1957) argued that, contrary to the neoclassical teachings, inequalities between countries and regions are increasing. Without correcting policies, the positive spread effects of development from the centre are weaker than the backwash effects of impoverishment of the periphery.

One needs to distinguish three notions. *Specialisation* refers to the share of an industry in a given location; that is, if this share in the production of a good (say, cheese) is larger than the share of the same good in other locations. *Concentration* is the local grouping (clustering) of firms from the same or similar industry. *Agglomeration* denotes the grouping of many and various industries in one location. In any case, and by any measure, economic output is spatially concentrated and agglomerated:

"Half of global GDP [gross domestic product] today is produced on just 1.5 percent of the world's land, which would fit comfortably into Algeria" (World Bank, 2009, p. 96).

Let's now turn to the principal theories of firm location. Acknowledging the arbitrary character of the classification and its limits, we shall look into the issue of how and where firms and industries locate in space, while ignoring the origin, control and ownership of the firm.

2.1. Initial Considerations

A starting point in location (spatial) theory and economic geography is the national endowment of primary factors of production. Adam Smith and David Ricardo referred to the geography of production and spatial economics in an indirect way. Their theory of international trade looked at countries just as points in space. A country specialises in the production of a good for which it has an absolute advantage in production. Hence, according to Smith, an absolute advantage determines the geographical location of production. Ricardo's argument is similar, but in his model a country specialises in the production of goods for which it has a comparative advantage in production. Regarding land, the fertile areas are put into cultivation first, while less productive land enters use only as demand for agricultural goods increases.

The models by Smith and Ricardo depend on constant returns to scale and on the assumption that competitive prices guide spatial resource allocation. Hence, free trade allows countries to gain, because they specialise in production that they carry out comparatively well. Such an assumption

is challenged by the view that in reality there are important market distortions such as monopolies; externalities (pollution); production linkages (backward and forward); GVCs; temporal relations between present choices about location and structure of production and future production opportunities; coordination failures; sanctions; and controlled prices and rigid wages that do not tell the truth about social costs. As such, they send wrong price signals to entrepreneurs and may lead to the conclusion that trade liberalisation may make matters even worse. Thus, some argued that protectionist policies were required, for:

- the security of jobs, as was the case during the 1930s, and ever since; or
- the safeguard of infant industries, as during the 1950s; or
- rent-shifting in oligopolistic markets, as during the 1980s; or
- the safety of the environment and products (are they toxic, or do they contain genetically modified substances?), as well as social (labour, human rights) standards and practices, as is currently the case[12].

In addition to these classical models, the standard neoclassical factor endowment or Heckscher-Ohlin model refers to locational fundamentals. The national endowment of factors (land, capital and labour) and their relative proportion decide the location of production. A relatively abundant national factor of production (compared with other factors) determines what a country produces and exports[13]. Conversely, a country will import goods whose production requires factors that are scarce in the national economy. Considering cases of 'localisation' of industry in one or a few countries, Ohlin (1933, p. 133) simply argued that

"it must be shown that costs of production on the basis of existing factor prices are lower than in other countries: certain factors are cheaper here than abroad, which accounts for that condition".

In an empirical study of trade patterns of over 100 economies, Leamer (1984) confirmed that the factor endowment model could explain these trade patterns rather well. This model elegantly assumes that there are no specific factors and that production functions are identical in all

[12] The old protectionist rationale of national security is not dying either easily or fast, as was the case with the Dubai Port World case in 2006, as well as others. A recent protectionists' invention is the carbon border tax. This is a new tariff on imports. It is set according to the quantity of greenhouse gases that were released during the production process of a good. The message is that if you do not have our clean (and expensive) technology in your location, we in the developed world will tax you for the damage that you do to your own environment.
[13] The assumption is that there are no factor movements. Otherwise, the factor abundance proposition has no meaning.

countries. As such, it does not give an answer to the whole issue, for it cannot explain intra-industry trade. This type of trade (within an industry), which includes a large part of trade among developed countries, is not based on differences in factor endowments among them. The Heckscher-Ohlin model is more successful in explaining locational patterns in aggregate sectors (agriculture, manufacturing and services) than in disaggregated industries within, for example, manufacturing such as production and trade in passenger cars.

The above classical theories of trade attempt to explain trade (and the spatial location of production) among countries that have different properties (North-South trade). Contrary to this, the 'new trade theory'[14] tries to explain trade among countries that are similar, that is, that have no or few differences in the endowment of factors and in technology, as in trade among the developed countries (North-North).

The factor endowment model has been one of the two principal single-cause models dealing with specialisation, trade and spatial location of production. The other principal approach, the 'new trade theory', takes account of differences in technology, economies of scale (in firms), love for variety (among consumers), trade costs (including the cost of transport) and learning by doing. It was conceived by Krugman (1979, 1980) and later developed and expanded both by him and by a host of others. In a nutshell, a variety of causes including history, expectations, competition and changes in technology (not only factor endowment, as before) influence a country's specialisation, spatial location of production and, hence, trade patterns. Differences in technology and increasing returns to scale change productivity levels. This in turn alters the bases for specialisation, spatial location of production and trade patterns. In the case where there are both increasing returns to scale (falling costs) and increasing transport costs, there is an incentive to concentrate production. By spatially concentrating production near its largest market, a firm can both profit from economies of scale and minimise transportation costs. Economies of scale have a stronger impact on the location of modern, knowledge-based production than was the case before the Industrial Revolution.

Sources of economies of scale (as well as of 'natural monopoly') include:

[14] In spite of its name, the new trade theory approaches its midlife.

- *High set-up or sunk costs*: in fixed capital, R&D and marketing. These are historic expenditures that cannot be recovered, but that yield benefits in the future.
- *Learning effects*: the more people learn how to use one Microsoft program, the easier it is for them to learn and adjust to other programs.
- *Coordination and network*[15] *effects*: the value of a good or a service to a consumer is not only linked to the properties of that good/service, but also to the number of others who use it. The more people use mobile phones, e-mails, Internet, Microsoft Windows, WhatsApp, Viber, Facebook, Instagram, TikTok, electronic banking or credit cards, or PayPal, the greater the utility of these to all network users.

While one cannot dispute the importance of factor proportions for the location of production and some processing of primary resources, this approach cannot explain the location of a footloose, mobile and internationally connected industry. In addition, this model does not consider market structure, economies of scale, demand conditions or trade costs. The Heckscher-Ohlin theory cannot explain the location of an industry in regions with high mobility of factors, labour in particular, such as the US and China, or in countries with a broadly similar endowment of factors (France and Germany). A common feature in the standard trade theory is that it considers states, while it utterly ignores firms or regions, that is, the spatial distribution of economic activity within a country. This gap is filled by spatial economics.

Since the 1970s, there has been a notable and highly publicised spatial reallocation of textile manufacturing from the developed market economies to unskilled and cheap labour-abundant developing countries and China. The US bilateral general trade deficit with China is particularly worrisome for American politicians. If one scratches below the surface of this 'Chinese deficit story' in the US, one may discover that behind it is a hidden reallocation of low-cost manufacturing from other Asian countries to China. China became attractive as both the location for production and the source of supply. Why has this mass relocation of manufacturing, 'hollowing out' of American, European or Japanese economies and exodus of jobs to Asia, and China in particular, taken place over the past three decades? Why not earlier, say before the 1980s, since wages had been very low in Asia for a long time in comparison to the developed countries?

[15] Networks are very old and enduring social institutions. They also include families, clans, castes, tribes, gangs and nations.

This complex problem involves various issues. These include Chinese openness to foreign firms since the 1980s; standardisation of production and the possibilities to geographically fragment, unbundle and offshore parts of the production process (the number of footloose tasks increased); lowering of trade costs, especially in the case of GVCs; and a reduction in the difficulty to coordinate production over space and time zones.

China is a country that emits more greenhouse gases than any other country in the world. This is because China has many more people (1.4 billion) than the US and EU put together (785 million). Much of the Chinese pollution comes from the fabrication of goods that are meant for consumption abroad.

"Once the pollution that goes into traded goods is assigned to the country that consumes them, the average Chinese person harms the planet less than does the average European and much less than the average American"[16].

Harris (1954) argued that manufacturing firms tend to locate in places where they have favourable and guaranteed access to a large market. He showed that the industrialised US regions were also the areas of high market potential: a large market where demand is high and transport costs are low. This attracts producers. The idea is that firms choose to locate in regions that have good market access. However, this access is good in areas in which many other firms already exist and have chosen to settle. There is a self-reinforcing, snowball or herding effect in which the size of the market attracts still more firms. Larger markets host more firms, more firms produce a larger variety of products, while competition increases productivity and reduces prices. Such areas became attractive locations as export bases.

The underlining feature of this approach, like in most models of spatial economics, is that it is based on some kind of economies of scale that rely on a spatial concentration (agglomeration) of certain production activities. Such agglomeration economies, also considered to be external economies of scale, represent increasing returns in spatial form. In addition, agglomeration (clustering) of certain activities in a few locations is one of the most distinctive features of production. The positive difference in profits between regions, however, is incompatible with the neoclassical spatial equilibrium in the long term. Inter-regional equalisation of profits may take

[16] The Economist (2015).

place in the neoclassical theory either through higher wages in the agglomerated region (this lowers profits of firms in this region) or through the reallocation of firms from the congested region to the periphery. This, however, did not take place on a large scale within countries. Between countries, firms relocated certain (initially manufacturing) operations to China. As wages increase in China, investors are looking for other low-wage locations such as Vietnam, Cambodia, Bangladesh and Africa for select business operations.

The 'make where you sell' strategy is also prompted by protectionism. The trade war that started in 2018 between the US and China prompted BMW to stop exporting its X3 model from South Carolina in order to avoid Chinese tariffs on goods made in the US. BMW can make that vehicle at its plant in China[17]. Conversely, to avoid US tariffs, as well as rising land and labour costs in China, Samsung announced that it would end production at its factory in Tianjin and rely more on its plants in Vietnam and India. Nikon, Panasonic and Suzuki announced ending production in China and moving to Southeast Asia and Mexico. Foxconn considered moving to Vietnam[18]. Leaving China and relocating production elsewhere is easier said than done.

The US has been running a huge trade deficit for years. Two things ought to be clear from the outset. First, deficits may spring from different comparative advantages. Second, and more importantly, overall national trade balance is not the sum of bilateral trade balances, but rather the relation between domestic demand and supply. The bigger the domestic private and public (deficit financed) demand, the bigger the impact on imports. High tariffs reduce the deficit in trade with China, but they increase the deficit with other countries. Trade remains similar, but the location of production changes. There is also a reasonable expectation that there would be retaliation.

Vietnam turned out to be the principal beneficiary of the US trade war with China. The country became the location for production that left China. Vietnam had the fourth largest trade surplus with the US of $105bn only in 2023 (the fourth largest after China, Mexico and the EU)[19]. Hence, Vietnam may easily become the next target for the excessive US tariffs. That is the reason why foreign investors, such as those from South Korea or Japan, may delay or reduce production and investment in Vietnam.

[17] EurActiv (2018).
[18] Geopolitical Futures (2018), December 27.
[19] Lakshmi (2024).

In the presence of economies of scale, monopolistic competition and transport costs, countries tend to specialise and export differentiated goods for which they have a relatively large domestic market (economic integration secures and enlarges the 'local' market). It is this large domestic demand or *home market effect* (Linder, 1961; Krugman, 1980) that induces local production, rather than domestic factor endowment. The response by firms to strong local demand for a product (and the possibility of making savings in trade costs) is so powerful that the country first produces this good for domestic consumption and then exports it. Countries have a competitive edge in the production of these goods and thus gain an advantage in foreign markets, while they import goods demanded by a minority of the home population. The force of the home market effect influences the spatial distribution of an industry. If this effect is strong, then a fall in the cost of transport may not considerably change the spatial distribution of an industry.

Strong potential domestic demand creates superb grounds to kick-start an industry. The Japanese authorities wanted to have their children well educated in music. Public schools were required to purchase a piano. Foreign pianos were dear, hence the domestic producers Kawai and Yamaha secured the large home market for their products. Production started and Japan became internationally competitive in the production of pianos. The same holds for copiers before the advent of word processors (typewriters were unreasonable for some 60,000 Chinese characters). Hand-written documents had to be copied. Strong domestic demand existed for copiers and the Japanese became global leaders (Porter, 1990, p. 403).

The factor endowment or comparative advantage model would not predict the home market effect, but economic geography would expect such results. The US, Japan and Germany have or, at least had, the greatest comparative advantage in goods for which their home market is relatively big. Empirical investigation by Davies and Weinstein (2003) found support for the existence of home market effects, as well as for the role and importance of increasing returns in determining the production structure in the Organisation for Economic Co-operation and Development (OECD) countries. The emergence of China as one of the principal world manufacturers (especially since the late 1990s) is somehow modifying this general observation, but one needs to note that China and India each have immense and growing domestic markets.

Even though the model may seem attractive on the surface, it is based on many theoretical compromises. The model is silent on how to mobilise resources and integrate resources in a

selected location. Are there intermediate goods? What is the nature and form of transaction networks? What is the structure and dynamic of the local labour market? Is the necessary technology available? Where from? Is innovation relevant? What is the role of new products that trigger new consumption? What is the role of substitution effects? What about education and learning effects? Even though the model is in large part a 'black box' and narrow in its scope, it points research and policy attention to an important issue: the relevance of economies of scale and the local market size for the location of firms and industries.

The home market effect and national demand may not explain all cases of business location and specialisation. For example, a small country such as Switzerland is highly specialised in, for instance, research and production of pharmaceuticals and huge marine diesel engines (Sulzer Wärtsilä Corporation) for container ships. The Swiss do not have a heavy national leaning towards diseases or hypochondria, nor is Switzerland a country with a significant navy or a strong merchant fleet, whose national demand would justify or explain the location and specialisation in these types of production. The size of national demand is not always translated directly into the location of production at the same place. The accumulated past knowledge and technology, and various expectations about its future path, prices and income, all play their role in the selection of location, together with the size and growth of the market.

2.2. Domestic Firms

The period between the 16th and 18th centuries was marked by discussions between the proponents of mercantilist and liberal ideas both in France and in England (and even in Germany). Mercantilists stood for and encouraged the spatial concentration of production. This created an unequal spatial concentration of economic activity and wealth. Liberals, on the contrary, favoured a spontaneous spatial distribution and location of economic activity according to natural and rational order in societies. This would create an economic equilibrium and social equality.

Richard Cantillon (1680-1734), an Irish-French-Dutch-English economist and a banker, was one of the early academic explorers of spatial location of firms. In his *Essai sur la Nature du Commerce en Général* (Essay on the Nature of Trade in General) (published in 1755, but written around

1730), Cantillon predicted in chapters 4 and 5, respectively, the theory of concentric circles in the location of economic activity. This was later formalised by von Thünen.

Spatial economic analysis began in the strictest sense in what is now Germany. At the heart of von Thünen's (1966 [1826]) location theory are differences in land rents and land uses. This model, published at the start of the Industrial Revolution in Germany, is primarily concerned with the *location of agricultural production*. His ideas were outside researchers' focus for a century. The reason is the Anglo-Saxon preference for time over space. Ricardo focuses on the difference in the fertility of land and mobility of goods between regions and countries. He excludes factor mobility. Contrary to this, von Thünen studies how different agricultural lands around a city (marketplace) bid for various uses: labour moves freely within a region, while goods move at a cost. It was hard and costly to move milk and vegetables during von Thünen's time. This was the era when animals provided power for transport services. The model assumes a given isolated city (one consuming centre) and a surrounding agricultural hinterland, a homogeneous land surface, free and costless mobility of labour, and identical tastes at all income levels. Land rents are highest in the city. From there, they steadily decline to zero at the outermost limit of cultivation (beyond that limit, land is a free good)[20]. Hence, the rent in a given location totals the difference between the value of its yield minus the sum of production and transport costs. Under the above assumptions, the model explains the kind of crop that would be grown at places located at different distances from the market. In other words, farmers decide on the type of production (expensive to transport vegetables, and cheap to transport grain) by taking into account land rent and cost of transport. As transport costs and yields differ among crops, the result is a spontaneous development of concentric circles of production around the city. Land is allocated among different crops in a way that minimises production and transport costs of different crops.

Von Thünen's model is based on careful study of farming practices in northern Germany in the early 19th century and is focused on the differences in the cost of transport among different locations (Ricardo's interest was the difference in the fertility of land). Such distribution of farmland is less obvious in the more intricate modern world, but there are certain remnants around large urban centres such as dairy production and gardening[21]. There is no answer to the

[20] Does the world have endpoints in this model?

[21] Europe imports a huge amount of flowers from Kenya. This production of flowers in East Africa and their air freighting to Amsterdam, Berlin, London, Paris and elsewhere may easily leave a much smaller carbon footprint on the environment than production of flowers in heated greenhouses throughout Europe. Another intercontinental air flight

question whether vegetables can be shipped in both directions: towards the city (in exchange for cloth) and towards the outermost limits of cultivation (in exchange for grain). This early model does not consider the role played by economies of scale and the housing market (urban modelling). The model, therefore, is suitable for the analysis of a pre-railroad and a pre-industrial society.

The problem with von Thünen's model is to find which agricultural good to produce in the given location. In the *optimal plant site* model by Weber (1962 [1909]) the branch of industry is given, so the problem is to find the spatial location for production. This is the era of industrialisation and fully blown expansion of railroads. Weber's model takes the spatial location of markets, raw materials and population as given and assumes that there can be only one location for production. The objective of an individual firm is to minimise the combined costs of production and delivery. When the production costs are independent of location, the locational problem relates to the minimisation of transport costs (TC) for inputs and output. This is shown in the 'triangle' equation (2), where m presents weight of goods in tons, t stands for cost of transport per tonne/km, d is distance in kilometres, while i represents a particular good in transport. The key condition in the decision-making process about the spatial location of a firm refers to the minimal TC[22].

$$TC = \min \sum_{t=i}^{3} m_i t_i d_i \tag{2}$$

When key resources are used on a large scale in certain industries and are highly localised, this may affect the location of the manufacturing industry. These industries would be tempted to locate near resource sites. Hence, this case is 'out of step' with the spatial uniformity and the underlying principle of central-place location for some, but not all, kinds of manufacturing. The problem with the Weber approach to the location of new firms is in the initial linear and static locational needs of a firm, which may alter over time as the firm, industry, technology, GVCs, consumers' needs, income and tastes, or markets change and evolve. Weber assumes fixed and identical production functions at all locations (elasticities of substitution are zero). But firms always combine inputs and compare alternatives and substitutes. Weber's closed economy model

and carbon trail involves cod caught off Norway's coast. Cod is flown to China for cleaning, and then flown back to Norway to be sold (World Bank, 2009, p. 34).

[22] In the case with one market and two deposits of resources, the optimal firm location would fall inside the triangular area which links these three different spots.

does not take into account important location-related elements which include the cost of land and labour in the considered locations. There is also the possibility of a historical accident which can be coupled with the economies of scale, lock-in effect and agglomeration.

Moses (1958) extended Weber's work and integrated location theory with the theory of production. Coefficients in production functions were allowed to alter. This approach allowed investigation of the relation between substitution of inputs and the geographical location of a firm. Hence, Moses's approach gave different location results from Weber's because it allowed for the existence of economies of scale. The problem with the Weber - Moses transport-only approach to the location of firms is that it does not consider market prices and profit of the firm in the selected location. This is relevant, as transport, logistics and overall trade costs often represent only a small share of total costs in most firms.

During the Renaissance, it took almost one month for a letter to reach Florence from London. Correspondence between London and colonies in Asia could take a year. Technology reduced this time and cost at an amazing speed, especially since the Second World War. This provided firms with improved chances to decide their business opportunities and location. Nonetheless, countries still differ because of endogenous reasons: natural resources, geography, mobility of factors, size of population and income, sophistication of both labour and management, level of development, economic policies, preferences ... in spite of the fall in trade costs.

Geographical remoteness can be a handy scapegoat for some who may wish to use it to explain why peripheral regions are marginalised in economic and social terms. Hummels (1999) argues that geographical factors remain an important determinant of both trade and location. He suggests furthermore that trade composition patterns and FDI decisions can be sensitive to changes in the relative cost of different transport modes and information transfer. His proposition that international transportation costs have not declined significantly finds some support from Finger and Yeats (1976), who demonstrate that with the process of tariff liberalisation overseen by the General Agreement on Tariffs and Trade (GATT) and the World Trade Organization (WTO), the effective rate of protection in the form of transport costs exceeds that from tariff barriers for many goods. Hummels (2006, p. 19) stated that:

"Distance impedes trade to a surprising extent ... typical estimates suggesting that doubling distances halves trade ... costs are substantially rising in distance, but this effect has diminished over time"[23].

Even though road haulage is not considered in Hummels's article, it is interesting to note that in the period 1970-99 the value of trade in all products increased 18 times, while the tonnage transported by sea and air combined more than doubled (ibid., pp. 21-22). Hence, international trade became on average 'lighter' weight-wise, while transported goods happen to be dearer.

Even though there was a drop in transport costs over time[24], these costs have a stronger impact on trade than tariffs:

"This provides a plausible explanation for one of the most robust facts about trade: countries trade primarily with neighbours. Roughly a quarter of world trade takes place between countries sharing a common border" (Hummels, 2007, p. 151).

In addition,

"Roughly half of world's trade takes place between countries located within 3000 kilometres of each other. This picture has not been changed significantly by the fact that long-distance transport costs decrease more sharply than costs for short-distance transport" (Kopp, 2006, p. 91).

Geographical remoteness is a shallow justification for a lack of competitiveness in many industries. Because of innovations in transport and communication technologies, these costs are rarely the most important determinant for the location of many businesses. In addition, the cost of 'transport disadvantage' on exports is already included in the local factor prices. Hence, neither is a peripheral or geographically unfavourable location an insurmountable obstacle, nor is a core or geographically favourable location a necessary condition for economic prosperity or the

[23] Similarly, Grossman (1998, p. 30) reported that, "*two regions separated by 500 miles will, all also equal, trade more than 2.67 times as much as two regions separated by 1000 miles".* Limão and Venables (2001, p. 453) argued that, "*a 10-percentage-point increase in transport costs typically reducing trade volumes by approximately 20 percent*". Brakman and Marrewijk (2008, p. 427) argued that, "*international trade flows are to a large extent determined by distance; a 10% increase in distance reduces trade by about 9%*". Similarly, Brakman *et al.* (2009, p. 21) stated that, "*a 10 percent increase in the distance to the German economic center results in a 10 percent drop in export flows from Germany*". Tharakan and Thisse (2011, p. 122) reviewed the issue and reported that: "*distance has a significant negative impact on trade flows: roughly, doubling distance halves trade flows. Even more interesting is the evolution of this relationship over a long period. As expected, the impact of distance was shown to decrease slightly between 1870 and 1950 but, more surprisingly, it started to increase since 1950*".

[24] "*Road transport costs have fallen substantially, by almost 40 per cent over the past three decades, despite higher energy and wage costs*" (World Bank, 2009, p. 175).

location of a number of firms and businesses[25]. This is demonstrated in the cases of countries such as Japan, Australia, New Zealand, Finland, Ireland and Switzerland. Conversely, nor is a favourable geographical location in itself – of Somalia or Albania – sufficient for economic success.

Christaller (1966[1933]) attempted to clarify and explain the rationale for the number, size and spatial distribution of cities (in southern Germany). The inductive analysis of this geographer is based on the idea of market threshold and transport distance. He suggested that cities form a 'hierarchy of central places'. This hinges on the supposition that larger cities can sustain a wider variety of activities and offer more social conveniences relative to smaller (low-order) cities and villages. Economies of scale are the source of such uneven distribution of production. The consequence of Christaller's model is that if there are economies of scale, then the size of the market matters for the location of business. The concern here, therefore, is about the most 'national' product, below which are other products.

The hidden idea in this model is based on the minimisation of transport costs by rational consumers who make multipurpose trips. In spite of its obvious value for the analysis of urban growth and distribution of services, this rigid application of the impact of market size ignores the consequences of unequal distribution of natural resources, changes in technology and negative externalities that come from agglomeration. Hence, this model may be an interesting application of geometry rather than a model that explains accidental distribution and location of cities.

Plander (a Swede) was the first economist out of Germany who contributed to modern economic geography. In his PhD thesis *Beiträge zur Standortstheorie* (Contributions to Location Theory), completed at Stockholm University College in 1935, he argued that progress in technology generates concentration in manufacturing activities, while the same progress spreads services. Improvements in technology (especially the replacement of coal and steam with electricity) allowed fragmentation in the production process. The actual spatial location of segments in the production chain is linked with the available local factors of production only if the transport of inputs is unfeasible. However, this also depends on the substitutability of factors of production

[25] "*Instead, a 'peripheral' region is the one which has few sizable markets nearby, whereas a 'core' region is the one which is located close to such markets"* (Begg and Mayes, 1994, p. 91). South-east England is a core region in spite the fact that it is at the geographical edge. However, Kansas located in the middle of the US can be treated in economic terms as a periphery; while, for example, areas around New York or San Francisco may be treated as core regions.

and inputs. If substitution is not possible, then the location of production depends on the spatial availability of inputs. If, however, substitution is an option, then there are multiple equilibria related to the location of economic activity. Production techniques and available inputs influence the spatial location of production.

Developing further the *central-place theory* and the issue of how the economy fits into space, the economist Lösch (1973 [1940]) started with a useful, but most unrealistic, assumption that there is a perfectly even distribution of raw materials and population. He starts his deductive consideration of the homogeneous economic landscape (flat world) with a 'local' good and seeks to minimise transport costs for a given density of central places. This is not about central place conglomeration, but rather about a different problem: how sellers of identical goods would distribute themselves over geographical space, and how firms selling similar but differentiated goods would distribute themselves over space. In such a space, centres specialise in different products, hence there is a diversification of the economic landscape. An efficient pattern of central places would have the shape of nested hexagonal (honeycomb) market areas with no empty corners[26]. This means that certain economic activities can be done only at a restricted number of locations. It was subsequently demonstrated that there is a wide range of geographical configurations of firms. The spatial arrangements that can satisfy the equilibrium condition include squares, rectangles and regular and irregular hexagons (Eaton and Lipsey, 1976, p. 91). Even though based on unrealistic assumptions, the model of central places need not be disregarded out of hand. A coherent general equilibrium model found some justification for the central-place theory (Krugman, 1993c, p. 298)[27].

While the background theory in Christaller's model of central places is implicit, Lösch's separate approach made it explicit. When they seek to find synchronised locations for many kinds of goods and services they use different arguments. Lösch's model is more appropriate for the analysis of the manufacturing sector, while Christaller's hierarchies are more relevant for the tertiary sector

[26] If one imagines a geometrically even distribution of centres across a flat and homogeneous surface and the corresponding circles (representing ranges of goods) around them, then if one wants to cover all of the space with circles there would be an overlap between the two adjoining circles. If one draws a straight line between the points where the two adjoining circles intersect, one would get a hexagonal market space of identical size around each centre without empty corners, as consumers would purchase goods from the cheapest (nearest) producer.

[27] If we put aside differences in the local availability of natural resources and climate (production of spices), a symmetric spatial distribution of economic activities may explain to a large extent the spatial configuration of production in the pre-Industrial Revolution era. The Industrial Revolution brought specialisation and economies of scale. Spatial distribution of economic activities among regions and states started to diverge.

(retail services). In any case, it would seem that Christaller and Lösch both deal with planning problems, rather than with considerations of market results.

The theory of central places points to the factors that need to be examined during the decision-making process about the location of an industry or a firm. These factors are sources of supply, intersections of traffic routes and the centre of gravity[28]. Hence, a planner may draw geometrical figures that describe spatial location of cities. In order to reduce inaccuracies in such a 'technical' process (suitable for the centrally planned system), it is necessary to consider additional elements that are in the 'parcel' of a market-based economy, including actions of other functionally related firms, competitors, consumers and government policies.

The classical German location theory dealt with the locational decisions of firms which are in essence reduced to two issues: homogeneous distribution of natural resources over a flat space, and optimum cost of transport. This literature was obsessed with the geometrical shape of market areas in an idealised landscape, or with the optimal production site with given resources and markets. It ignored the crucial issue of market structure and competition. This was

"doing things in the wrong order, worrying about the details of a secondary problem before making progress in the main issue" (Krugman, 1992, p. 5).

Organisational issues such as institutional reality and policies were put aside. Institution-free theoretical models avoid the problem of the impact of various policies on the location or reallocation of firms and industries.

Industrial concentration may be explained in theory by proximity to inputs, localised non-pecuniary externalities and the effects of market size in industries sensitive to economies of scale. If isolated, these forces produced clustered production geography. If combined, these factors may sometimes offset each other. Nonetheless, it needs to be kept in mind that

"everything is related to everything else, but near things are more related than distant things" (Tobler, 1970, p. 236).

[28] Paris and Madrid have relatively central geographical locations in France and Spain, respectively. The same holds for Munich in Bavaria. In 1998, the capital of Kazakhstan was transferred from Almaty to Astana precisely because the latter city lies at the intersection point of major north-south and east-west transport routes.

Brülhart (1998) recalled the fact that what economic geographers call 'industry concentration', trade economists call 'inter-industry specialisation'. A low level of intra-industry trade (IIT) in industries that are subject to economies of scale may indicate the existence of centripetal forces and industrial concentration in space. Conversely, a high level of IIT could point to dispersion of an industry. As far as the EU was concerned in the 1980-1990 period, Brülhart found that there was no further concentration of already clustered industries that are subject to increasing returns in the central regions; there was further concentration of textile-related industries at the periphery; and there were certain indicators of spread of 'high-technology' industries towards the EU periphery.

Brülhart and Trionfetti (2004) analysed the outcome of home-biased procurement policies on the location of manufacturing. They found evidence in the EU that, on average, a country would specialise in production of goods for which it has relatively large home-biased procurement ('pull effect'). A government's preference for domestic bidders (discrimination against foreign ones) causes a shift in profits from foreign to home-based firms. In addition, they found empirical support for the 'spread effect':

"industries that are subject to a relatively large share of public expenditure tend to be less concentrated across EU countries" (ibid., p. 877).

Biased procurement received increased international attention. This was not only in the EU where the Single European Market was supposed to reduce the impact, even eliminate this non-trade barrier (NTB), but also in the WTO under the auspices of the Uruguay Round. Still, fully open public procurement remained one of the principal remaining obstacles to a fully-fledged Single European Market.

Paelinck and Polèse (1999) tried to pass certain regions-related integration lessons from the EU to the North American Free Trade Agreement (NAFTA). Basically, a region's gain/loss from continental integration depends on its distance from 'the continental core' (economic heartland) and on its endowment of 'scale-scope, competitive and externality elements'. In simple terms, the location of production is expected to move over time towards the nation's principal trading partner. This would favour regions that are located close to the continental core. Applied to the case of Canada in NAFTA (rebranded in 2018 as the US-Mexico-Canada Agreement, or USMCA), the national economic core is contiguous to the continental core. Most of the Canadian economic

activity and population is located in a 150 km wide (or narrow) band along the border with the US. In the case of Mexico, the national core (Mexico City) is located at some distance from the continental core (the US). Integration within NAFTA could increase certain tensions between Mexico City and more dynamic regions in the north of the country, along the US border, towards which migrate certain (parts of) industries previously concentrated in Mexico City.

3. MULTIPLE AND RAPIDLY SHIFTING EQUILIBRIA

3.1. Non-Predictability

One of the principal features of the evolutionary economic geography is the issue of 'multiple equilibria'. They create a situation for welfare ranking and set up many temptations to try to pick winners by politicians and bureaucrats. These temptations appear whenever something new arrives that is commercially viable. The selection problem, however, remains unresolved as the entire system is unstable. Hence, a small historical accident, a tipping point (chance, arbitrariness, a serendipitous and non-repeating event, a small difference in timing), remains the unique deciding factor that can sway the final outcome and influence economic behaviour for a long time in the future (Arthur, 2002, p. 6)[29].

The tipping point is the point of no return. This rare and unusual moment in time can be recognised most often in retrospect. These chances are beyond the control of actors (individuals, firms, associations, governments). Chances create discontinuities and possibilities to have jump-starts and failures. There are several possible outcomes for the future at that chance-related moment. A lot of information, analysis, experience and luck are necessary to recognise that moment when it happens, what causes it, and whether these are only transitory fluctuations or sweeping changes. This episode may set the course of events for a long time in the future. At that critical moment – that is, the trigger point – the butterfly effect (a chance)[30] may initiate a hurricane thousands of kilometres away and the creation of a new quality that may last a long time. Hayek (1944, p. 51) hinted that:

[29] Alchian (1950) considered, among other issues, the role of chance and luck in achieving economic success.
[30] Albert Einstein remarked that "*A coincidence is a small miracle when God chooses to remain anonymous*" (see https://quoteinvestigator.com/2015/04/20/coincidence/. Accessed on 27 January 2025).

"It is the price of democracy that possibilities of conscious control are restricted to the fields where true agreement exists, and that in some fields things must be left to chance".

Curzon Price (2011, p. 9) goes further on this matter and explains:

"To say that 'in some fields things must be left to chance' is to place more faith in spontaneous competitive processes than in conscious control by authority, especially an authority so far removed from the citizens it claims to represent. My argument in favour of institutional competition between EU member states is therefore not based on static welfare considerations, but on democratic rights and the dynamics of competition. Here, indeed, 'in some fields things must be left to chance', but the probability is high that mistakes will be quickly corrected and better institutions speedily designed. Conscious control, or harmonisation at EU level, on the other hand, is both undemocratic and undynamic".

Hence, in line with Hayek's (1945, p. 527) idea that the market is a process of search and discovery, as well as on the grounds that diversity and risk-taking provide reasons for the dynamism of the market economy, Kay (2009, pp. 4-5) stated that:

"Market economies do not predict the future, they explore it ... Centralised systems experiment too little. They find the reasons why new proposals will fail – and mostly they are right in their suspicions, because most experiments do fail. But market economies thrive on a continued supply of unreasonable optimism".

Agglomeration (clustering) produces rents, which holds firms and factors in a certain place, even if policy action would lead to a geographical move. There is, however, a threshold. Once costs to firms and mobile factors outweigh the agglomeration rents, firms and mobile factors would move elsewhere. Here comes the principal and sharp difference with the standard neoclassical equilibrium model in which small changes in policy lead to small economic responses. Evolutionary models predict dramatic non-linear effects in such circumstances. A small change at a crucial point in time (once the threshold is reached) can lead to unexpected and shocking outcomes (a locational 'catastrophe'). Non-linear reaction to policy measures makes it hard to predict the actual effects of a given policy change.

If governments want to tip markets towards a preferred solution, then it is timing and, to an extent, instruments that are crucial. There are few and relatively short moments of enhanced locational freedom for new industries that set in motion new industrial trajectories. Hence, there

is only a narrow window of locational opportunity during which the policy may be effective. The light through this window of locational opportunity is rare.

This 'window of locational opportunity' view at times appears to conflate ex ante unpredictability with ex post inexplicability, and it is possible to advance a different interpretation of initial events and the 'birth' of new path-dependent industrial trajectories that allows local context (and hence local prior path-dependent development) a more significant role in determining where (and when) new techno-economic paths emerge (Martin and Sunley, 2006, pp. 425-426). In a number of cases there may be a relative variety of potential locations for the establishment of business. The winner location often wins by the skin of its teeth.

Many things cannot be explained if one excludes a chance, that is, an initially insignificant small historical accident (the window of locational opportunity). It is unpredictable where new firms will emerge in space (Sony, Apple, Microsoft, Facebook-X, Huawei, easyJet, Foxconn, Mozilla, Nvidia, ...) with new technologies (or significantly improved old ones). Silicon Valley emerged in the San Francisco area, but similar conditions existed in Dallas (Texas) or Phoenix (Arizona). Why in the San Francisco area? An element of chance may help in explanation. New businesses need new and rather liberal spaces and new institutions. The existing institutions often have embedded (subsidised and protected) clients and there is little space for the new and competing ideas and firms.

Brian Arthur (1989, 1990b, 1994a, 1994b), argued that certain models of production geography give weight to differences in factor endowment, transport costs, rents and competition. In such cases, the pattern of production locations is an equilibrium outcome. Hence, in these models history does not matter: the locational system is determinate and predictable (Arthur, 1994a, pp. 49-50). However, if one takes increasing returns and multiple equilibria into consideration, the new dynamic model of selection and adoption has four properties that create serious challenges and difficulties in analysis and policymaking (Arthur, 1989, pp. 116-117). These properties are non-predictability, absence of optimality, inflexibility and path dependence. One may also add time delays to Arthur's four properties. We now consider these five characteristics in turn.

Meaningful multiple equilibria exist when firms make independent profit-maximising decisions about output, prices and location. This means that there is a chance for the existence of more than one possible outcome. Sources of this multiplicity in a non-linear dynamic system are found in

pecuniary externalities that originate in (large) fixed costs[31], increasing returns and imperfect competition. This result, or these factors, undermine neoclassical welfare theorems and introduce a possibility for the role of government in the selection process to influence and, if possible, improve the location-related result. This is a highly subversive development regarding the neoclassical equilibrium model that wants only one equilibrium outcome. If several outcomes are possible, then the neoclassical equilibrium theory wants to know exactly what they are.

"Besides positions of stable equilibrium, there are theoretically at least positions of unstable equilibrium: but their only significance lies in the fact that they are the dividing boundaries between two positions of stable equilibria" (Marshall, 1890, p. 424).

The question is: which equilibrium gets established? Will it be stable in the long run? Since a particular geographical space has only a limited influence on the location of new economic activity (especially footloose industries), what are the options for policymakers in this situation? Krugman (1991a) reasoned, on the one hand, that there was a belief that the choice was basically resolved by history. Past events set preconditions that move the economy from one steady state to another. This reasoning, found in the traditional literature, argued that history mattered because of increasing returns to scale, lumpiness (inseparability), activity-specific knowledge, sunk costs and capital goods. On the other hand, there is a view that the choice of equilibrium is determined by expectations[32]. This observation is based on the belief

"that there is a decisive element of self-fulfilling prophecy" (ibid., p. 652).

The relative importance of history and expectations depends on the structure of the economy and on the costs of adjustment.

Ex ante knowledge, experience and preferences of firms, as well as the potential of technologies, may not be sufficient to predict the 'market outcome' regarding the (possible equilibrium) location of new activities. The outcome is indeterminate as equilibria are not locally unique.

[31] Just as a reminder, *fixed costs* deal with minimum outlays that are necessary to start a business; these are all costs before the first piece can be sold (machines, shelter, energy, permissions, material, publicity ...). *Variable costs* change (rise and fall) in proportion to the volume of output (materials, energy, labour). *Average costs* are the sum of all fixed and variable costs divided by the total quantity of output. *Marginal cost* is the cost of producing the last or the most recent unit of product. *Opportunity cost* is what you missed by not doing something.
[32] Schelling (2006, pp. 115-119) discussed a range of expectations.

There is a coordination failure[33] because of multiple equilibria. Such a dynamic or evolutionary system may always be out of equilibrium. Hence, when there is much uncertainty about the location of firms, especially new ones (because of their creativity, unpredictability and chance), comparative statics is largely a futile exercise in such circumstances. In addition, multiple equilibria may introduce the possibility of a locational 'catastrophe' (when all firms relocate and agglomerate in only one region, or leave it). Therefore, the distribution of economic activity in space is potentially quite unstable and unpredictable in the (very) long term.

3.2. Potential Inefficiency (Absence of Optimality)

Choice and actions by players under conditions of risk and strong endemic uncertainty follow different, often more unpredictable, courses than is the case in an idealised static neoclassical world with a stationary equilibrium as the final goal. Alternative technologies may compete passively (like certain species in biology) or actively and consciously. Active and strategic competition relies on sponsorship such as aggressive advertisement campaigns in which each rival emphasises their own superiority. Increasing returns (i-activities) may uplift the development of technology with inferior long-run aptitude as firms make irreversible investments under uncertainty. Perhaps, adaptation in the medium and long term, rather than maximisation, may be a response to unforeseen events and changing environment.

Geographically isolated countries, or countries whose economies operate under economic sanctions, create and foster a survival mindset. Necessity forces them to produce economic policies related to survival, and they develop a habit of adaptation which they transform into strategies[34]. Hence, markets anticipate developments in a vague, rather than a rational way.

[33] The coordination problem may be illustrated on the example of a small village of Palanpur in rural India. "*Palanpur farmers sow their winter crops several weeks after the date at which yields would be maximized. The farmers do not doubt that earlier planting would give them larger harvest, but no one, the farmer explained, is willing to be the first to plant, as the seeds on any one plot would be quickly eaten by birds. I asked if a large group of farmers, perhaps relatives, had ever agreed to sow earlier, all planting on the same day to minimize the losses. 'If we knew how to do that', he said, looking up from his hoe at me, 'we would not be poor'*" (Bowles, 2004, p. 24).
Another example of coordination failure is in Bangladesh, which is known for floods. Chicken farmers regularly lost their stock due to the floods. For a long time there was no cooperation or coordination between them to try and alleviate the hardship this caused. However, they did eventually find a solution to the coordination problem: they turned to breeding ducks. These domesticated birds can remain afloat and survive during floods.
[34] Switzerland and Japan were successful, Bhutan and Haiti were not.

The Russian economy has been operating under certain types of sanctions for over a century. For instance, the siege of Leningrad (now St Petersburg) during the Second World War, and Western sanctions because of the conflict with Ukraine (from 2014). Those sanctions forced Russia to transform its huge and sleepy economy and become much more self-reliant. Russia converted the crisis (not willingly) into an opportunity. It was 'forced' to adapt from being a significant global food importer to an important global food producer and exporter (in particular of grain to Turkey, Egypt, Mexico, and so on). Great internal land resources were used in a superior way to supply domestic and, subsequently, international consumers. The importation and planting of genetically modified organisms were banned. The opportunity was seized upon by domestic farmers and food producers who not only increased their production, but also diversified it. In addition, in 2016,

"Russia overtook the US this year to become the biggest exporter of wheat ... Russian food exports are now worth more than Kalashnikovs and all other military hardware combined"[35].

A famous example of being stuck on a local plateau relates to the QWERTY arrangement for typewriter keyboards. The name refers to the first six top-line letters on the keyboard. Christopher Latham Sholes, a Milwaukee mechanician, designed this keyboard at the end of the 1860s. He was more interested in a distribution of keys that would prevent jams, rather than in fast strokes. In fact, the QWERTY distribution of letters on keys was designed to slow down typing in order to prevent mechanical jams. Augustus Dvorak, a professor at the University of Washington, designed the most comfortable, user-friendly and efficient keyboard during the 1920s and patented it in 1936. The Dvorak keyboard, allegedly, offers possibilities to significantly increase the speed of typing. However, this keyboard still languishes in obscurity. Even though it has repeatedly been shown that the QWERTY distribution of letters is suboptimal, this system continues to be the standard for keyboards. This is a clear case of market failure where an inferior product 'locks in' for a long time and takes the market.

A superior technology may have bad luck in the situation with uncertainty in gaining few early adherents. For example, the US nuclear industry is dominated by light-water reactors because of the adoption of such a reactor to propel the first nuclear submarine in 1954. Engineering literature asserts, however, that a gas-cooled reactor would have been a superior choice (Arthur, 1989, p.

[35] A. Medetsky *et al.* (2016).

126). If the claim by engineers is that Sony's Betamax is a technically superior system for videotape recorders, then the market choice of the VHS (in about 1980) did not represent the best outcome[36]. Similar arguments could be used for the initial triumph of IBM computers, with their user-unfriendly DOS operating system, over Apple's Macintosh computers during the mid-1980s[37]. Market self-organisation can lock-in inferior products or technologies. Liebowitz and Margolis (2001, p. 237) considered these alleged lock-in types of market 'failure': they argued that there were no failures at all and that 'good products have won'.

This neoclassical view of how the world operates regards the QWERTY example as an anomaly. Otherwise, the neoclassical equilibrium model cannot work. The problem, however, is that there are too many QWERTY-type examples. They demonstrate that small events at critical times in a complex economy have crucial and long-term implications on business and its location in space. This is exactly what makes sense of many otherwise unexplained or neglected occurrences in economics.

3.3. Inflexibility (Lock-in Effect)

Once an outcome, such as a dominant technology or a location, begins to surface it turns out to be 'locked-in' (stuck) and it persists for a long period of time. This 'overcommitment' to a particular technology or a location may be the consequence of:

- increasing returns based on huge sunk costs;
- strong externalities (positive feedback);
- an early government procurement made before other possible outcomes were genuinely considered and evaluated; or

[36] Sony expected that its reputation, and favourable opinions by the experts about Betamax, were enough to impose this standard and to capture the consumer market. Sony came first in the market and expected that people would primarily use camcorders to make home movies (birthday parties, weddings, family reunions, holidays . . .). Sony was wrong. VHS (JVC) had one important advantage over Betamax: its longer playing time. VHS devices filled homes chiefly for watching pre-recorded films. For this Betamax was inferior. The market referendum made Sony abandon Betamax.

[37] Even though it was user-unfriendly, DOS computers initially succeeded on the market over Macintosh because of the following merits. First, they were cheaper. Second, it was easier to create programs for DOS computers than it was for Macintosh, which had copy-protected programs; diskettes had to be kept in the computer so only one program could run. Third, DOS updates on the screen were instantaneous, while Macintosh users had to wait for changes to appear on the screen. However, the switch from DOS to Windows was rather quick in the early 1990s. Even though it was costly, the advantages of the new system were so great that they more than compensated for the trouble.

- a risk-loving firm may market its new product with a low introductory price or a money-back guarantee, just in order to create a critical mass of clients necessary to tip demand and new standards its way.

This, however, matters if choice is irreversible, which is often the case over a longer period of time. In order to replace an entrenched product or technology, the new one must be very much superior to the existing and embedded one. It must be available at the right price, and sufficiently take into account three factors:

- culture (openness to change);
- investment of capital and time (learning) into the new good or service (once people learn to use something, they do not want to switch easily to something else); and
- convenience in use.

Once locked-in, the system may be in a 'conditional evolutionary equilibrium'. This transitory 'stable equilibrium' may not necessarily be optimal. However, even if it is optimal at a given moment in time, it is subject to redefinition by the system's internal (improvements in something in existence) and external (radical change in technology) dynamics of change (evolution) over time. Even the locked-in QWERTY may be challenged by the new technology that converts voice (even thoughts) into a type-written text.

As far as commitment to a mission or business is concerned (lock-in effect), consider the decision by Hernán Cortés in 1519. After unloading his 11 ships in Veracruz (now Mexico), Cortés ordered all his ships burned on the spot. Even though they did not know what was waiting for them in the new land, his 600 men with 15 horses (never seen before by the native people) and several cannons had 'no other option' but to go inland. Retreat was impossible. Courage and risk-loving may be also necessary in business. This is a different spirit from the one in the Roman Empire before its fall. It was impossible to defend the imperial idea of Rome from the attacks by the barbarians in the situation in which there was a deep internal sickness and decadence of the overtaxed society. For instance, there was widespread self-mutilation (cutting off of fingers) in order to avoid military service (Kalajić, 1971, p. 126).

"John Gutfreund, chairman of Solomon Brothers, one of the most aggressive investment banks of the 1980s, said successful traders must wake up each morning 'ready to bite ass of a bear'" (Kay, 2004, p. 315).

"While commitment does not ensure success, however, the lack of it makes success highly unlikely" (Porter, 1990, p. 116).

Continuous changes in market demand and technology create a tension linked with old path destruction, as well as a need for and prospect of the search for and creation of new paths. The Aarhus (Denmark) area is known for agriculture and maritime services. Technologies related to ploughs (turning ground) and propellers (especially three-blade solutions for turning water) were understood for a long time there, hence the link to the wind turbine blades (turning air). A similar industry developed in California, but the solutions were based on the developed path-dependent airplane technology: a downwind-pointing two-blade propeller. The Danish solution, a three-blade propeller facing the wind, turned out to be superior to the one in the US. This was based on the previous local understanding of ploughs and maritime propellers, and helped in the development of the local globally renowned wind-power industry in Denmark.

Certain firms are intent on escaping and breaking constraints that impose on them the path-dependent legacy not only in technical and production dimensions, but also regarding various social and political compromises made in certain locations. For example,

"When IBM developed its own personal computer [PC], the company located its fledgling PC capacity in Boca Ratton, Florida, way outside of the manufacturing agglomeration in the North-east Corridor" (Audretsch, 1998, p. 24).

Traditional explanations for this reallocation would include strong unions in the North and weaker ones in the South; lower labour costs in the South; economies of scale; and a better quality of life in the South compared to the American North[38]. Evolutionary explanations would add the elements of avoidance of the silos mentality and locked-in thinking, chance, and even whims by the managers. Government intervention, new entrepreneurs (even immigrants) and transnational corporations (TNCs) may also contribute to changes. Therefore, one needs to keep in mind that,

[38] Quality of life, it needs to be stated, is an important factor in location decisions. The number of Japanese living in Shanghai fell by 17 per cent in 2012, mostly because of pollution issues. Panasonic started offering 'smog pay'. High pay and benefit packages do not suffice any more. What is the point in making big money if pollution hurts the family (Rein, 2014, pp. 95-96)?

"Almost always the men who achieve these fundamental inventions of a new paradigm have been either very young or very new to the field whose paradigm they change" (Kuhn, 1970, p. 90).

Certain industrial regions may become victims of their own past success. Institutional sclerosis, vested interests of large firms, labour unions and public authorities may oppose changes and adjustment to new circumstances. For example, the German Ruhr region was led into the

"trap of rigid specialisation" (Grabher, 1993, p. 275).

There was too much local connectivity, not only among firms, but also with the local political elite. Such proximity could, but will not necessarily, create lock-in effects that may obstruct breakthroughs. The property of old industrial regions is that there is a deep involvement of the government in local economic and social matters. The Belgian Walloon region is in the same situation. Strong and relatively successful regional organisational structures may, over time, become a liability. Diversified regions may absorb shocks more easily because negative effects of a shock are scattered across various economic activities.

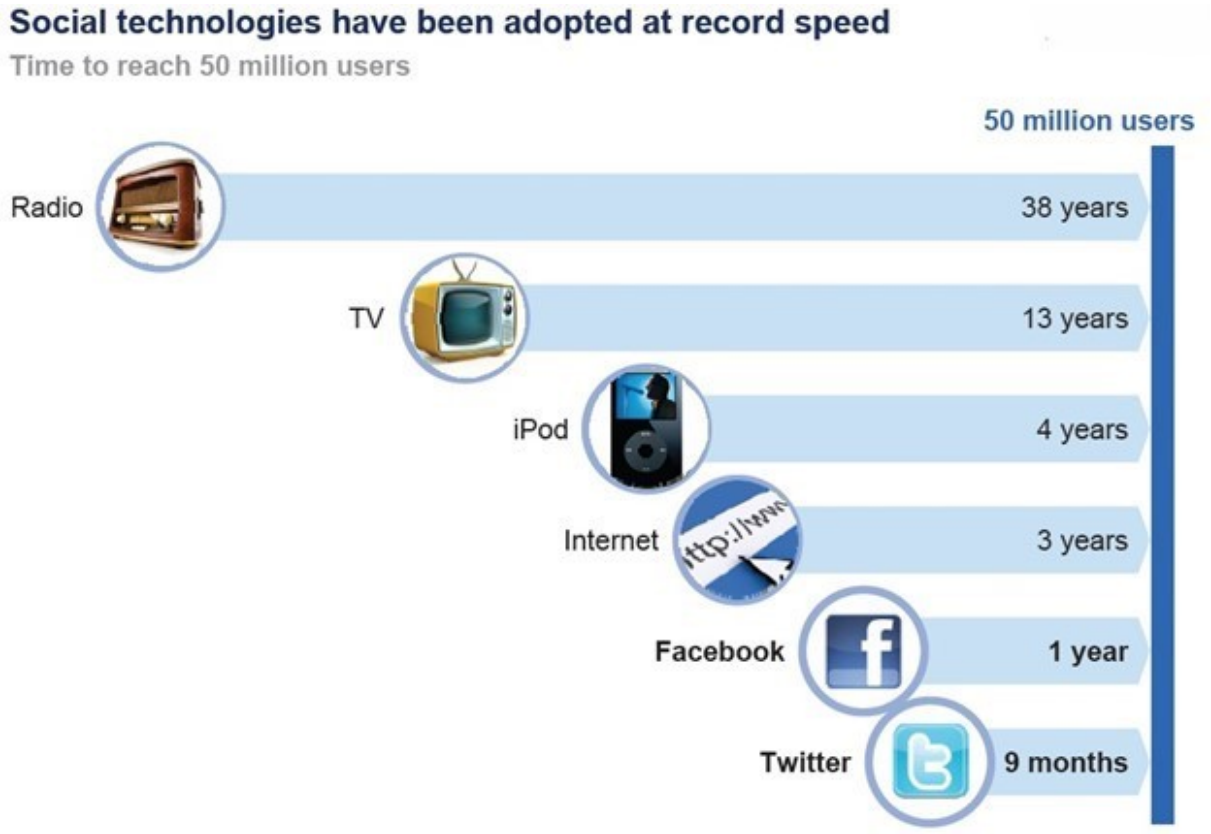
Figure 1 shows the rapidly dwindling time for the acceptance and use of novelty by consumers. While it took well over a generation and a half for radios to reach 50 million users, for the same number of users to harvest news on Twitter took just nine months.

3.4. Path Dependence (Memory), Strong History and Non-Ergodicity[39]

There are small, unpredictable, surprise, random and arbitrary events (chances, accidents, first-mover advantages, serendipity) at a micro level during crucial times in the past that may decisively and irreversibly challenge the established order and change our behaviour. These small and sometimes spontaneous past events (chances or 'accidents') in space 'chose' among various locational possibilities around which crystallisation starts. Hence, a small change and an

[39] An ergodic system (a pendulum; water in a glass on a table) ultimately returns to its original steady state, regardless of the mechanical disturbances between the starting and ending points in time. This system cannot shake free of its history, and yields a path-dependent outcome. Put loosely, the system is predictable. However, for predictions about the future of a non-ergodic system one needs to have certain information and knowledge about its past.

FIGURE 1 - *Adoption of New Technologies*



Source: McKinsey Global Institute (2012). The Social Economy: Unlocking Value and Productivity through Social Technologies, p. 22.

initial advantage (coupled with increasing returns to scale and sunk costs) may easily be turned into a large advantage in the future. In the real and second-best situation, and with increasing returns to scale in a non-ergodic system, these small events became magnified and self-reinforced, they tip the system towards a certain outcome. This inertia may set in motion mutations and routines in economic structures, at least for some time in the future[40]. Those two facts (chance and economies of scale) explain why there is no automatic economic convergence among regions if there is no certain policy action.

[40] This hypersensitivity to small events makes the mathematical model (equations) impractical for predictions about future behaviour. The problem is too complicated. Nonetheless, fathoming facts on how the system behaves and operates may be a great and useful achievement in itself.

Schumpeter was fully aware of elements such as chance and indeterminateness in economics. He wrote that:

"chance events may exert a powerful and lasting influence on the economic process and policy ... elements of random occurrence ... seriously limit our ability to forecast the future. That is what is meant here by 'a principle of indeterminateness'" (Schumpeter, 1991, pp. 441-442).

An example is IBM: when IBM wanted to launch its PC and enter the fast-growing market for Apple's products in 1980, it approached a small company in Bellevue, Washington. IBM wanted to consider the possibility that a small and flexible company called Microsoft would write a program for its PC. During the initial meeting, IBM asked Bill Gates, one of the founders of Microsoft, for his advice on the operating system that the new PC needed to run. Gates's suggestion was that IBM needed to talk to Gary Kildall of Digital Research who had an operating system that had become standard for hobbyist microcomputers. Kildall was doubtful about IBM's visit and when they came to meet him, he went hot air ballooning. Kildall's wife and lawyer, whom he instructed not to sign anything, met IBM's team. The frustrated IBM team turned back to Gates and asked if he would be interested in the operating-system project. Gates accepted the proposal. (IBM could have bought Microsoft, but did not want to. As a market leader in the production of mainframe computers, IBM could come under government pressure to break up the company. Hence IBM unbundled the software business.) While licensing the software product, Gates retained the right to sell his Microsoft Disk Operating System (MS-DOS) to non-IBM producers. As IBM was more interested in hardware than software, the deal seemed fine. The rest is history: a product complementary to a computer (that is, software) turned out to drive the market and to set the relevant standard. What would have happened if Kildall had not gone hot air ballooning? What if Microsoft had not got IBM's agreement to sell its products to other producers (Beinhocker, 2006, p. 327)? The effect of these small and seemingly irrelevant events for the long-term 'big picture' at that time when they take place is impossible to predict. Later on, they turn out to be a punctuating or a bifurcation point that decisively changes the evolution of an industry, even of a society.

3.5. Time Delays

The neoclassical equilibrium theory imagines that changes in the economic system are very fast, if not instantaneous. The real world, however, is full of imperfections such as a lack of inventories,

sophisticated personnel or insufficient production capacity. They all have an impact on the time delay between a policy action and the desired or expected feedback.

Time delays are inherent in dynamic systems that have multiple equilibria. Imagine that you are in a town that is new to you and that you check into a hotel that you do not know. Suppose that you want to take a shower in your hotel room. You turn the handle and you wait a little time between your action and the feedback from the water temperature. Once it starts flowing, it may be too cold, so you turn the handle in the direction of the red arrow. After a certain time delay, the water is too hot (the temperature overshoots the level desired by you). You then turn the handle in the opposite direction, and what happens? The water may be too cold (undershooting). Eventually, after some time you figure it out, the oscillations become smaller and, in the end, you have the temperature at the desired level (Beinhocker, 2006, p. 101).

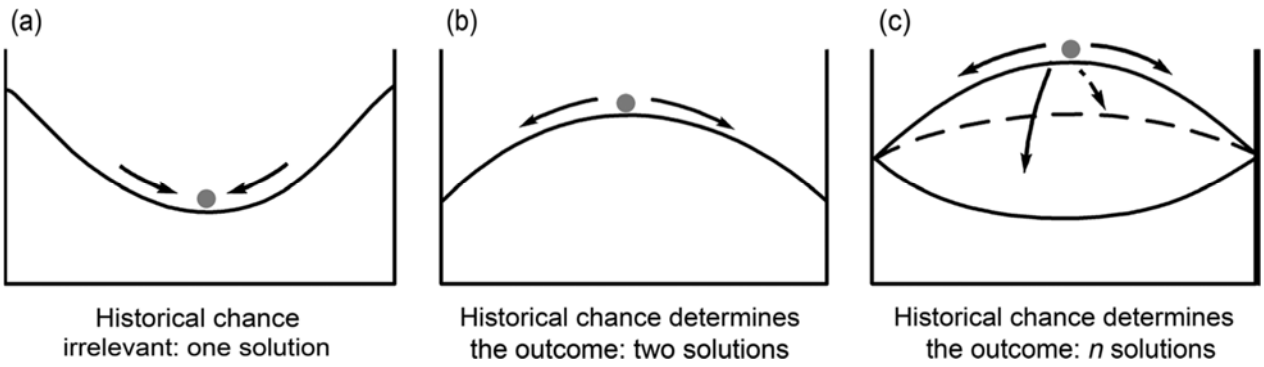
If there is a sudden surge in demand for a firm's product, and if this firm works at full capacity, the management may decide to invest in a new plant. This plant is built and starts operations after a certain period of time. But when this happens, the demand for the product may turn out to be temporary and one-time only. The firm is sunk with excess capacity and debt. Some period for reflection and analysis (a time delay) may be necessary before an investment is undertaken. However, if this period for reflection is excessive, other fast-moving firms may enter this business and skim the market milk. A forward-looking firm may wish to have a rather flexible production capacity (not too small and not excessively big), to try to link clients with longer-term contracts, to keep an eye on the current production capacity in its industry and how much is under construction, before deciding about new investment.

4. LOCATIONAL POSSIBILITIES AND OUTCOMES

The probability of the location of an industry resulting from a historical accident is shown in highly stylised, graphical form in Figure 2. If the distribution of potential locations of an industry in an imaginary space is concave (Panel (a)), with a single minimum and a corresponding single outcome, the location is not influenced by a historical chance. This type of distribution is exemplified by the mining and, sometimes, steel-making industries, which are usually located close to their source of raw materials. Mineral deposits (mines or oil wells or spas) determine such locations of firms. If, however, the distribution is convex (Panel (b)), with two minima, then there

are two potential outcomes, each resulting in a different location which may depend on the historical chance[41]. The third case represents a sphere in which there are *n* solutions for the location of footloose and mobile industries (Panel (c)). Hence, multiple equilibria make policy analysis conceptually difficult and uncertain[42]. This type of distribution is exemplified by corner shops, bakeries and petrol stations. However, firms in many industries need to be close to one another (that is, they tend to agglomerate and create clusters, towns and cities), not only to be close to common suppliers of inputs, but also to foster competition and to facilitate exchange of information, experience and knowledge, which can be hampered if firms are spatially dispersed.

FIGURE 2 - *Concave (a), Convex (b) and Spherical (c) Distributions of Potential Industry Locations*



If agglomeration forces based on increasing returns are unbounded, then a single geographical location monopolises the industry. Which region is selected depends on its spatial attractiveness and the historical accident of firm entry, in particular early preferences of first entrants. If, however, agglomeration forces that come from increasing returns are bounded, then various regions may share the industry as if agglomeration economies were absent (Arthur, 1990b, p. 249). Locations with a large number of firms cast an 'agglomeration shadow' in which little or no settlement takes place. This causes separation of an industry. Agglomeration occurs at a certain level of trade costs, while at another level of these costs the spread of activities takes place. With

[41] In order to keep transport costs down, the location of production of aluminium needs to be either close to bauxite deposits or near sources of electricity.

[42] "*When economic activities are not perfectly divisible, the transport of some goods between some places becomes unavoidable. In this case, the Spatial Impossibility Theorem tells us that no competitive equilibrium exists*" (Fujita and Thisse, 2008, p. 4).

bounded agglomeration economies, neighbouring locations cannot share the industry, but sufficiently separated regions can. In France, for example, Lyons lies between Paris and Marseilles. Bounded agglomeration economies caused separation and dispersion. Again,

"which locations gain the industry and which become orphaned is a matter of historical accident" (Arthur, 1990b, p. 247).

One of the most convenient, but among the most unrealistic theoretical assumptions, is that firms are identical. In reality, firms are heterogeneous even within narrowly defined industries, not only regarding their size and marginal production costs, but also regarding their trade, investment, R&D behaviour, technologies and the wages they pay. Firms move to be close to relatively large markets, not only because of clients and demand, but also because of the supply of inputs and marketing of output. Firms are more productive in the economic core than in the periphery, because they have lower marginal costs, and they tend to sell more. So, if a national regional policy subsidises reallocation of firms towards regions with a certain disadvantage (periphery), the implication is that the least productive firms from core regions (business agglomerations) may be attracted first by these (modest) subsidies. The opportunity costs of these low-productivity firms of leaving the agglomerated region are smallest. The selection effect of such a policy is that the most productive firms will move to (or stay in) the core regions, while the least productive firms will go to the periphery as they have the least to lose from the move from the core region. Modest subsidies have very little effect on regional welfare. They attract few firms and all of them are inefficient. Hence such a regional policy is ineffective in improving the competitiveness of remote regions (Baldwin and Okubo, 2005).

Problem-finding may in certain instances be as important and as hard as problem-solving. For instance, many great minds of the 18th century were occupied by the idea of how to turn lead into gold. Newton[43], Boyle, Locke and Leibnitz were among them. If such great minds were locked into such an intellectual dead end, what can one expect from other bright minds? How to filter the possible from the impossible? The mistake by great 18th century minds was the problem they chose (they were searching for the answer to the wrong question), not their method for solving it

[43] "*Many of his most creative years were devoted to alchemical experimentation*" (Mokyr, 2017, p. 100).

(Berkun, 2010, p. 129). George Bernard Shaw had a splendid thought on progress in his 1903 *Maxims for Revolutionists*[44]. The maxim on reason reads:

"The reasonable man adapts himself to the world: the unreasonable one persists in trying to adapt the world to himself. Therefore, all progress depends on the unreasonable man".

Entrepreneurs (sometimes seen as 'maniacs with a vision') often have genuine ideas, but many of them do not have the necessary knowledge about how to run a business. Some of them may know, but cannot persuade those with funds to support their projects[45]. Why is this so? Reasons include a number of cultural factors (risk-lovers[46] or risk-averse investors) and the lack of knowledge and understanding to recognise and seize the opportunity. Certain minimal material wealth is always necessary for survival, but wealth also exists in the form of knowledge and understanding. One of the principal defining features of humanity is the search for understanding, as much as the pursuit of happiness and the quest for beauty.

An example by Arthur (2002, p. 14) is quite instructive: a century ago a man went to a Belfast bank and asked for a loan to start a company there. He explained to the bank managers that his invention was a highly productive mechanical horseless device that would replace horses in ploughing. This was too risky for the bank, as it could not understand the technicalities and it could not foresee that the enterprise would be commercially viable. Hence this man, whose name was Ferguson, went to Canada. There he met another man by the name of Massey who understood Ferguson's idea. They teamed up, and the rest of the story is well known. The potential location of this business went from Ireland to Canada. Massey Ferguson became one of the major world producers of tractors and agriculture-related machinery. The moral of this story is that one needs to recognise, understand and have the guts to seize the opportunity. Carpe diem (seize the opportunity now; seize the day).

In order to profit from rare and exceptional opportunities, one needs to be endowed with both good knowledge, foresight and, yes, luck. Let us not forget that Harvard rejected Jack Ma, the

[44] See https://archive.org/stream/maximsforrevolut26107gut/26107.txt . Accessed on 27 January 2025.
[45] They need angel investors or 'super angel' investors. These are professionals with funds who understand new ideas that may be profitable and who are able, ready and willing to risk their money in start-up firms.
[46] Obtaining finances to start a small business has always been hard. However, Mike Markkula invested $80,000 to enable Steve Jobs and Steve Wozniak to start making computers. Markkula's choice to back Apple was spectacularly successful. He retired as a multimillionaire at the age of 32 (Kay, 2015, p. 167).

founder of the successful Alibaba Group, ten times[47]. Otto Wichterle, a Czech chemist, invented soft contact lenses at home in 1961. As such production was not part of any Czechoslovak enterprise plan, he sold his licence to a US firm. George Soros wrote that,

"There is very little difference between speculation and investment. Basically, the only difference is that investments are successful speculations, because if you successfully anticipate the future you make a speculative profit" (Soros, 2014, p. 84).

And there is the Chinese proverb:

"When the wind of change blows, some build walls, others build windmills".

5. CLUSTERS

Economic activity is not distributed randomly over space. Clusters of firms and people are not exceptions, but rather the rule. Clusters are a form of economic organisation of business that survives. Hence, clusters are an efficient or more efficient form of organisation in certain lines of business than are other configurations. Clusters are sometimes called industrial districts. They have various dimensions: geographical, historical, cultural and economic. This last dimension is the subject of this section.

Industrial clusters often determine what a country exports. Their success often establishes the competitive position of a country's goods and services on the international market. Prior to the Industrial Revolution, high transport costs splintered production into many units throughout geographical space. The effects of the Industrial Revolution reduced the cost of transport, but brought high fixed costs of production (and economies of scale) that contributed to the development of new agglomerations and the expansion of existing ones. In any case, clusters are often innovation incubators and depend on, and thrive when there is, demand for their output and when trade is (in general) liberal.

[47] Ma (2015).

5.1. Underlying Forces

A promising point of departure for the analysis of clusters is to consider two basic forces which, through their tension and interplay, influence the spatial distribution (concentration or spread) of firms and industries: the centripetal and centrifugal forces. These forces also help to explain the agglomeration and dispersion forces of the 'core-periphery' development model. A core region has the capability and potential to create, attract and employ resources and ideas. A peripheral region has its economic path determined chiefly by the developments in the core region(s). Let us consider the two background forces, centripetal and centrifugal forces.

Centripetal forces promote agglomeration (distance matters) and create long-term income divergence. These forces (linkages in production, thick markets and spillovers) form the 'trinity' of Marshall's external economies. They concentrate production and employment in specific geographical areas. These forces tend to be stronger during the early phases of an industry's life cycle because of tacit knowledge and non-existent standards (among other elements). Knowledge and experience are contextual and hard to codify. Transaction costs to transfer such vague knowledge and experience are quite high and it takes a long period of face-to-face interactions to establish and foster such a network of relations. Knowledge-based industries rarely fare well in isolation and away from clusters and big cities. Therefore, spatial proximity with face-to-face contacts and strong personal relations are necessary for staying in the business game. These 'snowball', 'me-too' or herding forces are based on certain complementarities among diverse players, and they include market size (concentration of firms and consumers); economies of scale; forward and backward linkages in production and distribution; trade costs; increasing returns in transport; existence of suppliers; limited spread of information and embodied knowledge; as well as a 'thick' labour market (especially for certain skills, so that employers can find workers and workers can find jobs). Workers can and do invest in industry- and firm-specific productivity-increasing skills, rather than in general ones that they may not need. Centripetal forces take the production system towards the spatial equilibrium.

Centrifugal forces (distance is dead) push the other way towards dispersion and test whether the equilibrium is stable. These dispersion (comparative advantage) forces discourage further spatial concentration of business. They favour a geographical spread of firms and include spatial factor immobility; regional or international wage differentials; relative cost of land rents; competition

for factors and consumers; commuting costs; pollution; congestion[48]; availability of parking and its cost; traffic accidents; crime; allergies; infectious and other diseases; lack of drinking water; sewage system; and waste disposal. However, these forces are still not enough to counter and compensate in full for the stubborn divergence created by centripetal forces.

The result on the location of business in space depends on the complicated balance between these two forces; on what grounds factors become mobile and on what grounds they become immobile; on barriers to the reallocation of resources; on demand and its change; and on public policy. In essence, three outcomes are possible:

- Economic activity may spread so that each region, industry and firm become more specialised in a certain activity. This type of clustering does not mean polarisation as is exemplified in North America.
- Activity may agglomerate in core regions, leaving others without production potential and without people.
- A long-run polarisation may split the country or an economic grouping into advanced regions with high incomes and low unemployment, and depressed regions with low income and high unemployment (Braunerhjelm *et al.*, 2000, pp. 29-30). This final result may turn out to be the creation of a set of growth poles and a set of growth sinks. One needs to distinguish, therefore, between 'good' and 'bad' agglomerations. Growth will be unbalanced, but that does not mean that it is not inclusive.

5.2. Meaning and Motives

A cluster is a relatively large group (a dense critical mass) of functionally related specialised firms in a particular and relatively small geographical area. Inseparably linked with this is a system of communication (formal and informal), collective learning, knowledge, coordination capacities, cost- and risk-sharing, skills and competences, as well as specialised institutions (standards-setting bodies, schools, universities, research institutes, trade associations, networks, rules and practices). These functional internal value-adding relations among firms may be downstream (suppliers), horizontal (competitors and collaborators), upstream (clients) and/or diagonal

[48] Traffic in the centre of megacities moves at the same speed as was the case more than a century ago when horses used to pull carriages.

through the circulation of accumulated knowledge, skills and competences by means of personnel turnover.

A concentration of functionally related business activities within a relatively small area (agglomeration, thick market effects, benefits of co-location, and non-ergodicity) provides firms with collective gains that would not be available if the firms operated in remote places or in isolation. These shared benefits or externalities are different from those that are created within and are available to a single firm. In essence, clusters create economies that are external to individual firms, but internal to a network of firms in a cluster. These economic gains are exchanged and enriched in a non-market way. Hence, a great deal of an individual firm's competitive advantage is outside the firm, but inside the location where it operates (for example, relations among suppliers, or gains from various types of partnership).

A firm locates in an area where there are firms from the same or related industry (in a cluster) because:

- it has production links with other firms;
- it may benefit from the already existing pool of suppliers;
- there are services such as finance, education, information, consulting and maintenance;
- there may be a pool of trained and experienced labour;
- firms may reduce the cost of transport;
- there may be a concentration of consumers (proximity to the major growing markets is often the most important reason for the selection of a particular location);
- there are competitors that may stimulate innovation in technology; and
- firms may jointly negotiate contracts with transporters and organise export promotion boards and the like.

Manchester (UK) in the 19th century provides an example of such a cluster specialising in textile manufacturing. Subsequently, the manufacture and repair of looms, and bleaching, dyeing and finishing facilities, also located there. In the course of time, the town set up a technical college in order to train people for the manufacture of machinery, textile design and other skills related to the local industry. All this was supported by marketing organisations such as the Cotton Exchange (Smith, 1981, pp. 6-61). However, even the most recent industries rely on the 'old concentration

rule' of geographical clustering. This is why there is a Silicon Valley in California and a Route 128 in Boston. The ingredients of success in these two clusters include a close proximity and interaction among entrepreneurs, some of the best universities, venture capitalists, as well as high labour mobility and immigration of highly qualified personnel[49]. To emulate that elsewhere in the short term would be quite tough as these ingredients and their combination may not be replicated easily.

A hint on the possible or approaching locational tipping point may be observed now in Silicon Valley. This small area from San Francisco to San Jose is the home of some of the finest engineering and entrepreneurial talent. It is also the origin of global corporate giants such as Apple, Facebook, Google and Hewlett-Packard, firms that changed the way we live and do business. Many have tried, many times, all around the world to clone the Silicon Valley success, but it was impossible. The area is not a simple assembly of various, many and rich talents and angel investors; it is also an idea, culture and spirit that may not be transferred like flu. Still, something may be changing in Silicon Valley. The cost of living is exorbitant. As a result, other locations became attractive for startups: Pittsburgh and Phoenix attract autonomous vehicles, Shenzhen hardware and London financial technology.

Is Silicon Valley becoming a victim of its own success? Strong centripetal forces ensure that certain spots remain 'sticky places in slippery space' (Markusen, 1996, p. 293), as agents that are near to each other exert more influence on each other than do the ones that are far away. However, high salaries often mollycoddle talents and may suppress entrepreneurial spirit and incentives. Newcomers may fill the possible void, but in other locations. Mixed with cuts in funds for universities and R&D, as well as an unfriendly policy stance towards immigrants, Silicon Valley may lose the fuel that made it thrive in the past. Locations where new ideas emerge and are put into practice may change in the future. For instance, billionaire Elon Musk decided in 2024 to move the headquarters of two of his most high-profile companies, rocket firm SpaceX and social media platform X, from California to Texas because of gender-related laws that affect and undermine the critical role of parents in raising children. High taxes also played a role in this relocation.

[49] "*ASML ranks among Europe's most valuable technical firms; about 40% of its 23,000 employees in the Netherlands are not Dutch*" (Kassam, 2024).

Agglomeration and the spatial clustering of firms and industries, the obsession in spatial economics, are motivated by efficiency considerations within a network. They are based on specialisation, economies of scale in production and transportation, as well as on transaction costs. For example, the Massachusetts General Hospital in Boston is among the largest in the world in terms of R&D funds. Industrial managers are attracted to the R&D department of the hospital because all the specialists and all the needed knowledge are 'within 20 minutes walking' (Lambooy, 1997, p. 298). This cluster of more than 400 firms linked to medical devices is invisible to an outsider as it is concealed in business categories such as electronic equipment and plastics products. Within a driving time of about 45 minutes is the area where labour market effects are the strongest (Venables, 2006, p. 10).

Agglomeration applies not only to private businesses, but also to certain public institutions. Bearing in mind the complexity of action and need for a fast response to changing and uncertain events, Simon Jenkins has noted that

"Britain is a centralised state. Two thirds of the top 100 powerful people in Britain work within 300ft of the Prime Minister and Chancellor of the Exchequer"[50].

Marshall (1890) considered the issue of why it is beneficial for producers from the same industry to locate together. Based on the analysis of the Sheffield metals industry in the UK, he offered three basic reasons why producers concentrate or agglomerate and form a cluster:

The first is *knowledge spillovers*: proximity eases exchange and spread of information. This increases efficiency in production through technical, organisational and control improvements. Workers learn skills rather fast from each other in a cluster, but this most often requires face-to-face contact and cooperative relations that take time to build.

The second reason is the advantages of *thick markets for specialised skill and inputs*: as the local demand for inputs is high, the unit cost of input is relatively low. Firms look at distance sourcing as a second-best solution. They have access to the first-best solutions locally. They may easily find

[50] Jenkins (2002). Regarding concentration of power, John Dalberg-Acton (1834-1902), First Baron Acton, has an often-quoted saying: "*And remember, where you have a concentration of power in a few hands, all too frequently men with the mentality of gangsters get control. History has proven that. All power corrupts; absolute power corrupts absolutely*" (see https://www.brainyquote.com/authors/john-dalberg-acton-quotes. Accessed on 27 January 2025).

the necessary labour, while workers may get promotion or work if the current employer does poorly; risk for both of them is reduced as the search and relocation costs are diminished. There is no need to create the human capital (skills, knowledge and health) from scratch. The 'mysteries of trade' (craft secrets, that is, very localised information and knowledge spillovers) are no mysteries at all as they are 'in the air' and children learn many of them unconsciously from their parents (vertical cultural transmission)[51]. Knowledge, experience and values are stored[52]. Intellectual capital is available. In addition, tacit and complex knowledge and shared culture of beliefs and practices is difficult to transfer without labour mobility, hence there is also social capital in the form of cooperation, support and trust which all enhance opportunities for innovation. In sum, the existence of a cluster hints that there is an opportunity where barriers to entry and starting a new business are lower than is the case elsewhere. Still, Boschma and Frenken (2018, p. 218) argue that in spite of the cluster-related accumulated knowledge that may be 'in the air' and physical proximity, not all firms have equal access to such local knowledge. Only the most connected firms have that access. Contrary to Marshall's opinion, this local knowledge and experience is less general and more selective.

The third reason why producers concentrate or agglomerate is the *backward and forward linkages* associated with large markets: links may clarify and explain a part of the concentration story only if there are economies of scale. Otherwise, a firm would set up a separate production unit to serve each distinct market. A concentrated industry provides a market for specialised local suppliers of components, as well as private and public suppliers of various services (institutional thickness)[53].

There are, however, other reasons for the spatial clustering of firms and industries, additional to the basic ones advanced by Marshall. They include the following:

- The past arm's-length and hierarchical organisation of a firm can be replaced by a flexible network of business organisations in a cluster, because there is a changed situation in competition and technology. Inter-firm competition is based on innovation.

[51] Learning from teachers is called oblique transmission, while learning from the members of one's professional group is horizontal transmission.
[52] When life expectancy is only 30 years, then the society is devoting most of its resources to pass the current knowledge from one generation to the next. Once life expectancy increases, there are resources in the society that may be devoted to an increase in the stock of knowledge (Boulding, 1991, p. 14).
[53] This institutional capital needs to be propped up by 'political capital', that is, capacity to carry out shared goals and visions; 'social capital' in the form of cooperation, support, trust and mutual gain; and 'intellectual capital' such as knowledge and experience.

- Firms may reduce transaction costs.
- The presence of one firm or industry creates a direct or indirect market for another firm or industry. One industry's output is used exclusively as another industry's input (for example, tyres and cars).
- Two or more industries use a common resource.
- Firms use common services (marketing, storage, accounting, repair, transport).
- Firms use common social infrastructure (schools, health care, roads).
- Firms may create entry barriers (for example, entry into the Geneva watchmakers' guild was very hard).
- Firms may have the intention to build a shared brand (for example, Solingen knives, Parmigiano-Reggiano cheese and Parma ham, Champagne, Scotch whisky, Californian wine).
- Labour markets of two unrelated industries may be complementary: one industry (metal or car assembly) uses male labour, while another (textiles or food processing) employs female labour.
- The more the firms potentially use certain new infrastructure, the more likely it is that infrastructure will be built. Each user in a cluster assists other users in this situation, hence the fixed cost of an infrastructure project (such as an educational or training centre, power plant, railway or airport) can be recovered and the project can come close to profitability. These are relatively standard projects, hence the government can readily be involved as specific local knowledge is not always essential for commercial success.

Competition over position in space and price is responsible for the clustering of firms. The clustering of firms selling identical goods, such as ice cream on the beach, was thought to be welfare-inefficient from a social standpoint (Hotelling, 1929, p. 53). However, the clustering of firms often occurs in response to consumers' desire to make comparisons between goods (for example, in shopping malls). Retailing exists because individual consumers are small in terms of wealth, and sometimes both ignorant and immobile. Consumers usually visit more than one store prior to making a purchase. Department stores often demand that certain space is made available in shopping malls to small shoe and clothing retailers, to assure clients that the offered goods are balanced by other competitive offers (that is, that a minimum differentiation exists), for comparison. Clustering of identical firms can serve a socially useful purpose because it lowers

transportation costs (Eaton and Lipsey, 1979, pp. 422-423). The same holds for the clustering of heterogeneous firms. A rational consumer would connect purchasing activities such as search, purchase and transport. Multipurpose shopping assists in the lowering of shopping costs (Eaton and Lipsey, 1982, p. 58).

If the firms are similar, they seek comparable if not identical features for the spatial location of business. If this is true, then the outcome may be a cluster in such an industry. For example, the US film industry started more than a century ago in California, around Los Angeles. The entertainment companies were all looking for comparable locational features such as dry weather and excellent daylight conditions. Although modern filming technology does not depend on natural light or weather conditions at all, the clustering of the film and entertainment industry in California continued its existence (path dependence) in the same location. The homogeneous initial needs of firms created the film cluster rather than the input-output production relations. Hence, functional production links and agglomeration economies are not on their own a necessary condition to create a cluster.

In spite of a few global TV stations (increasingly tabloid) such as the BBC, CNN, Russia Today, China Central Television (CCTV), Deutsche Welle and Al Jazeera, and newspapers such as the *Financial Times*, the *International Herald Tribune* and certain electronic media, most of the stories that editors publish or broadcast are still local. The local proximity of firms in the same industry increases both the visibility of the course of action of competitors and the speed of the spread of information. 'Popular luncheon spots are patronized by executives from several companies, who eye each other and trade the latest gossip. Information flows with enormous speed' (Porter, 1990a, p. 120). Sometimes such a 'pub-related' relaxed atmosphere gives more opportunities to exchange information and learn from others than formal events. This 'lunch club' or 'cafeteria effect' removes the danger of information blockage. It breaks the spatial boundaries to information flow and gives incentives for the creation of matching improvements by other firms in the cluster. It also offers a partial confirmation that there is a spatial limit to knowledge spillovers. The exact spatial scale where spillovers take place is rather obscure. Estimates range between 10 km and 300 km (Stam, 2010, p. 148). Distance matters for the spread and reach of knowledge and other spillovers; they decay fast with geographical distance. As such, this is harmful news for regional convergence.

Steppers are tools that that make computer chips. Japanese companies Canon and Nikon were globally dominant producers of steppers at the start of the 1990s, while the Dutch ASML had less than 10 per cent of the market. Twenty years later, ASML controlled 65 per cent of the market. How did this happen? ASML was too small to compete with the Japanese giants head on. The big Japanese companies had resources to do it all in-house. The small Dutch company redesigned the product and made it modular. It farmed out certain work to specialists such as Karl Zeiss in Germany. This enabled ASML to innovate more and faster than its Japanese competitors. But that was not all:

"ASML's openness took a more literal form, too. 'When a machine at Samsung broke down, 20 Japanese would come over and place a tent over it, so no one could see exactly what they did,' he says. ASML took the opposite approach, and showed customers the problem and how it would be fixed. Today, Nikon and Canon remain as closed as ever – and separate, even though merging their stepper businesses would make sense[54]".

The high speed of information diffusion is one of the major strengths of clusters. Entrepreneurs often prefer to enter or stay in a cluster even though they may be able to appropriate a higher return on their current innovation elsewhere. The reason for this preference is that the firms in the cluster are not only providers of information, but also recipients (Schmitz, 1999, p. 475). In fields where technology changes often and fast, personal contact may be the preferred way of communication rather than the less timely sources such as professional journals, fairs and conferences.

"Human capital accumulation is a *social* activity, involving *groups* of people in a way that has no counterpart in the accumulation of physical capital" (Lucas, 1988, p.19).

This all supports progress in technology and creation of knowledge.

Personal contacts and informality present a powerful and efficient social communication technology. They are essential for the exchange of complex information and tacit knowledge as they facilitate interpersonal transmission of imperfect and uncodifiable data (something that cannot be put into symbols: linguistic, mathematical, visual or vocal) and experience. Has anybody learned to cook a good French or Chinese dish merely by reading a cookbook? Or has anyone initiated and maintained a PC simply by reading an 800-page manual? One may learn

[54] The Economist (2009) November 5.

Greek grammar and words, but it is not possible to employ the metaphors in a reasonably appropriate context without face-to-face contact with the teacher. Personal contacts also ease learning and may create and sustain motivation. A few additional observations on the 'social glue' and 'social oil':

"*Tacit knowledge*, as opposed to *information* ... can only be transmitted informally, and typically demands direct and repeated contact. The role of tacit knowledge ... is presumably the greatest during the early stages of the industry life cycle, before product standards have been established and a dominant design has emerged" (Audretsch, 1998, p. 23).

Every firm in a cluster perceives its survival, growth and success in terms of collective growth and collective success. These firms learn and prosper together. The learning process in a cluster is not only interactive, but also cumulative, as it persists over time (virtuous circle). Once something is learned it seldom ceases to exist[55]; many people may have and enjoy the knowledge at the same time because it is non-rivalrous (many people may use the formula $E = mc^2$ at the same time for free and without consuming the formula), experience and discovery build on experience and discovery (everyone profits from the past improvements and changes, but not everyone benefits immediately from the current ones). Knowledge is an increasing-returns asset. Turnover of labour, technical staff and management among the cluster firms reinforces the transfer of tacit and compound knowledge, cross-fertilising research, the collective learning process and regional competitive advantage.

The underlying general operations within a cluster can be seen from other examples:

"New York City's garment district, financial district, diamond district, advertising district and many more are as much intellectual centers as is Columbia or New York University. The specific ideas exchanged in these centers differ, of course, from those exchanged in academic circles, but the process is much the same. To an outsider, it even *looks* the same: A collection of people in similar activities, each emphasizing his own originality and uniqueness" (Lucas, 1988, p. 38).

"... tacitness in relationship – routines and behaviour of 'uncertain imitability' – which cannot be replicated because no-one, not even the participants themselves, fully comprehend their nature" (Kay, 2006, p. 164).

[55] Fires in the libraries in ancient Greece, Alexandria, Rome and elsewhere may have rendered certain knowledge inaccessible and lost.

5.3. University

Universities are institutions that create, extend, store and transmit knowledge. This is done through education, research, conferences, publications and experts, that pass on the knowledge and experience throughout societies. The existence of a good university or a research institute is not enough by itself for the overall economic success of an area. Kent State University (Ohio) is the birthplace of liquid crystal displays. The Liquid Crystal Institute, founded in 1965, is not only the first of its kind in the world, but also one of the best. Still, the principal global producers of liquid crystal displays are now in China, including Taiwan, and in Japan. Boston (Route 128) and San Francisco (Silicon Valley) have excellent universities and innovative environment, but they also have socially attractive living, working and business environments. In spite of admirable universities in upstate New York (Cornell in Ithaca, and the University of Rochester) and government assistance to attract business there into science parks, these efforts were not rewarded with a noticeable economic success. Mark Andreessen, developer of Mosaic and founder of Netscape (if anyone still remembers this early search engine), studied at the University of Illinois at Urbana-Champaign. He left for Silicon Valley because there was the necessary infrastructure to support his business ideas.

The cluster of high-technology firms around Cambridge (UK) offers a different story. The common tacit code of behaviour among such firms in the region includes trust and cooperation. The cluster is based on two local 'collective agents': Cambridge University and consulting firms in R&D. In contrast to most other British universities which have formally regulated ties with the industry, Cambridge has rather liberal rules governing such links. In fact, faculty members are allowed to work part-time in the private sector if their teaching and research duties are fulfilled. Some of the extra revenue is used to finance research students. All this has had a strong and positive spillover effect on the regional cluster of high-technology firms. This self-grown cluster is different from others linked with universities. For instance, other renowned universities were financed by large external injections of funds from defence spending (Route 128 and Silicon Valley in the US) or from firms such as Eriksson in Stockholm or Siemens in Munich.

The city of Tsukuba in Japan was developed from the beginning of the 1960s. A university and 46 scientific research centres were established there during the 1970s. The Tsukuba Science City became operational in the 1980s. Another university and more research centres flocked to the city. Almost half of the public funds for research are spent in Tsukuba. This city became one of the

most important research centres in the world for high technology. There is, however, a problem in Japan related to the low mobility of labour, and especially of high-level professionals, compared to other developed economies. Professional mobility among the private sector, academia and public institutions is low because of the pension system. Continuous service is rewarded by generous lump-sum bonus at retirement. Those financial 'carrots' are the principal impediment to the low mobility of persons in Japan. The possibility to transfer pension benefits among employers would rectify this shortcoming.

5.4. Cluster Policy

Clusters emerge from various sources ranging from pure historical accidents to spin-offs, path dependence and conscious government planning. It is hard to establish general principles for their emergence (where, when, why and how) for a variety of reasons. 'Chaos' still prevails in the theory that attempts to explain the creation of clusters. Nobody knows at the beginning and in advance that a certain activity will gain significance and evolve later into a cluster. One may notice the existence of a cluster only once it is already 'there', and in retrospect discover the serendipitous or triggering event that started it all. Even those who are within the spontaneously emerging cluster may not realise that hot-spot development until it is 'too late'. Hence,

"there is a lot that we do not yet know about agglomeration economies" (Rosenthal and Strange, 2004, p. 2167),

and not only about agglomeration economies, but also about the appropriate policies, their timing, substance and institutional dynamics in clusters. Still, in spite of this state of flux and lack of knowledge about origin and patterns, once clusters are in place certain ideas exist about their life, transformation and demise.

Some may argue that the 'cluster policy' may be another name for the old and often unsuccessful public policy of 'picking winners'. Porter, however, noted that:

"Clusters often emerge and begin to grow naturally. Government policy had little to do with the beginning of Silicon Valley or the concentration of mechanical firms around Modena, Italy. Once a cluster begins to form, however, government at all levels can play a role in reinforcing it. Perhaps the most beneficial way is

through investments to create specialized factors, such as university technical institutes, training centres, data banks, and specialized infrastructure" (Porter, 1990, p. 655).

Governments and large TNCs often have nothing to do with the establishment of clusters such as Silicon Valley, clusters in Italy or the City of London financial district. These clusters were created out of 'thin air' and largely proceeded on 'autopilot'[56]. It took 30 years for the Silicon Valley cluster to become what it is now. In most other cases, the actual impact of the government policy was unclear.

"New technological trajectories often tend to emerge spontaneously and unexpectedly in space due to the importance of chance events, human agency and increasing returns" (Boschma, 2004, p. 1010).

If one wants to intervene and to assist in the creation of a cluster, then one needs, perhaps, to start to build on some of the existing local specialties and competences, rather than to create entirely new ones as the first step. This does not mean that the 'same old thing' needs to be perpetuated indefinitely, because of necessary adaptations to the industrial structure and adjustment to the changing market situation. However, something (vaguely) related to it or derived from it may be a good start. The old local structure needs to be 'disembedded', but certain and useful old local accumulated knowledge, skills and experience need to be 'rebottled' and reused. If the location in question has accumulated knowledge in the production of fertilisers, it would be sensible to start the production of pharmaceuticals rather than TV sets, for which there is no accumulated local knowledge and experience (although this may be created at a later stage). Scotland has a long history in fermentation and distillation of drinks. This accumulated local knowledge was used in the development of new activity such as biotechnology.

[56] "*The First Industrial Revolution proceeded largely without any systematic intervention from government. Parliament did not help to build a modern financial sector, fund transportation infrastructure in a widespread and coordinated way, provide additional incentives to adopt new technologies, manage fiscal or monetary policy with growth in mind, nor help to train a new labour force with appropriate skills for a new industrial nation. Nonetheless, Britain industrialized relatively quickly.*
Still, the use of the steam engine, and the rise of urban factories, created pressing policy issues. Many of these were local and specific to single uses of the engine rather than generic to steam power itself. Some were so large as to require broader policy changes. These broad policy issues are so numerous and diverse that we only illustrate them here with two cases. When it entered factories, the steam engine gave employment to men, women, and children, bringing them and their low wages to the attention of middle-class reformers who had been largely unaware of the rural poverty that had existed for millennia. Laws were passed to govern child labour and working conditions. Britain also saw the rise of powerful unions and the many laws governing the" (Lipsey *et al.*, 2005, p. 186).
Recall that century-long incremental improvements in textile machinery and subsequently the steam engine brought the First Industrial Revolution. The Second Industrial Revolution was strongly based on science, especially materials and energy (steel and electricity, as well as chemicals and internal combustion engines).

Once the concentration of business becomes too high, there may be negative externalities for work and private life such as pollution, sewage and waste disposal problems, congestion, allergies, crime and an increase in the price of land and rents[57]. This may have an impact on the spread and decentralisation of businesses and their shift to other regions as firms may wish to leave the 'threatened' regions. However, the EU has on average a less concentrated manufacturing geography and much more segmented markets than the US, as a long-term consequence of various non-trade barriers (NTBs) which increased trade costs.

Successful clusters normally build on existing local expertise. Their roots are in the unique local knowledge. They combine local traditions with new global trends. Governments usually contribute little to the emergence of a cluster, and

"when government did play a role in a successful cluster, it was only in the period after its emergence" (Hospers *et al.*, 2008, p. 442).

To conclude on the principal contributions of evolutionary economic geography to the issue of clustering of industries, Boschma and Frenken (2018, p. 217) state that:

"... the dominant explanation of industry clustering resulting from Marshallian externalities has been challenged: (i) clusters can emerge despite the absence of localization economies; (ii) clusters can emerge and exist because of a self-reinforcing process of local entry, in particular the entry of successful spin-offs; (iii) emerging clusters tend to be characterized by positive related-industry externalities; (iv) not all firms perform equally in clusters – some have better routines, partly owing to the pre-entry background of the entrepreneur, and firms differ in their ability to exploit positive externalities and cope with negative externalities in clusters; (v) emergent clusters produce new institutions or adapt existing institutions by

[57] In order to support the war effort in 1940, the British started developing aeronautics and related manufacturing and services industries in Bangalore (southern India). Subsequently the city became the Headquarters of Training Air Command of the Indian Air Force. This was fertile soil for the emergence of information technology and software industries. Technical schools, colleges and universities spread in the area. However, during the 1970s, the Government reduced and limited foreign ownership of firms in India. TNCs were leaving the country and a number of engineers left, many to the US. Then came two important events during the 1990s: first, the US made immigration more difficult; and second, new technologies significantly reduced telecommunication charges. The Indian diaspora in the US strengthened old links and created new ones with India, Bangalore in particular. The initial conditions and government policy (education) created favourable conditions for the return of top high-technology TNCs. Bangalore became an important world software and outsourcing hub. However, it may be the victim of its own success. The rapid growth of this city put a strain on its infrastructure. The population is about 6 million, roads are congested and power cuts are frequent. Local software firms are looking to expand elsewhere in India, possibly in Calcutta, Hyderabad, New Delhi and Pune.

the collective action of agents; (vi) declining clusters can revive and overcome lock-in, but not necessarily so".

6. Cities

6.1. Basic Ideas

Cities exist to connect both people and firms. These are the places where most economic activities take place. They are the result of forces that try to find equilibrium between agglomeration benefits and urban costs. If there were only benefits, everything and everyone would flock to one city. If there were only costs, everything and everyone would be spread over space. Geographic, economic, social, historical, military, political, cultural, technological, religious, ethnic and chance factors influence the creation and shape of cities. Taken together with strong path-dependent effects, those factors make cities complex systems that are challenging to study and understand.

Adam Smith (2005 [1776], p. 21) noted the importance of market size (consumers: households, firms and the government), specialisation and economies of scale for the formation of cities. In particular:

"There are some sorts of industry, even of the lowest kind, which can be carried on nowhere but in a great town. A porter, for example, can find employment and subsistence in no other place. A village is by much too narrow a sphere for him; even an ordinary market-town is scarce large enough to afford him constant occupation. In the lone houses and very small villages which are scattered about in so desert a country as the highlands of Scotland, every farmer must be butcher, baker, and brewer, for his own family".

According to Fujita *et al.* (1999, p. 131), cities exist because firms locate at a cusp in the market potential function made by a concentration of other firms. The size of the cities differs (there is a hierarchy) because of differences in industrial externalities and transport costs. Finally, natural advantages (such as the existence of harbours) help as they create natural cusps in market potential. Arthur (1994b, p. 109) argued that the observed pattern of cities cannot be explained only by economic determinants without considering chance events. Cities also grew up where

immigrants with certain skills settled, where Captain Cook found anchorage, where politicians decided to build railways and canals or dams[58], or where trains stopped for the night.

Just like clusters, cities exist to reduce the cost of transport of goods, people and ideas across space. There is a pool of labour, suppliers and consumers. Hence, the positive aspect for the formation of a city is linked to cost reduction and economies of scale. One needs to recall that the location factors for a city are different from those that support its subsequent growth.

According to an extreme and unqualified version of the neoclassical equilibrium theory, cities should not exist. All economic activity needs to be spread more or less evenly across space, that is, flat geographical space. This extreme version requires constant returns, no fixed costs, no agglomeration of resource inputs, no differences in the productivity of land, and so on. Nonetheless, neoclassical economists were not such fools. Fixed costs are in the standard theory of the firm. Yet people and businesses are not spread evenly across space. Certain agglomerations are facts and business. Increasing returns in production may be among the principal man-made reasons for this development. It is compounded by the convenience and advantage of spatial proximity for relations among various agents. Finally, the time dimension is also relevant. Accumulation of knowledge takes time, and it also takes time to transfer it to the future generation. This accumulation may decay at a slower rate if it is concentrated than if it is widely spread in space.

6.2. Urbanisation

Cities are settlements that generate their own growth (Jacobs, 1969, p. 141). Hence, the affluence and attractiveness of a city life is not the result of design, but rather of 'obliquity'. In economic terms, a city may be thought of (with certain qualifications) as the absence of space among people, firms and institutions, as well as the presence of invisible communication channels. Diversity and a certain tolerance are also integral parts of cities. Even though rents are, in general, higher in

[58] The Hoover Dam was built on the Colorado River (Nevada) from 1931 to 1936. The Bureau of Reclamation awarded the largest (at that time) public contract for the project. People, desperate for jobs, flocked to work there. A new federal company settlement, Boulder City, was constructed 10 km from the project site in 1931. It offered safety, security and a full belly. No alcohol or gambling were permitted. Those interested in the consumption of such goods and services had a chance to go to Las Vegas (32 km away). The sale of alcohol was permitted in Boulder City only from 1969, while it remains the only place in Nevada where gambling is still outlawed. The settlement is highly praised as an attractive place to retire.

cities than in the countryside; even though certain streets in some older cities are in general narrower than is the case in rural areas; and even though there is more pollution and fewer agreeable green landscapes than in the countryside, people create, develop and flock to cities. To preserve the cosy urban lifestyle in Manhattan, Jane Jacobs, a respected urban activist, fought fiercely a decade-long battle (she won) with Robert Moses, the New York City planner, when he proposed to cut a road through Greenwich Village in 1955.

Cities usually emerged on crossroads, near religious sites or near administrative power. Such old cities are different from the emerging megacities around the world, especially those in developing countries (Lagos, Mumbai, Jakarta, Kinshasa and Sao Paulo). Dirty slums, pollution, diseases, drugs, crime and violence surround them. They have a different feel from Paris, London, San Francisco or Seoul. Growth of the new megacities is limited by what the local infrastructure can support. Many people flock to cities because they think that there they may find a way out of poverty. The problem is that many move from poor hinterlands (or countries) to poor slums in the cities. Hence, the challenge to planners and policymakers is to manage those developments, rather than to capitulate. Singapore, Hong Kong, Shenzhen and Seoul may serve as examples of how megacities may avoid being monsters and how they can turn into 'smart cities'.

The question of why many economic activities are agglomerated in a relatively small number of places (typically cities) raises a number of complex issues. For example, a consumer organises their own multiple store shopping itinerary to minimise the total cost of purchases, including transport costs. This is a particularly difficult combination problem to solve, while the final equilibrium becomes very problematic. In any case, the reasons for agglomeration include: (1) externalities under perfect competition; (2) increasing returns to scale under monopolistic competition; and (3) spatial competition under strategic interaction (Fujita and Thisse, 1996). Even though rents are often higher in cities than in the countryside, people pay higher rents in cities to be closer to other people. This is especially pronounced among skilled people who live in areas that offer platforms for their connection. The same holds for many firms: externalities create centripetal forces; there is a functionally indivisible production (strategic complementarity; economies of scale) in certain types of production such as chemicals or microelectronics; firms want to be closer to their principal market; and agglomeration is a cumulative, accelerating and self-enforcing process. For instance, even though there was a twin shock in Ireland (the 2016 Brexit vote, and the adverse tax ruling against technology giant Apple

in 2016[59]), foreign companies continue to flock to Ireland. Technology links and agglomeration continue to have a strong snowball effect.

Regarding the location of cities, clusters and industries, spatial economics distinguishes between first- and second-nature geography. 'First nature' refers to exogenous conditions. It deals with the preference for certain locations over others because of their natural endowments of factors (fresh water, access to waterways, climate, deposits of coal or minerals) or proximity to rivers, estuaries, coasts or natural harbours. This was particularly relevant for the location of production before and during the First Industrial Revolution. 'Second-nature' geography refers to the man-made activities to improve and extend the 'first nature'. This deals with economies of scale and interaction among economic agents. Agglomeration tendencies in modern industries may be much more influenced by second-nature forces compared to the impact of first-nature (natural) advantages.

6.3. Foundation Mechanisms

The most attractive spatial locations for setting up cities over many centuries included:

- crossroads;
- estuaries;
- safe natural harbours (with a rich hinterland) on the sea and navigable rivers;
- defendable hilltops;
- places with important religious functions (shrine, abbey, monastery, bishops' site);
- places with ample fresh water (for drinking and power) and rich soil; and
- dry cool highlands (free of malaria-carrying mosquitoes).

In the past, the location of a firm or a city was influenced both by the endowment of immobile local resources and by flows of mobile factors. Once the development of a business activity starts in an area, and if the economic system is flexible, this area attracts other business activities to the region. This model of uneven development is based on the possibility of meaningful multiple equilibria in the presence of external economies. The point is evident: with external economies,

[59] The European Commission ruled that Apple's tax deal with Ireland amounted to illegal state aid to the tune of €13 billion and must be collected (Boland, 2017).

the return on resources in a particular industrial activity is higher when more resources are committed to it (proximity–productivity effect) (Krugman, 1991, p. 651). The property of a modern firm is high mobility in its search for profitable opportunities, not only within its region or within country, but also internationally. For a footloose firm, the advantages of one location in relation to others are much more man-made than subject to resource endowments.

Many cities used to have economic advantages (production and market access) for a long time just because they were located in the US. However, this advantage is losing its strength and importance in a globalised world. Similarly, footwear production is moving from Italian clusters into Romania (and elsewhere): design and sophisticated production is still kept in Italy (the Chinese are fast learning Italian skills), while much of the raw production is done in Romania, often in Italian-owned firms. A new wave of deglobalisation and nearshoring may somehow alter such a spatial spread. In addition,

"Another strategy for coping with uncertainty is nearshoring – relocating supply chains to nearby countries, especially friendly ones (friendshoring). But we find little evidence that firms are embracing this approach, either. On the contrary, in industries where automation is an option, countries have been reshoring even from neighbors with which trade barriers are unlikely to emerge. Germany is a case in point: far from shifting production to its fellow EU members in Central and Eastern Europe, where labor costs are lower, it has moved production from those countries back onto its own territory. US firms have also reshored production from Mexico – though, again, having the option to use robots, rather than expensive domestic labor, is essential"[60].

Locations of cities were influenced in the past by the cost of transport. Archaeologists noted that distances between regional capitals in ancient Egypt were similar. The reason appears to be the capacity to transport and store grain. Shipping of grain became so costly beyond a certain distance that it paid to develop a new settlement. The First Industrial Revolution reduced the cost of transport, extended markets, but also increased the size of plants. Ships and railways reduced the cost of transport and extended markets. Origin and destination places became favoured, to the detriment of intermediate positions (Combes *et al.*, 2008, pp. 44, 48).

Cities and urban agglomerations are based on three microeconomic mechanisms (Duranton and Puga, 2004):

[60] Marin (2025).

- *Sharing*. Sharing indivisible facilities (production facilities, markets and public goods); sharing the gains of a wider variety of inputs and suppliers; sharing the gains of narrower specialisation and economies of scale; and sharing risks.
- *Matching*. Urban agglomerations with a large, even increasing number of heterogeneous agents (firms and workers with various skills, as well as intermediate goods and services) increase both choice (quality and diversity) and the probability of matching. The incomplete information problem is potentially reduced in cities[61]. There is also the possibility of higher-quality matching among agents. If the economic environment is uncertain and changing fast, cities offer a variety of matching gains. This diversity within cities (as opposed to specialisation) is linked with the interaction between urban location costs and the cost of trade with other locations. Diversity of opportunities (jobs and life amenities) is an important feature for a city to attract talent and human capital.
- *Learning*. Cities bring together a large number of people with their ideas and skills. Knowledge is created, accumulated, stored, extended and diffused. Proximity among players reduces their communication costs.

In the consideration of the composition of industrial systems in cities, their growth and innovation, the professional literature often referred to three basic views:

- *Localisation economies (specialisation)* (Marshall, 1890; Arrow, 1962; Romer, 1986). Most urban and agglomeration externalities arise from the transfer of knowledge and various spillovers within an industry. Local specialisation, increasing returns, intense competition and knowledge spillovers among firms in the same industry are the leitmotif according to this approach. These intra-industry transfers – that is, leaks – are obvious in Silicon Valley or the City of London financial district, or within the Basel pharmaceutical cluster, or in Geneva for high-class watches, or in Chicago and Rotterdam regarding transport. Barriers to knowledge spillovers are quite low, while competition is high. Gurao, in the Guangdong province of China, is a one-industry town specialised in underwear. Clustered suppliers and workers with relevant skills are available locally to lubricate the value chain. Fabric-dyeing plants and congestion, however, had a toll on the environment. Wages and costs of production are rising. Consumers are unwilling to pay

[61] Technology may also reduce in part the information 'search and find' problem. If something or someone cannot be 'Googled', well, to certain searchers, that information or target object or person may not exist.

more for products; hence, Vietnam, Cambodia and Myanmar are attractive places to relocate production. The Gurao firms are trying to change technology and upgrade production. They use less labour and more robots, for instance, to make seamless laser-cut underwear[62].

- *Urbanisation externalities (diversity)* (Jacobs, 1969). Most externalities take place across different industries. A diversified and heterogeneous industrial structure of a city may make it innovate more and grow faster than is the case with specialised cities. Sometimes the blessing of serendipitous meetings and contacts among people that belong to different, but open networks in a diversified city may make this happen (cross-fertilisation of ideas). Existing in a cluster or in a city makes such face-to-face meetings possible and likely. They are often more fruitful than exchange of mails, papers, phone or video calls. Hence, unplanned and multiple social contacts and exchanges make life in a city rich and attractive. Jacobs' externalities are also relevant in rapidly changing high-technology industries. As diverse businesses are not in direct competition with each other, they may be quite open to new ideas and willing to interact and cooperate with other businesses. This feature may be stronger than is the case in the Marshall-Arrow-Romer situation where firms may perceive other firms as their rivals that may poach business ideas, clients and labour. In sum, diversified cities do not depend on the fortune of one or two industries as specialised cities do. Even though specialisation may enhance growth in specialised cities in the short term, these cities are less stable in the longer term than diversified cities. Chinitz (1961, p. 281) emphasised this difference between Pittsburgh, specialised in steel, and diversified New York. Subsequent works, including Jacobs's, confirmed Chinitz's ideas.
- *Competition* (Porter, 1990). The power of competition within cities intensifies the value of various externalities. This supports innovation. If a firm were not up to the standards of the technological frontier, it would leave the market, as clients are immediately able to find another supplier. Most cities continue to grow and sprawl in spite of both globalisation of economic activity and an alleged general reduction in the cost of transport and communication.

[62] The Economist (2016), April 16.

Empirical evidence about predominance of one or another view may be rather inconsistent. Results are based on availability and comparability of data, particularly longer-term time series; general or specific-industry effects; manufacturing that produces tradable goods (production can be concentrated) and services, many of which have a non-tradable character (they have to be spread out and produced close to consumers); the maturity or development phase of each industry; external shocks; and developments in costs of transport.

Benjamin Disraeli, Karl Marx, Friedrich Engels and their contemporary observers were impressed by the UK city of Manchester and the stunning efficiency of its immense textile mills during the 1840s. Manchester was the most advanced among cities at that time and the envy of the world. Birmingham, on the other hand, with few relatively large manufacturers, no obvious business speciality and with its small-scale household trades in guns, steel, glass, leather, toys and jewellery seemed to be outmoded. Nobody was nominating Birmingham as the city of the future. However, over time, highly specialised Manchester stagnated, while diversified Birmingham became the city of the future. And London too:

"London is the safest place to base a company, even if living in some of its inner suburbs requires reckless courage and Kevlar undergarments. Follow the money, in other words"[63].

New possibilities and changes such as the use of the Internet permit modifications in working time and place of work for select professions. Certain key employees can work at home in the environment that they enjoy best (presumably, non-congested countryside). Working from home became widespread globally since the COVID-19 pandemic (2019-2023). In this situation, some office spaces in cities have become liberated. These spaces may be partly transformed into entertainment areas, while it may be quite demanding to convert them into apartments.

One should never forget the attraction of a location for people's social lives, in particular to creative people and their spouses[64]. They also need to have a favourable local natural and social milieu that may inspire them for innovation in production and management, as well as to keep them in the area. They need to meet their peers not only for business or research ventures, but

[63] Guthrie (2007).

[64] "*The most important factor in a location decision: a woman*", is the title of a box in the volume by de Meirlier (2006, p. 34). He reproduced an interview with a reporter: "*What, in your opinion, is the most important factor in a location decision? My reply: Women. We had come across a number of cases where a woman played a decisive role in the site decision*" (ibid.). Wives and/or daughters of chief executives had a decisive role.

also for family matters and an agreeable social life. Florida (2002) emphasised that the 'people climate' needed to be a complement to the traditional 'business climate' in a location.

6.4. Size and Sprawl

Why are cities of different sizes? Henderson (1974) argued that there is a trade-off between economies of agglomeration of industries specific to that city, and general diseconomies (negative externalities) such as costs related to commuting and high rents which (apart from pollution) do not depend on the structure of the local industry. The optimum city size depends on the maximum welfare of participants in the economy. It does not make much sense to put industries without spillovers (say, steel production and publishing) in the same city. However, it does make sense to locate the steel and textile industries in close proximity as they employ, respectively, mostly male and mostly female labour. Needs and potentials of dual-industry and dual-career couples need to be taken into account. Cities need to be specialised in one or a few industries with related external economies. These external economies, however, vary a great deal across industries. There are still certain co-agglomeration tendencies for different industries to locate near each other. Because of strong links internal to the industry, a financial centre may do best if it includes virtually all national financial institutions. The same is not fully true for the textile or the food industry. Hence, the optimal size of a city depends on its role (Fujita *et al.*, 1999, p. 20). In addition:

"... in equilibrium, each city has a well-defined size that depends on the type of firms it accommodates. As cities vary in their industrial mix, they have different sizes because industries differ in the external economies they are able to create. However, in this approach, cities are like floating islands because nothing is said about city locations" (Fujita and Thisse, 2008, p. 16).

Although Henderson dealt with spatial aspects, this type of model is surprisingly 'aspatial', because it says little about the actual location of a city, either in space or relative to other cities. Cities may look like 'floating islands' in space. In addition, national trade policy can influence the size of a city, as exemplified by Mexico City. An unintended by-product of the Mexican import substitution policy was the expansion of the capital city because of production linkages and economies of scale. Once Mexico started to open its economy during the 1980s, there was a relocation of certain firms away from Mexico City, mainly towards the northern frontier. Nonetheless, even when the Latin American governments lost faith in import substitution policies during the 1960s and beyond, the policies reluctantly continued. The governments were

unable to undo those policies. Subsidised vested interest were able to capture policymaking even though their social contribution was meagre. Distortions lasted for a couple of decades and the well-being of whole countries suffered as public resources were directed towards economic activities that were socially underachieving.

Urban areas may have two theoretically fixed relations. The first is 'Zipf's law' (Zipf, 1949)[65] that refers to a special static and rather stable case of the relation between city population size and its rank among other cities. If this power law with exponent 1 holds, then the largest city is k times as large as the kth largest city. This implies that the largest city is twice as large as the second-largest, three times as large as the third one, and so on. The law depends on the existence of labour mobility, but it says nothing about the geographical location of cities. The distribution of US city sizes is close to the rank-size rule, while this relation is less obvious in Europe (Midelfart *et al.*, 2003, pp. 860-861; Thisse, 2015, p. 28). Hackmann and Klarl (2018, p. 7) stated for the case of US cities that

"since 1960, the scaling exponent significantly drops year by year until 2016, indicating more evenly distributed city sizes and a departure from Zipf's law".

Elsewhere, the Second World War had a profound and lasting impact on the city size distribution in Germany:

"The overall city size distribution does not adhere to Zipf's Law" (Bosker *et al.*, 2006, p. 26),

but for firm size distribution, Bottazzi *et al.* (2015, p. 592) found some support.

While in countries such as the US, this power law may be more or less obvious, this does not apply to countries such as China. This country has a list of 'top' megacities such as Guangzhou, Shanghai, Chongqing, Beijing, Hangzhou, Wuhan, Chengdu, Tianjin, Xi'an and others that may be quite comparable, but they do not create a hierarchy. It does not apply to Japan, Russia, France or Mexico, and many others.

[65] George Zipf, an eccentric linguist, calculated that 'the' is the most frequent word used in English (7 per cent of all words), 'of' follows (3.5 per cent), then comes 'and', followed by 'was', and so on. Transposed to the size of cities, one would expect that the second-largest city is half the size of the largest, that the third-largest city is a third of the largest city and so on. The practical soundness of Zipf's law on urban matters is rather controversial.

The second relation relates to 'Gibrat's law'. It claims that all cities grow in parallel as they have similar growth rates; that is, population growth rates in cities are independent of the initial city population size. However, this does not mean that economic policy (intervention) may not influence a city's size, growth and its structure. A city's place in the hierarchy of cities, its success or failure, is influenced by astute and imprudent policies. Shenzhen, for instance, had a meteoric rise since 1979 when it became a Special Economic Zone in China. Gibrat's growth relation is not fixed over time. It alters in response to big changes in technology (productivity, transport, communications and air-conditioning) or labour relations (unionisation). Millions of Americans moved from the northern Rust Belt to the southern Sun Belt states.

The desire to realise a new lifestyle in a suburban environment, outside the inner city, stimulates the spread of cities. Urban sprawl takes place when the rate of land-use conversion exceeds the rate of population growth. This means that urban sprawl is a product of changing lifestyle, rather than a growing population. A century ago, people were walking to work. Today most of them use public transport or drive cars (especially in the US). Sprawling cities require more utilities, more roads, and more energy and water supplies. Citizens travel more, pollute more and contribute to traffic and parking congestion and climate changes.

7. The international firm: theory

Ownership and control of the firm was neglected in the analysis due to an implicit assumption that this does not matter, or the supposition that all assets and structures are domestically owned. The presence of TNCs increases the mobility of capital, expands the availability of information and new products, widens marketing and trade networks, changes the competition structure and alters the substitution of labour for capital. A TNC has different locational considerations compared to a comparable national firm engaged in the same type and scale of activity. A tendency is that strictly national firms expand where they already are, while TNCs enter where they think they may profit from access to the largest and growing market, availability and favourable costs of inputs, transport and/or taxes and subsidies. TNCs have 'organisational capital', that is, a special culture, a common set of rules, practices, routines and values, which help them to overcome various spatial barriers through an internal network while operating in different geographical, social, cultural, legal and other environments. However, if they expand too fast and too widely, there is a chance that, somewhere along the road, they may make a costly mistake.

The role and importance of the growing-market argument as a driver of FDI is exemplified in car market developments. Toyota, for instance, makes its Tundra pick-up truck in San Antonio, Texas because roughly one in seven pick-up trucks sold in the US are sold in Texas[66]. Stagnant Western European car markets prompted TNCs to search for fast-growing market outlets in Russia before harsh sanctions in 2022.

A number of theories explain why firms engage in trans-border business activities (locate abroad) and become TNCs. All of them are based on the existence of structural or transactional market failures. These include imperfections in goods and/or factor markets, product differentiation, economies of scale, contracting friction and asymmetric information. We consider these theories in turn.

First, in an early study, Weber (1962 [1909]) offered two basic reasons why firms 'go to produce abroad': the primary determinant is the achievement of *lower labour and transport costs*, while the secondary element is the benefit of large-scale production.

Second, the motivation to control foreign firms may not come from the mere need or desire to employ assets and structures in a prudent way in foreign markets according to market or technological efficiency criteria, but rather to remove competition from other enterprises. Hymer (1976) advocated such a *market-power approach* by TNCs. Reuber (1973, pp. 133-4) argued in a similar vein that long-term strategic factors for FDI include the desire to eliminate competitors from foreign markets, to be within a protected foreign market, to secure a low-cost source of supply and to lock in the target country to a specific technology for a long time. Such a longer-term strategic view overshadows possible short-run variations in the profitability of FDI.

The problem with this argument is that most TNCs (measured by their number) are small or medium-sized. This shows that, to locate business abroad, and hence to become a TNC, a firm need not be a monopolist or an oligopolist at home that tries to exercise its power abroad. If there is strong competition in the market for differentiated goods and services, and if there is a high degree of substitutability between products (perfumes, soaps, watches, clothing, vehicles, passenger air transport on certain routes, to mention just a few examples), then the market-power argument for the transnationalisation of business is weakened.

[66] Reed and Nakamoto (2007).

Third, while the market-power model excludes potential rivals from competition, the *internalisation theory* holds that an arm's-length relation among individual firms is in some cases less efficient (for example, trade in technology) than an intra-firm cooperative association. Internal transfers of knowledge offers a better control of secrets and lowers the resource costs of transactions. Intra-firm trade is trade between enterprises that belong to the same group, but are located in different countries. Profits may be maximised by means of efficient and friendly intra-firm trade in intermediaries that eliminates sometimes excessive transaction costs (middlemen, exchange rate risk, infringement of intellectual property rights, bargaining costs) that arise when the business is conducted through the market. In these circumstances a hierarchical organisation (an enterprise) may reward parties better in the longer term, as well as curb bargaining and incentives to cheat, than markets and external contractors. The problem with this vantage point is that it focuses on cost minimisation, while it neglects the value creation that may come from locating abroad.

Internalisation may expand until the cost of an additional transaction within a firm equals the cost of the same transaction on the open market. The problem, however, is that the larger the firm, the higher the costs of coordination and management; internal inflexibility may increase; and the labour unions may become too strong. So, outsourcing certain operations or splitting them and allocating them abroad may reintroduce flexibility within a firm.

While Reuber and Hymer conceive TNCs as vehicles for reaping monopoly profits and for the internalisation of pecuniary externalities, the internalisation model looks at TNCs as a mode of business organisation that reduces transaction costs and internalises non-pecuniary externalities. This model of FDI may be convincing in some cases, but it may not explain the structure and location of all FDI flows. In addition to the internalisation possibilities, there should be ownership-specific and locational advantages for FDI.

Excessive internalisation leads firms to diversify into unrelated technologies. This may provoke an increase in costs of production as they venture into businesses that are outside their distinctive core competence. These core capabilities or competences within a firm (unmatched and unmatchable by most competitors) are normally very few in number and individual in nature. Mastering new technologies and areas of business that are marginal to the core business diverts time, efforts and funds from further specialisation in the core business competence. Therefore,

buying from other specialised firms (outsourcing), rather than producing certain goods and services in-house, may be the preferred course of business action to achieve and maintain a greater degree of efficiency and competitiveness. Business strategists should ask whether a new activity extends the core competence of a firm or not. If it does not, they and their investors and shareholders need to be wary of entering into this new business. Nonetheless, the evolutionary approach would suggest that they do not need to be negative about it from the outset. There is another reason for outsourcing certain activities: gardeners, security officers, cafeteria workers and cleaners are employed through contractors as they may hire cheap, non-unionised labour that does not participate in the firm's (generous) in-house retirement or health plans.

Markusen (1984) tried to integrate TNCs into general equilibrium trade models and locational patterns of firms that integrate horizontally or vertically across national borders. Intangible assets, as sources of multi-plant economies, are often firm-specific rather than plant-specific. The intangibility of assets limits the extent to which they may be licensed to independent firms without risk of losing control over quality. These assets include, and spread over, organisation, management, control, R&D, advertising, marketing and distribution. Many of these activities are centralised (finance) and present a 'joint input' across all production units. A TNC avoids replication of these activities that would be necessary if these units operated as independent national firms. This brings a technical advantage to a TNC. However, if these advantages are transformed into market power, then the welfare effects of TNCs may not be clear. In any case, a firm will locate in two or more countries if trade costs are sufficiently higher relative to the disadvantage that comes from fixed costs linked to the operation of two or more plants that serve two or more local markets in different countries (Markusen and Venables, 2000, p. 221).

Fourth, the '*eclectic paradigm*' (Dunning, 1988, pp. 42-45; 1999, pp. 1-3) explains the trans-border business activities of TNCs in terms of the following analytical framework. Trans-border business activities stem from a joint mix and interaction of three independent factors – ownership, location and internalisation (OLI):

- *Ownership* (O) (firm-specific). In order to locate production abroad and be commercially successful, a TNC must have or control internationally mobile income-generating ownership-specific advantages, assets, structures, capabilities or skills. Seen in evolutionary terms, the firm endogenously and actively generates, accumulates and

perpetuates its own advantages; it does not passively respond to demand, costs, market imperfections or public incentives and restraints, as assumed in neoclassical equilibrium theory. These firm-specific advantages include tangible and intangible advantages such as knowledge capital, better technology, brand name, access to wide markets, monopoly, competence of managers to organise and control, ability to innovate and so on, that are superior to those available to local firms (including other TNCs) in the potential target country. That is, a firm needs to operate either on a different production function from other firms or at a different point on the same function. In any case, TNCs should have or control certain advantages over their competitors in the target location, otherwise they would be unlikely to start a costly and risky business at a distance, with resultant disadvantages.

- *Location* (L) (country-specific). Locational (non-mobile) advantages refer to the comparative or location-specific advantages of the target country. They refer to the size of the market and the geographical distribution of resources, those created by the government, as well as costs of transport[67]. In addition, potential positive spillover effects and benefits from an existing cluster of already settled firms might act as a locational magnet.
- *Internalisation* (I). There must be opportunities for the internalisation of ownership-specific advantages (management and quality control, protection of property rights, avoidance of contracting friction and uncertainty from buyers, and so on). It should be in the interest of the firm to transfer these advantages abroad within its own organisation, rather than sell the right to use them to other firms located in the country of intended production (because of the risk linked to asset dissipation). Fixed exchange rates or a single international currency (if created and operated properly) provide a degree of stability necessary for longer-term business planning with a high degree of confidence.

The eclectic paradigm claims that the exact mix of the OLI factors facing each potential investor depends on the specific context. If a firm possesses or controls ownership-specific advantages, then it may use licensing in order to penetrate foreign markets. If it has both ownership-specific

[67] For example, over a certain period of time governments may change the availability, quality and cost of domestic factors. The disposable tools for this policy include training of labour and education of management, R&D, science, transport and communication infrastructure, and tax policies. The host governments may also sort out labour relations. These aspects of liberal or stringent labour-relations rules and high/low quality of education and the labour force may be considered among other factors that influence the location of TNCs.

and internalisation advantages, such an enterprise may use exports as a means of entering foreign markets. Only when a firm is able to make use of OLI advantages simultaneously does it employ FDI as a means of locating and operating in foreign markets[68].

The OLI analytical framework does not apply to diversified and vertically integrated TNCs (Caves, 1996, p. 5). It does not explain FDI from the developing countries to developed ones, either. In addition, an emphasis on OLI factors, internalisation advantages and cost reduction postpones the longer-term strategic decisions for FDI. Varieties of business operations abroad undertaken by TNCs (licensing, franchising or subcontracting), and the downsizing of hierarchies through outsourcing certain activities, question and relax some of the internalisation arguments.

Fifth, the *product-cycle model* reasons that mature (and, perhaps, environmentally unsound) lines of production of goods (there is no explicit reference to services) are passed on to developing countries (Posner, 1961; Vernon, 1966)[69]. Earlier, when the focus of academic research was on the impact of technological gaps in international trade, Posner and Vernon directed their attention towards production. According to this model, the relation between trade and FDI has the spirit of substitution. Their proposition on spatial reallocation of production was based on the experience of Anglo-Saxon firms during the 1950s and 1960s and depended heavily on low factor costs and weak environmental standards in the target location. A snag with this model is that it is too hierarchical regarding the innovatory and technological capabilities of firms in different locations. Following environmental disasters such as the one in Bhopal (in 1984) and subsequent lawsuits, TNCs may easily be inclined to implement stricter and higher environmental standards than domestic firms. 'Clean' TNCs may replace 'dirty' home firms. Another problem is that this model may not be applied to products such as new software that is marketed throughout the world at the same time.

[68] Sub-Saharan Africa may be an interesting case to study. Why have TNCs avoided it to such an extent for such a long time? Apart from countries rich in oil (Nigeria and Angola), and relatively developed South Africa, TNCs generally used to avoid this continent. Poor education and infrastructure, weak governance, government instability and policy uncertainty, possible civil unrest, AIDS, and vulnerability to volatile shifts in commodity prices were the principal reasons stated by a number of Western TNCs for avoiding sub-Saharan Africa as a location for FDI. China does not see these as insurmountable obstacles, and spreads FDI and various operations across Africa.

[69] In spite of allegedly lower environmental standards in developing countries, a massive transfer of polluting business activities to these countries has not taken place. This is the consequence of the constant focus of TNCs on the developed countries and the adherence of TNCs to the unilateral application of environmental standards (often the same as in their home country) that are superior to those prevailing in the host developing country. TNCs fear legal challenges in their home countries for damage done abroad.

The product-cycle argument as the major explanation for the location of business abroad and, with a rather vague timetable, for this spread of production cannot pass the test of recent developments. Asian firms, however, do not replicate this pattern of location of production abroad on a large scale. The Japanese auto companies invested at home and in the US, Europe and elsewhere at about the same time in a similar type of production. There is a heavy concentration of FDI in developed countries, while the majority of developing countries have been relatively neglected in global FDI flows. In addition, countries now start investing abroad at a much earlier stage of their development than before. The newly industrialised countries (South Korea and China) are already investing abroad. In many cases these investments are in the developed world. Such developments may be prompted by the desire:

- to be present in the developed countries' markets (closer to wealthy customers);
- to be near the source and cluster of the principal technological and innovation-related developments in manufacturing, distribution, management and already existing infrastructure (to have a foreign 'listening and learning post');
- to participate in host-country R&D programmes;
- to avoid the dangers of protectionism in target countries;
- to win public contracts; and
- to exploit the strength of the host country's domestic currency.

Sixth is the '*follow-your-leader*' ('me-too') hypothesis. Oligopolists are risk-minimisers. They would like to protect their own market position and avoid destructive competition. Therefore, they typically try to minimise risk and follow each other into the new (foreign) markets (Knickerboker, 1973, p. 100). The technological gap between the home and targeted foreign location plays no role in this model. An overview of timing of FDI in Europe by American TNCs in manufacturing seems to support this snowball or herd behaviour. The Japanese TNCs in automotive and consumer electronic industries were 'following their domestic leader' when they located their manufacturing facilities in the US and the EU during the 1980s and 1990s. This type of business location decision was not limited to the Japanese vertical *keiretsu* (business group), but also to horizontal groups. In addition,

"French firms are attracted to regions where there are other French firms" (Mucchielli and Puech, 2004, p. 55).

Similarly, South Korean TNCs in China mostly flock to the coastal Shandong province (Debaere *et al.*, 2010). Unfortunately, this model of 'bunching up' firms does not say why the first company moves abroad.

While relatively low labour costs in China could have been the reason to start to locate certain operations there at the start of the 1990s, the situation has changed. It was reported that China's Pearl River Delta attracted $1 billion of FDI per month. Microwave ovens are produced in Shunde. Just one of its giant factories produces 40 per cent of global output. Shenzhen produces 70 per cent of the world's photocopiers and 80 per cent of artificial Christmas trees. Dongguan has 80,000 people working in a single factory making running shoes for the world's teenagers. Flextronics is a Singapore electronics-maker that produces for Microsoft, Motorola, Dell and Sony. The manager of its Chinese plant in Doumen said:

"It is a myth that companies are coming here just for the cheap labour. It is the efficiency of the supply chain that drives them here as more and more of worldwide demand is consolidated in this area"[70].

Apart from establishing itself as the global factory, China is also an enormous market. Its middle class of over 600 million grows both in size and wealth, and it is much larger than the entire EU market. China is also an expanding location for R&D. In pharmaceuticals, for instance,

"Chinese R&D sites are opening or growing almost as quickly as European and US sites are closing or shrinkin"[71].

Moreover,

"Between 2012 and 2021 foreign firms increased their collective Chinese research personnel by a fifth, to 716,000.

... Add investments by local firms and China now matches Europe's R&D tally. Only America splurges more. In 2022, despite harsh covid-19 lock-downs, 25 new foreign R&D centres opened in Shanghai. ... Western R&D centres in China have been re-engineered, from places to learn about the domestic market into hotbeds of innovation whose fruits can be found in products sold everywhere. ...

Last year [2023] Volkswagen invested more than $1bn in an innovation centre in the inland city of Hefei. Bosch, a fellow German firm that supplies parts to Volkswagen and other car giants, is building its own $1bn R&D outpost in nearby Suzhou. ...

[70] Kynge and Roberts (2003).
[71] Waldmeir (2012).

A big reason for doing lots of R&D in China is the country's surplus of young engineers and scientists. ... This is a giant talent pool in which foreign multinationals can fish. ... One multinational's China boss recons he gets 30% more working hours out of his research staff in China than his company managers to coax from similar workers in Europe.
A lot of this work is focused on D rather than R. In many areas China still produces less basic research than America but, by many accounts, more applications"[72].

Seventh, the competitive international industry model for the location of business abroad refers to oligopolistic competition and rivalry within the same industry. This is basically an *exchange of threats* (tit-for-tat strategy) regarding business moves by foreign rivals (Graham, 1978). Large firms keep an eye on the actions of their rivals; that is, they act strategically: they pay attention to the likely reaction of their competitors to their own actions. What Texaco does in Europe, Shell will (try to) do in the US. Competition is not cut-throat, but rather stable and gentlemanly among several oligopolies. Other examples of this rivalistic trend in business include FDI in the manufacturing of cars and tyres, or supply of services such as hotels and advertising. Small and medium-sized enterprises (SMEs) such as a petrol station in the middle of nowhere may act independently in their business. However, SMEs in a cluster keep a watchful eye on the actions of their competitors.

Eighth is the *diversification of portfolios model* of foreign investment (Brainard and Tobin, 1992). This approach considers uncertainty. Fluctuations in the rate of return on capital invested in various countries introduce an element of risk. This inconvenience may be reduced by a diversification of portfolios. Following the Plaza Accord of 1995, the Japanese yen was forced to appreciate sharply. This provoked Japanese manufacturing firms to seek ways to reduce costs and maintain the competitiveness of their output. One of the ways to achieve this was to move production to South-East Asian countries[73].

Ninth is the *strong currency argument*. Firms from countries that have strong currencies are sources of FDI. These firms buy firms in countries with weak currencies (FDI hosts). This theoretical model of FDI neglects costs of operations, available resources and labour relations in

[72] The Economist (2024), July 20.
[73] Similarly, TNCs such as BMW and Volkswagen wanted to hedge against a sharp appreciation of the euro against the dollar in 2008, and expanded their FDI in the US.

the target country. FDI may easily flow from countries with weak currencies towards countries with strong currencies. Capital may wish to flee from instability.

Firm-specific assets and exchange rates is the tenth basic theoretical reason for FDI. Suppose that there is a target firm in the US with an innovation (a firm-specific asset) that can make the acquiring firm's assembly line 10 per cent more productive (10 per cent more output for the same level of input). If a US firm wants to acquire the target firm, then a change in the rate of exchange makes no difference, as its gains will continue to be denominated in dollars. If, however, the acquiring firm is from Japan, the gains will be denominated in yen. A depreciation of the dollar relative to the yen would increase the Japanese firm's reservation bid, while the US firm's bid would remain unchanged. If everything else remains the same, it is more likely that the Japanese firm would acquire the asset in this situation.

The actual evidence about this taking place is rather mixed. The price of US assets need not matter, only the rate of return: when the dollar depreciates, both the price of a US asset and its rate of return will go down. Blonigen (1997) found a connection between exchange rate movements (weak dollar) and higher levels of Japanese acquisitions in the US in industries that involve firm-specific assets. This refers to Japanese acquisitions in the US from 1975 to 1992. However, in their analysis of relations between exchange rate movements and FDI flows from the US to 20 countries during the 1980 to 1995 period, Chakrabarti and Scholnick (2002, p. 19) found that devaluation in the preceding year does not have a robust positive impact on FDI inflows.

In addition to the above basic theories on why firms locate abroad, three other dimensions are relevant for coming to grips with the issue of such locational choices: *cost minimisation*, available technology, and taxes. First, Kravis and Lipsey (1982, p. 222) argued that the location of foreign affiliates of TNCs is decided on the basis of cost minimisation. However, the intensity of this determinant varies from industry to industry. Second, Yamawaki (1993, pp. 19-20) did not dispute the importance of a relative difference in factor costs, but the *availability of technology* in the target country is an additional and equally important factor for the location of Japanese FDI in the EU. Yamawaki gives the example of a Japanese TNC from a certain industry which decided to locate in the EU country that has a particular advantage over other EU countries in the same industry. The UK is preferred by Japanese TNCs for the location of production of cars and electrical and electronic equipment; Germany for precision instruments and machinery; Belgium

for stone, glass and clay products; while TNCs from the chemical industry prefer Germany, the Netherlands, Spain and France. China makes similar FDI inroads in the EU. Sweden is preferred for car manufacturing, while Germany is the target for robotics. The third dimension refers to *differences in taxes and fiscal incentives*. This last dimension will be considered in a separate section.

8. How desirable are tax incentives?

"The economic desirability of locational incentives is not clear, particularly if they detract from building competitive capabilities and encourage bidding wars" (UNCTAD, 2003, p. 126).

Tax incentives for investment, in particular FDI, are conventionally not recommended. This is also the usual stance taken by the World Bank, the IMF and other international bodies that advise on tax matters. To sum up, the United Nations Conference on Trade and Development (UNCTAD, 2015, p. 208) repeated its earlier warning:

"Policy advisors in international organisations have long warned against the dangers and downsides associated with incentives. Ongoing work by the IMF, OECD and WTO on incentives for G20 adds to the negative policy advice on incentives. The World Bank's research and advisory work has long focused on the cost of incentives and on the redundancy of many schemes for attracting investment – with good reasons: many schemes have indeed been found to be inefficient and ineffective".

Tax incentives are bad in theory and bad in practice. They are bad in theory as they introduce distortions. Investment decisions by entrepreneurs are made differently from cases that do not feature special tax stimulus. They are bad in practice because of their ineffectiveness: tax considerations are rarely the principal determinant of the location of FDI. They are also inefficient as their cost may well exceed any benefit that they may bring[74]. They are difficult to administer as there is a lack of transparency and they are subject to abuse and corruption by the 'old boys' club'. Finally, they are not equitable as they benefit certain investors, but not others (Easson, 2001, p. 266). In addition, on ideological grounds, poor domestic taxpayers subsidise rich foreign owners of capital. If various tax privileges are offered only to foreigners, then there is a possibility of 'round-tripping'. Domestic capital leaves the country in order to return in the guise

[74] How should one measure these benefits? What should be measured? Should it be done in fiscal terms? Or in social terms?

of FDI, so as to profit from better treatment in the home country. This has been the case in both Russia and China from the early 1990s. Among the leading 'foreign investors' in China are Hong Kong, the British Virgin Islands, Western Samoa and the Cayman Islands, while Cyprus was an important investor in Russia. In these cases, recorded FDI is not so much 'foreign', but rather 'round-tripped' domestic investment.

More than 100 countries worldwide offer tax incentives for FDI. The type of incentive that is most commonly employed is the tax holiday, which is the worst in almost every respect. Administration of these incentives is amazingly complicated; there are opportunities for abuse and avoidance; they may attract only short-term FDI to benefit from the tax holiday while it lasts; and (particularly in developing countries) they are often beyond the capacity of tax administrations to manage and monitor. Therefore, it is not surprising that tax incentives are often inefficient and ineffective (Easson, 2001, p. 375; Tanzi and Zee, 2000, p. 316). Tax holidays do not matter much for the first-time location of a TNC in a badly governed country.

In spite of the above arguments, tax considerations have recently become an increasingly important factor for the spatial location of investment and business. Why is this so? Why are tax incentives becoming more and more generous? Tax considerations do not feature highly in the initial strategic decision by TNCs to invest abroad or not. However, after the decision to locate business abroad is reached, differences in taxes between regions in the target country or differences among countries tend to play a significant role.

Traditional analysis generally assumed: (1) that TNCs did not base locational decisions on tax considerations; and (2) that there was little competition between locations in the developed and the developing worlds. Both of these assumptions are challenged by more recent research, which found that TNCs are becoming more responsive to locational incentives and that competition between locations in the developed and the developing countries is increasing. The responsiveness by TNCs to locational incentives is also on the rise. Developing countries may not normally match the tax incentives offered in developed countries, just as the EU cannot match US subsidies because of the eurozone rules.

The US Government created in 2021 major new programmes through the Infrastructure Investment and Jobs Act ($550 billion), the CHIPS and Science Act ($280 billion) and the Inflation Reduction Act ($394 billion) that authorise $1.2 trillion of spending on subsidies. These

relate to the 'green economy', renewable energy and semiconductors. These subsidies are available to all firms, both domestic and foreign, that intend to locate production in the US and have their intended effect materialised. The 2025 US Government headed by Donald Trump shifted priorities away from the 'green economy' towards fossil energy and artificial intelligence.

Instead of offering tax incentives, developing countries may better use those scarce resources to improve their infrastructure, property rights, education, effectiveness of the judiciary and information-gathering and dissemination, which may benefit both TNCs and domestic firms alike (Moran *et al.*, 2005a, p. 382). Focusing on tax incentives (corporate welfare) diverts public attention and resources away from the real problem of making serious longer-term policy reform. What makes things worse is that the lost tax revenue makes it even more difficult to pay for the profound reforms necessary for economic development.

In order to attract Continental (a German TNC), the Government of Serbia gave subsidies of €19,000 for each employee. Continental accepted and located in Novi Sad, a city with a strong information technology cluster. Those that get jobs in Continental get a company bonus if they bring in another experienced engineer. The local cluster that employs approximately 20,000 engineers is furious. Domestic firms in the cluster pay taxes[75] that the Government used to attract a foreign competitor who poaches the best and most experienced local engineers. Those engineers could be replaced neither easily nor fast[76].

Even though corporate taxation has a certain impact on locational decisions in certain conditions, if one focuses only on tax-related matters in home and host countries and puts aside other matters and policies, one may seriously overestimate and distort the effect of taxes on locational decisions. This may have an unhelpful and misleading impact and relevance for economic policy. One should keep in mind that:

"On the whole, taxation would seem to be a relatively minor factor affecting the locational choices of multinational enterprises as compared to policies affecting the ease of entry for foreign firms, their labour costs and the functioning of product markets in the host country" (Hajkova *et al.*, 2006, p. 6).

[75] The Government did not understand that the task of a good tax collector is to pluck as many geese as possible with the minimum possible noise.
[76] Telesković (2018).

Large and growing local markets such as in China attract TNCs in spite of inefficient formal institutions and regardless of local corporate taxes. Simply put, location decisions by TNCs are complex. The theoretical potential of a location to attract TNCs is one thing; its actual performance is quite another. Similarities and differences between China and India may present a good example regarding the attraction of FDI.

The optimal solution in theory to the issue of enticements related to FDI may include an international agreement among countries to eliminate all tax incentives for investment (or to limit them in a uniform way). In the absence of such an agreement, and according to the prisoner's dilemma concept, few countries would risk acting unilaterally in this subsidy-elimination way. In conclusion:

"While some degree of tax competition is healthy and can have positive economic effects, the main objective of any tax system is to raise a certain amount of revenue in order to finance public services and transfers. This should be done as effectively as possible. Tax systems can entail costs either in the form of reduced economic efficiency – due to tax distortions – or through compliance and administrative burdens on tax payers and tax administrations"[77].

9. HISTORY AND COMPETITION AMONG EXPECTATIONS

9.1. Background

The Heckscher-Ohlin theory cannot explain why industries locate in regions with high mobility of factors (the US and China) or in countries with a broadly similar endowment of factors (France and Germany). Patterns of regional specialisation and location of firms and industries are often created by a historical accident. In 1894, Karl Marx wrote about capitalist production in *Capital* (Vol. III, Ch. 48, III):

"But in reality this sphere is the sphere of competition, which, considered in each individual case, is dominated by chance; where, then, the inner law, which prevails in these accidents and regulates them, is only visible when these accidents are grouped together in large numbers, where it remains, therefore, invisible and unintelligible to the individual agents in production"[78].

[77] EU Commission (2005).
[78] See http://www.marxists.org/archive/marx/works/1894-c3/ch48.html . Accessed on 27 January 2025.

Ohlin noted that:

"Chance plays a significant part in determining the localisation of industry ... A different distribution of inventions would have caused a different localisation" (Ohlin, 1933, p. 137).

More recently, Krugman wrote:

"I at least am convinced that there is a strong arbitrary, accidental component to international specialization; but not everyone agrees, and the limitations of the data make a decisive test difficult" (Krugman, 1992, p. 9).

As already noted, economics is a field that is known to be especially slow at opening itself up to new ideas that are different from those accepted by the mainstream. In the evolutionary world,

"equilibrium if it exists at all, is not unique and historical accidents matter" (Lipsey *et al.*, 2005, p. 41).

In any case, evolutionary dynamics regarding the spatial location of firms in the future are based to a large extent on chance (and human agency).

When there are multiple equilibria, various spots for the location of business are substitutable ex ante ('putty'). A firm is free to locate its business wherever it wishes. After the investment is made, these locations are not easily substitutable for quite some time ('clay'). Agglomeration may take place. Once a business is established at a specific geographical location, it is then 'locked in' through learning, path dependence, sunk costs, value chains, circular and cumulative causation[79] effects. In this sense,

"history matters in a way that it does not in neo-classical theory" (Eaton and Lipsey, 1997, p. xxv).
"The firm's present position, and alternatives for the future, depend on its past history and its present location. *For it, history matters*" (Lipsey *et al.*, 2005, p. 54).

Two questions are relevant here:

[79] Cumulative causation mixes the causes and effects of an event. They are combined in a chain reaction that is increasingly circular, snowballing, herding or perpetually accumulative. This type of self-reinforcement has different labels in economics, which include economies of scale, path dependency, virtuous and vicious circles, as well as threshold effects. The sources of this process are large sunk costs, learning, and network and coordination effects. Cumulative causation is reinforced by the following four elements: favourable demand conditions, good factor-related setting, strong market competition and good access to supporting and related industries.

- Are there inherent differences among locations that create predestination for certain activities?
- How can a small historical accident, a chance (something that is beyond the prior knowledge or control of an investor), alter the economic fate of an industry, region or country?

On the one hand, there is a belief that the choice of location of firms and industries is basically resolved by history. Past events set preconditions that move the economy from one steady state to another. This reasoning, found in the traditional literature, argues that history matters because of increasing returns and lumpiness (inseparability) and activity-specific knowledge and capital goods (Eaton and Lipsey, 1997, pp. x-xi). Various locations (regions, cities, clusters) may relentlessly compete to attract the same capital, firms, industries and labour, and they may also struggle to penetrate the same markets[80]. Some locations are more successful in attracting business and people than others; they are more 'competitive', but this local ability evolves and changes over time. Occasionally quite rapidly.

The 'lock-in idea' is also behind promotional campaigns or 'location tournaments' (David, 1984) among countries and/or regions to attract international footloose capital (TNCs). The impression is that once an activity starts in a place, it is self-perpetuating and related businesses gravitate towards it for a long time. However, this

"self-reinforcing aspect of foreign investment begins to operate only after a certain development threshold has been reached" (Wheeler and Mody, 1992, p. 71).

That is, spillovers on the local economy are not a direct and automatic consequence of the location of a TNC. Gains that come from these externalities can be achieved only if the local firms have taken on, or can be motivated to take on, new (foreign) technologies, skills and links with foreign investors. The structure of the local economy, attitude of the private sector and policies of the government that promote change play key roles. In addition, there is evidence in the case of the US foreign investors that

[80] If this competition among locations to attract firms is relentless, then there is no equilibrium economic landscape. The neoclassical equilibrium model is silent about this evolutionary possibility and reality.

"past investment in the country was a strong predictor of new investment. That persistence was attributed to the favourable effects of agglomeration" (Mody and Srinivasan, 1998, pp. 780-781).

History mattered in this case too.

There is, on the one hand, a proposal that history sets the equilibrium location of firms and industries; while on the other hand is a view that the choice of equilibrium is determined by expectations (especially about future earnings). This observation is based on the belief 'that there is a decisive element of self-fulfilling prophecy' (Krugman, 1991, p. 652). Let us consider both these views in turn.

9.2. Historical Lock-in Effect

The non-linear probability theory can predict with some certitude the behaviour of systems subject to increasing returns. Suppose that balls of different colours are put on a table. The probability that the next ball will have a specific colour depends on the current proportion of colours of balls. Increasing returns occur when a red ball is more likely to be added when there is already a high proportion of red balls (Arthur, 1990a, p. 98). Equilibrium depends on the initial point and later arrivals. Hence, history as a series of (random) arrivals sets the final result.

The national rate of growth of capital stock (without foreign direct investment and foreign loans) depends on home savings and investment[81]. Suppose now that one region or country initially accumulates more capital than the other. In the following period both regions grow, but the one with more capital grows faster than the one with less capital. As manufacturing capital grows, the relative prices of manufacturing goods fall. After a certain period of time, there is a point where the lagging region's industry cannot compete internationally and it begins to shrink. Once this process begins, the new theory of trade and strategic industrial policy suggests that nothing can stop agglomeration for a long time. Economies of scale may drive prices down in the capital-abundant region, and at the same time the lagging region's manufacturing industry disappears. In

[81] General sources of economic growth include market size (division of labour, economies of scale and innovation); investment (an increase in capital per worker); and technological change. Technology is taken to mean knowledge of everything (products, processes and organisation) that can create economic value (Lipsey *et al.*, 2005, p. 10). In addition, Arthur (2007, p. 276) defined technology as 'a *means* to fulfil a human purpose'. This purpose may be explicit and direct (for example, to power an aircraft). It may be multiple, changing or cloudy (use of cloth). A personal computer has no single and precise purpose.

this model, relatively small beginnings can have large and irreversible final consequences for the manufacturing structure of a country, its trade and the competitiveness of its output (Krugman, 1990, pp. 99-100).

The dynamics of capital accumulation ensure that the region that starts with a higher capital stock (than the other regions) ends up with a dominant industrial position. If this is reinforced by a learning process and cumulative causation (strong internal production links where extension of one activity increases the profitability of others), then the existing pattern of comparative advantage is reinforced over time even if the overall structure of the economy has changed. This process-dependent development adds new layers of firms and industries onto the inherited production structure. If output is concentrated within a relatively small area, firms can benefit from economies of scale and linkages (growth of one activity increases the profitability of others). If this area is close to a larger market, there are additional benefits in the form of lower trade costs (including transport). Hence, in this example, the current state of the economy influences, even determines, its future shape.

The investment decisions and trade policy of a country in the current period will have a lasting impact on the shape and direction of the national economy in the future periods. For example, at the end of the 19th century, Argentina and Sweden were relatively comparable low-income farming-based economies. At about the same time, Argentina invested in the education of lawyers and priests, while Sweden invested in the education of engineers. The long-term impact of such choices (coupled with an important problem of weak governance in Argentina) on the material standard of living of the two countries is obvious.

Many developing countries are not necessarily poor because the West is rich. However, the poor countries would be poorer if the West were poorer. Reducing poverty is not necessarily the fairer redistribution of the currently produced wealth. Confrontation with poverty needs to be dealt with by an improved use of the main resource that a country has: its human capital. Switzerland is an obvious example.

Once the structure of an economy becomes unsustainable, there are certain critical branching points (bifurcations) at which the qualitative behaviour of the economy changes (off-path change). New production geography either evolves or is triggered. The long-term and dynamic picture of evolution of a relatively open market economy is not like a ball that rolls smoothly down

a slope in a bowl with a predictable direction and speed (towards a new and predetermined steady state). It is more like a cube or, better, a polyhedron that 'rolls' down a slope: full of unexpected developments, novelties, turbulences, accommodating adaptations to a dangerous environment, amplifications, 'births and deaths', triggers, replications and sudden turns in uncharted directions. Quite a bumpy ride. Hence, the location strategy of a firm is like a lizard's separable tail: predators (competitors) may have it, but the animal without it may continue its life without any major disruption for some time in the future[82].

Changes, in particular in technology, have been (many would argue) the principal sources of growth since the Middle Ages. For example, when water wheels were the principal source of power, the location of many production lines was along rivers and they were spatially limited by the length of the drive shaft. The steam engine permitted a huge spatial reallocation of production towards cities, but here again the length of the shaft restricted the size of the plant. Electricity permitted flexible location of production and the creation of a number of new goods and services.

More than a century ago, Alfred Marshall spelled out the idea of 'backward-looking dynamics' (or 'external economies' in the modern jargon). In his analysis, factors of production are moving towards those industries in which they earn the highest current rate of return. If there are several meaningful equilibria in which the returns would be equalised, then the initial conditions determine the final location-related outcome. History matters, together with factor endowment, tastes and technology (Krugman, 1991, pp. 653-654). As in all laws that deal with motion, the initial position and conditions matter.

Marshall also described the concentration of specialised industries in particular localities in the following way:

When an industry has once chosen a locality for itself, it is likely to stay there long: so great are the advantages which people following the same skilled trade get from near neighbourhood to one another ... if one man starts a new idea it is taken up by others and combined with suggestions of their own; and thus becomes the source of yet more new ideas (Marshall, 1890, p. 332).

[82] Schumpeter (1983, pp. 62-63) commented that: "*static analysis is not only unable to predict the consequences of discontinuous changes in the traditional way of doing things; it can neither explain the occurrence of such productive revolutions nor the phenomena which accompany them*".

Contemporary jargon refers to these processes as 'externalities of innovation'[83].

Recent research has provided some evidence that general government policy can indeed influence the location of firms and industries. From 1947, a number of manufacturing industries in the US migrated to southern states because of the wish to escape the northern unions. This was, however, only one of the reasons for the change in the location of certain businesses in the US. Others included innovation and change in transport (substitution of truck haulage for rail freight which contributed to a spread in production), as well as the advent of air-conditioning that

"made the climate in the South relatively more attractive than the climate in the North" (Holmes, 1998, p. 670).

The Southern Sunbelt (sun and sprawl) was quite attractive to some compared with the Northern Snowbelt (cold and dense) residence patterns and lifestyle. The economic map of the US was redrawn. Still, Facebook went from the southern states of Virginia and California to Prineville in north-western Oregon in the US (and to Lulea, northern Sweden, in Europe). The reason was favourable locations for the cooling systems of their data centres (server rooms), which consume a lot of energy. Another reason was a favourable tax system. Oregon is one of the five states that has no sales tax. This is important, as servers cost a lot of money and need replacement every three years[84].

Firms cluster together in order to benefit, among other reasons, from the availability of a close network of suppliers. They usually, but not necessarily, cluster in locations with a large local demand[85]. This demand will be large in the areas where most producers choose to locate (a process of circular interdependence or cumulative causation).

"There is a degree of indeterminacy in the location of activities – firms locate where they do because of the presence of other firms, not because of underlying characteristics of the location" (Venables, 1996, p. 57).

[83] Because of the availability of wood, Sweden developed its huge pulp and paper industry. Strong links with the suppliers contributed to a similar success for the machinery involved in the paper production processes.
[84] The Economist (2012), October 27.
[85] The daily activities by Muslims involve flexing their knees. As a consequence, some may have problems with their knees. With this in mind, the American Hospital in Dubai developed knee implants that may bend 150 degrees.

For example,

"Japanese business firms operating in Germany have an unexplainable attraction to Düsseldorf rather than Frankfurt" (Beckmann, 1999, p. 61).

9.3. Competition among Expectations

Resources move gradually from one location and/or industry to another in response to differences in current earnings. If this shift is only gradual, then there must be certain barriers that increase the cost of the move. If there are costs, then the resource owners will be interested in expected returns in the future, rather than only in current returns. However, future returns also depend on the decisions of owners of other factors and their expectations about future earnings. For instance, when a new technology is introduced, multiple equilibria cause competition among expectations of potential action from competitors. In this model, expectations (rather than history) determine the future shape of the economy and the location of production.

A river may, over centuries, enrich its bed and reinforce the work of natural forces. However, a strong earthquake or a tsunami (a bifurcation point) may instantaneously ruin such a long history and create a new geographical shape. Applied to economics, such was the case with the mechanical and, later, electrical cash registers that were made obsolete by digital ones. Subsequently, optical scanners that read bar codes replaced digital cash registers. These changes were all based on very different technologies (mechanical, electrical, electronic and optical).

A coal-mining town may cease to exist after the arrival of affordable liquid fuels (innovation-driven shock). Once mobile phones arrived, the need for phone booths waned. The same happened with ice-houses when refrigerators appeared on the market. Photo cameras based on celluloid films were 'attacked' over decades by various alternatives to the standard 35 mm film format. They failed. However, in about 2002 digital cameras based on totally different technology gave a decisive blow to the widespread use of 35 mm film cameras. Mobile phones became platforms for reading newspapers: paper-based newspapers. This had a decimating effect on Finland's paper industry. The Internet introduced and accelerated disruptions in retail (starting from books and ending with almost everything): newspapers and advertisements went onto social media and other Internet-related information-spreading platforms; photographs went onto photo-sharing

platforms; and computers went from huge mainframes to pocket ones such as smartphones. Similarly, the demand for rail transport began to fade away once cars and lorries appeared after the arrival of new transport technologies. This was reinforced by the arrival of air travel (another new transport technology), particularly cheap air travel. The underlining competitive advantage of a firm or a location may remain intact; however, the markets, consumer preferences, technology and expectations may evolve in a way that render these original advantages less and less relevant.

Those who expected and invested in the winning technology ('got it right'), kept the production in their old locations. Alternatively, if the winners relocated to other locations, such as China, they kept the ownership and control of the production process in their head office and let others elsewhere produce and potentially pollute their local environment and/or increase congestion there.

A good encyclopaedia has always been a must if parents want their children to be well educated. The door-to-door salesman selling the *Encyclopaedia Britannica* was replaced by Microsoft's cheap CD-ROM encyclopaedia *Encarta*, first produced in 1993. Wikipedia and the Internet encyclopaedias then found their place with many users. In addition, Amazon is taking business from real-world retail booksellers, and retailers of everything else. Faxes and e-mails damaged and significantly diminished the rents available to postal services, but online shopping increases demand for postal and courier deliveries. Google is entering into Microsoft's market, while elsewhere YouTube is a preferred Internet video search engine. For now.

Just as you climb to the top within an industry and become complacent with your geographical and hierarchical location of business because you 'got it right', out of thin air, somewhere, someone else may appear who is fitter than you are in the evolutionary world. The shelf life of each advantage in the evolutionary world is ephemeral and finite. If your advantage works well for you, you have to keep in mind that this is only an experiment, more or less successful in the fight against entropy. Hence, the pressure and the need to try out something new or better may always be present. One of the most powerful companies ever, Westinghouse Electric, met its demise in 1999; earlier, the British East India Company dissolved in 1874 after a long and splendid business life. Evolution changed the economic environment, while these companies' adjustments were insufficient to fit into the newly evolved situation. These business giants were not flexible enough,

or not innovative enough, or not fast enough, in comparison with their rivals. The old business locations suffered, while the new ones gained.

The Red Queen, from Lewis Carroll's *Through the Looking-Glass*[86], said:

"Now, *here*, you see, it takes all the running *you* can do, to keep in the same place. If you want to get somewhere else, you must run at least twice as fast as that!"

There is no such a thing as winning the Red Queen's race in the evolutionary world in the longer run. The race never ends and there are no winners. If there is a suggestion to a businessman, it is to run, run always, and do so faster than others that compete, as there is no finish line.

"In a flat world ... you can't hide" (Friedman, 2006, p. 185).

The above hypothesis emphasises insecurity. It also reflects the possibility that workers would be increasingly paid for what they are worth on the international market. According to this view, someone or something such as a new technology would come from somewhere and take your job or business. This may add to other general hypotheses[87] of why certain countries keep on growing wealthier than others[88]. This emphasis on insecurity should not come as a surprise because many or most economists worry or sell worries for a living. Those 'economists' who always predict a persistently splendid future are either paid lobbyists or politicians, or those who need to reduce their consumption of cocaine.

Distance matters for the location of economic activity and it will do so for quite some time in the future. Incomes per capita in various locations continue to differ, while their conversion is highly debatable. There are certain suggestions that income *between* countries is weakly converging, but there is no consensus about what happens with the convergence of income per capita *within* countries (Brakman and van Marrewijk, 2008, p. 420). The world is not 'flat', that is, the environment is not similar throughout the world. Facts do not corroborate the neoclassical idea

[86] See http://www.gutenberg.org/files/12/12-h/12-h.htm/ . Accessed on 27 January 2025.
[87] Other hypotheses refer to geography (resources, climate), history and culture.
[88] Institutions play an important role as they may be extractive or inclusive (Acemoglu and Robinson, 2013). *Extractive institutions* concentrate power in the hands of a narrow elite with few constraints to exercise this power to extract wealth from the rest of the society. This may be justified in select cases (such as post-war or revolution periods, or in times after natural or other disasters) and only for a limited time when the concentration of resources is necessary for a targeted intervention to trigger prosperity. Undemocratic rulers and monopolies are examples of such institutions. *Inclusive institutions* spread economic and political power and benefits throughout the society, hence they may lead to economic progress and the reduction of poverty.

that free markets equalise income. FDI is still largely concentrated in the developed world (market-seeking). Developed countries still invest more in other developed countries than in the developing world. The level of consumers' incomes, rather than the low wages of workers, is still one of the predominant incentives for FDI.

A newly introduced good (or service) has to be vastly superior to the well-established and embedded product, be it in convenience of use, price, size, shape and/or speed, in order to dislodge the entrenched product or service from the market. Innovation is threatening the existing firms from every side; hence the lock-in effect is not for ever. Being top dog at a given time t is no guarantee at all that this position will be retained at time $t+1$ and beyond. Achievement and continuous safeguarding of the top position and competence is linked in an evolutionary framework with sustained learning, adjustment, innovation and change. High business concentration and the reliance of a cluster or a town on a single line of production can leave it highly vulnerable to external shifts in demand and technology. Rational entrepreneurs should profit from the 'natural course' of business evolution. Being a key player and staying on top is not a state, but a process. They need to be ready for the next expected and unexpected phase in business.

A clear example of how competition among expectations shapes the national geography of future production can be found if one considers the case of debt in the developing countries[89]. In the 1980s, these countries found themselves unable to repay the huge debts owed to the big international banks. Had they all gone into default together, they would have threatened the global economic system. US Secretary of the Treasury Nicholas Brady developed a scheme for refinancing this debt, with the IMF playing the role of the financial policeman of developing countries by linking lines of credit to severe austerity programmes. This was taken to be a victory for the IMF and international institutions. It could have been, but only in the short run.

During the 1970s, the world saw a massive increase in the price of oil and other internationally traded commodities. The conventional wisdom (expectation) was that commodity prices would only go higher. The Club of Rome and other similar observers of history (and predictors of the future) pointed out that the world was running out of scarce resources (this was their expectation). The world was compared to a spaceship: resources were being exhausted by growing

[89] Stratfor (1999), November 15.

populations and intensified industrial use. It followed that the price of commodities such as oil, copper and wheat could only increase (a similar situation evolved on the market in 2007). If this were the case, a rational expectation was that the best business area to invest in was commodities. Correct?

The primary producers of commodities were developing countries that lacked manufacturing capability, but controlled natural resources much demanded by the industrial world. Any sane investor in the 1970s knew that investing in industries that purchased raw materials was foolish, while investing in production of raw materials was smart. So everyone (or very many), particularly the international banking community and the World Bank, began investing billions of dollars in ventures designed to produce raw materials in developing countries from Mexico to Nigeria to the Philippines. As the crucial commodity, namely oil, was expected to cost $40, $50 or even $100 a barrel in the future, the cost of production was not a critical element in decision-making about investment. The price of commodities was going up, and it was important to get into this business early. All of the technocrats simply knew this and the entire international economic system became skewed towards investing and lending to commodity producers in the developing countries[90].

As a consequence of the above process, the inevitable happened. It turned out that while the world may have a finite amount of oil, iron or copper, there were still huge untapped reserves and progress to be made in technology (which saves on the use of resources). In addition, there was a shift in the production structure in the developed world from foundries and metal-based production towards services. Together with the development of substitutes, this shift towards services curbed the demand for commodities. When the megaprojects in the developing world began to operate, production increased and the price of commodities collapsed[91]. When prices fell below the cost of production, projects became unsustainable and bankruptcy ensued. The outcome was the debt crisis. Nicholas Brady and the IMF stepped in to rectify the trouble. The debt crisis, arising from a belief in commodity scarcity (expectation), led to an avalanche of investment decisions that has left a legacy of misery. The simplistic and linear projection of the

[90] Increasing prices of assets (houses, Bitcoin, and so on) is most often a sign to sell, rather than to buy assets. Only the successful understand those market signals.

[91] On the advice of the World Bank, Vietnam started exporting coffee. It rapidly became the number two world producer after Brazil. That extra production "*flooded the market and drove the price down from 70 cents a pound to around 40 cents*" (Fletcher, 2010, p. 193). The World Bank stated that it financed only a small part of this production of coffee.

future in which commodity producers dominated industrial commodity consumers was rendered false by the collapse in commodity prices, and made irrelevant by another phenomenon in the early 1980s: this new phenomenon is related to Microsoft and similar companies that emerged from thin air[92].

Microsoft, and the endless number of other software and related companies that appeared during the 1980s (many of them disappeared later), altered the equation that had obsessed the World Bank and most other serious economic thinkers. The emergence of computing technologies and 'brain imports' in the US meant that it was possible to increase economic growth without having a similar increase in commodity consumption. Microsoft and creators of ideas, after all, produce wealth without consuming commodities in proportion to growth.

The Internet, for example, may create value as it reduces time and costs for finding and transmitting information (similar in principle to the reduction in transport costs by the advent of railways). Therefore, the new economy may increase investors' hopes (expectations). But too many hopes may create 'irrational exuberance'. When translated into plain English, this cryptic term used by Alan Greenspan in 1996 means that growing markets might become dangerously overvalued (such as overinvestment in telecommunication networks). After this 'new gold rush', stock markets fell sharply in 2002. This had a negative impact on earlier great hopes linked with the new economy. Perhaps too much was expected from the new information and communication technology. New technologies may

"rejuvenate the growth process; they do not necessarily accelerate it ... Typically, several decades are required for a GPT [general-purpose technology] to make a major impact – and that impact may then stretch over more than a century ... electricity is the prime example of this trend" (Lipsey *et al.*, 2005, pp. 112-113).

The overall impact of the new economy based on the new information technology is rather mixed. Does the new economy then look much like the old economy?

[92] Honda, Google, Amazon, Facebook, easyJet, Huawei and ZTE, for instance, were also companies that emerged from 'nowhere'. Some of them were based on new goods and services, while others such as easyJet were doing old things in a new way. Stelios Haji-Ioannou spotted consumer needs in a dormant oligopolistic airline industry in Europe, and stepped in with 'easy' everything. In addition, Toyota was initially an unimportant maker of textile machinery in Japan. However, it grew to become one of the top world-class car manufacturers.

"It is also important to remember that the case for the New Economy rests not just on faster productivity growth, but also on the remarkable behaviour of inflation and employment, and a level of overall economic stability not seen since the 1960s" (Temple, 2002, p. 249).

In spite of obvious benefits brought by the new economy such as increased access and flow of information, speedy electronic transactions, increased transparency and lower prices, the dark side of the information economy increases risks. These nemeses refer to and include serious problems such as identity theft, financial fraud and risk linked with the loss of privacy.

A sharp acceleration in productivity growth and increased competition based on knowledge and innovation is at the heart of the 'new economy of permanent prosperity'[93]. In this situation, it was thought in the mid-1990s that 'old rules' no longer applied. The economy could be closer to full employment without acceleration in inflation. The new situation may have a greater impact on the production of services than on goods. Investments by US firms in shorter-lived assets such as information and communication technology (supposedly important contributors to an increase in productivity) were very high from 1995 to 2000. After 2000, these investments shrank, but productivity growth accelerated. What is the reason for such growth? Are studies of productivity flawed? These studies may arbitrarily include or exclude depreciation. They may also assume that new computers have a direct and full impact on production from day one, but in reality it takes quite some time to master new equipment[94]. Or, are invisible and intangible activities such as formal and informal computer-related training and reorganisation of business methods more important than measurable investments in new equipment, that show only the tip of the iceberg?

New intangible assets helped American companies to produce more output with fewer production workers. As far as Europe is concerned, it had the worst performance in industries that are heavy users of information and communication technology, especially the retail trade (where US productivity is growing much more strongly, as is evidenced by retail chains such as Wal-Mart)[95]. In addition, US energy consumption in 1997 and 1998 was almost unchanged, while the US economy grew by 9 per cent during those two years[96]. Those economies that are successful in

[93] Some even thought that economic cycles were dead.
[94] The performance of computer chips at a given price doubles every 18 months (Moore's law). However, this does not mean that the impact of computers on productivity follows this trend in a linear way. It depends who uses computers, how, when, and for what purpose: management of stocks or computer games.
[95] Gordon (2003).
[96] The Economist (2000), August 19.

reducing energy consumption per unit of output (under the expectation of energy shortages) can, comparatively, benefit a lot from increases in the price of energy in the future; competitors that have not made similar adjustments will suffer.

Technology was changing rapidly over the 20-year-long waves. During the 1970s and 1980s integrated circuits (tiny processors and memory on microchips) assisted in the serious expansion of computational backing of business and lives. In the 1990s and 2000s, the connection of digital processes and the Internet made everything converse with everything else. Fragmentation of production (GVCs), its spatial spread and coordination, was made possible globally. The current morphing began roughly in the 2010s, with omnipresent sensors and surveillance. Arthur (2017, p. 3) also argues that:

"We have radar and lidar sensors, gyroscopic sensors, magnetic sensors, blood-chemistry sensors, pressure, temperature, flow, and moisture sensors, by the dozens and hundreds all mashed together into wireless networks to inform us of the presence of objects or chemicals, or of a system's current status or position, or changes in its external condition. These sensors brought us data – oceans of data – and all that data invited us to make sense of it".

This is a demanding task.

The extraordinary growth of the US economy has many causes. There is no doubt that the persistent growth of productivity in the US is due to the improvement in efficiency introduced by computing and 'brain imports'. Even Federal Reserve Chairman Alan Greenspan acknowledged this (during his term in office), while also recognising that it is hard to calculate the impact. This much is apparent. At a time when productivity should be falling, inflation and interest rates soaring and the economy moving towards recession, productivity (driven by the effects of computing, as well as new ideas) is maintaining a long-term growth trend[97]. In addition, the American economy has become much more energy-efficient compared with the past. Demand for electricity, for instance, grew roughly 1 per cent a year over the first decade of the 21st century.

[97] The Wal-Mart effect (better use of information technology; a hire-and-fire employment and management culture) transformed the American economy as much as it changed the landscape. The European employment and spatial planning laws and practices prevented productivity growth from taking off in retail services. This may be the result of European producers' power at the expense of (poorer) consumers.

This was far below the roughly 3 per cent annual rate of growth for the economy. Over the preceding decade, the annual rate of growth for power stood at 2.5 per cent[98].

The war in Ukraine and the Western sanctions against Russia from 2022 are other events that have boosted the US manufacturing industry. Most trade with Russia ceased, energy prices in the EU skyrocketed (which reduced the competitiveness of output), and the US offered exorbitant subsidies to the new 'green industries', hence the new 'green' production migrated and expanded in the US and China. Europe deindustrialises, while the US reindustrialises.

At the start of the 1980s, Japan produced cars and cameras at lower prices, smaller in size and of better quality than the US. Two decades later the ability to produce cars, copiers and cameras is much less rewarding profit-wise than the ability to write software. The structure of production has changed almost beyond recognition. The hardware that runs a web server is much less valuable than the non-material intellectual property that resides on the server. Perhaps the Japanese industrial targeting (expectation) was superior during the age of fax machines, while the US liberal (unplanned) economic system and alert financial markets that force firms to adjust rapidly to shifting markets and changing technology are more appropriate for the age of the Internet (Krugman, 2000, p. 174). To put it bluntly and in terms of expectations, the Japanese bet on hardware, while the Americans bet on software. The decoupling of value from physical production – its shift to intellectual production – is a millennial shift whose full meaning will not unfold for many generations.

German car-makers were betting on the bright long-term future of diesel cars. And they were technically superb and leaders in that technology (even though there was a 'bit' of cheating regarding the emission and volume of polluting substances)[99]. However, the highest German administrative court ruled in 2018 that cities have the right to ban diesel cars, meaning those cars may disappear from German (and other) roads. The Germans invested too much and for too long in technology that apparently lacks a great future. The 'silos mentality' made the Germans think for too long in terms of pistons, gaskets, crankshafts or mufflers, and they became superb at this

[98] Mongoven (2007).

[99] In the period 2007-2015, VW equipped 10.7 million vehicles with impermissible software to gain an economic advantage. In 2018, Germany fined VW €1 billion for this criminal scandal. 'The €1 bn fine is the largest penalty seen in Europe so far, though it pales in comparison to the more than $25 bn Volkswagen has paid for damages in North America, including a $2.8 bn criminal fine ... Volkswagen pleaded guilty in the US in early 2017 to equipping cars with software that cheated tests and emitted up to 40 times the permitted level of nitrogen oxide in the real world' (McGee, 2018).

type of manufacturing. Their electric cars are somehow poor versions of their liquid fuel cars. They are too expensive (compared to their Chinese competitors) and have little appeal for buyers. Volkswagen announced in 2024 for the first time in its 87-year history the closure of at least three plants in Germany with tens of thousands of redundancies. So,

"The issue is not only that Germany specialised on the wrong technologies, but that German companies have no incentives to move out of them. The consequence of this is that a high-tech industrial sector can only happen through new companies, not old ones. Yet European and especially German industrial policy is focused on the protection of existing commerce"[100].

The Chinese started from a 'clean sheet of paper' in the construction of electric cars and created a decade-long advantage *vis-à-vis* all others[101]. China avoided the German 'mid-technology trap'. This Chinese bet on electric cars created advantages that will be hard to match any time soon. Chinese, and to an extent, US producers have become leaders in electric and self-driving cars. They do not necessarily compete with prices (still an important item); more importantly, the Chinese successfully compete with advanced electric-cars-related technology. This does not mean that in the future the cars will not be made in Germany or Europe, this means that the location where profits will be made is elsewhere.

The firms that will define the next 20 years of economic history and production geography are probably completely unknown to anyone today. Uncertainty and risk are very high and the way forward is not always clear. Therefore, it is very difficult to predict with a high degree of confidence where the world's economy is really going. Profound changes that occur little by little (such as the transformation in the economy and in our lives that was brought about by Microsoft) may be noticed only long afterwards. Therefore, the greatest danger and mistake is in linear thinking: the belief and expectation that what exists today will also exist tomorrow, but with a higher degree of intensity. Consider what the world looked like in 1980. No scholar ever dreamed or imagined then how the general situation would turn out in reality in 1991 or 2000. Or go back to 1900-1920, or to almost any other 20-year period, in particular during the 20th or 21st centuries, and you would be able to observe these great changes in retrospect.

[100] Eurointelligence (2024), May 14.
[101] A standard liquid-fuel car has approximately 32,000 different parts which compares with approximately 7,000 in an electric car.

The only certain thing is that there are uncertain changes. What the world will look like in 20 years from now, even the greatest thinkers of today cannot say. This is the same situation as in the Ancient Greek and Roman age when the greatest minds were unable to predict the end of serfdom, the arrival of Christianity or the attack by the barbarians.

10. GLOBAL VALUE CHAINS

The fragmentation of the production process is not new[102]. Adam Smith explained the specialization (fragmentation, unbundling) of production in a pin factory. There were 18 distinct operations performed by 10 specialized workers (Smith, 2005 [1776], p. 11). Today, what is new, is the much wider spread and scope of unbundled production occurring across international space[103]. It is powered by changing technology: reliable, fast and cheap communication; data processing that makes the spatial coordination of complicated activities possible; the global spread of knowledge; the trade and investment liberalization that opened up foreign markets; economic integration; and improved and affordable transportation. Various tasks or jobs existed for years or even generations (i.e. how to make a car). What has been changing is the process, or how firms solve the linking of tasks.

The creation, evolution and spread of GVCs changed trade, production and investment. National economic advantages can no longer be found in the production of final goods or services, but rather in specific tasks and intermediate goods along GVCs. Analysis and policy actions also changed. The GVC model altered earlier import substitution policies and reinforced the decisions of many developing countries to abandon inward-oriented policies in favour of continued openness for trade and FDI. Countries that are more open to trade and FDI can benefit from economies of scale in specific kinds of niche production, which could provide the country with an additional growth bonus, no matter how small it may be.

The internalization of operations (reduction in *internal* costs) explains the existence of firms, according to Ronald Coase. A new dimension is currently being added to this explanation for the

[102] For example, almost from the start, motor car production was fragmented among parts producers. The difference is that they had to be located not too far away from the assembly line in order to ensure coordination and the arrival of parts when necessary.

[103] This spatial spread of manufacturing is evident in the production of electronics and computing equipment. It is scattered throughout South-East Asia. The 'centrality' of chemicals production is still strong, with Germany and the US as the global production hubs.

existence of firms. The tipping point here is when the volume of the transaction cost internal to the firm is higher than the cost of managing a GVC 'external' to the lead firm. The external costs of organization have fallen dramatically in line with changes in transport and coordination costs (computers and containers, in particular). Therefore, firms' costs became more compartmentalized than they had been. This is pure Coase, but operating in reverse as *external* costs fell.

A GVC is a continuous and smooth range of coordinated domestic and foreign activities that a lead firm takes on to create a product from an idea and a blueprint, and to facilitate fabrication, assembly (according to the agreed common standard), marketing, distribution and delivery to final consumers. Then comes after-sales service and end-of-life management once final use and consumption have taken place. GVCs provide a new opportunity to include many countries in the global investment, production and trade process, and are a potential avenue for diffusing technology and novelty.

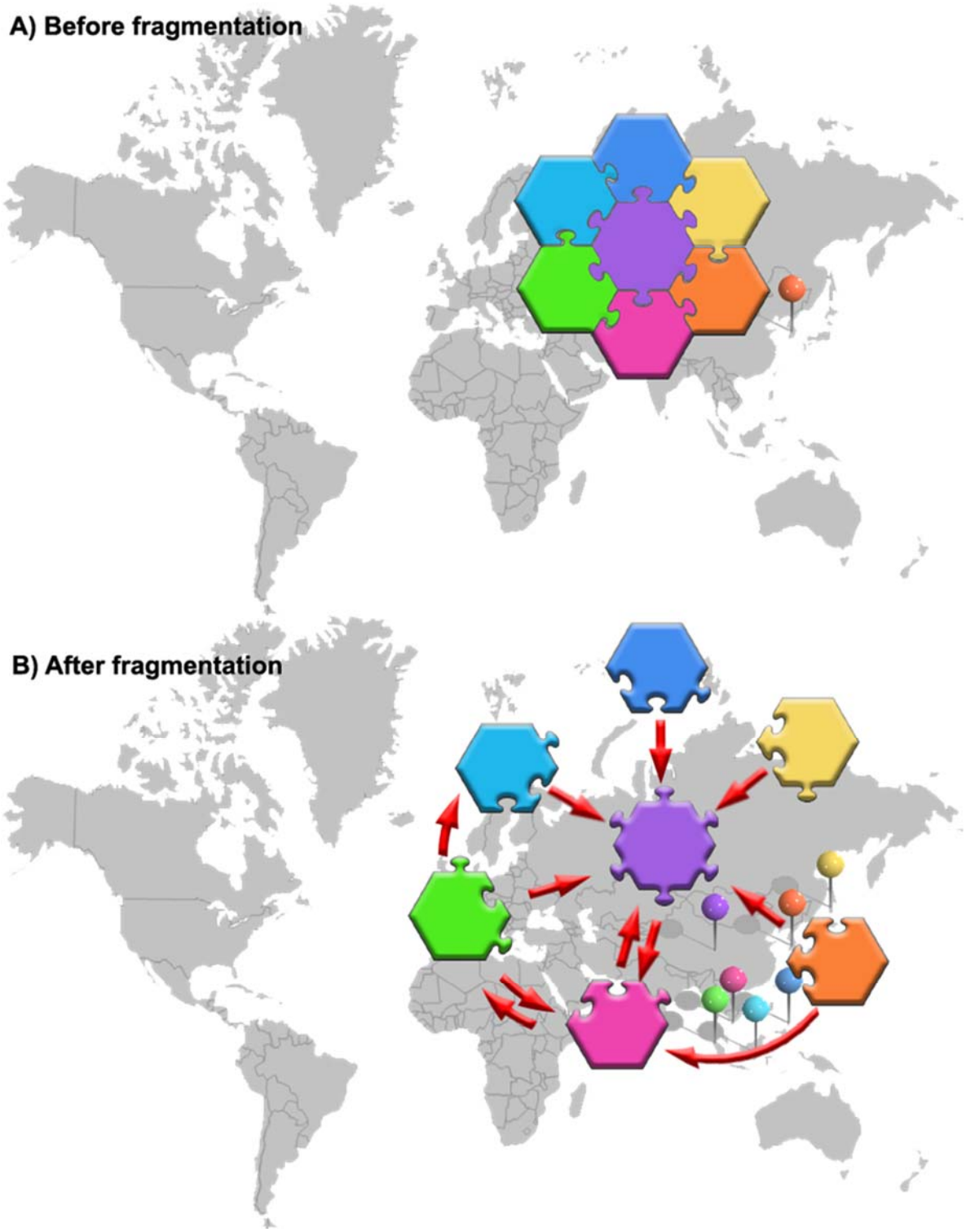
There are two basic patterns of GVCs.

"Snakes are processes whose sequencing is dictated by engineering; spiders involve the assembly of parts in no particular order" (Baldwin and Venables, 2013, p. 245).

In practice, the basic patterns are often a complex mix of the two. New business possibilities are on offer, however, as there are many participants with potentially different objectives and priorities, and so there are potential coordination costs and failures. Organizers of GVCs need to be on top of the administration of such a complex process. They also need to have developed contingency plans in place. If there is a high risk and cost of organizational failure, then GVCs have the potential to be rather narrow spatially, and there would be little offshoring.

Global value chains provide a new platform for industrialization in developing countries (see Figure 3). This new industrialization podium has an advantage over the one that prevailed during the 1950s and 1960s. When South Korea industrialized in the 1950s, almost all of the production process had to be located in one geographical area, but not necessarily within one firm (Figure 3 – panel A). In the 1950s and 1960s, organizational 'snakes' followed the predetermined technical order. There was an 'industrial champion'.

FIGURE 3 - *Contours of Production*



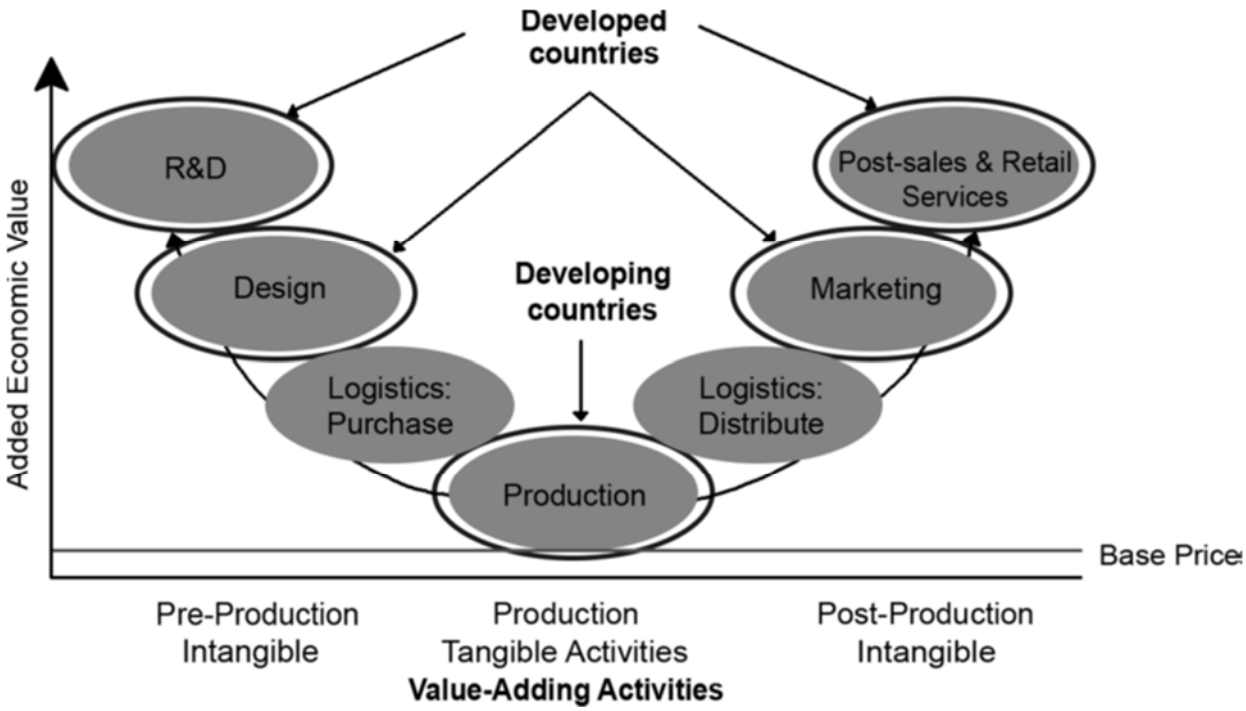
The new GVC era is based on communication, progress in technology, the management of scale, and transport and logistics. Spatial fragmentation (unbundling) of production offers the possibility for a country, especially a developing one, to specialize in only one or just a few segments (tasks, blocs) that are included in the final product, rather than in the entire manufacturing process (Figure 3 – panel B). From the 1980s, the organizational 'spider's' web enabled firms – especially those of small and medium size – to specialize in production segments and market niches, rather than in the whole production process. In the production process of many goods (high-technology ones, in particular), everything is related to everything else in the GVC. To fit properly, each segment in the chain must be of an identical high standard.

The firm's usual dilemma of 'make or buy' does not appear in panel A of Figure 3. Everything is made within a firm at one location. However, the dilemma appears after unbundling (panel B). There are a number of opportunities for where to locate (at home and abroad) the fragmented value-chain segments. Quality standards and cost considerations (lowering costs) play principal roles.

The average size of firms has been reducing since the mid-1970s. This is a great development-related opportunity, and a challenge for developing countries both big and (especially) small. National champions have little role to play in the GVC world. Malaysia and Thailand, and even China, are countries that initially used the GVC model in their modern development. Rather than catching up in the development process in a linear way, the GVC model provided grounds for an economic leapfrog start in select segments. This non-linear process puts a sudden pressure on a country's educational system, which needs to follow the highest standards. Special and differential treatment for developing countries does not have any justification in this regard. If any country does not follow the highest standards that are requested by GVCs, it is eliminated from the GVC by default. Countries may target the highest value-added segments of a GVC to attract business. If a country fails to do so, it may be destined for a simple low-profit 'screwdriver' GVC segment. Education and technical training are therefore of crucial importance, both for technological leapfrogging and for inclusion in GVCs. The low labour-cost advantages of developing countries play little role in decisions related to the location of a GVC's business. The substitution of low-cost labour for technological disadvantage does not work in the GVC production and trade model.

Where is the biggest chunk of profit located? Is it in the making of goods, or in creating and selling ideas? These are questions that face many investors and managers. The methodological and statistical differentiation is quite blurred, as are the boundaries between manufacturing and services. Is Apple a manufacturing or a services-based firm? It designs and sells its gadgets, but does not own the factories that manufacture them. The competitive advantages enjoyed by Apple and similar corporations are based on the architecture of relations between creators, suppliers, assemblers and retailers. Nonetheless, if a country does not have developed services, manufacturing is a great starting point, as the experience of South Korea shows.

FIGURE 4 - *Smile Curve of Value-Added Activities in Global Value Chains*



Source: Adapted from Gereffi (2016, p. 5).

Flat value-added relations between various segments of the production chain, where each part adds a similar share of value, are increasingly turning into U-shaped relations. Stan Shih, the founder of Acer, proposed the 'smile curve' explanation of value-added appropriations along the production chain (Figure 4). Manufacturing adds value, but it is at the bottom of the smile curve. Services at either side of the smile curve (before and after fabrication), with concept on the left

and logistics on the right, add relatively more value than pure manufacturing. Even though manufacturing is highly important in the GVC process, it is the least value-adding activity in a GVC, especially in high-technology business. Big money is in the services provided by the innovative segments of the chain. The further a segment is from the manufacturing phase of the smile curve on either side of the curve, the higher the potential to make big profits and keep highly-paid jobs. If a firm in the production chain aims at a larger market, its smile curve will be flatter. However, if it targets niche markets, the same curve will be deeper.

The biggest game changer in the evolution of the GVCs comes from new technology: the digitalization of production and 3D printing. Offshoring may become a less attractive option in industrial organization in the longer term. The location of production in the developed world may become more attractive, meaning the future, spread and length of GVCs would be quite different from its past.

Depending on the industry and the technologies involved, GVCs may be long and short, both in space and in time. Distance means more management issues and a decline in spontaneous coordination. The optimal length of a GVC depends on the balance between its internal and external costs of operation. The public policy challenge is to create conditions in GVC segments that attract lead firms to those locations and keep them there, while attempting to provide high value-adding activities and high-quality jobs. Continuous education is therefore essential, and the catering of firms to these types of needs plays a crucial role in that regard.

11. ECONOMIC SANCTIONS, TRADE WARS AND LOCATION OF PRODUCTION

Economic sanctions are actions such as the restriction or refusal of trade, investment, payments, or transport by a country, group of countries or an international organisation. They can include asset seizure and are meant to harm, increase costs and penalise the economic (and other) interests of the target country, its officials or citizens, to create pressure to force change in the economy or politics. The globalised world with GVCs was supposedly expected to make such actions unlikely. However,

"Hyper-globalisation has made us so interdependent that we have become limited in our ability to slap sanctions on any country that forms a critical part of global supply chains[104]".

War is, unfortunately, a usual state of affairs as various participants find ways to compete via the use of 'muscles'. Small and weak countries try to find some shelter in international law. Wars are always dirty, bloody, tricky, chaotic, uncertain and expensive. They may have more or less defined goals: you fight to kill, conquer and defeat the enemy. Economic wars are messy. They include sanctions and counter-sanctions (if the target country can retaliate). Targets are often vague or incorrect and there are indiscriminate and random collateral victims, as well as various overreactions. Hence, economic sanctions are the contemporary equivalent of the sieges of castles and towns in the Middle Ages.

Economic sanctions are used as apparently 'bloodless' blunt policy instruments that replace 'hard' wars. Paper and fountain pens seem to be tidier and smarter alternatives than guns and bombs. However, the sanctions seen by Germany and Japan (no free access to crucial raw materials including oil) were among some of the important triggers for world wars. In those cases, sanctions added oil to the fire. Germany (1939) and Japan (1931) gunned their way to raw materials and invested in artificial materials at home.

As a latecomer in the colonial era and, later, because of the post-First World War defeat, humiliation and external sanctions, Germany had the 'incentive' (was forced by events and circumstances) to produce at home various chemicals as substitutes for natural inputs that were not available either in its colonies or through 'normal' trade. Self-sufficiency has often been of vital national importance. The German chemical industry was based on domestic deposits of salt and coal, as well as skills. And trouble and necessity too: there was a shortage of soap in Germany during the First World War. This gave incentives to chemical engineers to invent the first washing powders in 1916. Germany's technological success in various substitution dimensions was remarkable. The country developed world-leading chemical industries and all the related products and services that go with that, such as pumps and sophisticated precision measurement and control instruments. Hence, the isolation of Germany provided grounds for local advances in technology, but the economic 'quarantine' of Germany made the country more dangerous for others. The country transformed itself in four short years, from hyperinflation into the industrial

[104] Münchau (2022).

heart of Europe. This is a clear sign that targeted economic policies have an impact on economic development – a strong one. Can any lessons be found here for the consequences of isolation in Iran, North Korea, Russia, Syria, and others?

Just as Germany (and Japan) invested in and created new artificial materials, the 2014 and 2022 sanctions against Russia, triggered a long overdue reform of the Russian economy. Russia basically exported oil, gas, raw materials and military equipment before sanctions. It was simpler to export those products and import everything else, than to produce them at home. Western sanctions and Russian counter-sanctions forced Russia to replace imports, employ vast domestic resources and rely much more on domestic firms and talent. This sanctions-forced reform and economic transition turned 'sleepy' Russia from a significant global importer of food to one of the principal global exporters of grains. Russia has for years earned more from exports of food than from exports of military equipment. Gas and oil remained, of course, the primary foreign currency earners[105]. To be correct, one needs to recall the Russian import-substitution policy of 2012 which got a boost and was 'facilitated' after the 2014 sanctions.

The 2014 Western sanctions (trade, finance, technology) and Russian counter-sanctions (no imports of food from the West) turned Russian agriculture into an attractive and profitable business. This trend continues. Russia became a net exporter of grains, potatoes and vegetable oils. There is still an unused pool of land suitable for agriculture in Russia (existing farming land may be extended by a third). Domestic consumers changed their habits and turned towards domestic foodstuffs (chicken, potatoes, bread) that were cheaper than the previously imported ones (beef, dairy). The domestic production of copies of foreign cheeses flourished. The quality of these products slowly improved over time. Hence, the Western sanctions on Russia had unintended consequences: instead of injuring Russia, they converted Russia from an important global importer of food to a significant global exporter of food. Russian farmers are in their best shape since the times of Tsar Nikolai II.

One should not be under an illusion that it would be only Russia that suffers from sanctions. Sanctions cut both ways. Both industries and households are hit in the West. The spiralling prices of grains, fertilisers, paper, cement, glass, aluminium, cobalt, copper, nickel, palladium, platinum

[105] Medetsky *et al.* (2016); Astrasheuskaya (2021); "*The INF reckons from 2014-2018, the oil price had a negative impact three times larger than the sanctions. When world oil prices recovered, so did Russian GDP*" (Beattie, 2022).

and steel, to name just a few of those that are most exposed to an increase in the price of gas, is putting into jeopardy reasonably priced everything, from food to the protection of the environment. An increase in prices for these items would just add to the acceleration in inflation in many countries.

Russia has a long history and experience of being under sanctions and sieges, such as with Leningrad (St. Petersburg) or Stalingrad (Volgograd). In fact, Russia (including the former Soviet Union) has been under perpetual sanctions for over a century (with a stint during the 1990s which was economically disastrous for the country). For some of those sanctions there were no big preparations; for others there were. In any case, no matter how the conflict between Russia and Ukraine ends, sanctions on Russia shall remain for a very long time[106]. Hence, Russia is and will remain in many economic and other aspects, a self-sufficient country[107].

Most of the countries in the world, that cover 85 per cent of the global population, do not participate in the Western sanctions against Russia. Quite a few are involved in sanctions-busting activities. They wonder if they might become targets of the West in the future, and want to have alternative economic partners.

The times of manageable, low-cost and trouble-free sanctions is behind us. In the situation as a whole, there are losers in Russia, Ukraine and the rest in Europe, especially Germany (the EU economic engine) and the EU countries that traded a lot with Russia and imported energy, the price of which exploded[108]. Companies started to close down as they could not maintain profitable output with such high prices for energy, commodities and other inputs. Those that continue to operate will find their outputs costlier and less competitive. This strongly works in favour of the US and Chinese producers of high technology goods and related services. The EU economy suffers heavily as there is an exodus of industries towards the US and China. Skidelsky[109] argued that:

"There is, after all, another peaceful model for bringing about improved international and domestic behavior. Montesquieu, Adam Smith, Richard Cobden, and many others argued that the exchange of goods, people, and ideas has pacifying effects. Sanctions work directly against this. When proponents of these

[106] Foy (2022).
[107] A promising theme for a PhD thesis may be: Operation and transformation of an economy under perpetual sanctions.
[108] In spite of strong EU efforts to reduce or eliminate energy imports from Russia, Russia became again the principal supplier of gas to the EU. Only in the month of September 2024 the EU imported gas from Russia worth €1.4 billion. Comfort above all. (Nova News, 2024) November 21.
[109] Skidelsky (2021).

punitive measures claim that commerce is possible only between civilized people, they ignore the civilizing effect of commerce itself".

Once inflation – which accelerated well before the crisis in Ukraine – as well as the cost of food, fertilisers and energy start to bite the general population's real income[110], the question which the population must ask is: with this sanctions-related war, friendly fire and economic self-flagellation, what's in it for us? How long may this last? This also refers to consumers in the EU. Rather than taking a 'wider view' including anger *vis-à-vis* Russia, voters may look 'inside' their own nations and request local action and solutions as their purses suffer. Households in less developed EU regions (the south, Poland and the Baltics) spend a larger part of their income on food and energy. No wonder the inflationary pressure is higher in those areas than in the rest of the eurozone. This bites into the eurozone's structure and unity. How can this be squared with the much more expensive energy sources that are the alternatives to the 'classic' ones that come from Russia? Still, the costly alternative sources of energy may be acceptable as the 'basic' ones become much more expensive. New energy companies may be glad and profitable in these new conditions, but how about the consumers of this expensive 'new' energy in both industries (competitiveness) and in households?

Trade wars and sanctions create other problems and solutions. When the target country or a company is forced into a corner, it turns creative if it wants to survive. It searches for other options. The case in question here is when there was a Western ban or a limitation on exports of state-of-the-art computer chips to China, China created DeepSeek with a new software that works on the 'older' chips in manner superior to the similar services available on the newest chips. Consider this example:

"Nvidia is facing its first real competitor in China. US export controls, designed to cut off China's access to advanced chips and chipmaking gear, were supposed to ensure that no domestic rival could emerge. But the very sanctions meant to shut down China's chipmaking sector have instead fuelled it, accelerating the rise of an unexpected challenger: Huawei.

The paradox is clear – had the US never imposed chip export bans, the Chinese conglomerate would have continued to rely on Taiwan Semiconductor Manufacturing Company for its chips. Chinese chips would have probably remained second-tier, reliant on foreign technology with little urgency to innovate. Instead,

[110] Prices of foodstuffs had an increasing trend well before the energy crises of 2022. Climate change and the COVID-19 pandemic contributed to increases in food prices years back.

by sanctioning Huawei and cutting it off from advanced US chips, Washington has become the greatest driver of the technological self-sufficiency it sought to prevent.

Huawei together with Chinese chipmaker SMIC – which is also under US sanctions – has made a key breakthrough in chipmaking, improving the yield of its latest AI chips to about 40 per cent, doubling from 20 per cent a year ago.

...

But the broader battle over chips extends far beyond Huawei. China, the largest chip consumer in the world, is a market Nvidia cannot afford to lose. Analysts estimate that last year alone, Nvidia made $12bn from 1mn H20 AI chips sold to China. That a single product generated revenue equivalent to nearly a tenth of the company's annual total underscores how critical the Chinese market remains to Nvidia.

Yet Washington's greatest miscalculation may not be underestimating China's chipmaking capabilities, but rather overlooking the forces that drive technological progress. History has shown that every industrial power that has tried to suppress a rival's technological rise has, at best, delayed it – and at worst, accelerated it. Chips are no exception. The chip war is far from over, but in the long run, the US may have ensured that it is a war China cannot lose"[111].

... so ...

"Beijing responded by imposing its own restrictions in a field in which it holds a distinct advantage: critical minerals. China is responsible for almost 70 percent of the mining of rare earth elements. And with its large domestic reserves and long-established mining concessions in Africa, it is responsible for more than 90 percent of all rare earth processing. For some minerals, China has a near monopoly. The U.S. produces just 12 percent of global supply and relies on Chinese machinery for extraction. And so, after Washington introduced the chip bans, Beijing placed restrictions on rare earth extraction and separation technologies. Later, in 2024, it banned the export of some rare earths required for the manufacture of semiconductors"[112].

... and to roundup ...

"Historically, the US achieved technological leadership while also lifting up many other countries. America stayed on top in innovative activities such as research and design, but the gap between it and the rest of the world narrowed. In recent years, however, the strategy has shifted to one in which the US remains on top by pushing others down. Yet as the Vice President Kamala Harris argued at the recent presidential debate: 'The

[111] Yoon (2025).
[112] Geopolitical Futures (2025), March 7.

true ... leader actually understands that strength is not in beating people down, it's in lifting people up'. This applies not only to people, but to countries"[113].

The EU and its member countries admit that there is a disparity between the home technology and what can be obtained from abroad. Hence,

"European companies are increasingly doing deals with Chinese rivals to prevent them from falling behind in the core areas — software, batteries and autonomous vehicle systems — that will drive the future of the automotive industry. Volkswagen, Mercedes-Benz, Stellantis and BMW have all signed agreements with Chinese groups to get access to technology"[114].

However, there is a new tendency in Europe to rearm. Still, there is a question related to an important choice: is the future, for instance in Germany, to produce cars and trains or tanks? Or something else? Global power was, until the 21st century, in the hands of those who controlled seas. From then, this global power is in the hands of those who control space. Will the future armed battles be done by tanks or by satellites and laser-type weapons? Is the investment and location of production of tanks a smart choice or not?

Trade is not a zero-sum game. David Hume wrote in his essay of 1752, 'Of the Jealousy of Trade'[115], a message that is at least as relevant and misread today as it was during his time and much earlier:

"I will venture to assert, that the increase of riches and commerce in any one nation, instead of hurting, commonly promotes the riches and commerce of all its neighbours; and that a state can scarcely carry its trade and industry very far, where all the surrounding states are buried in ignorance, sloth, and barbarism . ... I shall therefore venture to acknowledge, that, not only as a man, but as a British subject, I pray for the flourishing commerce of Germany, Spain, Italy and even France itself. I am at least certain, that Great Britain, and all those nations, would flourish more, did their sovereigns and ministers adopt such enlarged and benevolent sentiments towards each other".

Countries grow rich and prosperous together, not at each other's expense. Otherwise, if at the end there is only one winner, they will find nobody to talk to. A certain sharing of wealth and prosperity is necessary for growth and progress. While the balance of trade matters to an extent,

[113] Goldberg (2024).
[114] Inagaki *et al.* (2025).
[115] https://www.econlib.org/library/LFBooks/Hume/hmMPL.html?chapter_num=34#book-reader (accessed on 15 March 2025).

the mere balance does not show a country's gain or loss from trade. What matters for the gains from trade to materialise is the volume. The larger the volume of trade, the larger the international impact on competition, allocation of resources and welfare (Jovanović, 2015, p 614).

12. Conclusions and policy implications

Spatial elements and the punctuated economy are some of the hardest nuts to crack in analysis and policymaking. What works in one place does not necessarily work in the same place but at a different time, nor in other places and/or at other times. One size – such as von Thünen's concentric circles or Weber's triangles or Lösch's hexagons – does not fit all, at all times and places. Many elements in spatial economics are not easily measurable, hence generalisations are difficult outside theoretical models or specific composition of selected points. Rather than having a classical conclusion, this article ends with thoughts on what was discovered and what needs further enquiry.

Where a firm should locate and stay is not a question with a straightforward answer. Rich nations fear that firms and industries will move to low-wage nations; poor nations worry that the production of goods and services will migrate to developed countries; small nations are concerned that businesses will go to large countries; countries that are not integrated with others are afraid that production will be transferred to the integrated countries; while all countries are troubled because of a large-scale reallocation of manufacturing towards China, especially following the Asian financial crisis of 1997, and to the US following Donald Trump's 'Make America Great Again' policy of 2025. The Chinese worry about the domestic lack of primary resources, hence they locate some of their investment abroad in search of these resources.

The general worry about this exodus of manufacturing jobs to China needs to be mitigated, on two counts. First, China is a grand processing economy. Its weakness lies in the lack of raw materials and energy, as well as increasing labour costs. However, China attempts to mitigate those weaknesses by heavy investments in R&D and innovation. Second, the big money does not lie in manufacturing. It lies in innovation, product design, branding and marketing, where the developed countries still call the tune. At least for now.

During the time and age of 'globalisation' of business, one might expect the importance of firm location to be diminishing. Some argue that distance is dead in terms of production and business

because of astonishing changes in communications and computers.

"Time zones and language groups, rather than mileage, will come to define distance" (Cairncross, 2001, p. 5).

But this may be relevant only for certain manufacturing and services activities that are simple, routine and with codified knowledge. However, the local proximity (clusters) of firms that produce similar, competing and/or related products together with supporting institutions still matters for complex activities, with no or lightly codified knowledge and where market changes and uncertainty require a rapid reaction. Economies of scale, activity-specific backward and forward linkages (indivisible production), accumulated knowledge, innovation, lock-in effects, existence of sophisticated customers and a fall in transportation costs play relevant roles in the protection of clusters and the absolute advantage of certain locations. 'Global' competitiveness often depends on highly concentrated 'local' knowledge, capabilities and a common tacit code of behaviour which can be found in a geographical concentration (a cluster) of firms.

The neglect of spatial economics (geography of production) in mainstream economic theory was not because this research field was uninteresting. It came about because the issues were traditionally regarded as intractable. Ricardo's comparative advantage model provided insufficient answers to the problem. New research tools such as increasing returns, production linkages (presence of intermediate goods), multiple equilibria (with centrifugal and centripetal forces) and imperfect competition were introduced in the field of spatial economics in the early 1990s. These methodological tools also assisted in explaining why firms form clusters. Such new developments did not mean the birth of a new subject. Spatial economics has always been important. Even though the ever-increasing demand for quantitative severity is making considerations harder, the introduction of new analytical tools assisted spatial economics in finding its proper place in mainstream economics and to become a hot research topic, especially at times of trade wars, sanctions, 'green production' and deglobalisation.

The wide-raging US tariffs of 2 April 2025 may provoke the global economic slowdown and an 'economic pneumonia'. Would this aggressive protection of domestic industries assist or harm them? Would it mollycoddle them or support adjustments to turn them more competitive and would that bring lower prices? How long and how far this would go depends on the pressure by the domestic consumers (inflation); domestic US corporations that over the past 50 years heavily

invested and located manufacturing abroad to export to the US (was there something left to outsource?); and countermeasures by other countries. Following the barmy US trade war against everybody, it seems that China may offer more to the global south (rich in critical minerals) than can offer the US. China became a beacon of free trade in an uncharted world order.

Even though globalisation is slowing down, one needs to search for explanations under the bonnet:

"What matters is *why* it's slowing in the first place. One factor is geopolitics. Asset manager PGIM argues that globalisation has entered a "dual-track era". It finds deglobalisation in items with national security implications, such as artificial intelligence, high-end semiconductors, critical minerals and military technology. (This captures most media and political focus.)

But outside the limelight, it finds continued, high-speed globalisation for goods and services, which account for the remaining *75 per cent* of global GDP. This includes professional and IT services, entertainment, consumer electronics and luxury goods"[116].

The new economic geography differs from the traditional model in several important dimensions. The new model makes a case that production specialisation in a given locality is not only based on certain comparative advantages, but rather on a self-reinforcing lock-in effect, path dependence, accumulated knowledge, agglomeration, clustering and linkages (indivisible production). In addition, while the traditional models reason that a reduction in trade costs among locations favours local specialisation, the new economic geography claims that the effect on local specialisation is ambiguous. The final outcome is industry specific and depends on the functional intra-industry production linkages, market structure, consumer preferences (homogeneity of tastes) and factor market (availability and mobility of factors and flexibility in prices). In general, the choice of location for a new investment or a new firm that produces tradable goods and services depends on a complicated interplay of at least three elements: factor intensity; transport intensity; and each in relation to the already established firms and their activities and (re)actions. Remote places do not necessarily need to be poor locations for firms as their remoteness is already reflected in their factor prices (Venables and Limão, 2002, pp. 260-1).

[116] Parikh (2025).

The analysis of spatial economics depends on special assumptions. It is often a detective-type study and a collection of particular cases. Generality is often abandoned for discovery. This should not come as a surprise as spatial economics is simultaneously charged with some of the most difficult problems in economic theory: economies of scale, externalities, imperfect competition (which kind?), linkages and multiple equilibria. Nonetheless, many useful things can be learned from exceptional cases. The examination of the issue is still elastic, fluently sceptical about neoclassical equilibrium economics and more highly suggestive (in particular for how a historical accident may shape the production geography) than it is conclusive, convincing and general. A coherent theory on the subject is not yet in sight. However, there are various approaches that contribute to the raising of new questions and understanding of the issue. This leaves the topic subject to further theoretical and empirical analysis[117].

Economists rarely agree on anything. That is not because economics is a bogus discipline, but rather because it is complex and difficult. Hard physical methods and straightforward laws are inappropriate in social disciplines. This creates 'physics envy' in certain mathematics-prone economists. The reputations of economists, especially their abilities to predict events, have been especially damaged and low since 2008 and the Global Financial Crisis. Not only did the majority of them not foresee the crisis, but also many (not all) expected the immediate economic collapse of the UK if it voted in 2016 to leave the EU (Brexit). Despite the vote to leave the EU, economic collapse did not happen and important economic indicators such as employment improved[118]. The same holds for the strong UK attraction of FDI, placing the UK just after the US and above China and the EU in 2025[119]. Nevertheless, credibility and damage to the profession remained. It may take a long while to re-establish overall respect and confidence in the discipline.

Even though predictions are hard in economics and linked with many unknowns and uncertainties, there are matters that economists know, but to a more limited extent than the textbooks might make one think. We know that droughts (if there is no reliable and efficient irrigation) often lead to crop failures. This increases the prices of crops and often increases the

[117] Topics for further research include: spillovers, externalities and linkages as they are so poorly understood; local interactions; monopolistic competition; cross-fertilisation with industrial organisation and urban economics, trade and growth; cost benefit analysis that can include linkages; services, as research was mainly concentrated on the manufacturing sector; the speed (exponential?) of weakening of the effect of spillovers and information relative to distance, especially in developing countries; the impact of new technologies and energy; deglobalisation and ruptures in supply chains; and the agglomeration and spread of innovative activities.
[118] Allen (2016); Jackson (2016); Mikhailova (2018).
[119] Stewart (2025).

incomes of those who are lucky enough to have a crop to harvest (because the market demand is inelastic). We know that in this situation the people would spend more of their income on food and less on other items, savings and investment. We also know that major monetary expansions to cover government budget deficits lead to inflation and, if persisted in, to hyperinflation[120]. We also know that fast-changing technology, automation and moves towards services alters job markets, and that many lost routine jobs in coal mining and retailing (such as in department stores) would not be restored, because of new clean energies and consumer behaviour such as online purchases, respectively. To address the issue, we know that public intervention in the form of investments in education and infrastructure is more welcome than cuts in taxes to the rich in the hope that they would invest, on their own, on a socially desirable scale. We know that if an economy is depressed, austerity measures lead to more unemployment, even slower growth and more pain and suffering to the people. We also know that innovation (Schumpeter's view) together with widespread education and continuous transformation of economic structure stimulates economic development. The problem is that many innovators and entrepreneurs are misunderstood.

A superior mode for supporting new green and high-tech industries without costs to domestic taxpayers is to open up to FDI. This would invite and spread know-how among countries. With this in mind, *The Economist* suggested the following[121]:

"America and the rest of the West have fallen behind China and other Asian countries in some crucial areas, including large-scale chipmaking and clean technologies such as solar power and electric vehicles. The way to catch up is to let leading firms in those areas open factories in the West. TSMC, a Taiwanese company, is building what will probably be America's most advanced chip factory in Arizona, even if Intel struggles. CATL, a Chinese firm and the world's largest EV-battery-maker, is investing in Germany and Hungary; LG Energy Solutions, a South Korean firm, is now the biggest maker of lithium-ion batteries in Europe. Asia learned from the West by welcoming its best companies. Now the West needs to learn from Asia".

Time will tell how the West will respond to this challenge. There is, on the one hand, the slow-moving, regulations-based and conventional EU, and the completely opposite US of Donald Trump, on the other.

[120] I am grateful to Richard G. Lipsey for these comments.
[121] Economist (2024) November 30.

For the invisible hand[122] of the market to work and to 'behave properly' (to 'get it right'), one must assume that there are many competitors in each business and that returns to scale are diminishing, not increasing. This 'miracle of the market' has been the principal workhorse in economic theory for two centuries. Hence:

"This led many students to draw the inference that the 'miracle of the market' could do the whole job without any human assistance. That such an inference was wrong was forcibly illustrated by the disastrous consequences of the marketisation of the former Soviet Union's economy in the absence of many of the needed institutions" (Lipsey, 2013, p. 37).

Firms in the past knew their principal competitors. This has changed. Firms now may often be unaware of their competitors until it is rather late. Competition is coming from all sides, and the newcomers appear on the business radar screen only once they reach a critical size. A new set of rules has entered the competition game: availability of information at low cost; fast reaction to changes in technology and demand; and acceptance of lower profits, at least initially. Nobody can predict in advance the time and place of the appearance from 'thin air' of firms such as Microsoft, Apple, Google, Honda, Dell, Facebook, Huawei, Embraer, easyJet, Nvidia, Uber or others, and their lasting effects on economic growth and our lives. The neoclassical equilibrium model would argue in favour of the removal of all market barriers. The evolutionary model may, however, offer certain other policy advice. At least some freedom for general action is most welcome, but additional astute intervention is necessary and welcome. Many would question the available freedom, particularly political freedom, in China. This country, in some dimensions, may be a growth model to follow for many countries in the developing world.

Economic policies were tested in China step by step from the 1980s. Responses by entrepreneurs, domestic and foreign, were positive in general (in spite of occasional corruption)[123]. Even if some firms failed, the society at large gained. So, careful intervention followed. This does not make China a development and growth disaster. Privatisation, for instance, was not emphasised. Nonetheless, a skilful government provided for a stable macroeconomic situation, openness to trade and investment, strongly encouraged the creation of human capital, stimulated (forced) the transfer of technology and its local development, as well as motivated savings. Many firms observe their presence in the huge and growing Chinese market as a sign of success, even survival as is the

[122] The invisible hand of the market is often conducted by the visible hand of lobbies and politics.
[123] Do countries choose to afford the rule of law only after they grow rich?

case with the German car-making giants. South Korea earlier had a similar policy stance. Free markets that were supposed to 'get prices right' (to realise static comparative advantages), as the key ingredient of the neoclassical Washington Consensus, did not feature highly at all. Select intervention to develop dynamic local comparative advantages was the rule. These dynamic advantages in China and South Korea could, perhaps, be achieved without intervention much later (if ever). Hence, this growth took place in countries with strong (not weak) and competent governments that intervened in the economy. How long this may last, is another issue. The evolutionary model suggests that policies – context-specific policies – need to be evaluated and adjusted continuously in a situation with uncertainty and endemic disequilibrium.

An East Asian development model does not exist. Countries differed a lot. While Singapore applied intervention, Hong Kong was closest to the neoliberal model. South Korea and Taiwan intervened a lot in the economy. South Korea invested in heavy manufacturing industry and R&D, as well as supporting the creation of private conglomerates (*chaebols*) to internalise the costs and risks of complex technologies. Taiwan was open to FDI, supported R&D, but did not back heavy industry or giant conglomerates. SMEs are the backbone of its domestic economy. The Government supported their collaborative efforts, especially regarding innovation, R&D and new technologies. Plastics, consumer electronics and apparel were targets during the 1960s. Once those lines of production and exports lost competitiveness, industrial restructuring took place from the 1980s towards high-technology production.

Successful companies also have different paths of development and ways of doing business. Take a look at two Chinese companies. Huawei gradually and incrementally built on internal capabilities and invested slowly, first in less-developed and marginalised countries. In contrast, Haier accepted more risk and invested in advanced countries early on in its internalisation process. The intention was to expose the firm to tough competitors from which Haier can learn and improve fast (Buckley *et al.*, 2018, p. 16). Such openness to the influx of successful firms from Asia is the path for the EU and local firms if they want to remain globally relevant. In fact, the car-making giants such as Mercedes argue in favour of welcoming Chinese car factories in the EU as they would stimulate competition and avoid a trade war with China[124].

[124] Bounds and Inagaki (2025).

"I never think of the future, it comes soon enough", once observed Albert Einstein[125].

While managers may not do much to change trends in demography or global warming, they may anticipate certain changes and profit from adjustment to them. Some spotted and anticipated touch-screen technology and applied it to smartphones (Samsung, Apple); others failed to do so and were left far behind (Nokia). This had a significant long-term impact on the location of production and creation of wealth. In addition,

"It is difficult to avoid the conclusion that boosting manufacturing employment is like chasing a fast-receding target. The world has moved on, and the nature of manufacturing technologies has changed irrevocably. Automation and skill-biased technology have made it extremely unlikely that manufacturing can become the labor-absorbing activity that it once was. Whether we like it or not, services such as retail, care work, and other personal services will remain the primary engine of job creation. That means we need different types of good-jobs policies, with a greater focus on fostering productivity and labor-friendly innovation for services"[126].

A small change, such as the 'butterfly effect', at a crucial point in time (once the threshold is reached) can lead to shocking outcomes, which may lead to 'locational catastrophes'.

"The possibility of non-linear responses to policy makes it much more difficult to forecast the effect of a given policy change . . . it can be seriously misleading to base expectations of the effects of future policies on linear approximations from the past" (Baldwin *et al.*, 2003, p. 229).

Trade liberalisation, wars, economic integration and other policy changes may have very non-linear effects on industrialisation, location of firms and growth in an evolutionary framework. The evolutionary road is not only unpredictable, but also bumpy.

Even though the barmy US trade war discredited the WTO and hit hard globalisation already hurt by a messy web of bilateral trade and investment agreements, the globalisation process is not dead. The winner locations in the global economy shall be the ones that are able to create (educate), attract and keep talented people. In addition, chaotic situations create great opportunities for sweeping changes and reforms. The WTO may be the first one in line.

[125] Bisson *et al.* (2010).
[126] Rodrik (2024).

The sheer number and complexity of possibilities for the location of production in an evolutionary framework is enormous. As there is an important lacuna in our understanding of evolutionary spatial economics, it is of little wonder that there are many disagreements, and only several agreements in theory. Hence, generalisations and synthesis in the form of policy recommendations are coupled with huge problems and risks. In some cases, inputs in the decision-making process are limited and insufficient. In others, there is an avalanche of information that needs to be filtered and consolidated. The list of conditions for the application of policy recommendations is long, their voicing quite delicate, while final policy suggestions are few and narrow at best. Hence, policy advice on most issues (location, trade, technology, investment, taxes...) must be based

"on a mixture of theory, measurement and subjective judgement" (Lipsey *et al.*, 2005, p. 515) and certain practice:

"... economics will forever have to contend with the biggest X factor of all: people. As Mr. Solow notes: 'You feed people with poison, and they die. But feed them a subsidy and there is no telling what will happen. Some will use it wisely, others perversely and some a mix of both.' As a result, a certain amount of psychological guesswork is part of an economist's job"[127].

There are said to be only three golden rules on how to conduct successful economic policy and spatial location of business. Unfortunately, nobody knows them. Nonetheless, there are certain general principles (lessons learned for economic policy) that one has to keep in mind and consider during the search for a solution to a particular spatial problem. Here they are:

- Starting a new activity in a given location requires building on and improving already existing accumulated expertise and specialisation of that location. One may initially need to stick to the production process that one understands in general terms. That is found in vaguely related industries. This embedded general knowledge, experience and specific competence is an important input (seed) into the creation and evolution of a new industry[128]. This bottom-up specialisation needs to be extended, but this does not mean at

[127] Segal (2010).
[128] If one knows how to make radios, it is supposed that one may learn to make TV sets, and later computers, more easily than if one initially produced cheese or shoes. The Danes understood ploughs (turning ground), then maritime propellers (turning water) which subsequently helped them to perfect wind turbines (turning air). Scotland has been proficient in whiskey-making technology. This subsequently provided grounds for its world-class blossoming

all that the earlier production structure has to be continued[129]. Something (vaguely) related to it or derived from it may be the starting point which, later on, needs to be altered, sometimes fundamentally[130].

- Using policy intervention to influence the location of firms is most effective only very early in the process. Evolutionary economics argues that very little intervention is necessary to produce a successful locational result. There are few and relatively short moments of enhanced locational freedom for new industries that set in motion new industrial trajectories. If there is a choice of potential locations for the establishment of business, the winner location often wins by the skin of its teeth. Hence, there is only a narrow locational option window through which a short, rare and slim beam of opportunity light can pass, for the policy to be effective. Subsequently, and in the medium and longer term, market forces may be the strongest socio-economic technology for resource allocation in the situation of disequilibrium. Others would argue that markets need not guide decisions, but rather that their role is to serve them.
- Local education and vocational training have a strong positive impact on the location and attraction of firms and industries.
- Depending on the industry, economies of scale and functional linkages in production and its servicing (strategic complementarity) have a certain impact on the location of firms and industries.
- Policy should be reviewed periodically and have elements of flexibility.
- Industrial success or failure of a region or a policy sometimes depends on uncontrollable factors and events that are themselves located outside the region.

O LORD, correct me, but with justice;
Not in Your anger, lest You bring me to nothing.
Jeremiah 10:24

biotechnology. Similarly, Sheffield was known for the transformation of iron ore into steel. Once this business faded, the city created its material science and engineering centres. You cannot teach innovation; you have to experience it.

[129] The 2025 failure of the Swedish giant-to-be battery producer Northvolt offers a clear lesson for Europe - stick to what you are good at – such as green industries where Europe does not start from scratch (The Financial Times, 2025 – March 12).

[130] Birmingham was one of the prime manufacturing areas in England. It started with metalworking and railways equipment in the 18th century. As technologies and demand changed, the local industry transformed and diversified into other mobility-related businesses such as bicycles and cars in the 20th century.

REFERENCES

Abboud, L., B. Hall and H. Foy (2024), "Europe Faces Mortal Threat, says Emmanuel Macron", *Financial Times*, April 25.

Acemoglu, D. and J. Robinson (2013), Why Nations Fail, Profile Books: London.

Alchian, A. (1950), "Uncertainty, Evolution, and Economic Theory", *Journal of Political Economy*, 58(3), 211-221.

Allen, K. (2016), 'The post-referendum economy: data defies gloomy UK predictions', *The Guardian*, September 21.

Arrow, K. (1962), "The Economic Implications of Learning by Doing", *Review of Economic Studies*, 29(3), 155-173.

Arthur, B. (1989), "Competing Technologies, Increasing Returns, and Lock-in by Historical Events", *Economic Journal*, 99(394), 116-131.

Arthur, B. (1990a), "Positive Feedbacks in the Economy", *Scientific American*, 262(2), 92-99.

Arthur, B. (1990b), "Silicon Valley Locational Clusters: When do Increasing Returns Imply Monopoly?", *Mathematical Social Sciences*, 19(3), 235-251.

Arthur, B. (1994a), *Industrial Location Patterns and the Importance of History*, in: B. Arthur (Ed.), "Increasing Returns and Path Dependence in the Economy", University of Michigan Press: Michigan, pp. 49-67.

Arthur, B. (1994b), *Urban Systems and Historical Path Dependence*, in: B. Arthur (Ed.), "Increasing Returns and Path Dependence in the Economy", University of Michigan Press: Michigan, pp. 99-110.

Arthur, B. (2002), "How Growth Builds upon Growth in High-Technology", Annual Sir Charles Carter Lecture, Northern Ireland Economic Council: Belfast.

Arthur, B. (2007), "The Structure of Invention", *Research Policy*, 36(2), 274-287.

Arthur, B. (2017), "Where is Technology Taking the Economy", *McKinsey Quarterly*, October 2017, pp. 1-11.

Astrasheuskaya, N. (2021), 'Russia starts to sow seeds of 'wheat diplomacy'', *Financial Times*, September 2.

Audretsch, D. (1998), "Agglomeration and the Location of Innovative Activity", *Oxford Review of Economic Policy*, 14(2), 18-29.

Baldwin, R., R. Forslid, P. Martin, G. Ottaviano and F. Robert-Nicoud (2003), Economic Geography and Public Policy, Princeton University Press: Princeton.

Baldwin, R. and T. Okubo (2005), "Heterogeneous firms, agglomeration and economic geography: spatial selection and sorting", Graduate Institute of International Studies, Geneva, mimeo.

Baldwin, R. and A. Venables (2013), "Spiders and Snakes: Offshoring and Agglomeration in the Global Economy", *Journal of International Economics*, 90(2), 245-254.

Beattie, A. (2022), "Germany's Nord Stream 2 Decision Shows Sanctions on Russia Haven't Worked", *Financial Times*, February 23.

Beckmann, M. (1999), Lectures on Location Theory, Springer: Berlin.

Begg, I. and D. Mayes (1994), "Peripherality and Northern Ireland", *National Institute Economic Review*, 150(November), 90-100.

Beinhocker, E. (2006), The Origin of Wealth, Random House: London.

Bekar, C. and R. Lipsey (2001), "Clusters and Economic Policy", Simon Fraser University: Vancouver, mimeo.

Berkun, S. (2010), The Myths of Innovation, O'Reilly: Sebastopol.

Bernstein, J. and D. Weinstein (2002), "Do Endowments Predict the Location of Production? Evidence from National and International Data", *Journal of International Economics*, 56(1), 55-76.

Bisson, P., E. Stephenson and P. Viguerie (2010), "Global Forces: An Introduction", *McKinsey Quarterly*, June.

Blonigen, B. (1997), "Firm-Specific Assets and the Link between Exchange Rates and Foreign Direct Investment", *American Economic Review*, 87(3), 447-465.

Boland, V. (2017), "Ireland Enjoys High Foreign Investment Despite Shocks; Jobs Growth", *Financial Times*, January 4.

Boschma, R. (2004), "Competitiveness of Regions from an Evolutionary Perspective", *Regional Studies*, 38(9), 1001-1014.

Boschma, R. and K. Frenken (2018), *Evolutionary Economic Geography*, in: G. Clarke, M. Feldman, M. Gertler, D. Wojcik (Eds), "New Oxford Handbook of Economic Geography", Oxford University Press: Oxford, pp. 213-229.

Bosker, M., S. Brakman, H. Garretsen and M. Schramm (2006), "A Century of Shocks: The Evolution of the German City Size Distribution 1925-1999", CESifo Working Paper No. 1728.

Bottazzi, G., D. Pirino and F. Tamagni (2015), "Zipf Law and the Firm Size Distribution: A Critical Discussion of Popular Estimators", *Journal of Evolutionary Economics*, 25, 585-610.

Boulding, K. (1991), "What is Evolutionary Economics?", *Journal of Evolutionary Economics*, 1, 9-17.

Bounds, A. and K. Inagaki (2025), "EU should Welcome Chinese Car Factories, Says Mercedes Chief", *Financial Times*, January 19.

Bowles, S. (2004), Microeconomics, Russell Sage Foundation: New York.

Brainard, W. and J. Tobin (1992), "On the Internationalization of Portfolios", *Oxford Economic Papers*, 44(4), 533-565.

Brakman, S. and C. van Marrewijk (2008), "It's a Big World after all: on the Economic Impact of Location and Distance", *Cambridge Journal of Regions, Economy and Society*, 1(3), 411-437.

Brakman, S., H. Garretsen and C. van Marrewijk (2009), The New Introduction to Geographical Economics, Cambridge University Press: Cambridge.

Braunerhjelm, P., R. Faini, V. Norman, F. Ruane and P. Seabright (2000), Integration and the Regions of Europe: How the Right Policies Can Prevent Polarization, CEPR: London.

Brittan, S. (2007), "It is Time to Jettison the Forecasts", *Financial Times*, December 20.

Brülhart, M. (1998), "Economic Geography, Industry Location and Trade: The Evidence", *The World Economy*, 21(6), 775-801.

Brülhart, M. and F. Trionfetti (2004), "Public Expenditure, International Specialisation and Agglomeration", *European Economic Review*, 48(4), 851-881.

Brundsen, J. (2016), "Junker Admits EU Faces 'Existential Crisis'; State of the Union", *Financial Times*, September 15.

Buckley, P., J. Clegg, H. Voss, A. Cross, X. Liu and P. Zheng (2018), "A Retrospective and Agenda for Future Research on Chinese Outward Foreign Direct Investment", *Journal of International Business Studies*, 49, 4-23.

Cairncross, F. (2001), The Death of Distance 2.0, Texere Publishing: London.

Cantillon, R. (1755), Essai sur la Nature du Commerce en Général (Essay on the Nature of Trade in General), (written in about 1730, published in 1755). Available at http://la.utexas.edu/users/hcleaver/368/368cantillonessaytable.pdf. Accessed on January 27, 2025.

Caves, R. (1996), Multinational Enterprise and Economic Analysis, Cambridge University Press: Cambridge.

Chakrabarti, R. and B. Scholnick (2002), "Exchange Rate Expectations and Foreign Direct Investment Flows", *Weltwirtschaftliches Archiv*, 138,1-21.

Chinitz, B. (1961), "Contrasts in Agglomeration: New York and Pittsburgh", *American Economic Review*, 51(2), 279-289.

Christaller, W. (1933), Die Zentralen Orte in Süddeutschland, Gustav Fischer: Jena. [Translated from German by C. Baskin (1966) "Central Places in Southern Germany", Prentice-Hall: Englewood Cliffs].

Chu, A., J. Smyth and S. Jones (2024), "German Industry Unlikely to Fully Recover from Energy Crisis", *Financial Times*, April 10.

Combes, P-P., T. Mayer and J-F. Thisse (2008), Economic Geography, Princeton University Press: Princeton.

Connolly, K. (2025), "Germany Election: Merz Says it's 'Five to Midnight' for Europe", *The Guardian*, February 24.

Curzon Price, V. (2011), *Institutional Competition in the European Union: Causes and Consequences of the Drive to Harmonise'*, in: M. Jovanović (Ed.), "International Handbook on the Economics of Integration: Competition, Spatial Location of Economic Activity and Financial Issues", Edward Elgar: Cheltenham, pp. 3-20.

David, P. (1984), "High Technology Centers and the Economics of Locational Tournaments", Stanford University, mimeo.

Davies, R. and D. Weinstein (2003), "Market Access, Economic Geography and Comparative Advantage: An Empirical Test", *Journal of International Economics*, 59(1), 1-23.

Debaere, P., J. Lee and M. Paik (2010), "Agglomeration, Backward and Forward Linkages: Evidence from South Korean Investment in China", *Canadian Journal of Economics*, 43(2), 520-546.

de Meirlier, M. (2006), Location, Location, Location, Technical Publishing: Chapel Hill.

Dunning, J. (1988), Explaining International Production, Unwin Hyman: London.

Duranton, G. and D. Puga (2004), *Micro-Foundations of Urban Agglomeration Economies*, in: V. Henderson, J. Thisse (Eds), "Handbook of Regional and Urban Economics", Elsevier: Amsterdam, pp. 2063-2117.

Easson, A. (2001) "Tax Incentives for Foreign Direct Investment", International Bureau of Fiscal Documentation Bulletin, (Part I) pp. 266-274, (Part II) pp. 365-375.

Eaton, C. and R. Lipsey (1976), "The Non-Uniqueness of Equilibrium in the Löschian Location

Model", *American Economic Review*, 66(1), 77-93.

Eaton, C. and R. Lipsey (1979), "Comparison Shopping and the Clustering of Homogeneous Firms", *Journal of Regional Science*, 19(4), 421-435.

Eaton, C. and R. Lipsey (1982), "An Economic Theory of Central Places", *Economic Journal*, 92(365), 56-72.

Eaton, C. and R. Lipsey (1997), On the Foundations of Monopolistic Competition and Economic Geography, Edward Elgar: Cheltenham.

EU Commission (2005), Communication from the Commission to the Council and the European Parliament, 'Implementation of the Community Lisbon Programme – The Contribution of Taxation and Customs Policies to the Lisbon Strategy', COM(2005) 532 final, p. 4.

EurActiv (2016), "Renzi compares EU to Sinking Titanic's Orchestra", February 11.

EurActiv (2018), "BMW to Build €1 Billion Car Factory in Hungary", August 1.

Eurointelligence (2024), 'Technology is Why we are Losing', May 14.

Finger, J. and A. Yeats (1976), "Effective Protection by Transportation Costs and Tariffs: A Comparison of Magnitudes", *Quarterly Journal of Economics*, 90(1), 169-176.

Fleming, S., B. Hall and H. Agnew (2025), 'Davos Hits Peak Pessimism on Europe as US Exuberance Rises', *Financial Times*, January 23.

Fletcher, I. (2010), Free Trade Doesn't Work, U.S. Business & Industry Council: Washington, DC.

Florida, R. (2002), "The Economic Geography of Talent", *Annals of the Association of American Geographers*, 92(4), 743-755.

Foy, H. (2022), "Western Powers Oppose Linking Ukraine Peace Talks to Sanctions Relief for Russia", *Financial Times*, March 31.

Friedman, T. (2006), The World is Flat, Farrar, Straus and Giroux: New York.

Fujita, M. (2005), *Introduction*, in: M. Fujita (Ed.), "Spatial Economics" - Volume I, Edward Elgar: Cheltenham, pp. ix-xxxv.

Fujita, M., P. Krugman and A. Venables (1999), The Spatial Economy, MIT Press: Cambridge.

Fujita, M. and J-F. Thisse (1996), "Economics of Agglomeration", *Journal of the Japanese and International Economies*, 10(4), 339-378.

Fujita, M. and J-F. Thisse (2008), "New Economic Geography: An Appraisal on the Occasion of Paul Krugman's 2008 Nobel Prize in Economics", CEPR Discussion Paper No. 7063.

Geopolitical Futures (2018), "In China, Manufacturers Feel the Heat of the Trade War", December 27.

Geopolitical Futures (2025), "The Coming Battle for AI", March 7.

Gereffi, G. (2016), "Global Value Chains and Upgrading: Export Promotion is FTZs", World Free Trade Zones Organization Bulletin No. 5, pp. 3-11.

Goldberg, P.K. (2024), "What a Successful Industrial Policy Needs", Project Syndicate, September 17.

Gordon, R. (2003), "America Wins with a Supermarket Sweep", *Financial Times*, August 19.

Grabher, G. (1993), *The Weakness of Strong Ties: The Lock-In of Regional Development in the Ruhr Area*, in: G. Grabher (Ed.), "The Embedded Firm", Routledge: London, pp. 255-277.

Graham, E. (1978), "Transatlantic Investment by Multinational Firms: A Rivalistic Phenomenon?", *Journal of Post Keynesian Economics*, 1(1), 82-99.

Grossman, G. (1998), *Comment*, in: J. Frenkel (Ed.), "The Regionalization of the World Economy", Chicago University Press: Chicago, pp. 29-31.

Guthrie, J. (2007), "How the Old Corporate Tortoise Wins the Race', *Financial Times*, February 15.

Hackmann, A. and T. Klarl (2018), "The Evolution of Zipf's Law for U.S. Cities", CESifo Working Paper No. 7232.

Hägerstrand, T. (1952), The Propagation of Innovation Waves, The Royal University of Lund: Lund.

Hajkova, D., G. Nicoletti, L. Varia and K. Yoo (2006) "Taxation, Business Environment and FDI Location in OECD Countries", OECD Economics Department Working Papers No. 502, OECD: Paris.

Harris, C. (1954), "The market as a factor in the localization of industry in the United States", *Annals of the Association of American Geographers*, 44(4), 315-348.

Hayek, F. (1944), The Road to Serfdom, George Routledge: London.

Hayek, F. (1945), "The Use of Knowledge in Society", *American Economic Review*, 35(4), 519-530.

Henderson, J. (1974), "The Sizes and Types of Cities", *American Economic Review*, 64(4), 640-656.

Holmes, T. (1998), "The Effects of State Policies on the Location of Manufacturing: Evidence from State Borders", *Journal of Political Economy*, 106(4), 667-705.

Hospers, G., F. Sautet and P. Desrochers (2008), *Silicon Somewhere: Is there a Need for Cluster Policy?*, in: C. Karlsson (Ed.), "Handbook of Research on Innovation and Clusters", Edward Elgar: Cheltenham, pp. 430-446.

Hotelling, H. (1929), "Stability in Competition", *Economic Journal*, 39(153), 41-57.

Hummels, D. (1999), "Have International Transportation Costs Declined?", University of Chicago, Graduate School of Business, mimeo.

Hummels, D. (2006), *Transportation Costs and Trade over Time*, in: "Transport and International Trade", OECD/ECMT: Paris, pp. 7-26.

Hummels, D. (2007), "Transport Costs and International Trade in the Second Era of Globalization", *Journal of Economic Perspectives*, 21(3), 131-154.

Hymer, S. (1976), The International Operations of National Firms: A Study of Direct Foreign Investment, MIT Press: Boston.

Inagaki, K., E. White and P. Nilsson (2025), "Europe Helped Teach China to Make Cars, Now Tables are Turning", *Financial Times*, April 16.

Isard, W. (1956), Location and Space-Economy, Chapman & Hall: London.

Jackson, G. (2016), "Number of EU Workers Rises as Jobless Rate Reaches 11-Year Low", *Financial Times*, November 17.

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

Jenkins S. (2002), "Cherie Mauled by the Tiger Tony couldn't Tame", *The Times*, December 11.

Jovanović, M. (2009), Evolutionary Economic Geography, Routledge: London.

Jovanović, M. (2015), The Economics of International Integration, Edward Elgar: Cheltenham.

Jovanović, M. (2019), "The Supply Chain Economy: How Far does it spread in Space and Time?', *Economia Internazionale/International Economics*, 72(4), 393-452.

Jovanović, M. (2020), Evolutionary Spatial Economics, Edward Elgar: Cheltenham.

Jovanović, M. (2022), "Economic Sanctions: Disappointing Old Wine in New Bottles", *Economia Internazionale/International Economics*, 75(4), 545-576.

Kalajić, D. (1971), Uporište (Stronghold), Delo: Belgrade.

Kalajić, D. (2015), Smak sveta (Doomsday), Mir Publishing: Belgrade.

Kassam, A. (2024), "Dutch Ministers Trying to Stop Tech Firm ASML Moving abroad over Foreign Labour Fears', *The Guardian*, March 6.

Kay, J. (2004), Culture and Prosperity, HarperCollins: New York.

Kay, J. (2006), The Hare & the Tortoise. An Informal Guide to Business Strategy, The Erasmus Press: London.

Kay, J. (2009), "The Rationale of the Market Economy: A European Perspective", *Capitalism and Society*, 4(3), 1-10.

Kay, J. (2015), Other People's Money, Profile Books: London.

Knickerboker, F. (1973), Oligopolistic Reaction and Multinational Enterprise, Harvard University: Boston.

Kopp, A. (2006), *Summary of Discussions*, in: "Transport and International Trade", OECD/ECMT: Paris, pp. 87-102.

Kravis, I. and R. Lipsey (1982), "The Location of Overseas Production and Production for Export by U.S. Multinational Firms", *Journal of International Economics*, 12(3-4), 201-223.

Krugman, P. (1979), "Increasing Returns, Monopolistic Competition and International Trade", *Journal of International Economics*, 9(4), 469-479.

Krugman, P. (1980), "Scale Economies, Product Differentiation and the Pattern of Trade", *American Economic Review*, 70(5), 950-959.

Krugman, P. (1990), Rethinking International Trade, MIT Press: Cambridge, MA.

Krugman, P. (1991), "Increasing Returns and Economic Geography", *Journal of Political Economy*, 99(3), 483-499.

Krugman, P. (1992), Geography and Trade, MIT Press: Cambridge, MA.

Krugman, P. (1993), "On the Number and Location of Cities", *European Economic Review*, 37(2-3), 293-298.

Krugman, P. (2000), "Can America Stay on Top?", *Journal of Economic Perspectives*, 14(1), 169-175.

Kynge, J. and D. Roberts (2003), "The Pearl River Delta is Attracting $1 bn of Investment a Month amid one of the Fastest Bursts of Economic Development in History", *Financial Times*, February 4.

Kuhn, T. (1970), The Structure of Scientific Revolutions, University of Chicago Press: Chicago.

Lakshmi, A. (2024), "Why Vietnam Risks Ending up a Big Loser from Donald Trump's Tariffs", *Financial Times*, November 18.

Lambooy, J. (1997), "Knowledge Production, Organisation and Agglomeration Economies", *GeoJournal*, 41, 293-300.

Leamer, E. (1984), Sources of International Comparative Advantage, The MIT Press: Cambridge, MA.

Le Monde (2024), "In EU Speech, Macron Says Europe is 'Mortal' and 'can Die", April 25.

Liebowitz, S. and S. Margolis (2001), Winners, Losers & Microsoft, The Independent Institute: Oakland.

Limão, N. and A. Venables (2001), "Infrastructure, Geographical Disadvantage, Transport Costs, and Trade", *World Bank Economic Review*, 15(3), 451-479.

Linder, S. (1961), An Essay on Trade and Transformation, Almquist & Wiksells: Uppsala.

Lipsey, R.G. (2013), *Some Contentious Issues in Theory and Policy in Memory of Mark Blaug*, in: M. Boumans, M. Klaes (Eds), "Mark Blaug: Rebel with Many Causes", Edward Elgar: Cheltenham, pp. 31-62.

Lipsey, R.G., K. Carlaw and C. Bekar (2005), Economic Transformations: General Purpose Technologies and Long Term Economic Growth, Oxford University Press: Oxford.

Lloyd Parry, R. (2024), "Japanese Nappy Maker Switches to Adult Incontinence Pads", *The Times*, March 27.

Lösch, A. (1940), Die Räumliche Ordnung der Wirtschaft, Gustav Fischer: Jena [Translated from German by W. Woglom and W. Stolper, The Economics of Location (1973), Yale University Press New Haven:].

Lucas, R. (1988), "On the Mechanics of Economic Development", *Journal of Monetary Economics*, 22(1), 3-42.

Ma, J. (2015), "Harvard Rejected me 10 Times", World Economic Forum, September 14.

Marin, D. (2025), "The End of Globalisation as we Know it", Project Syndicate, March 5.

Markusen, J. (1984), "Multinationals, Multi-Plant Economies and the Gains from Trade", *Journal of International Economics*, 16(3-4), 205-226.

Markusen, A. (1996), "Sticky Places in Slippery Space: A Typology of Industrial Districts", *Economic Geography*, 72(3), 293-313.

Markusen, J. and A. Venables (2000), "The Theory of Endowment, Intra-Industry and Multi-National Trade", *Journal of International Economics*, 52(2), 209-234.

Marshall, A. (1890), Principles of Economics, Macmillan: London.

Martin, R. and P. Sunley (2003), "Deconstructing Clusters: Chaotic Concept or Policy Panacea?", *Journal of Economic Geography*, 3(1), 5-35.

Martin, R. and P. Sunley (2006), "Path Dependence and Regional Economic Evolution", *Journal of Economic Geography*, 6(4), 395-437.

McGee, P. (2018), "VW Fined €1 bn over Role in Diesel Scandal", *Financial Times*, June 14.

Medetsky, A., M. Campbell and Y. Fedorinova (2016), "Putin is Growing Organic Power one T-34 Tank-Tomato at a Time", Bloomberg, June 7.

Midelfart, H., H. Overman and A. Venables (2003), "Monetary Union and the Economic Geography of Europe", *Journal of Common Market Studies*, 41(5), 847-868.

Mikhailova, A. (2018), "Project Fear's Brexit Predictions were Wrong by £100 billion, New Report Says", *Telegraph*, April 22.

Mody, A. and K. Srinivasan (1998), "Japanese and U.S. Firms as Foreign Investors: do they March to the Same Tune?", *Canadian Journal of Economics*, 31(4), 778-799.

Mokyr, J. (2017), A Culture of Growth, Princeton University Press: Princeton.

Mongoven, B. (2007), "The US Energy Debate: Whether to Bet on Future Technology", Stratfor, August 2.

Moran, T., E. Graham and M. Blomström (2005), *Conclusions and Implications for FDI Policy in Developing Countries, New Methods of Research and a Future Research Agenda*, in: T. Moran, E. Graham, M. Blomström (Eds), "Does Foreign Direct Investment Promote Development?", Institute for International Economics: Washington, pp. 375-395.

Moses, L. (1958), "Location and the Theory of Production", *Quarterly Journal of Economics*, 72(2), 259-272.

Mucchielli, J. and F. Puech (2004), *Globalization, Agglomeration and FDI Location: The Case of French Firms in Europe*, in: J. Mucchielli, T. Mayer (Eds) "Multinational Firm's Location and the New Economic Geography", Edward Elgar: Cheltenham, pp. 35-58.

Münchau, W. (2022), "After Pestilence and War Comes Famine", Eurointelligence, April 25.

Myrdal, G. (1957), Economic Theory and Under-Developed Regions, Duckworth: London.

Nova News (2024), "In September, EU Countries Bought Russian Gas Worth 1,5 Billion Euros", November 21.

Ohlin, B. (1933), Interregional and International Trade, Harvard University Press: Cambridge.

Paelinck, J. and M. Polèse (1999), "Modelling the Regional Impact of Continental Economic Integration: Lessons from the European Union for NAFTA", *Regional Studies*, 33(8), 727-738.

Parikh, T. (2025), "Globalisation will Triumph over Donald Trump", *Financial Times*, March 30.

Porter, M. (1990), The Competitive Advantage of Nations, The Free Press: New York.

Posner, M. (1961), "International Trade and Technical Change", *Oxford Economic Papers*, 13(3), 323-341.

Reed, J. and M. Nakamoto (2007), "How Toyota Went Native in America", *Financial Times*, August 6.

Rein, S. (2014), The End of Copycat China, Wiley: Hoboken.

Reuber, G. (1973), Private Foreign Investment in Development, Clarendon Press: Oxford.

Rodrik, D. (2024), "America's Manufacturing Renaissance will Create Few Good Jobs", Project Syndicate, April 16.

Romer, P. (1986), "Increasing Returns and Long-Run Growth", *Journal of Political Economy*, 94(5), 1002-1037.

Rosenthal, S. and W. Strange (2004), *Evidence on the Nature and Sources of Agglomeration Economies*, in: V. Henderson, J. Thisse (Eds), "Handbook of Regional and Urban Economics", Elsevier: Amsterdam, pp. 2119-2171.

Rugman, A. and R. Hodgetts (2001), "The End of Global Strategy", *European Management Journal*, 19(4), 333-343.

Schelling, T. (2006), Micro Motives and Macro Behaviour, W.W. Norton: New York.

Schmitz, H. (1999), "Collective Efficiency and Increasing Returns", *Cambridge Journal of Economics*, 23(4), 465-483.

Schumpeter, J. (1983), The Theory of Economic Development, Transaction Books: New Brunswick, NJ.

Schumpeter, J. (1991), *American Institutions and Economic Progress*, in: R. Swedberg (Ed.), "The Economics and Sociology of Capitalism", Princeton University Press: Princeton, pp. 438-444.

Segal, D. (2010), "The X Factor of Economics", *New York Times*, October 17.

Skidelsky, R. (2021), "The Case against Economic Sanctions", Project Syndicate, October 19.

Smith, D. (1981), Industrial Location – An Economic Geographical Analysis, John Wiley: New York.

Soros, G. (2014), The Tragedy of the European Union, Public Affairs: New York.

Stam, E. (2010), *Entrepreneurship, Evolution and Geography*, in: R. Boschma, R. Martin (Eds), "The Handbook of Evolutionary Economic Geography", Edward Elgar: Cheltenham, pp. 139-161.

Stewart, H. (2025), "UK Second Most Attractive Country for Investment", *The Guardian*, January 21.

Stratfor (1999), "IMF's Camdessus Misses the Point", November 15.

Tanzi, V. and H. Zee (2000), "Tax Policy for Emerging Markets: Developing Countries", *National Tax Journal*, 53(2), 299-322.

Telesković, A. (2018), "Kontinental: ne radimo iza leđa domaćoj IT industriji" (Continental: We are not Working behind the Back of the Local IT industry), *Politika*, March 27.

Temple, J. (2002) "The Assessment: The New Economy", *Oxford Review of Economic Policy*, 18(3), 241-264.

Tharakan, J. and J. Thisse (2011), *Market Integration: Trade versus Economic Geography*, in: M. Jovanović (Ed.), "International Handbook of the Economics of Integration: Competition, Spatial Location of Economic Activity and Financial Issues", Edward Elgar: Cheltenham, pp. 119-147.

The Economist (2000), "What the Internet cannot Do", August 19, p. 9.

The Economist (2009), "Invisible but Indispensable", November 5.

The Economist (2012), "A Sense of Place – Special Report", October 27, p. 18

The Economist (2015), "Climate Change – Special Report", November 28, p. 9.

The Economist (2016), "Bleak Times in Bra Town", April 16.

The Economist (2024), "Research Developments", July 20, pp. 49-50.

The Economist (2024), "Lessons from Northvolt's Failure", November 30, p.12.

The Financial Times (2025), "Northvolt's Fate Teaches Europe a Lesson: Stick to What you're Good at", March 12.

Thisse, J. (2015), "Skilled Cities, Regional Disparities and Efficient Transport: The State of the Art and a Research Agenda", CEPR Discussion Paper No. 10790.

Tobler, W. (1970), "A Computer Movie Simulating Urban Growth in the Detroit Region", *Economic Geography*, 46(S), 234-240.

UNCTAD (2003), World Investment Report – FDI Policies for Development: National and International Perspectives, United Nations: New York.

UNCTAD (2015), World Investment Report – Reforming International Investment Governance, United Nations: New York.

Venables, A. (1996), "Localization of Industry and Trade Performance", *Oxford Review of Economic Policy*, 12(3), 52-60.

Venables, A. (2006), "Shifts in Economic Geography and their Causes", London School of Economics, Jackson Hole Symposium, mimeo.

Venables, A. and N. Limão (2002), "Geographical Disadvantage: A Heckscher-Ohlin-von Thünen Model of International Specialisation", *Journal of International Economics*, 58(2), 239-263.

Vernon, R. (1966), "International Investment and International Trade in the Product Cycle", *Quarterly Journal of Economics*, 80(2), 190-207.

von Thünen, J. (1826), Der Isolierte Staat in Beziehung auf Landtschaft und Nationalökonomie. Perthes: Hamburg, [Translated from German by C. Wartenberg, von Thüunen's Isolated State (1966), Pergamon Press: Oxford].

Waldmeir, P. (2012), "China Offers a Taste of R&D to Come", *Financial Times*, November 14.

Waterfield, B. (2024), "Macron: Too much Regulation and not enough Cash will kill EU", *The Times*, October 3.

Weber, A. (1909), Über den Standort der Industrien, J.C.B. Mohr: Tübingen [Translated from German by C. Friedrich, Theory of the Location of Industries (1962), University of Chicago Press: Chicago].

Wheeler, D. and A. Mody (1992), "International Investment Location Decisions: The Case of U.S. Firms", *Journal of International Economics*, 33(1-2), 57-76.

World Bank (2009), Reshaping Economic Geography, World Development Report, The World Bank: Washington.

Yamawaki, H. (1993), *Location Decisions of Japanese Multinational Firms in European Manufacturing Industries*, in: K. Hughes (Ed.), "European Competitiveness", Cambridge University Press: Cambridge, pp. 11-28.

Yoon, J. (2025), "The US has Spurred the Chinese Chip Industry", *Financial Times*, March 5.

Zipf, G. (1949), Human Behavior and the Principle of Least Effort, Addison-Wesley: Cambridge MA.

