One could reflect hours or even days on the title of this lecture proposed by the organiser: Did geography at any time play a major role in economic theory, at least in mainstream theory? If not, what than actually got lost, and when? Furthermore: has modern economic theory really found geography? Can we detect important geographical aspects in most recent economic mainstream theory? And what is “geography”?

I will not investigate into this latter question and will leave this tricky problem to geographers. As an economist I will deal with geography as the location and distribution of economic activity across space from an economists’ point of view. I will investigate into the geographical aspects in the economic literature and completely ignore the economic aspects in the geographical literature. I will argue that geographical factors – at least in my definition – dominated economic development before economic theory existed as a science. At the time when economics arose as a science in the work of the great classic economists’, that is in the late 18th century, the industrial revolution had produced a bipolar world – Britain and the Rest of the World –, depleted of any geographical aspects. This reality was modelled in contemporary classical theory; it dominated the dominating Anglo-American economic theory up to the eighties of our century at least, at a time when the real world had changed and differed drastically from the models’ world. In the middle of the 19th century German economic policy had attacked the classic’s models bipolar view of life – in reality and – consequently – in economic theory. International success of this alternative theory, however, failed to appear for more than a hundred years at least. Since than, however, several factors brought geographical determinants back into world-wide reality: The specialisation of industries within countries due to higher knowledge intensity, European integration, the loss of dominance first of Britain, than of the United States, as well as policy efforts to ward off suspected negative effects of globalisation. In the 1980’s economic theory followed suit – slowly.

I will first sketch the varying importance of geographical factors for economic development and propose an answer to the question, why economic theory has not found geography in mainstream theory for more almost two centuries. In the second part I will try to investigate, to which extent modern economic theory has already found geographical aspects.

1 Geographical factors dominate early economic development

---

Footnote 1:
“I suspect that geographers proper will not be entirely pleased at what they see and may deny that the kind of stylized models that economists find appealing are part of their field.” (Krugman 1991b, xi). However, “partly because of the novelty, partly because of the inherent sexiness of the stories ... these models are a lot of fun to work with.” (Krugman 1998, 172). “... the technical tricks that characterize the new economic geography have opened the door to theoretical modelling of the sort of inherently exciting spatial economic issues that economists had previously neglected.” (Krugman 1998, 168).
Until the heydays of industrial revolution economic development depended almost exclusively on geographic determinants. In the very early days climatic conditions and availability of resources dominated: Agriculture was formed in the mountains of the Golden Triangle north of Mesopotamia as in this very place the ancestors of sheep and goat lived and the early forms of grain grew wildly – geography did matter. After the domestication of animals, the cultivation of plants and after experience had been gained with the cultivation of (heavier) soil, civilisation centred in the fertile valleys of Euphrates, Indus, Nile, etc. The dominant centres, however did not form at the sites of the resources. As administrative and/or trading centres they arose at sites characterised by intersections of trading routes, change of the means of transportation, or crossing of rivers. Merchants met each others at those places; quite often not only the means of transportation had to changed at those places, quite often a different merchant as well served the further part of the route, so that the goods were frequently sold in these places. From an economic point of view therefore two important facts can be deducted: Firstly, that geographical factors dominated early economic development; and secondly, that not the factors of production dominating Classical Theory – labour, capital and soil – worked out as the dominating determinants of actual economic development in these early times, but trade-induced accumulation as well as condensation of knowledge and information; modern economic theory has become aware of this important aspect.

The dominance of trade over production as the main determinant of agglomeration was one of the characteristics of economic development not only in early times but in antiquity and in the Middle Ages as well. One of the most famous examples is the well-known Silk road: Due to a large number of facts – wars, position and influence of states, transport technology, etc. – this “road” shifted there and back in time, between Siberia and Indian Ocean (Haussig 1983, 1988), but wealth and economic development was wherever the “road” actually was, the route relevant for the time being. Geography mattered and trade mattered and this is true for inner-European development as well: In Austria, for instance, European centre of iron production in the 16th century, agglomeration took place and wealth accumulated at the trading-places (Bruck/Mur, Steyr, Waidhofen), not were the iron was found, where it was melted or at the centres of production of iron goods. The period between 1500 and 1750 is known as the age of merchant capitalism. Despite very high land-transport costs, inner-European trade in grain, cattle, salt, metal and textiles was important, and considerable and profitable trade took place with Asia. European regions specialised in their economic activities, and the big trading houses and trading companies created an European market: “The Whole World as to Trade, is but as one Nation or People, and therein Nations are as Persons.” (Dudley North, 1691, cited in Glamann 1971, 452). The centre of this one world of European Trade however shifted from the South of Europe to its North, from the Mediterranean to the North Sea – another aspect in which geography mattered in the real world of economic development of those times. However: No well defined and elaborated economic theory existed at that time: Macroeconomic thinking was unfamiliar in this age of mercantilism, contemporaries saw no difference between nations and persons, as the quotation of Dudley North demonstrates, thus susceptible for the fallacy of composition. Both, nations as well as persons were

2 Along the land-route of the silk-road Turfan, Dunhuang, Kashgar and Barygaza (at the estuary of the Narmada) should be mentioned as such type of trading places, Kiev and Trebizond in the North, and Sri Lanka, Aden and Damascus along the sea-route (Haussig 1988).

3 This is in complete accordance with models of New Economic Geography: Fujita and Mori (1996) demonstrated that such nodes generate cusps in the market potential function.

4 In the late 16th century the transport of grain from Krakow to Venice caused an fourfold increase in price (Glamann 1971, 455).

5 Quantitatively extra-European trade was small – in 1700 Britain imported 359 000 tons (tonnage of shipping) and exported 182 000 tons, while the East-Indian trade accounted for 5000 tons (Glamann 1971, 454). More than half of the imports, however, was timber and two fifth of the exports coal. East-Indian imports in contrast were spices, tobacco, tea, coffee, luxury textiles, much less bulky but much more expensive.
considered to become wealthy by accumulation of gold. Economic theory could not find geography in this period, because no elaborated economic theory existed.

2 Geography “lost” in the real world of Industrial Revolution

The real world changed drastically in the course of the Industrial Revolution. Any differences among nations and regions, typical for the pre-industrial period, were dwarfed by the differences between industrialised Britain and the Rest of the World. Britain produced and exported industrial goods, the Rest of the World extracted raw materials, food and other inputs for the British industrial complex. Principal British exports goods in 1831 were industrial products as cotton yarns and textiles (24 percent), woollen yarns and textiles (7 percent), linens (3 percent), and cutlery and hardware (2 percent), principal import goods cotton (19 percent), sugar (15 percent), tea and coffee (11 percent), grain (9 percent) and flax and hemp (5 percent)(Woodruff 1979, 664, 672). In stark contrast to the time before, production displaced trade as the primary source of wealth, at least in dominating Britain: British coal production increased from 11 million metric tons in 1800 to 50 million metric tons in 1850 and 230 million metric tons in 1900, British production of pig iron from 0,2 million metric tons (1800) to 2,7 (1850) and 8,8 million metric tons (1900). From 1750 to 1800 British industrial production increased by 160 percent, from 1800 to 1850 by 360 percent and from 1850 to 1900 by a further 200 percent (Mitchell 1979). The geographically differentiated, multi-polar world of mercantilism was displaced by a bi-polar division into industrialised Britain and non-industrialised Rest-of-the World. A bi-polar world, however, leaves no room for geographic differentiation. Economic theory as a science originated in this very situation of reality: Britain as the leading and unrivalled industrial power of the world. Adam Smith in his famous “Inquiry into the nature and causes of the wealth of nations” (1776) emphasised that neither gold nor resources nor trade constitute the basis of the wealth of a nation but its productive capacity. He opted for free trade as a “vent for surplus”. David Ricardo (Principles of political economy and taxation, 1821) emphasised that trade and competitiveness rely not on absolute but on relative, on comparative advantages, thus calming the anxieties of Britain’s less industrialised trading partners, feeling unprotectedly exposed to the competitiveness and power of the dominating nation: Everybody gains from free trade – according to Ricardo’s theory. No doubt: The main elements of classical economic theory were custom-tailored to the British situation and to British interests, to the interests’ of the leading power in world production and world trade. In its developed form, the (neo)classical theory of trade is the famous 2x2x2-theory: Two countries – Britain, later the U.S., and the Rest of the World –, two factors of production – labour and capital –, and two goods – manufactured ones and others (inputs). Labour and capital are considered as immobile.

---

6 World production of pig iron increased from 1 million ton in 1820 to approximately 20 tons in 1900 (Woodruff 1979, 682), production of ferrous metals in general increased sixty-five times between 1820 and 1910 (Woodruff 1979, 660).
7 World trade in value terms increased nearly nine-fold between 1820 and 1880 (Woodruff 1979, 658).
8 The intention of this paper is not – and cannot be – to elaborate the depth and richness of the different theories. Its intention is to confront the importance of geographical factors in reality and in economic theory at different times; it therefore refers only to those parts of economic theory which are relevant for this very purpose.
9 Kurz 1990a, 244f
10 Notice that Ricardo’s famous example in chapter VII refers to cloth as an export good of Britain (and an import good of Portugal) and wine as import good of Britain (and export good of Portugal).
within this model, an assumption not without a corollary in the real world of the dominant economy. Britain protected her comparative advantage by highly restrictive practices: The severe restrictions on the migration of skilled artisans were not lifted before 1825, the restrictions on the export of machinery and important techniques not until 1843 (Landes 1969, 148; Woodruff 1979, 678). Similar restrictions are known to exist in the leading power to follow, the United States: In that case the restrictions referred to the qualification of immigrants and to the exports of high-technology – for the rest of the goods the U.S. fights (or fought at least as long as its dominance was undisputed) for free trade as well.

Geographic elements completely lacked in the classical models for two reasons: First the structure of a model (or a world) consisting of two regions – Britain and the rest of the world – is not rich enough to model geographical influences, and secondly the model promised quick convergence of any differences anyhow: Free trade was assumed to quickly equalise prices among nations, assure worldwide full employment even of immobile factors of production, and equalisation of factor prices.

The free-trade-convergence model of classical economic theory served the interests’ of the leading economic powers well – first Britain, then United States – and it so became the workhorse of Anglo-American theory and the basis for the free-trade ideology until today. The high level of the models’ abstraction shielded it against critique, theoretical as well as practical one. Geography could not enter into this model, neither internationally nor nationally and it did not enter, because it was of no importance in the real world of Adam Smith and David Ricardo.

3 Geographic elements “found” in the German model of industrial catch-up

“... the Crystal Palace Exposition in 1851 marked the apogee of Britain’s career as the ‘workshop of the world’. ... Her merchandise dominated in all markets of the world; her manufacturers feared no competition” (Landes 1969, 124); British income per capita was approximately two-and a-half times the German one This outstanding position had resulted from Britain’s head-start in the Industrial revolution and the big cost advantage of sea transport. Both advantages, however, turned out as temporary. Germany worked hard to catch up, using methods very different from the British ones, and the fast diffusion of railways quickly reduced the cost-advantage of sea transport: The cost of land transport halved in the eighty years between 1834 and 1914 (Woodruff 1979, 690).

The countries of Continental Europe in general had a very slow start as the dominating British industry destroyed the weak starting points for industrialisation and overtook their markets. Nevertheless the countries worked hard to improve their material infrastructure of roads, waterways and railroads, with Germany at the forefront. The immaterial infrastructure of expertise in new industrial technology was slowly improved by employing British artisans, and, especially in Germany, by government efforts to create technical schools, embedded in a larger system of advanced education, non-teaching academies, museums and technical expositions (Landes 1969, 151, 340 seqq). The formation of the Zollverein enlarged the economic area in 1834, liberal policies in the fifties and sixties relieved industries from restricting regulations (Gewerbefreiheit). Increasing scale and the forces that brought it about combined to recast the economic map of Europe in the form of localisation, spatial concentration, relocation and the rise of new centres of industrial activity. Geographical aspects reappeared in economic development with the transition from a bi-polar to a multi-polar world.

In contrast to the complete lack of international mobility of labour and capital, perfect mobility was assumed within countries. The nation thereby was implicitly reduced to one point, so that regional factors – geography within the country – were excluded as well.
Britain’s share in world industrial output shrunk from one third in 1870 to one fifth in 1900 and one seventh at the eve of World War I.

Two German particularities must be mentioned: Firstly the early transition from family- to professional management, the predilection for large enterprises and the political acceptance of mergers (Fukuyama 1995, 252). The particularity, however, most relevant not only for the catch-up but for a more distant future as well, was the heavy weight government put on *compulsory education* and the *establishment of a science base*; in this field Germany was far ahead of Britain (Landes 1969, 340), and this advantage proved essential in the coming period of marriage of science and technology. As a result the German system was characterised by institutionalised innovation, while the British manufacturers preferred empiricist tinkering and refrained from innovations not immediately successful (Landes 1969, 352). Above all the German chemical industry was famous for its scientific laboratories, and it was the German chemical industry which proved to be the most successful. “Pure and applied science were brought together in an extraordinary way. So much so, that on the eve of war the Western industrial nations had begun to make what they wanted out of what they had – a very different situation from that which had prevailed in the nineteenth century. International trade and development then had depended on the close inter-relatedness and reciprocity of the industrial with the primary producing world. Europe had enriched itself (none so much as Britain) from this arrangement. But Europe’s new wealth ... had weakened some of the economic ties that bound Europe to the primary producing parts of the world.” (Woodruff 1979, 685). The bi-polar world – Britain and the Rest of the world – had come to an end, a multi-polar world re-emerged, *geography stood at the brink of regaining importance*.

German economic *theory* reflected this situation, completely different to the British one. German economic reflection had always been different from the Anglo-American one, historical-national rather than abstract-analytical; it attempted to explain and predict the rise and fall of the nation, to explain the *national* system of political economy, not the causes of the wealth of nations in general. *Spatial* aspects had an important place in German economic thinking, e.g. in the location theory of Thünen (1826). The most important opponent to the British theory of the classical free-trade-convergence model, was List (1841). He was most interested in explaining the changing position of nations in history, and in devising strategies for a backward nation to catch up with the leading countries, the situation of his home-country Germany at that time. He correctly emphasised, that free trade would preserve the advantage of the leader, that it would even increase the inequality among nations. Free trade in goods could be accepted by backward nations if accompanied by free movement of expertise at best, a claim heavily opposed to by Britain. List suggested four policy options for Germany: *Investing in education* to promote an adequately trained work-force; creating a network of infrastructures to allow the *dissemination of the most important resource, know-how*; creating economic ties among countries, such as *customs unions*, and to *protect infant industries* to allow them to develop the expertise needed to face international competition. List has become famous for his last proposal – Erziehungszölle – but this was only one of several ones, all reflecting closely what Germany had to do – and actually did – to improve her backward position. Gerschenkron (1962) emphasised that industrial catch-up occurs by no means automatically, but requires a significant amount of effort and “institutional instruments, for which there was little or nor counterpart in an established industrial country” (Gerschenkron 1962, 7).

Translated into the language of modern theory List – and Gerschenkron – emphasise that the route leading each nation to build her technological competence is highly path-dependent. More than one technological avenue leads to the wealth of a nation, more than one single equilibrium distribution of production among nations exists; each nation has her own characteristics and specialisation: Britain traded with the world, Germany with Europe at that time, and recent econometric tests confirm that the differences among industries within most countries are smaller than the differences among identical industries in different countries (Costello 1993; Bayoumi and Prasad 1995; Helg et al 1995). In summary: *Geographical factors certainly were of importance in the Europe of the late 19th century*
they could play a role within the theory sketched by the German school of those days. But this branch of theory was not accepted by the mainstream.

4 Geography “lost” in the closed national state from 1918 to the sixties

“The growing need of the most industrialised nations in 1913 was to find markets for their industrial goods. In sharp contrast to the relations that existed earlier ... the world’s major exporters of manufactured goods were increasingly the world’s major importers of manufactured goods (Woodruff 1979, 685; my italics G.T.). The struggle for economic dominance set in: World War I was one of the consequences and it changed the European economic system. The break-down of the gold standard, the economic troubles of the Great Depression and the devaluation races of the thirties marked the end of the seven-decades-long period of a multi-centred world-economy. Countries believed to fare best as closed national economies with their own national demand policy. Keynes provided the theoretical basis for this approach, considering demand as the key element for the state and the development of an economy. Contrary to Say’s law of classical theory, demand was assumed to automatically create its matching supply, but market forces were not believed to guarantee sufficient demand and full employment. Full employment and growth therefore became the dominant targets of economic policy.

After the Second World War this model continued to dominate economic policy and – slightly modified to the Neoclassical Synthesis – economic theory as well, even if the underlying conditions had changed considerably: Far from the depression-conditions of the thirties the industrial countries approached full employment. Sufficient demand was nevertheless considered as the crucial factor for the wealth of the nation, and national demand policy was held responsible for full employment. Foreign trade entered this model as an important demand component: As politicians – following the tradition of the thirties – regarded insufficient demand as the main threat, they considered export led growth as an important remedy. The dominant nation of that time – the United States – pressed for free trade – for obvious reasons – and the European countries did not object to liberalising foreign trade and even capital movements, as long as foreign trade soared – as it actually did. The countries did not realise that they thereby undermined their capacity to ensure full employment by demand management: Germany was the first country to find “Außenwirtschaftliche Absicherung” – external safeguard – of national demand policy as the critical element in the early sixties; other countries soon experienced the same problem. Two types of theoretical models were developed to deal with the problem of external “leakages”: Two-country models and the model of the Small Open Economy (SMOPEC). Both were devoid of geographical aspects: The two-country models – “We and the Rest of the World” – returned to the bi-polar world, familiar from classical and neo-classical theory, even if they emphasised demand rather than supply. The SMOPEC-model concentrated on national elements and introduced foreign relations as leakages only, as restrictions to domestic demand policy. Evidently geography could not matter in both types of models.

In the late sixties full employment was attained in most countries and the demand model evidently did not fit any more. Empirical economists posed the question “Why growth rates differ” (Denison 1967), and supply-side considerations reappeared – in the economic policy of Thatcher and Reagan earlier than in economic theory. Supply-side theory in general, however, was a superficial interlude. The incentive for a fundamental change in economic theory rather stemmed from foreign trade: Due to liberalisation, full employment and high growth foreign trade soared, but in a way very different from traditional theories’ expectations: Not trade with dissimilar goods among countries with different factor endowments soared, as was the case in the 19th century and as the Heckscher-Ohlin hypothesis had predicted, but trade with similar, industrial goods between similar, industrialised countries (Frankel 1995, 73)! This was a big challenge to the dominating neoclassical theory and the off-spring

5 Geographical elements in recent economic thinking

The classical model was a-geographical as it modelled the bi-polar state of the economic world of the late 18th and early 19th century. Subsequently reality and theory diverged: With the industrialisation and specialisation of the Continental European countries, geographical elements re-emerged in reality, but in economic theory the neoclassical model dominated, became more and more abstract and exempt of any elements not suitable for mathematical treatment. This style of rigid modelling gave rise to economics as a science but eliminated any connections to geography. Based on constant returns to scale, perfect competition and rid of any externalities, the model could not deal with concentration and specialisation (spatial as well industrial); with labour and capital assumed as perfectly mobile within countries but completely immobile in-between and with firms and countries devoid of any dimensions, the model could not deal with spatial aspects as well. Devoid of any structural aspects it was unable to deal with differences among countries. Foreign trade was modelled to bring about price and factor equalisation based on comparative advantage; the model did not incorporate transport costs and rested on partial immobility: If goods and factors of production are mobile, comparative advantage can no longer be defined. Neoclassical theory furthermore neglected aspects of distribution, as the model implied strong convergence.

The neoclassical model proved to be the workhorse of mainstream-economics up to the eighties; wherever economists had to deal with spatial aspects they were forced to use different models – in regional economics or in location theory. In regional economics models of Keynesian spirit dominated, implying quantity adjustment rather than market clearing by competitive price adjustment; the export-base concept was the most applied one. Polarisation models stressed cumulative causation and therefore divergence: Some of the determinants of growth are locally determined (e.g. local productivity is a function of local growth), some globally (e.g. the wage rate). Regional divergence therefore is more likely than convergence. Classical location theory rested on partial models: Factors of production were assumed perfectly mobile, scale economies and imperfect competition prevailed – allowing economies of agglomeration – and transport costs did matter. Firms, however, were assumed identical, market structure and motives of firms were ignored. “Much of the literature ... has ... been obsessed with geometry – with the shape of market areas on an idealized landscape …” (Krugman 1991b, 5). In general many of the new elements that nowadays have been incorporated in a more geographic-minded economy could have been found in regional or location economics. Economists of these branches were regarded as second class, however, and their articles did not find their way into the core journals. The adaptation of economic theory, therefore, was endogenous, resulting from the development of new mathematical tools, able to deal with new, more complicated assumptions.

The changeover to a new model wouldn’t have been that fast, however, hadn’t the real questions been so much at odds with the esoterical questions the neoclassical theory dealt with. Already in the early fifties Leontief’s paradox – America exported labour intensive goods and imported capital intensive ones which contradicts neoclassical trade theory – forced economists to reconsider the assumptions about factors of production, differentiating between skilled and non-skilled labour, vintages of physical capital, later to introduce technical progress or human capital as additional factors of production. The uneven economic development of countries posed the question “Why growth rates differ” (Denison 1967); theory’s inability to give an answer induced empirical growth accounting

12 Between 1950 and 1980 world trade increased by more than 12 percent per annum (6,8 % real).
13 For geographic patterns in European trade see Tichy (1992).
(Abramovitz 1956), finding the “residual” i.e. the unexplained factor as the main but unconvincing explanation. For other pressing problems of those days – growing concentration, rise of the multinational corporation, increasing specialisation, optimum size and composition of customs and currency areas (Tichy 1997a) or consequences of globalisation – neoclassical theory was not the appropriate instrument as well. It predicted convergence, but the actual debate oscillated between “delinking” of the poor South from the rich North and the fear that industrialisation of the developing countries destroys industrial countries’ full employment. At the same time the question of the optimality of world-wide free trade rose again – in Europe as a result of regional integration, in the United States as a consequence of dwindling own hegemony (McCulloch 1993), rise of Japan, the Southeast-Asian tigers and even Europe (end of the so-called Eurosclerosis). The win-win interpretation of free trade gave way to an awareness of its distributional aspects – rent shifting.

The elements of a new model were developed in the New Theory of International Organisation, in New Growth Theory, in innovation theory, in New Trade Theory and in consequence a New Economic Geography resulted. In the course of the seventies and early eighties they all switched to the new, more realistic assumptions of increasing economies to scale, imperfect competition and externalities and elaborated respective mathematical models. In the late eighties the profession became aware of the spatial dimension of the new paradigm: 1988 Lucas, shifting to a new and potentially more relevant field of analysis, remarked when looking at the widely differing and fluctuating growth rates: “I do not see how one can look on figures like these without seeing them as representing possibilities. In there some action a government of India could take that would take the Indian economy to grow like Indonesia’s or Egypt’s?”(emphasis in original). Lucas’ question was right, his method, however, highly inadequate. 1990 Porter investigated into the competitive advantage of nations, and in the same year Krugman (1991b, 1) “suddenly realized that I have spent my whole life as a international economist thinking and writing about economic geography, without being aware of it.” Geography had been found in economics – at least conceptually.

Several characteristics distinguish the new type of models from the traditional ones. Most important and most often emphasised, are increasing economies to scale and imperfect competition, which need no further description. Externalities assumed are specialised skills, provision of specialised not-easily traded inputs and technological spillovers. Countries and regions have a spatial dimension and transport costs, therefore, prevail. Technical progress is endogenous in most models, while the assumptions on the mobility of the different factors of production vary: Physical and financial capital are assumed mobile in most cases, human capital and/or knowhow mostly mobile among regions of one country, less so among countries. Labour is regarded as mobile in a few models only. Countries differ, especially as to their legal and financial systems, their education systems and their starting points. Path dependence and multiple equilibria are typical for all these new models, so that a wide differentiation and geographical diversity can result.

The several assumptions are combined in an already large number of different models with different ambitions and goals, unsuitable for a consistent description as to their spatial results. The price these models have to pay for their greater realism is that they cannot be solved analytically in most cases, and that their calibrations and simulations are extremely sensitive to parameter variations. It is not easy to deduct generalised results with broad applicability. The following paragraphs will therefore be

---

14 Krugman (1998, 164) coined the slogan “Dixit-Stiglitz, icebergs, evolution, and the computer” as a short-hand characteristics. Dixit and Stiglitz (1977) elaborated the standard model of monopolistic competition, the "iceberg"-assumption, is a trick to model transport costs (see footnote 16), “evolution” implies the conditions guiding to one of several possible equilibria, and “computer” refers to the fact, that models normally can’t be solved analytically but need the computer for calibrations and simulations.


16 For reasons of modelling transport costs frequently are assumed in the form of “iceberg"-costs, implying that the transported good smelts away in proportion to the transport distance.
restricted to the general insights that the new approaches can provide into aspects of economics related to potential geographical aspects.  

5.1 New growth theory

The New Growth Theory retains the neoclassical model as much as possible but attempts to endogenise technological progress by the introduction of externalities in the form of learning by doing or learning by producing. As these are external to the firm but internal to the economy the model is able to produce potential polarisation in an otherwise neoclassical model, a result highly interesting under geographical aspects. Because of the model’s challenging solution on the one hand, and its highly abstract character on the other hand, unrealistic as an explanation of real world development (Kurz 1998), a lively debate arose whether polarisation can actually be observed. Sala-i-Martin (1990, as in Sala-i-Martin 1996, 1327) introduced two concepts: $\beta$-convergence and $\sigma$-convergence. $\beta$-convergence is a negative relation in cross-sections between the growth rate of income per capita over a certain period and the initial level of income. $\sigma$-convergence implies a falling dispersion of real per capita income among groups of economies over time. A host of studies investigated into $\beta$-convergence among countries or regions within the neoclassical model to demonstrate a convergence rate of 2 percent per year (Barro and Sala-i-Martin 1991; Sala-i-Martin 1996), thus conserving neoclassical results even under the new assumptions.

The results are subject to at least two types of criticisms, theoretical and empirical ones. Already a decade before Barro and Sala-i-Martin Abramovitz (1986) had criticised the neoclassical convergence theory of productivity catch-up caused by unexploited technology: It can neither explain late comers to forge ahead, leading countries to fall behind nor changes of rank within the bulk, which frequently occur. Furthermore it is well known, that Germany, for instance, did not forge ahead of Britain because of catching up in the technological fields dominated by Britain and applying British methods, but by choosing a completely different approach in very different industries – as chapter 3 tried to demonstrate. As to empirics Abramovitz criticised that the evidence for catch-up primarily by closing the productivity gap is weak, even for the advanced industrial countries. The criticism of weak empirical confirmation of convergence is true for the data of Barro and Sala-i-Martin as well: Detailed analysis reveals polarisation results at least for the poorest European regions (Tichy 1997b, 25seqq). Quah (1996a, 1357seqq) demonstrated that $\beta$-convergence can result from statistical invariance and doesn’t say anything about how the poorest 10 percent of the distribution are catching up with the richest 10 percent: In effect concentration took place at the upper as well on the lower end of the concentration with a consequent weakening of the center (Quah 1993; Quah 1996a; Pritchett 1997), consistent with the theory of convergence clubs. Empirical evidence therefore suggests that both, convergence as well as polarisation forces are in action, with neighbouring regions usually belonging to the same club, irrespective of belonging to the same country or not. Similar results show up in studies investigating the differences as to the character of the shocks hitting the regions and the consequent reactions: 20 to 60 percent of the shocks are idiosyncratic (Decressin and Fatás 1995) and the countries react distinctly different even to common shocks (Fuss 1997). Today’s world is differentiated and it will remain so.

---

17 The New Theory of Industrial Organisation and the New Trade Theory will be somewhat neglected in this paper, partly for reasons of time and space, partly because most of their geographically interesting aspects are dealt with under the other headlines.

18 In country studies conditional convergence was investigated, i.e. the convergence with a set of additional variables held constant.

19 Quah, D.T, 1993; Pritchett 1998

20 .48 states of U.S., 47 prefectures of Japan, 90 regions of EU, 10 provinces of Canada.
5.2 Innovation Theory

As far as standard economic thinking considered geographical aspects at all, the availability of (natural) resources found most attention, thus reflecting the state of the world in the late 18th century. In general concentration of knowledge, be it trading knowledge or administration, appears to have determined economic development much more forcefully than mineral wealth. Modern theory has accepted this facts and emphasises the accumulation of human capital, knowledge networks, innovation and specialisation. Germany was the first country to recognise the importance of this additional factor of production in the early 19th century. Having realised the importance of a broad-based technical education and of in-house-R&D laboratories appears to have been one of the social innovations most important for further economic development, decisive in its contribution to a continuous flow of radical innovations.

Modern innovation theory stresses the importance of National Systems of Innovation (NSI), a concept introduced by Lundvall (according to Freeman 1997, 24) and first applied to Japan by Freeman (1987); the idea however can be easily traced back to List (1841). It says that the technological abilities of a nation’s firms are the main source of her technological prowess; the abilities are national and can be built by national action. Archibugi and Pianta (1992, 148) point to the striking diversity not only in the nature of the different NSI and the different ways of introducing innovations, but also to the different outcomes of these activities at the sectoral level. The national profiles in science and technology are not only strikingly diverse, their diversity and specialisation is still growing. Most industrialised countries therefore appear now present at the technological frontier in their selected field of specialisation.

This is in stark contrast to the neoclassical mainstream, assuming knowledge to be a free good, available worldwide and it raises the question: Why is technological knowledge so little mobile? This question is still under discussion, but externalities, economies to scale, learning effects (accumulation), slowness of learning new things and strategies of multinational corporations may be of greatest importance.

Nation-specific factors playing a crucial role in shaping the technological character of a nation are its institutions, especially education and research facilities, direct and indirect (via defence) support of innovation, and the way industries’ co-operation is organised, in combination with Marshall’s famous externalities: specialised skills, provision of specialised inputs – both clearly dependent on the education and science network –, and local technological spillovers, mainly via “transfer by feet”, i.e. mobility of skilled personal in-between local firms. Much of the nationally accumulated knowledge is of tacit character; its importance is best characterised by Foray’s (1991) phrase that in those industrial districts “the secrets of industry are in the air” – off course the secrets of that very industry branch or technology in which the district is specialised.

Multinational firms were supposed to quickly erode the national specialisation by transferring research, development and production to whatever place relatively best suited, thus quickly creating a homogenised undifferentiated world. As far as research has shown up to now, this is not the case, at least for the multinationals based in large countries. Strategic research and headquarter functions remain concentrated in the home country: The theory of strategic asset selection (Caves 1991) defines the multinational firm as a bundle of heterogeneous intermediate products, technologies and skills producing a bundle of products with national specifics. Acquisition of foreign firms and direct

---

21 See e.g. Piore and Sabel (1984) on North-Italy.

22 It is somewhat astonishing that one has to emphasise the importance and non-transferability of specialised knowledge: Everybody is aware that ham from Parma or Camembert cheese cannot be imitated, even if it may appear not too difficult to feed pigs, smoke ham and add the appropriate ferments to milk. Only in Parma or in France, however, one can learn the secrets out of the air of a long tradition. Industrial knowhow surely is not easier to transfer than ham- or cheese-knowhow.
investment in foreign countries serves to increase skills, especially the skills necessary to serve the foreign market as well; it normally does not distract strategic functions or activities from the headquarter. R&D is done at home and is significantly influenced by home-country characteristics (Patel and Pavitt 1991). Archibugi and Michie (1997b) proposed to distinguish within technological globalisation between „global exploitation of technology“, „global technological cooperation“ and „global generation of technology“. In the first and in the second case innovative activity remains were it had been before; global generation of technology is the only form implying locally decentralised R&D. Casson (1991, 272) emphasised that „the story of globalised R&D is the story of a fairly small number of very large firms carrying out research in a small number of leading industrialised countries.“ Patel (1997, 207) restricted it to multinationals with their headquarter in small countries: The patents of Swiss and British subsidiaries (at the U.S. patent office) comprise two fifth of their mother’s patents, while only 8 percent and 1 percent for U.S. and Japanese firms. In addition global generation of technology shows up in traditional industries rather than in high-tech and it appears to stem from the acquisition of research intensive firms rather than from deliberate decentralisation of R&D (Cantwell 1997, 230). This is due to the fact, that decentralised research and development in the foreign subsidiaries is less efficient than in the headquarter (by 10 percent according to Adam and Jaffe 1994).

In fact most countries have increased their specialisation in technological activities, small countries are forced to higher levels of specialisation than large ones, and countries with a strongly specialised activity in selected sectors tend to show better industrial and technological performance (Archibugi and Pianta 1992). Multinationals even contributed towards this specialisation as far as they utilise the comparative advantages offered by different National Innovation Systems.

In using the words “innovation” and “specialisation” one has to be extremely careful, however: Innovation is not restricted to the high-tech sector and specialisation concerns very specific products in most cases. The United States and Japan, which specialise to some extent on R&D-intensive industries, lost world-market shares in high-tech products to less advanced Southeast-Asia. This suggests that less-advanced countries can easily produce at least part of the so-called high-tech goods. The EU, specialising more on “mainstream” industries gained world-market shares in goods not normally considered to be optimally produced in advanced countries. Evidently these goods can be produced less easily elsewhere, probably because they afford as much innovation as high-tech goods (even if less research) and specialised skills (“industrial air”) additionally, not available in less-advanced countries. Specialisation in Europe frequently does not refer to branches or broad categories of goods, but is restricted to very specific products: On the NACE 2-digit level little specialisation among the EU-countries can be detected, while there is substantial specialisation at the 3-digit level (European Commission 1999, chapter 2). In the U.S. otherwise considerable “macro-specialisation” can be detected with the car industry in the Mid-West, textiles and apparel in the South, etc. Krugman (1991b, tables 3.2 to 3.4) and Decressin and Fatás (1995, 1629seq) clearly demonstrate the difference in the type of specialisation between European countries and comparable U.S.-regions. Insofar a very differentiated picture (of up to now) unexplained geographical abundance emerges, challenging economists and geographers as well. What differentiates countries or regions is, probably, not so much their methods of production in certain industries, but much more their relative strength’s or weaknesses in very specific types of products.

6.0. New Economic Geography proper: Economies of scale and transportation costs

---

23 The opposite pattern, however, has emerged in scientific research.
24 For a definition see Peneder 1998.
The offspring of New Economic Geography was the combination of economies to scale with transport costs. The most important results showed up already in the very first paper: the likelihood of multiple equilibria and potential shifts from periphery to core and back again if transport costs are lowered in several steps. In his seminal paper Krugman (1991a) assumed two regions and two sectors: agriculture with constant returns (perfect competition) and geographically fixed factors, and manufactures with increasing returns, produced by mobile factors. The economies to scale are purely internal to the firm. Even in absence of externalities, scale economies, transport costs and factor mobility, manufacturers tend to clump together because of forward and backward linkages; they want to concentrate production near markets (demand linkage). An opposite centrifugal tendency results from producers desire to move away from competitors. As a result multiple equilibria arise: With high transportation costs economic activity is evenly divided among the two regions; low transport costs allow three equilibria: even distribution or concentration in either one of the regions, whichever gets a head start.

Later, more sophisticated models allow more complicated dynamics: Krugman and Venables (1995) assume two (almost) identical countries, each with a perfectly competitive agricultural and a monopolistically competitive manufacturing sector, producing consumption goods and intermediaries (input-output linkage). Labour is immobile. At high transport costs both regions are essentially self-sufficient. Gradually declining transport costs lead first to a concentration of production in one of the countries, which for any (stochastic?) reason has a slight initial advantage of a larger manufacturing sector or more intermediate goods. Scale advantages dominate. A core forms and the periphery suffers a decline in real income. A further fall in transport costs and increasing wages in the core overcompensate scale economies so that “the world” tends towards convergence. The innovative point is the U-pattern of adjustment, the – temporary – polarisation even without cumulative causation.

The result is in startling contrast to the neoclassical model which implies rising costs of agglomeration. The new models reveal, that demand linkages, input-output linkages or technological (supply) linkages (limited geographical reach of spillover) can overcompensate the neoclassical mechanism. Whether this implies a friendly or an unfriendly story, depends on further details: If the periphery learns, as Germany did in the 19th and Southeast Asia in the 20th century, convergence may actually occur. If, however, elements of cumulative causation are added – accumulation of knowledge, advantages of richly specialised inputs, large market size, accumulation of capital, etc.) permanent polarisation will turn out as a more likely result. If transportation costs are – correctly – defined as a very wide concept, including all hindrances to trade and direct investment (market access, legal system, financial system, regulations, uncertainty, etc), the models’ results conform to the theory of convergence clubs: The rich industrialised countries and/or regions converge; due to low interaction costs they find themselves on the upward-sloping part of the U. The poor countries and regions, however, rather diverge, as they are on the downward-sloping branch of the U, due to high interaction costs. The authors themselves, however, warn against over-interpreting the results, as the model is still rather simple, e.g. completely neglecting capital or the simple geography of assuming two countries only.

Similar considerations turned out seminal in models of localisation theory as well, in which labour is usually assumed as mobile. Christaller (1933) and Lösch (1940) had already used economies of scale to explain how central places develop; Pred (1966) emphasised the contribution of initial advantages. The New Theory could base on these assumptions, but progressed from partial to general equilibrium models. Similar results show up as before: Multiple equilibria are frequent, polarisation quite likely, and the results frequently depend on initial conditions and the character of the disturbance: A flat distribution of industrial locations for instance, if randomly disturbed, can easily change into a structure in which all manufacturing is concentrated in only two regions (Krugman 1998, 169). Cumulative processes in a general equilibrium model of a circular economy with 12 locations may form one, two or three cities depending on parameter values (Krugman 1993). In an economy with

25 Transportation costs in this theory is a synonym for all types of impediments to trade.
multiple industries and scale economies circular causality through forward (real income) and backward linkages (specialisation) develops a system of central places similar to the static Christaller-model (Fujita et al 1997 according to Krugman 1998). Transportation infrastructure chanellises space, as firms choose a location at specific points, which may cause agglomeration or decentralisation depending on transportation cost (Thisse 1993).

In evaluating the results one must not forget, that the New Economic Geography is less than a decade old, and that the models quickly get mathematically very complicated, not to be solved by analytical methods. Only rather simple combinations of assumptions therefore have been modelled up to now. On a less formal basis it is obvious, however, that a combination of the new impulses of New Growth Theory, Innovation Theory and Economic Geography will provide new insights not only in the spatial dimensions of economic processes but in their intrinsic character as well. This is especially true for the processes of regional concentration and clustering. *Regional concentration* results from the combination of economies to scale and transport costs with various elements of circular causation as intermediate-goods industries following final-goods industries, thus creating an input- and skill-infrastructure liable to attract further final-goods producers, growing market size attracting even more producers, and so on. Rising wages normally don’t stop agglomeration as they expand the market; rather they trigger off a process of concentration on high-value products. The European experience of the sixties suggests that the periphery more likely gains from full employment of labour in the centre, enforcing massive outsourcing, than from large wage differentials.

The interaction of elements proposed by Economic Geography and innovation theory can explain why *clusters* form and how they do so. Porter (1990) demonstrated for Germany, Japan, Switzerland and the United States that the industries with the highest degree of international competitiveness are regionally strongly clustered. Spatial nearness enables them to internalise externalities. This has to be combined with product-cycle considerations: Up to the middle phase of the product-cycle these externalities combine to fast growth, agglomeration and concentration on those very products, benefiting most from externalities. Other products are driven out of the region, and the “shadow” of the agglomeration (Arthur 1990, 247) prevents other clusters to form in the neighbourhood. In the late phase of the product cycle, however, when the growth of the dominating product fades away, the cluster and the region are likely to age and it is not easy to rejuvenate them (Tichy 1998a). A combination of these elements of new theories can thus tell a story of the rise and fall of regions – if probably not of nations – which neoclassical theory cannot tell.

Quah (1996b, 3f) pointed towards the importance of regional linkages: “By contrast, in the empirics below, I show that a regions’ economic wellbeing can be predicted by that in both surrounding regions and the host state. Both factors matter for regional distribution dynamics. Thus, by either consideration, European regions are not isolated islands. More than does the host country’s economic performance, that of surrounding regions helps us understand the inequalities across European regions. Thus for explaining regional distribution dynamics, physical location and geographical spillovers matter more than do national, macro factors.” Even if Quah’s analysis is not without contra-evidence – he studied regional per-capita income while industry-studies find inter-country differences still larger than inter-industry ones – it points clearly towards the increasing importance of geographical elements even in standard economic analysis.

### 6.0.0.0.0.0.0.0.0 Elements of a theory of endogenous factor migration

The neoclassical mainstream model rested on rather simple assumptions about possible and actual mobility of factors of production: Labour was assumed completely mobile within a country, reacting sensitively to wage differentials, but completely immobile among countries. Capital was assumed to move freely within countries as well as in-between them, equalising interest rates. Brain capital, i.e.
technological knowhow was not dealt with in the traditional model and assumed as a fully mobile public good lately. All these assumptions were considered as more and more unsatisfying by most modern students, at odds with reality and barricading the way towards an explanation of the most pressing problems.

Labour was at no times immobile in economic reality. International migration of labour is not only a problem of our days, it has always been a problem, from the earliest days, when the Jews started to find the Promised Land. In the 1980’s about 100 million people were living outside their countries of birth or citizenship, 45 million in Europe and North America, approximately 2 percent and 7 percent of the respective population (Pries 1997a, 15). In the second half of the 19th century the absolute number of migrants was off course much smaller, but probably not their share in population (Castles and Miller 1997). Labour, therefore, is not at all immobile between countries (and much less then fully mobile among regions), and more than solitary exogenous factors are responsible for international migration. Econometric estimates suggest income differences as only one of several determinants of today’s migration. The probably most important causes are habit (significant lagged dependent variable and country-specific coefficients in the equations), labour demand in the receiving country26 and the level of income in the sending country in an inverted U-pattern: If income rises from a very low basis, emigration increases due to better information and less binding financial restraints, but falls back with increasing wealth (Faini and Venturini 1993, 441).

As to the mobility of capital one has to distinguish financial capital from physical one: Financial capital appears to be internationally highly mobile indeed and following interest-rate differentials, even if even the most sophisticated empirical tests of interest arbitrage leave unexplained residuals, indicating omitted explanatory variables. International mobility of physical capital on the other hand cannot be easily explained by interest rate differentials, nor is it true that it moves from (capital-)rich to poor countries: 76 percent of world foreign direct investment during 1988/93 flowed to developed countries, of which 43 percent to EU-countries and 21 percent to the U.S.A; more than half (58 percent) of direct investment flowing to EU-countries had its origin in other EU-countries, with the richer countries as the main recipients. A rather similar pattern can be found for mergers. These facts suggest that the movements of physical capital follow a much more complicated pattern, similar to the one suggested by modern trade theory and the theory of the international firm with specialisation, scale economies, utilisation of local skills, etc. as the dominant explanations. In general, however, one should not overemphasise the mobility of capital, so much publicised by the financial press: Feldstein and Horioka (1980) found that national saving is the most important determinant of national investment, and this has not changed in the meantime, notwithstanding the booming literature on efficient capital markets (Gordon and Bovenberg 1996).

Technical knowledge, finally, is a complicated factor of production, as has been demonstrated in the section on innovation theory. Far from being a public good at free disposal to anybody in need of it, it evidently has national characteristics and implies externalities. Total factor productivity depends not only on domestic R&D but also on foreign one (Coe and Helpman 1995), especially in small countries. Technology is transmitted by international trade in intermediate and investment goods, but definitely needs receptivity on the side of the receiving country, that is an adequate level of education, technical knowledge and at least some own R&D. The theoretical concept of pure imitation has no counterpart in the real world: Any imitation implies at least some minor innovation (Tichy 1991).

As to the (re-)integration of geographical elements into mainstream economic theory, the elements of an endogenous theory of factor mobility and migration appear to be of utmost importance. They demonstrate that national borders are not negligible but of varying importance, depending on the subject under investigation. In some examples a regions’ economic wellbeing can be better predicted by that in her surrounding regions than by that the host state (Quah 1996b, 3f), in other ones the

26 Freeman (1993, 448) emphasises “receptivity to immigration” in general as the missing element in migration analysis.
National Innovation System turns out to be of utmost importance. Political-administrative space, economic space, social space (Pries 1997a, 16ff) and geographic space appear to interact in a kaleidoscopic fashion, as European integration and the (overemphasised27) globalisation strongly bring to our minds. A broad field of problems of apparently geographical character arises in economics.

The problem of the political-administrative, economic and social space falling apart extends far beyond geography and economic to geography and political science: The transfer of sovereignty to the EU, implying concentration of power, and the secession of states earlier belonging to the Soviet Empire, Yugoslavia or African states formed by the colonial powers, brought about research investigating into the causes of size and number of nations. A trade-off was found between the economies of scale of large jurisdictions and the costs of heterogeneity of large and diverse populations. Democratisation and economic integration allow secession at lower political and economic costs28 and increase the number of independent countries, if no appropriate redistribution systems are provided (Alesina and Spolaore 1995). EU’s member countries do not show much awareness of this important result.

6 Summary

This paper tried to demonstrate that geographic factors in general are very important for economic development; by chance – or bad luck – this was not the case in the late 18th century when economics formed as a science. This congenital defect characterised mainstream economic theory for about 200 years. Sidelines only – German economics in the second half of the 19th century, or unavoidably location theory and regional economics – considered geographic factors, but never found their way into the top-journals or best-selling text books. The concentration of mainstream economics on an abstract, a-structural and a-geographical model contributed towards the uprise of economics as the most exact and most “scientific” among the social sciences, but left many problems and developments of the modern world unexplained.

In the late seventies the mathematical tools were available to deal with models, better suited to explain the economic problems of our days: Models based on increasing economies to scale, imperfect competition or externalities, partly internal, partly external to the firm. Contrary to the unequivocal convergence path of the a-structural neoclassical model, the new class produced a rich diversity of results, depending on a host of factors – initial conditions, structure, size of scale economies and transport costs, size of coefficients, but to some extent stochastic elements – chance – as well. The national and regional diversity gives room for geographic considerations, determining the specific outcome within the general economic framework. In addition geographical aspects gain importance as to increasing number of states – due to secessions –, and the growing importance of regions – caused by dwindling importance of national economic policy as a consequence of institutional and factual integration.

Articles including geographical aspects in top-economic journals 1994-1997

<table>
<thead>
<tr>
<th>Journal</th>
<th>“geographical” articles</th>
<th>all articles</th>
<th>Percent geographical</th>
</tr>
</thead>
<tbody>
<tr>
<td>American Economic Review</td>
<td>19</td>
<td>674</td>
<td>2.8</td>
</tr>
</tbody>
</table>

27 See Tichy 1998b.
28 More interactions with surrounding foreign regions, less dependency on the host country.
All this should imply a growing number of articles dealing with geographic aspects. This however is not the case. Krugman (1998, 172) deplores that the seminal theoretical articles have led to little validating empirical work, and in the top-economic journals not much work on geographical aspects can be detected anyhow, as the table clearly demonstrates. As before articles dealing with spatial aspects appear in the specialised journals on regional economics or urban planning. Only 3 percent of the articles in four leading journals dealt with geographic themes. This is not very much, even conceding that the field still is in its early teens. So one has to conclude: Economics has found a methodology apt to find geography and it has produced a considerable number of seminal articles; but it has not yet found geography in actual work.

<table>
<thead>
<tr>
<th>Journal</th>
<th>Articles</th>
<th>Pages</th>
<th>Authors</th>
</tr>
</thead>
<tbody>
<tr>
<td>Economic Journal</td>
<td>6</td>
<td>310</td>
<td>1.9</td>
</tr>
<tr>
<td>European Economic Review</td>
<td>20</td>
<td>521</td>
<td>3.8</td>
</tr>
<tr>
<td>Zeitschrift für Wirtschafts- und Sozialwissenschaft</td>
<td>4</td>
<td>86</td>
<td>4.7</td>
</tr>
<tr>
<td>Insgesamt</td>
<td>49</td>
<td>1591</td>
<td>3.1</td>
</tr>
</tbody>
</table>

29 One must not forget, however, that reality changed considerably: Anecdotal evidence suggests, that firms still cannot evaluate the costs of producing in peripheral countries, locating and relocating factories as they adjust and re-adjust their estimates of relative cost-advantages. This makes it fairly impossible to collect data as an input for empirical models.
Literature


Adam, J.D., Jaffe, A.B., 1994, The span of the effect of R&D in the firm and industry, Mimeo.


Beinsen, L., Kurz, H.D., Hg, 1998, Ökonomie und Common Sense, Graz: Leykam.


Christaller, W., 1933, Central places in Southern Germany, Englewood Cliffs, NJ: Prentice Hall.


Freeman, Ch., 1997, The ‘national system of innovation’ in historical perspective, in Archibugi and Michie 1997a, 24-49.


Kurz, H.D., Hg, 1990b, Adam Smith (1723-1790). Ein Werk und seine Wirkungsgeschichte, Marburg: Metropolis.
List, F., 1841, Das nationale System der politischen Ökonomie,
Lösch, A., 1940, Die räumliche Ordnung der Wirtschaft, Jena: Fischer.
Pries, L., 1997a, Neue Migration im internationalen Raum, in Pries 1997b, 15-44.
Pries, L., Hg, 1997b, Transnationale Migration, Soziale Welt Sonderband 12, Baden-Baden: Nomos.


