LN-14 From pin factory to endogenous growth II. Tracing the roots of endogenous growth theory.

Introduction

We started in the previous lecture with Paul Romer's article *Endogenous technological change* (JPE 1990) which established the concept of endogenous growth theory as different from the growth theory of Solow 1956 with its exogenous, and accordingly unexplained, representation of technological growth. Endogenous growth theory claims to explain long-run growth as emanating from economic activities that create new technological knowledge. After Romer's 1990 article there has been hectic research activity at many research centres towards a more complete and powerful growth theory.

Our focus was, however, backwards in time following Romer's search for earlier ideas in the history of economics starting with Adam Smith's famous passage on the increasing returns in the pin factory as an allegory of the mechanisms which have allowed growth in income per capita to increase over the last two centuries.

From Adam Smith we traced what Arrow later called the underground river of ideas through 19C from Ricardo and Malthus via Marx and Mill to the marginalists and Marshall. In 20C we picked up important tributaries in articles by Allyn Young and Frank Ramsey (even in the same issue of Economic Journal 1928). The emergence of Keynesian macroeconomics which renewed and invigorated economics in several ways did not contribute in this regard.

Contemporary with Keynes was John Hicks who offered a more modern version of general equilibrium in his *Value and Capital*. Hicks was, perhaps, not really a Keynesian but wrote the most influential paper popularizing the Keynesian message, largely through Hicks' skilful use of Marshallian type diagrams. The equilibrium analysis in *Value and capital* did not, however, offer any openings for increasing returns.

Another idea that would play a most important role in Romer's attempt to explain growth was 'monopolistic competition'. The term was coined by the Harvard economist Edward Chamberlin (1899-1967) who published *The Theory of Monopolistic Competition* in 1933. Although it seemed paradoxical to some to combine the idea of monopoly and competition it made perfect sense to

Chamberlin who could watch at close range the rise of a number of big brand names competing in the market place. The importance of brand name in the market place had in fact been noted already by Allyn Young as early as 1908. A commodity, whether soap or oysters, with a brand name was different from just soap or oysters. The brand name or trademark was something in addition, and that addition had a monopolistic character as it was the property of the seller, no one else could use it. Veblen had also touched upon this idea (as reflected in the later works of his most influential follower in USA, John Kenneth Galbraith).

Producers with a trademark would be observed to have 'selling expenses', which as Young expressed it, were "incurred, not in producing things people want, but in inducing people to want the particular things the entrepreneur has for sale". Chamberlin became convinced that many producers were in the situation that they had a monopolistic edge because their product dominated the market or shared dominance with a small number of rivals with whom they often could collude. Such producers were not forced to sell at a price that just covered the marginal cost but could choose a combination of quantity and profit that would seem to be more profitable. Increasing returns, as in railways, provided a major reason for monopolistic situations to occur and were called natural monopolies, often also allowing the possibility of charging different prices to different customers and thereby enhancing the monopolistic advantage. But how different was the pin factory from railways? The idea of the *monopolistic competition* had come to stay.

Chamberlin's book was published just after Joan Robinson's *The Economics of Imperfect Competition* appeared in England pursuing similar ideas. Robinson's argument was more Marshallian, while Chamberlin was more concerned with the product itself as the source of monopolistic power. Robinson coined *monopsony*. Chamberlin argued for the necessity of some degree of monopoly for the businesses to cover fixed costs. Allyn Young's two most important students, Frank Knight in USA and Nicholas Kaldor in England both got involved in following up Young's ideas.

The end of the 1930s was the time for USA taking over the dominance of economics. At about the same time signs were visible for indicating that economics at the highest level in the future would definitely be a mathematical discipline. Both Chamberlin and Robinson argued in the traditional literary style. The new times was marked by Paul Samuelson (1915-2009) who was both a Keynesian and an innovator and enhancer of the entire range of established theory, rewriting the core of it in a way which set a high standard for precise

mathematical expression in his *Foundations of Economic Analysis* (1947), based on his doctoral dissertation from 1941.

Samuelson held Chamberlin in very high regard. In the second edition of *Foundations* published in 1983, Samuelson wrote a new introduction in which he took note of the hollowing out that had occurred after *Foundations* appeared and stated: “More can be less. Much of the mathematical economics in the 1950s gained in elegance over poor old … Edward Chamberlin. But the fine garments sometimes achieved fit only by chopping off some real arms and legs.” The next wave of mathematical techniques had produced remarkable advances, “but they seduced economists away from the phenomena of increasing returns to scale and … technology that lie at the heart of oligopoly problems and many real-world maximizing assignments”. (This echoed criticism made in Presidential addresses by Frank Hahn /ES 1970/ and Wassily Leontief /AEA 1971/.)

Before we proceed to the post-WWII period there is one more influence we have to include, namely that of Joseph A. Schumpeter (1883-1950). Schumpeter had early focused on technical change and the role of the entrepreneur as the crucial elements in the explanation of growth. This message came first in *Theorie der wirtschaftlichen Entwicklung* in 1911, translated and published as *The Theory of Economic Development: An inquiry into profits, capital, credit, interest and the business cycle* in 1934, shortly after Schumpeter had moved permanently to USA. Schumpeter’s theory was closely related to business cycles and also to his view on how technologies replaced each other through what he called ‘creative destruction’ (‘schöpferische Zerstörung’).

Schumpeter had a strong belief in the use of mathematics in economics but he did not describe his system mathematically. Neither did he write anything explicitly about increasing returns. Schumpeter’s theory emphasized, like that of Marx, change over time. He called his theory ‘dynamic’ (=emphasizing change) as different from static (=emphasizing equilibrium). Schumpeter (and Marx) shared also with Marshall the focus on change over time but did not draw on the Marshallian devices of ‘neighbourhood effects’ and ‘spillovers’. Schumpeter’s ideas were also mediated in his grand two volume *Business Cycles: A theoretical, historical and statistical analysis of the capitalist process* (published just at the outbreak of WWII in 1939) and *Capitalism, Socialism and Democracy* (1942).

It was the latter book that introduced the term ‘creative destruction’ (it has a chapter called *Creative destruction*) which is a classic in the tradition of the

underground river. The book described the innovative entry by entrepreneurs as the force that sustained long-term economic growth, even as it destroyed the value of established companies that enjoyed some degree of monopoly power. Because of the significant barriers to entry that monopolies enjoyed, new entrants would have to be radically different: ensuring fundamental improvement was achieved, not a mere difference of packaging. The threat of market entry would keep monopolists and oligopolists disciplined and competitive, ensuring they invest their profits in new products and ideas. It was this innovative quality that made capitalism the best economic system in Schumpeter's view.

The 1942 book reiterated the themes of the 1912 book about the key mechanisms by which economic growth takes place: the appearance of new goods, new markets, new methods of production and transportation, new forms of industrial organization, usually in clusters and usually as sudden bursts of activity after periods of comparative calm. Schumpeter wrote in 1912 on the background of the literature setting out the new doctrines of marginalism, ignoring completely the topic of the growth of knowledge. Schumpeter harboured no fear in his 1942 books of the earth's resources running out: "It is one of the safest predictions that in the calculable future we shall live in an *embarras de richesse* of both foodstuffs and raw material, giving all the rein to expansion of total output that we shall know what to do with."

The modernizing of economics sidelined Schumpeter. His works ended up being read fairly widely but not by the up-and-coming within the profession. It didn't help Schumpeter's reputation much that he rejected the Keynesian approach more or less completely. Schumpeter retired into the history of economics.

World War II had a considerable impact in USA not only on the economy but also on economics. A large number of economists, including several future Nobel Laureates, took part in the war effort in one way or another. The war resulted directly and indirectly in enormous technical advances, laying the basis for post-war growth. The traditional way of doing economics went out of fashion. New ideas and tools for economics came out of the war, such as game theory, linear programming, macroeconometric models. The leading new practitioners such as Samuelson, Arrow, Friedman, positioned themselves at leading institutions. initiating a pattern still recognizable. The Cowles Commission, which had existed since 1932, came to play a key position in new developments in Chicago from 1943 and until it moved to Yale in 1955. But

issues of concern in our context such as pin factories and the significance of increasing returns disappeared underground.

The postwar years turned into a sustained boom, unlike the tumultuous years after WWI. If the explanation of the boom was a puzzle by itself, few worried about it. Companies were bigger, competition perhaps less intense; the growth was higher than ever. Most economists agreed that Keynes had shown the way to both stability and prosperity.

Roy Harrod (1900-1978) who was close to Keynes discussed just before WWII whether the Keynesian stability formula could be recast as a formula for stable growth. The Polish born Evsey Domar (1914-1997) did something similar immediately after WWII. They reasoning was a little different but both reached the same formula, namely that steady growth was possible with the growth rate $g = s/v$, where s is the savings rate and v is the (marginal) capital-output ratio. Their efforts would later be grouped together and denoted the Harrod–Domar growth model. An uncomfortable feature of the Harrod-Domar model was the knife edge character of the solution.

Then Robert Solow (1924-) entered the scene. Solow began his studies at Harvard College in 1940, volunteered for the army two years a later and came back to Harvard after three years army service and completed his education with a PhD at Harvard a few years later. Solow had worked on linear models, linear programming and related issues when he decided in the early fifties to reconsider the problem of stable growth. He was then well familiar with the Ramsey model and knew about Harrod-Domar. Solow's philosophy was to make a model as simple and transparent as possible without more fancy mathematics than necessary and his instinct as a well trained economist told him that substitution was a crucial element. The result was the Solow model, *A Contribution to the Theory of Economic Growth* (QJE 1956), with its production function $Y = A(t)\,F(K,L)$ and constant savings rate, which all economists are familiar with, including the graphical illustrations provided by Solow. $A(t)$ was assumed to grow steadily with the passage of time. Solow was well read both in economic history and knew John Stuart Mill's exposition of diminishing returns, which he didn't disagree with.

The backdrop for Solow's article was a research revolution going on. The first satellite was launched just as the paper was published. The surprising implication of the Solow model was that the savings rate didn't really matter for the growth rate. Capital deepening would only have a transitory effect on the growth rate. A nation couldn't save itself into a higher rate of growth. Only population growth, which also was exogenous, and the rate of technological

change could do that. Solow followed up in 1957 with the article *Technical Change and the Aggregate Production Function*, in which he showed that the capital growth which was explained in the model, explained only on eighth of the growth from 1909 to 1949. The rest was due to the population growth and the technical change. By this Solow drew a lot of attention to the *residual*.

"Technical progress" as measured by the residual was the creator of wealth, while labour and capital accumulation was of lesser importance. Students of Smith, Mill, Marshall and Schumpeter knew well enough that technical change was real, important, and quite different from conventional inputs. Solow's model made the point in modern mathematical language. And although the source of increasing returns lay outside the model, Solow had succeeded in setting it up such that the contribution of knowledge could be measured.

For some, Solow's results about the role of the residual proved that Schumpeter had been right after all. But there were other reactions to the residual. The residual was "the measure of our ignorance" said the Stanford economist Moses Abramovitz. "Naming is not explaining", said the Chicago econometrician Zvi Griliches. The econometricians' approach, as represented foremost by Griliches and Dale Jorgenson in the mid-1960s, was to "endogenize" technical change, i.e. to explain it in strictly economic terms and thus make the residual disappear. Whatever was left was soon given a new name: *Total Factor Productivity* or *TFP*.

Kenneth Arrow pursued almost any important problem that came up. After his pioneering dissertation work on social choice, his fundamental contribution to general equilibrium theory in the 1950s, and work on asymmetric information in markets, Arrow had been a frequent visitor to the RAND Corporation, a large research institution in California set up after WWII by the US Air Force. Sometimes called a "university without students", RAND managed to gather a considerable number of the "best and the brightest" working on problems of defense.

At Rand Arrow started to think about the characteristics of the production of knowledge around 1960. The production of knowledge was intrinsically uncertain. Knowledge was not "appropriable", meaning that the person who created and paid for it couldn't necessarily expect to benefit exclusively from it. And new knowledge was "indivisible", meaning that it entailed a certain fixed cost before its benefit could be enjoyed. The 'indivisible' term was in used in related meanings. It could mean that benefits were freely available for all such as police protection, the signal from a lighthouse or a radio broadcast. Or that something was "lumpy", you couldn't buy just a little bit, you have to pay for it

all. You couldn't buy half a piece of new knowledge any more than you could buy half a bridge. Then once you had it you could use it again and again. It was this quality of indivisibility, Arrow said, which meant that investment in knowledge "obeys the law of increasing return".

Arrow proceeded to build a model of accumulation of knowledge through experience. Arrow distinguished between 'practice' (which took time), 'research' (which required financial resources) and 'experience' (which was simply a side effect of production undertaken for its own sake. Arrow denoted this form of knowledge accumulation *learning-by-doing*. This was at the outset something very close to Marshall's externalities of spillovers. Arrow relied on empirical evidence that learning-by-doing seemed to be the way the world worked. In several kinds of activities such a building airplanes and ships it had been observed that productivity increased steadily for 15 years without additional investments. It was known as the Horndal effect, after a Swedish steel mill.

To this idea Arrow added rational expectations, implying that manufacturers would be successful in finding out what their neighbours already knew. Just as in Marshall's system the external increasing returns made something for everyone. Spillovers would increase with scale as the industry grew. They ensured that no firm would use its own learning to try to build a monopoly. Arrow presented his model in 1962 and thus increasing returns was finally made respectable by modern formalization. After being received with enthusiasm Arrow's model nevertheless faded as it was found that the process was not stable, a small change could throw it off track. It didn't become part of the economists' tool kit; instead it became part of the underground river of thinking about externalities and increasing returns.

Arrow's succinct summary of the characteristics of knowledge accumulation remained intact, namely that knowledge was inappropriable, that is was indivisible (meaning that it generated increasing returns) and that its production was intrinsically uncertain.

In the 1960s optimal growth became a very much discussed issue, partly due to the lead Arrow had introduced. Two of his best students, Karl Shell and Hirofumi Uzawa, gave important contributions and among the younger students involved were George Akerlof, Joe Stiglitz, William Nordhaus and others. Control theory was taken into use, almost as soon as it had been developed by Pontryagin. It was about optimal growth but with links to policy, how could growth be speeded up through policy. New concepts, such as 'turnpikes' and 'golden rules', flourished. The 1960s, which gave us "Easy

Rider" and "The Wild Bunch" (both in 1969), also marked the longest sustained business expansion ever recorded in Europe, Japan and USA.

Then in the 1970s came the controversies and also new tools. Dynamic programming, developed by Richard Bellman already in the 1950s, was taken into use by Robert Lucas. Lucas and Prescott published *Investment under Uncertainty*. The economic turmoil of the 1970s broke up the broad support for the *neoclassical synthesis*, and gave birth to *Neo-Keynesian economics* (later to be followed by *New Keynesian economics*), the *new classical macroeconomics*, and even other schools. The 1970s was a decade of worries for economists – about inflation, unemployment, resource scarcity and the productivity slowdown. Enough to make anyone *Dazed and Confused*, as the title of a Led Zeppelin song from that period.

The 1970s also had the worry about Japan whose trade record was unsurpassed and took over the dominance of one product after another, assertedly using a protected home market as practice field, before conquering the global market. Paul Krugman (1953-) was graduate student at MIT in the late 1970s thinking about the Japanese strategy as he studied international trade theory, only to get the impression that in trade theory all the interesting work had already been done. Krugman reflected on the Japanese export success but also on the surprising case of Sweden having become a car exporter and the fact that USA and Germany exported airplanes and cars to each other. The latter phenomenon would soon be known as the "puzzle of intra-industry trade".

A bunch of other young economists were at the same time working on related questions. In the 1970 a new subfield known as *Industrial Organization* had emerged with the involvement of the signaling and screening models of George Akerlof (1940-), Michael Spence (1943-), Joe Stiglitz (1943-) and others. Another young bunch comprising David Kreps, Paul Milgrom, John Roberts and Robert Wilson worked on other related problems following the leads of the first generation of game theorists.

Thus questions about how single firms might come to dominate their markets were being asked in increasingly formalized ways and with some success. The approaches had many aspects but at the centre of it all were problems of increasing returns. The new models showed how firms could earn increasing returns by expanding the variety of their products, burnishing their brand name or using various means of blocking their competition. The problem of the pin factory was about to come to life. Krugman also studied the Dixit and Stiglitz article *Monopolistic Competition and Optimum Product Diversity* (AER, 1977),

before he pursued the idea of trying to make a monopolistically competitive trade model.

Krugman worked out a model, it appeared as *Increasing returns, monopolistic competition, and international trade* (J. of Int. Econ. 1979) and it received a lot of attention. With the new models Krugman could show how increasing returns and general equilibrium could coexist. If one country got a head start in mass production it might keep it. Geography didn't need to have anything to do with it, it could be that whatever country got a good start, could keep the lead. But the model had loose ends and also inconvenient implications such as the possibility of multiple equilibria, an almost subversive result.

Monopolistic competition swept through trade theory in the 1980s. It seemed to fit well the facts of global trade, which came to be seen increasingly as consisting of two tiers. Tier one was trade in commodities and services in terms of perfect competition and comparative advantage, while tier two was monopolistic competition with great multinationals helped by government subsidies mounted assaults on each other's markets and specialization was determined by market size. In 1985 Krugman gave his new ideas broader scope when he published with Elhanan Helpman as co-author *Market Structure and Foreign Trade: Increasing Returns, Imperfect Competition, and the International Economy*.

We go back to Paul Romer (1955-) again. He had studied mathematics at undergraduate level, then became a graduate student at MIT. He took a year off away from studies but when he was ready to go back for completing the PhD studies he surprisingly shifted to Chicago. His idea was to build a new model of economic growth which would encompass an element of falling costs, explained somehow by the growth of knowledge. Chicago's department of economics was divided once again, it was the old and more literary school with Milton Friedman, Arnold Harberger, Theodore Schultz, George Stigler and others who were about to surrender the scene to younger economists who would become known as *freshwater* macro guys, some of them with cutting edge skills in mathematics. There were also a group of labour economists, comprising Gary Becker, Jim Heckman, Sherwin Rosen, and others. It was Rosen who suggested to Romer to look at Allyn Young's 1928 paper.

Romer wanted a model in which growth could continue indefinitely, unlike the Solow model. Romer's problem was not least to figure how to model the accumulation of knowledge. The general idea was to have increasing returns to knowledge and decreasing returns to land, labour and capital. But the problem was to avoid that a single firm could take advantage of increasing returns to

knowledge to monopolize its markets and destroy the assumption of competition. He followed a suggestion from Lucas and his first model came close to the Marshallian idea that new technology could not be appropriated, "trade knowledge that cannot be kept secret", although Romer hadn't read Marshall yet. But then he came up against the problem of multiple equilibria, which was essentially what Krugman had struggled with and eventually solved it, but Romer didn't know about that either. Finally, he also had to struggle to avoid the "knife-edge problem" that Arrow had met with and also others. Romer's PhD dissertation was completed in 1981 titled *Dynamic Competitive Equilibria with Externalities, Increasing Returns and Unbounded Growth*.

We are back were we started – almost. It was after the dissertation that Paul Romer reviewed the old literature more systematically. It led him surprisingly to rethink the mechanism of his model. In the end he decided to recast it as a model of specialization with no spillovers. Instead of a given set of goods, as Krugman had in his model, there would be an increasing set of goods. As Krugman, also Romer got a lot out of the ideas inherent in the Dixit-Stiglitz model of 1977. Romer wrote out a small model describing an economy turning out a steady succession of new goods. None of the goods would be a perfect substitute for any other. There would be pins, sure, but also all other kinds of fasteners that could be designed and produced. Every firm would act as a small monopolist, setting the prices of its fasteners well above cost, hoping there would be a stream of profits sufficient to pay for the design and enter the business in a big way.

This model would later be called neo-Schumpeterian, one that depended on the introduction of new goods to make it work. Romer's model was far from perfect. It had creation but no destruction. The old goods never disappeared. The measure of a country was the size of its population. Implying that a big nation like China should grow much faster than a small closed economy. But the bottom line was that specialization, meaning new goods, and the increasing returns that came with them was the key to rising output. This was a big step away from what had been asserted in the dissertation.

This new model was about to come into final form. It had to be presented in the key meeting for discussion and review. In the meantime Romer had to pick up on a so far neglected duty, namely to squeeze publishable articles out of the dissertation. He published two papers from it, *Cake Eating, Chattering and Jumps: Existence Results for Variational Problems* (Econometrica 1986) and then, with some difficulty, the more essential summary of his dissertation work as *Increasing Returns and Long Run Growth* (JPE 1986), embodying the ideas he

had by then surrendered. Romer left a clue in the article, where he wrote: “Formally, increased specialization opens new markets and introduces new goods. All producers in the industry may benefit from the introduction of these new goods, but they are goods, not technological externalities.”

Four years later Romer published a paper based on the new idea he had conceived in 1986 as *Endogenous Technological Change* (JPE 1990). And this is where we started. There is a story, surely, to be told about the active exchange and controversies among economists both in the 1986-90 interval and after the publication of the 1990 paper. This exchange is still going on and the controversies too.