Cantner, Uwe; Vannuccini, Simone

Working Paper
A new view of general purpose technologies

Jena Economic Research Papers, No. 2012,054

Provided in Cooperation with:
Max Planck Institute of Economics

Suggested Citation: Cantner, Uwe; Vannuccini, Simone (2012): A new view of general purpose technologies, Jena Economic Research Papers, No. 2012,054, Univ. u.a., Jena

This Version is available at:
http://hdl.handle.net/10419/70135

Terms of use:
Documents in EconStor may be saved and copied for your personal and scholarly purposes.
You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licenses), you may exercise further usage rights as specified in the indicated licence.
A New View of General Purpose Technologies

by

Uwe Cantner
Simone Vannuccini

www.jenecon.de

ISSN 1864-7057

The JENA ECONOMIC RESEARCH PAPERS is a joint publication of the Friedrich Schiller University and the Max Planck Institute of Economics, Jena, Germany. For editorial correspondence please contact markus.pasche@uni-jena.de.

Impressum:

Friedrich Schiller University Jena  Max Planck Institute of Economics
Carl-Zeiss-Str. 3  Kahlaische Str. 10
D-07743 Jena  D-07745 Jena
www.uni-jena.de  www.econ.mpg.de

© by the author.
A New View of General Purpose Technologies

by Uwe Cantner and Simone Vannuccini*

Abstract

The economic literature started to recognize the heterogeneity characterizing the nature of different technologies, introducing the concept of General Purpose Technologies. In this paper, we offer a “new view of General Purpose Technologies”, building on the historical as well as on the recent literature, enquiring more in deep the definitional problems related to the GPTs and the conditions for their emergence, together with the characteristic for their prevalence and pervasiveness. A Schumpeterian and evolutionary view pointing at the micro and meso level of analysis – that of the dynamics of firms and industries –, is in our view the privileged perspective economists need to adopt in order to revitalize the theoretical and empirical study of GPTs. The similarities with the emergence of dominant designs and the relations with dynamics of increasing returns and path dependency in the choice between alternative technologies offer us a set of tools well suited to study the establishment of GPTs as a process unfolding in time, more than as a single homogeneous shock.

JEL: E32, L16, O30, O33, O40

Keywords: General Purpose Technologies; Long Waves; Business Cycles; Dominant design; Pervasiveness of technologies; Neo-Schumpeterian economics.

1. Introduction: General Purpose Technologies

Technologies are not all alike. Some of them add incrementally to the economic and productive system; other technological innovations, instead, have a revolutionary impact: they impose on the economy a new structure of dependencies and complementarities (rejuvenation, in Carlota Perez’s (2004) terms) and exploit physical phenomena in new ways (Arthur, 2009). The economy continuously reconfigures itself and its working logic around such technologies, producing as a result an

* Uwe Cantner, Friedrich-Schiller-University Jena, Faculty of Economics and Business Administration, Carl Zeiss Str. 3, D-07743 Jena and University of Southern Denmark, Department of Marketing and Management, I2M Group, Campusvej 55, DK-5230 Odense M; Simone Vannuccini, Graduate College “The Economics of Innovative Change”, Bachstrasse 18k, D-07743 Jena. The present paper is published as a chapter in the Festschrift for Prof. Dr. Oppenländer (2012).

open-ended evolutionary process of change. Recently, the economic literature started to recognize the heterogeneity characterizing the nature of different technologies, introducing the concept of General Purpose Technologies (GPTs hereafter); the aim of this paper is to critically guiding the reader into the topic, offering also what we think is a novel perspective in this field of studies.

The recognition that major technological changes are the main determinant of cyclical and non-linear patterns in the evolution of an economy is not a monopoly and exclusive right of the literature on GPTs; conversely, the idea is at the core of the long-standing research and debates on the existence of Long Waves (Silverberg, 2003) and dates back to end of the Nineteenth century, when scholars started to abstract from the case-specific theories of economic crisis, generalizing formal models of trade cycles (in the English tradition), business cycles (in the United States) and Konjunktur (in Germany) (Besomi, 2010). If already Shakespeare, cited in Jevons’ Political Economy chapter (XIV) about the periodicity of Industry, wrote that “there is a tide in the affairs of men, Which, taken at the flood, leads on to fortune”, we can agree upon the constant relevance of the topic over time. The work of Kondratieff (Kondratieff and Stolper, 1935) and Schumpeter (1939) – especially the latter often misrepresented and re-invented hypothesis on clustering of innovations and creative destruction – are milestones in this sense, paving the way for a wide range of theoretical and empirical attempts to identify long-wave patterns in economic history.

From the brief perspective just outlined, the research on GPTs appears more as a contemporary endeavor to empower endogenous growth theory, with the analytical tools capable to explain economy-wide fluctuations, than a conceptual novelty. However, we find the GPT “instantiation” of the more general topic quite interesting and important for innovation scholars, since the focus of GPTs theories, instead of explaining the wave in itself or stressing the systemic consequence of techno-economic paradigm changes (Perez, 2004), is narrowed to the nature of technology and to its effect on productivity dynamics, capital accumulation and innovative activities, such as the return on R&D investments.

Before starting with the most diffuse definition of what GPTs are, however, it is worth recalling two more issues. The first helps us to frame the GPTs models in the literature: it is clear the similarity between the concept of general purpose technology with that of radical innovations, macro-inventions (Mokyr, 1990) or shifts in the technological paradigm (Dosi, 1982). Therefore, as theorists found a tricky challenge in trying to identify a clear-cut boundary between macro and micro or radical and incremental innovations (and the same problem holds in the innovation-studies literature distinguishing between process and product innovations), a similar shortcoming affects the appro-

---

2 For a comprehensive overview, see Silverberg (2003).
3 For the sake of clarity, in the paper we call endogenous growth theory the two sets of models, one inspired by the AK approach and Paul Romer’s contributions (Romer, 1990; 1986) the other collecting under the label of Schumpeterian growth models are the quality-ladder and R&D-based models such as Aghion and Howitt (1992), together with the quasi-endogenous literature started by Dinopoulos, Segerstrom and others (Dinopoulos and Sener, 2007).
4 Alternatives to GPT-based modeling of economic fluctuations and Long Waves are, for example, Jovanovic and Rob (1990), formal account of Schumpeterian cycles, technological opportunities, extensive and intensive search.
5 For a complete taxonomy see Coccia (2003, p. 11).
appropriate criteria for identifying the technologies that actually are GPTs. We argue more about that in the next pages, since this is a relevant point for our argument.

The second premise is methodological, and we stress it because it should be clear that GPT-based modeling brings together the analytical framework of neoclassical growth theory, that is linear in nature since – following Solow (1997) – a theory of growth should not explain short term fluctuations but only the long-term potential trajectory of an economy, and the cycling behavior of economic aggregates, traditionally a feature of heterodox growth theories (Setterfield, 2010). Such a “refinement” with respect to the simpler approach to technological change adopted by mainstream growth theory, where a generic (a scalar) stock of “ideas”, “knowledge” or technology interact with a production function, is not the only modeling strategy available to economic theorists. As Goodwin (1951) puts clearly in his dealing with the (non-linear) accelerator principle as the determinant of cycles, “[a]lmost without exception economists have entertained the hypothesis of linear structural relations as a basis for cycle theory […] whether we are dealing with difference or differential equations, so long as they are linear, they either explode or die away with the consequent disappearance of the cycle or the society. One may hope to avoid this unpleasant dilemma by choosing that case (as with the frictionless pendulum) just in between. Such a way out is helpful in the classroom, but it is nothing more than a mathematical abstraction […] Mention should also be made of the fact that there exists an alternative way out of the dilemma – that of an impulse-excited mechanism. There are two basically different classes of such mechanisms to be distinguished. (a) There are the synchronized systems of which the most familiar is the ordinary pendulum clock. […] The wider system […] is a particular type of nonlinear oscillator since it is autonomous and maintains a uniform cycle independently of initial conditions. (b) Significantly different is a system subject to random shocks. Here the mechanism itself is damped, but an outside, unexplained source keeps it going, and in this sense it is not a complete theory, for the source of maintenance lies outside the theory […]”

GPTs models fall under the point (b), since a new general purpose technology – also in advanced models inspired to Goodwinian Lotka-Volterra dynamics (Fatás Villanfranca et al., 2011) and even more in the “classical” modeling approach to GPTs (Helpman, 1998) – arrives from the outside of the system, both when it is modeled directly as an exogenous variable (in a deterministic or stochastic fashion, as we deepen further later) and when it is an indirect result of endogenous knowledge accumulation. GPTs are then just shocks revitalizing an economy characterized by the tendency to “relax” in a steady-state equilibrium growth.

As it will made clear in the paper, we argue that in addition to Goodwin’s choice (a), that of using non-linear systems, it is possible to frame a more complex point (c), where the emergence of a GPT is the result of localized and directed knowledge interactions, the exploitation of technological opportunities and the coordination in production across heterogeneous and evolving industries and firms. What we propose is a real evolutionary and Schumpeterian account of GPTs, where the inno-
enerative change comes from within, producing differential growth (Metcalfe et al., 2006; Metcalfe and Foster, 2010).

2. Defining and Identifying GPTs, Engines of Growth

The strand of literature dealing with GPTs has been initiated by David (1990) and especially by Bresnahan and Trajtenberg (1995), where general purpose technologies are defined as key technologies, fully shaping a technological era, “characterized by the potential for pervasive use in a wide range of sectors and by their technological dynamism” (Bresnahan and Trajtenberg, 1995, p. 84). GPTs execute some generic functions such as “continuous rotary motion” or “binary logic” and act like platforms, “enabling mechanisms” for complementary innovations in downstream sectors, whose development leads to the transformation of the economic system as well as to generalized productivity gains. Rosenberg and Trajtenberg (2004) identify more precisely the properties of a GPT in their historical case-study of the Corliss Steam Engine in the U.S. (italics added):

“first, it is a technology characterized by general applicability, that is, by the fact that it performs some generic function that is vital to the functioning of a large number of using products or production systems. Second, GPTs exhibit a great deal of technological dynamism: continuous innovational efforts increase over time the efficiency with which the generic function is performed, benefiting existing users, and prompting further sectors to adopt the improved GPT. Third, GPTs exhibit “innovational complementarities” with the application sectors, in the sense that technical advances in the GPT make it more profitable for its users to innovate and improve their own technologies.” (Rosenberg and Trajtenberg, 2004, p. 65)

Therefore, on the “input side”, what makes a technology a GPT is its i) general applicability, ii) technological dynamism and iii) innovation spawning. Quite similar features are listed in alternative definitional exercises, to be found in the collection of papers by Helpman (1998) and in the studies of Lipsey, Carlaw and Bekar (2005), Guerrieri and Padoan (2007) and Jovanovic and Rousseau (2005). The latter, in particular, testing empirically similarities and differences between two popularly recognized GPTs, electrification and ICT, add six other “symptoms” of a GPT derived from theoretical models and holding also (even with different magnitudes for the two technologies) from data evidence (Jovanovic and Rousseau, 2005, p. 1203-1204); these can be meant as “output side” characteristics of GPTs: i) productivity slowdowns, due to learning effects and to the allocation of productive resources to develop new compatible and complementary capital required to use the GPT; ii) rise in the skill premium (the increase in demand for skilled labor should facilitate and shorten the learning process); iii) rise in entry, exit and mergers as a measure of reallocation of re-

---

6 General applicability is sometimes replaced in the literature by the term “widely used” (more precisely, see Bresnahan and Yin (2010)), showing a tendency in the theoretical analysis of GPTs to loosen the concept, in order to integrate a wider range of technologies under the definition of “general purpose”. An additional criticism that we will not address here concerns the definition of general applicability itself, which in the GPT-literature is always conceptualized in relation to the number of application sectors that use the GPT. Alternatively, general applicability can be interpreted as the feature of “doing nothing in particular” (Simon, 1987), therefore pointing more to the breadth of functions a technology can potentially operate, abstracting from the connections with other sectors or technologies.
sources; iv) initial fall of stock prices, due to the acceleration in the rate of obsolescence of old capital vintages caused by the adoption of the new GPT; v) changes in market shares favoring young firms; vi) rise in the interest rate and worsening of trade balance, since assets reallocation, reducing output, push demand and consumption to search for foreign markets.

Productivity slowdowns on the one hand and, on the other, the consequent time-lag needed for the new GPT’s productivity improvements to show up in the data, can be seen as one of the explanation for the so-called Solow paradox (Solow, 1987; Basu and Fernald, 2007); this is also the main dynamic generated by the first cohort of GPT-based growth models, built around the concept of “the time to sow and the time to reap” (Helpman and Trajtenberg, 1994). Fluctuations in productivity, together with the acknowledgment that technological progress is uneven, “comes in bursts” (Jovanovic and Rousseau, 2005, p. 1221) and is pervasive with different degrees, can be considered the main motivation leading to the development of GPTs literature.

We envisaged earlier in the paper that the problem emerging from this kind of definitions is one of identification. The issue is problematic from an ex-ante point of view (can we infer the GPT nature of a technology since its very introduction in the market?) as well as from an ex-post perspective (that of classifying under the label GPT what has been recognized in a generic way as a radical innovation). Which technology is a GPT and which one, instead, is not? The knife-edge distinction, here, is between those scholars who recognize only two or three GPTs since the industrial revolution (the steam engine, electrification and the more questioned ICTs) and see them as singularities or extreme cases of radical innovations (“epochal innovations”, as Rosenberg and Trajtenberg (2004) rename them), and those who expanded the list to a much more wide range of technologies. As we will show in the next paragraph, the first generation of GPT-based growth models (except the very first model of Bresnahan and Trajtenberg (1995)) – employing one GPT per period – is closer to the first interpretation, while recent models tend to a more generous interpretation of the notion. The empirical literature made some steps forward in solving the identification puzzle; the results can be considered useful only for what concerns measurement issues but without being able to distinguish clearly between GPTs and just “radical innovations”.

David and Wright (1999) stress precisely the ex-post identification point when, after another enumeration of the properties characterizing a GPT, they criticize the growing number of technologies labeled “general purpose” by growth and innovation scholars:

7 In the literature, a GPT is never seen as an “emergent property” of market and technological interactions.

8 For example, Jovanovic and Rousseau (ibid.) cite the study of Cummins and Violante (2002) who – clearly choosing a capital-embodying perspective on technological change – “classify a technology as a GPT when the share of new capital associated with it reaches a critical level, ad if adoption is widespread across industries”. Another empirical choice is that needed to identify the beginning of a GPT era; again Jovanovic and Rousseau set it as “the point in time when the GPT achieves a one-percent diffusion in the median sector”. Other empirical analysis make use of patent data to “uncover” GPTs and to forecast new potential “candidates” for this role (see for example Hall and Trajtenberg, 2004; You-tie, et al., 2008; Feldman and Yoon, 2011).

9 For example, Carlaw and Lipsey (see references later), referring to their extensive work on GPTs, suggest five technological classes into which to group different GPTs; the classes are materials, ICTs, power sources, transportation equipment and organizational forms.
“One has only to consider the length of such proposed lists of GPTs to begin to worry that the concept may be getting out of hand. History may not have been long enough to contain this many separate and distinct revolutionary changes. On closer inspection, it may be that some of these sweeping innovations should be better viewed as sub-categories of deeper conceptual breakthroughs in a hierarchical structure. Alternatively, particular historical episodes may be fruitfully understood in terms of interactions between one or more GPTs on previously separate historical paths.”

Although subscribing the David and Wright comment, we have to admit that a heuristic, a rule or a principle to discriminate between GPTs and non-GPTs is still to be found. Nevertheless, in the fourth paragraph we will make an attempt to characterize some of the sources and conditions that can lead both to the ex-post prevalence and pervasiveness of a GPT.

3. Modeling GPT-based Economic Growth

It is useful to distinguish between a first and a second generation of GPT-based formal models. The first generation includes, in addition to the Bresnahan and Trajtenberg seminal paper (Bresnahan and Trajtenberg, 1995), the models collected and reprinted in Helpman (1998), in particular the two contributions by Helpman and Trajtenberg and that by Aghion and Howitt, who introduce GPTs into a modeling framework à la Helpman and Grossman (1991). After an interval of approximately five years, the research on GPT restarted with the models of van Zon et al. (2003), Carlaw and Lipsey (2005; 2006; 2011), Guerrieri and Padoan (2007), Harada (2010) to end up with the recent contributions of Bresnahan (2012; 2010). The list can be enlarged by another model, that of Fatás-Villafranca et al. (2011), which deals with major innovations, cycles, Long Waves and technological eras though without explicitly referring to GPTs. In what follows, we review only some of the models, who we think are the most representative and important.

The rationale for distinguishing between the two different cohorts of models is both conceptual – considering the diverse perspectives on the nature of GPTs – and chronological, since the second generation of models is a tentative reprise of the topic after its rapid success and even faster decline at the end of the 1990ies.

The Bresnahan and Trajtenberg (BT from now on) paper (Bresnahan and Trajtenberg, 1995) cannot properly be considered a growth model, since it stems from a micro/industrial organization-framework and it represents the interaction between two kinds of sectors, the GPT sector and a number of application sectors (AS), as a strategic game leading to Nash-equilibria. The focus in this model is not on the GPT itself, but mainly on the pure incentive-based “dual inducement mechanism” between the two types of sectors, where an increase in the quality of the GPT (what we called “technological dynamism”) incentivizes the AS to increase their technological level (the “innovation complementarities” property of GPTs), and this, in turn, induces the GPT sector to advance its technology. There are two kinds of externalities that, from a welfare point of view, lead to a social rate of return greater than the private rates of return: one is a vertical externality, related to the connected and hierarchical payoffs-structure of GPT and AS as well as to the role of imperfect information flows and appropriability between the sectors; the other is an horizontal externality focusing on
the role of demand, since the more AS exists in an economy, the more valuable is the GPT; this kind of externality raises the importance of public subsidies and public demand. The implication of the model is that “the coordination problem between technology-innovating and technology-using industries” (Hall and Trajtenberg, 2004) cannot be solved optimally in a decentralized market system.

The Helpman and Trajtenberg (HT hereafter) model (Helpman and Trajtenberg, 1994) draws on BT insights on GPTs and inject them into a fully-fledged endogenous growth model assuming agents’ perfect foresight, where GPTs become the main determinants of long-run macroeconomic dynamics. Output is produced with a GPT and a continuous set of components that have to be compatible with the general technology and that are produced by innovators in a monopolistic competition framework. New GPTs arrive in a deterministic way at predetermined time intervals of equal length, generating a symmetric cycle with two (or three, in a special case) sub-phases; in the first one, the old GPT is used to produce final output while resources and labor are allocated to R&D in order to develop components for the new GPT. In the second phase, starting after a minimum threshold of components has been produced, the development of components continues but the new GPT is adopted, fostering productivity. During the first phase of the cycle real GDP declines as wages increases; the other way round is found in the next phase. This mechanism has been successfully summarized by the expression “a time to sow and a time to reap”, and has been extended by the same authors in a follow-up paper (Helpman and Trajtenberg, 1996) that maintains the formal structure but allows for the existence of many sectors. The order of adoption of the GPT across the sectors between early adopters and laggards, so the diffusion process of the technology, can lead to multiple long-run equilibria. Policy implications are derived from the model, in particular the advice to act in order to shorten the first phase of the cycle, but its empirical verification results problematic.

The Aghion and Howitt model (Aghion and Howitt, 1998, shorten as AH) starts from HT’s basic formulation, stressing its limited empirical relevance in terms of the representation of the size of the slump (“all of the decline in output is attributable to the transfer of labor out of manufacturing and into R&D. But since the total amount of R&D labor on average is only about two and a half percent of the labor force, it is hard to see how this can account for change in aggregate production of more than a fraction of a percent.” (Aghion and Howitt, 1998, p. 55)) and the timing of the slow-down, that in the previous model follows immediately the arrival of the new GPT. Therefore, AH add to the HT model both a Schumpeterian flavor, making the arrival of GPTs stochastic realizations of a Poisson process, and a more realistic representation of the adoption process, involving social learning (imitation, in evolutionary terms).

The AH model divides the cycle in three phases, instead of two: “first, the economy wide GPT must be discovered. Second, a firm in that sector must acquire a “template”, on which to base experimentation. Third, the firm must use this template to discover how to implement the GPT in its particular sector” (Aghion and Howitt, 1998, p. 63). The role of social learning is relevant here: a firm (an industry) can move from phase zero to phase one – the acquisition of the GPT template – via independent discovery (depending on a Poisson process) or through imitation, whose likelihood to occur is a probability given by a cumulative binomial distribution. Transition from phase one to phase
two, then, requires the allocation of labor resources to R&D activities, with a rate of success depending on another Poisson distribution. In addition to that, AH provide some extensions of the model, considering skill differentials, wage inequalities and their relationship with technological change (and with the size of the slump generated by the arrival of the GPT), as well as the effect of the innovation-wave arrival on capital obsolescence.

Despite conceptual or analytical differences, the first generation GPT-based models share a common assumption: the GPT is recognized ex-ante as a general purpose technology. In BT it is the “first mover”, that then incentivizes application sectors to exploit innovation complementarities and starting what has been called a dual inducement mechanism. In HT – as well as in AH – an explicit assumption is made about the impossibility to develop new components before a new GPT has arrived. By having ex-ante knowledge about the existence of a new GPT, economic agents are left only with the possibility to decide, on the base of their expected profit, on the allocation of resources to research. The picture is quite simplified with respect to a reality of continuous technological change, with competition (Arthur, 1989), diffusion and selection happening in an uncertain environment, which opens rooms for the role of risk-taking entrepreneurs.

The model of van Zon et al. (2003) is thus assigned to the second generation of models, although it is just a modification of the Romer model (and so it may appear to belong to the first category), not only for a chronological reason, but mainly because it departs from the assumption that GPTs are identified ex-ante. It is also the first model that allows for co-existing GPTs. The model assumes two types of stochastic (Poisson) R&D processes: a basic R&D sector, which produces “core” technologies (the GPTs), and an applied R&D sector, producing “peripherals”, corresponding to HT components. Both R&D sectors are subject to decreasing returns, so after the arrival of a core technology the economic incentive – and so the labor force – switches to the production of peripherals, and the other way round. The fundamental novelty in the model, showed in a simulation study, relates to the possibility that some cores become “failed” GPTs if few or no components are developed for them. Failed GPTs remind us that “during the innovation process, the actual pervasiveness of an innovation when and if it arrives can only be guessed at” (van Zon et al., 2003, p. 8-9). A GPT is an “ex-post mental construct”, deriving from the evidence that a particular technology is capable to execute a wide range of (old and new) productive functions in the economy; assuming otherwise can lead to a limited comprehension of the GPT-based economic growth.

Carlaw and Lipsey (2006) (CL from now on) extend the idea of the van Zon et al. model, proposing an out-of-equilibrium three (competitive) sectors model, where the appearance of a GPT is driven by an endogenous mechanism. The three sectors, each represented by a specific production function, are: i) a fundamental research sector that accumulates a stock of basic knowledge and produces the GPT; ii) an applied R&D sector and iii) a consumption sector. The latter sector produces consumption goods with a productivity level derived from a share of the knowledge generated by the applied R&D sector; in turn, this one accumulates knowledge with an effectiveness that depends on the stock of knowledge available in the fundamental research sector. Finally, the fundamental research sector creates GPT-related knowledge with a productivity depending on the share of applied knowledge that is not directed to the production of consumption goods. The arrival of a new GPT is again stochastic and it is modeled around a slightly more complicated mechanism than the Poisson
process just seen in other models: two beta distributions generate two random values; the first is compared with a threshold and, if bigger, the GPT appears. The second serves to weight the share of new fundamental knowledge affecting as a productivity term the applied sector. Closing the model with assumptions on consumers’ expectations, maximization problem and resources allocation (made by a social planner), CL can provide a simulation study. Their model is then further developed in a succession of studies ending-up with multiple and coexisting GPTs being active in the economy (Carlaw and Lipsey, 2011): the fundamental research (GPT) sector is divided into different technological categories, while the applied R&D sector is represented by many research facilities. The picture of the economic evolution offered by the model becomes quite realistic, even if it is worth recalling the early warning we quoted from David and Wright: modeling plenty of GPTs, so revolutionary technologies, can appear as a forcing of the theoretical concept, with the consequence of losing the innovative feature of this strand of research. Moreover, the promising idea of the van Zon et al. model to challenge the assumption of an ex-ante identification of GPTs is lost in the CL formulation, which returns to the simple stochastic modeling of the GPTs arrival.

Before concluding this overview paragraph, another critical point should be added: in the latter models, as well as in the Goodwinian model of Fátás-Villafranca et al. (2011), the arrival mechanism of a new GPT depends on the accumulation (due to the optimal allocation of resources to research, or to routinized decision-making) of a certain quantity of knowledge: technological eras follow one each other because of the collection of a generic, non-well specified commodity called knowledge. Once we realize that the evolution of knowledge is something more complex, localized, purposeful and somehow sticky, the modeling strategy used in the existing literature to represent the arrival of GPTs results too much stylized. A new view of GPTs should deal also with this issue.

In spite of some of the subsequent theoretical developments, we think that the initial approach followed by Bresnahan and Trajtenberg (the BT model) is still the most promising starting point to deal with GPTs, especially in the more generalized forms recently proposed by Bresnahan (2012) and Bresnahan and Yin (2010). The latter paper deals with the role demand plays in the interaction between GPT and growth, interpreting the dynamics of GPTs replacement as the overcoming of “growth bottlenecks”, generated by the inertial (locked-in) trajectory of technical progress and the presence of un-served demand. The former study returns to the supply-side to analyze the conditions for the emergence of a “GPT cluster” (a particular GPT connected with several AS), suggesting that the market knowledge available and the entrepreneurial knowledge characterizing the innovators affect both the incentive to introduce a new technology and the expectations about its value. Three stylized situations are highlighted: i) planned initiative, the classical hierarchical interpretation of GPT, where the introduction of a general purpose technology induces the development of complementary innovations; ii) technological convergence, where specific technologies are invented first – even lacking the knowledge about their potential linkages – and that raises the expected profit of inventing a GPT that connects the already existing components; iii) inversion, when a specific innovation increases the value of inventing a GPT whose introduction, in turn, augments the incentive to introduce a new specific technology.

The last two contributions, even if they don’t tackle the definitional issue (a GPT is identified as such ex-ante in all the three different mechanisms just outlined), pave the way for an analysis built
on micro and meso arguments. An uneven and self-reinforcing (or self-reducing, as it could be possible in the case of a vicious circle of dis-incentives for innovative activities both in the GPT and in the AS) interaction between a hierarchy of technologies is by definition a perfect start for a fully evolutionary account of GPTs, that has to be qualified with insights and categories coming from innovation economics. In addition to that, the recognition of the role of technological specificities/opportunities (in the case of convergence and inversion), together with the part demand plays in the “coordination game” of GPTs’ introduction and diffusion, takes us very close to the industrial dynamics literature.

4. The Microeconomics of GPTs – Prevalence and Pervasiveness

To summarize what we have been discussing so far, we quote again Bresnahan (italics is ours):

“One goal [of studying GPTs] lies in growth macroeconomics, to provide an explanation of the close link between whole era of economic growth and the innovative application of certain technologies, called GPTs, such as the steam engine, electric motors, or computers. Another goal is in the microeconomics of technical change and proceeds by differentiating between innovations of different types. The incentives and information related to the invention of GPTs themselves, may differ from those related to the invention of applications; another example would be the incentives and information related to an established GPT with successful applications in contrast to earlier stages. A third goal links the macro and the micro. Can we understand the linkages between aggregate economic growth and the incentives and information structures related to particular inventions and to their application to particular uses and sectors? (Bresnahan, 2010)

In this paragraph we will study GPTs from a microeconomic point of view. As the discussion of GPTs in the growth literature has shown, in modeling, their appearance is taken as rather exogenous and their influence on other industries and sectors in an economy and hence their pervasiveness is taken as given. Certainly, to analytically proceed in this direction can be justified in two ways. First, it is for the purpose of modeling convenience allowing for an analytical solution. Secondly, the discussion of Long Waves of economic development has repeatedly highlighted the occurrence of fundamental technologies. The emergence of these fundamental technologies (as well as the approach of Long Waves in general) still is a phenomenon not well understood, despite several attempts in the 1980ies and 1990ies (Weidlich and Haag, 1983; Weidlich, Haag and Mensch, 1987; Mensch, Haag and Weidlich, 1991) searching for an explanation. In view of that state of the art, especially the assumption about the exogeneity assigned to GPTs in macro modeling seems to be not too far-fetched.

Our stance in this discussion is, however, to go further and to highlight some directions allowing to better grasp and understand the phenomenon of GPTs. For this purpose we combine the view by van Zon et al. of a GPT as a “ex-post mental construct” with considerations on early indicators of the emergence of GPTs. Hence we attempt to come closer to the “origins” of a GPT. By this we certainly follow Arrow (1991) in admitting that “…it is hopeless to develop a model which will genuinely predict innovations” and in claiming that those models and the considerations behind them
will provide “…some useful idea of the average rate of technological change, of the degree of fluctuations and the kinds of surprise that we may find in the future. We cannot, of course, predict a surprise; that is a contradiction in terms. But we can predict the kind of surprises that might occur”.

On this basis we suggest to enrich the approaches to GPTs by a micro and a meso level analysis and draw attention to approaches which address path dependency, technology competition and dominant designs, as well as collective innovation and sectoral interdependencies. A look at the characterization of GPTs from a microeconomic point of view again already indicates some avenues to follow.

An attempt to summarize various definitional exercises on GPTs in the literature (Bresnahan and Trajtenberg, 1995; Lipsey, Bekar and Carlaw, 1998; Jovanovic and Rousseau, 2005) leads to the following three characteristics:

1. Pervasiveness: A GPT should have an impact on technical change and productivity growth across a large number of uses / industries;
2. Improvement: A GPT should experience a wide scope of improvement and elaboration in its own industry;
3. Innovation spawning: A GPT should lead to product and process innovation in a broad range of uses / application sectors.

These characterizations are based on an implicit assumption, namely that the technology under concern, the GPT, is a prevailing technology, it exists for a longer period of time, it is accepted on a broad scale and for these reasons it impacts on an economy in a pervasive, improving and innovation generating way. Hence, the conditions for a broad usage of a GPT in an economy are of interest but neither her sources nor the conditions of her prevalence. What are the sources of prevalence and of pervasiveness so much at the core of GPTs? What can be said beyond addressing an exogenous source?

**Conditions for prevalence**

The prevalence of a technology is given by her persistency over time. In other word, such kind of technology is unlikely and difficult to be challenged by new alternative technologies – it seems to be incontestable, at least for some time. Instead of taking that feature as given we have to think about mechanisms and determinants just providing for prevalence. To accomplish that let us first look into the reasons for a technology to be incontestable.

According to David (1985; 1987) this prevalence is the result of a coordination of agents’ choices on a specific technology. The outcome then is a specific allocation which is rather stable over time. Three conditions lead agents to coordinate their choices and also lend persistence to the resulting allocation:

1. the technical interrelatedness of system components;
2. quasi-irreversibility of investment (or, more generally, switching costs);
3. positive externalities or increasing returns to scale.
The technical interrelatedness of a system (1) appears to be an aspect very much out of the economic realm. Chemical and physical laws as well as engineering types of relationships presumably determine which kinds of technologies fit together, which ones may be substituted, and which complementarities cannot easily be challenged. If we consider GPTs as the core of such kind of a system then the explanation for their prevalence is a rather technical one.

The quasi-irreversibility of investment and related – often very high – switching costs (2) extend the previous argument and translates it into economic cost terms. The technical interrelatedness (1) could as well be expressed in cost terms: the switching costs, related to the resources required for exploring new chemical and physical laws or engineering relationships, which allow for breaking up the interrelated system, are very (if not infinitely) high. In other cases, it is the systemic dimension of the supply of the goods and services related to a certain technology which protects against the challenges of new invader technologies – the combustion engine for automobiles and the accompanying system of fuel stations and fuel logistics just being a point in case. Combined with these investments are mutual dependencies – not only of a technical nature but also in terms of relative prices – which contribute to the prevalence of the core technology. As long as relative factor price changes remain in a certain range, switching costs to new alternatives prevail high and secure the persistency of the existing technology.

Another aspect of a technology, partly related to the aforementioned systemic aspect, addresses the existence of positive externalities or increasing returns-to-scale (3) in the use of a certain technology. These increasing returns may arise either on the supply side of a market as a result of learning effects (learning by doing or by using) or on the demand side as a result of positive network (or agglomeration) externalities that raise the benefits of a technique, product, or location for each user as the total number of users increases. Learning-by-doing, allowing the efficiency a certain technology to increase the more it is used, sustains competitiveness and dominance of the dominant technology. The positive network effects related to customers’ adoption of a certain technology are based on the idea that the individual benefits a customer enjoys depends positively on the number of other customers – as is the case in telephony, video system or in computer software. These supply and demand side based externalities already protect the established technology against possible invading new alternatives.

At least the second and third of the conditions above – irreversible investment and increasing returns – indicate already that the effects on prevalence are realized not at a single point of time but rather dynamically. In either case it is a positive feedback from the macro state of the system (high level of front-up investment, high number of adopters) to the choices of individual agents, eventually resulting in de facto standardization on a single technique. In the words of David (1987; 1987), effects of path dependence are prevailing here.

Taking into account the dynamic dimension of the conditions for prevalence quite naturally leads to considerations about the sources of prevalence. Hence, the mechanisms that sustain prevalence are also the ones leading to dominance. Or to put it in reverse order, observing a technology that in competition appears to dominate due to positive feedbacks allows the conclusion that this technology will stay dominant for a certain longer period.
The dynamics towards prevalence

For understanding innovation and technology competition insights from industrial dynamics and the inherent innovation, learning and knowledge dynamics are useful. This literature informs about possible mechanisms, conditions and structural dynamics on which the development or the appearance of a standard or a dominant design is based.

At the core of the further discussion are approaches in industrial dynamics. This field of research established by formulating major criticism to the approaches in industrial organization tackling the traditional issues related to the Neo-Schumpeterian hypotheses, namely the question whether large (monopoly) firms or small (competitive) firms are the major drivers of innovative activities and the resulting economic development. The inconclusive empirical evidence on these hypotheses induced a research agenda that considers the analysis of industries and the innovative activities herein to necessarily take into account innovation and technology dynamics. The resulting competition between differently innovative firms or different technologies is a major driving force in shaping industry and market structures.

A GPT in this context can be considered the result of competing new technologies or the successful challenge of an old (GPT?) technology. The mechanism behind is path dependent, presumably leading to the establishment of a standard or a dominant design – both resembling the characteristics of dominance over alternatives.

A dominant design in a product class is, by definition, the one that wins the allegiance of the marketplace, the one that competitors and innovators must adhere to if they hope to command significant market following. (Utterback, 2007)

Compared to a GPT, a dominant design is just defined more from the generating side, in the sense of where it comes from and under which circumstances it appears. About its further effects not much is stated except that it is useful as well as widely accepted and used. In this sense, it complements the concept of GPT just from the generating side and combining both may help understanding better the appearance of GPTs.

As to the emergence of a dominant design, it is a competitive process of trying out the basic features of the dominant design:

Prior to the appearance of a dominant design many of its separate features may be tried in varied products which are either custom designed or designed for a particular and demanding market niche. (Suàrez and Utterback, 1995, p. 118)

The industry life cycle literature – dealing with the long-run development of industries – just applies the concept of a dominant design (Utterback and Suàrez, 1993) to explain the transition from a phase in which the number of firms in an industry is increasing to the appearance of the so-called shake-out, during which the net entry is negative and a sharp decline in the number of firms is observed – then eventually leading to a phase in which the number of firm is constant (an oligopoly).
The appearance of dominant design is here seen as the outcome of the competition between various different technological solutions during the phase of expansion.

Among the factors reinforcing this development of competitive selection – best technology compromise; cooperation; combination of sociological, political, and organizational dynamics; economies of scale in Frenken and Murmann (2006); similar firm level related factors as well as environmental factors in Suarez (2004) – sources of dynamic externalities or positive dynamic returns to scale are important in our context. In some literature these effects are related to path dependencies. Those may be connected to the size of the firm and the accumulated production experience or to demand side effects. In principle these externalities provide for additional benefits for a technology which is leading in terms of produced or sold or used units. Hence, a first mover advantage related to size (production volume) contributes to the dominance of a certain technology.

The effects of economies of scale (Klepper, 1996; 1997) and related learning economies draw from the fact that the more units have been produced or the more often a technology has been applied the higher is the contribution to productivity which materializes in lower unit costs (and hence lower prices) or higher product quality (and hence a higher quality/price ratio). Both contribute to the benefit a user can reap and hence contribute to the dominance of the technology concerned.

An equivalent argument can be formulated for the demand side. David (1995) and Arthur (1989) highlight the benefit of using/consuming a specific product depending on the number of other users. Due to random factors one technology will gain a larger share in the population of consumers accompanied with comparatively higher returns. The likelihood that next consumers will select just this superior technology increases and subsequently without further larger random shocks the technology with the small lead will win a dominant position. Over time, the technology which accidentally grasps a certain lead in market share will in the end come to dominate the market or industry. Deviations from this outcome are either due to major stochastic shocks or the trespassing of a certain critical mass of adopters of a competing alternative (Witt, 1997).

Taken these arguments together, approaches addressing dominant designs inform about the competitive environment within which technologies strive for market dominance. One of the mechanisms or factors behind this process, namely dynamic economies of scale, coincides with the mechanism and conditions which provide for the prevalence of GPTs. On these terms, the emergence of a GPT and its prevalence seem to be intimately connected via the same kind of mechanism.

**Conditions for pervasiveness**

The other core characteristic of a GPT, the pervasiveness in application and for further innovation, is related to (a) her broad impact into other sectors and branches and (b) her relevance for further innovative activities there.

Pervasiveness appears on a first sight to be related to technological interrelationships, one of the conditions for a stable allocation formulated by David, and hence to be an exogenous factor. However, research into the innovation activities of firms, research institutes and other actors has delivered collective invention and innovation (Allen, 1993; von Hippel, 1987) as the mode of organiz-
ing these activities which has become more and more frequent over time. The basis for cooperating in innovation is risk and costs sharing on the one hand and knowledge exchange, access and sharing on the other. The context within which this collaboration appears is characterized by dispersed knowledge and competences which for the sake of coming to new solutions are required to be combined and interacted. This aspect becomes the more relevant the more complex is a technology to be developed and pursued.

Addressing again the literature on dominant designs\(^{10}\) Liebowitz and Margolis (1995) show that collaboration in innovative activities as well as related strategies such as licensing of new ideas is conducive to the emergence of a dominant design. One may assign this property also to the appearance of GPTs. Moreover, the participation of several actors and their agreement on the best technological compromise (Abernathy and Utterback, 1978; Christensen, Suárez and Utterback, 1998) induces a broad acceptance which contributes to pervasiveness. In this context, the multidimensional nature and high development costs of many complex products rather naturally requires that several parties agree on cooperating and negotiating (Cowan, 1990; Tushman and Rosenkopf, 1992) in the technology to be pursued collectively.

Establishing a dominant design collectively brings about a certain pattern of further innovative activities. According to Anderson and Tushman (1990) the appearance of a dominant design is followed by innovative activities which are of a rather incremental type. These incremental steps allow refining the basics of the dominant design and exploiting specific sectoral, industrial or agent specific opportunities the dominant design offers. As long as these opportunities do not get exploited completely, the pervasiveness of the dominant design / the GPT (as well as her prevalence) appears to be secured.

Certainly, the foregoing discussion of the emergence of a dominant design as a collective outcome and its persistence over time is quite neutral with respect to the breadth of this design. With breadth we mean the number of industries affected by the design, using it or working with it. In fact, the literature on dominant design is mainly focusing on industries and the competition among firms and/or technologies in these industries – hence it is an intra-sectoral analysis with the inter-sectoral dimension not taken on board. Applying these findings to our understanding of the pervasiveness of GPTs requires certainly taking on board the inter-sectoral dimension. To this end further research needs to be performed which will be informed by analyses on the technological relations between sectors (Cantner and Hanush, 1999), the dimensions of related technological variety between sectors (Bürger and Cantner, 2011), as well as the policy designs in fostering collaboration in innovation (Cantner and Pyka, 2001).

5. Conclusions

The success of the GPT “category” in economic theory is probably well explained by the need to gather all the insights coming from a rich research trajectory on Long Waves and Business Cycles and to condense them into a workable “mainstream” modeling exercise. However, both the very no-

\(^{10}\) For example, Cusumano, Mylonadis and Rosenbloom (1992); Khazam and Mowery (1994).
tion of GPT and the simplifying assumption accompanying it in the macro-models can be questioned. In this paper, we offered a “new view of General Purpose Technologies”, building on the historical as well as on the recent literature, enquiring more in deep the definitional problems related to the GPTs and the conditions for their emergence, together with the characteristic for their prevalence and pervasiveness. A Schumpeterian and evolutionary view pointing at the micro and meso level of analysis – that of the dynamics of firms and industries –, is in our view the privileged perspective economists need to adopt in order to revitalize the theoretical and empirical study of GPTs. The similarities with the emergence of dominant designs and the relations with dynamics of increasing returns and path dependency in the choice between alternative technologies offer us a set of tools well suited to study the establishment of GPTs as a process unfolding in time, more than as a single homogeneous shock.

Explaining GPTs is thus not a matter of technology “arrival”, nor does it require explicitly some inherent “radicality” of the technologies. General purpose technologies can be introduced in particular niches of the market or in specific industries and there they can be “cultivated” or developed till, thanks to industrial interactions, demand pressures and technical competitions, they assume the role of core technologies, shaping the general configuration of production and of the whole economy.

As Mokyr (2010) points out, “major discoveries rarely arise de-novo, and what seems to us a breakthrough was only the last step in a long intellectual journey”. Our task as economist is to try to intercept the trajectory of this intellectual journey so to understand the possible alternatives an economic system can have to generally increase the total welfare of the Society.

6. References


Arrow, K. (1962b), Economic welfare and the allocation of resources for invention, in The Rate and Direction of Inventive Activity, Princeton University Press and NBER.

Arthur, W.B. (2009), The nature of technology: What it is and how it evolves, Free Press.


