Bureau d'économie théorique et appliquée (BETA)
UMR 7522

Documents de travail

«Pervasive Enough? General Purpose Technologies as an Emergent Property»

Auteurs

Vladimir Korzinov, Ivan Savin

Document de Travail n° 2016 – 49

Novembre 2016

Faculté des sciences économiques et de gestion

61 avenue de la Forêt Noire
F-67085 Strasbourg Cedex

Secrétariat du BETA
Géraldine Del Fabbro
Tél. : (33) 03 68 85 20 69
Fax : (33) 03 68 85 20 70
g.delfabbro @unistra.fr
www.beta-umr7522.fr

Pervasive Enough? General Purpose Technologies as an Emergent Property*

Vladimir Korzinov†[1] and Ivan Savin‡[1,2]

[1]Chair for Economic Policy, Karlsruhe Institute of Technology, Germany
[2]Bureau d'Economie Théorique et Appliquée, UMR 7522 Université de Strasbourg - CNRS, France

Abstract

We propose a novel model of knowledge discovery shedding light on the emergence of General Purpose Technologies (GPTs), the process which has been largely neglected in the literature on technological change. We demonstrate that GPTs emerge only when certain conditions with regard to the following techno-economic factors are met: knowledge diffusion, coordination on technological trajectories and volatility in the rank of expected returns on products. Furthermore, our model provides intuitive explanation for technological lock-ins, S-shaped curves of technology adoption, temporal clustering of innovations in time and replicates distinct features of empirical networks of relatedness among technologies and products.

Keywords: *general purpose technology, technology networks, pervasiveness of technologies, knowledge diffusion, innovation*

JEL Codes: *C63, D83, D85, L16, O3*

*Financial support from the Helmholtz Association (HIRG-0069) and Projex CSES, Initiative d'Excellence, Université de Strasbourg. Thanks are due to very helpful comments and suggestions from Robin Cowan, Patrick Llerena and Ingrid Ott. This work has benefited from presentations at workshops in Jena, Bochum, Karlsruhe, Strasbourg, Pisa and Essen as well as from CEF conference in Bordeaux and ISS congress in Montreal.

†Corresponding author: Rüppurrer Str. 1a, Haus B, Room 5.22, D-76137 Karlsruhe, Germany. Tel.: +49-721-60843073, Fax: +49-721-60845255, `Vladimir.Korzinov@kit.edu`

‡Ivan.Savin@kit.edu

1 Introduction

It is a common knowledge nowadays that innovations are vital for the process of economic growth (Romer, 1990; Aghion and Howitt, 1992, 1998; Helpman, 1998). A specific type of drastic innovations called General Purpose Technologies (GPT) were introduced as one of the forces to explain this growth and its cyclicality (Bresnahan and Trajtenberg, 1995; Bresnahan and Yin, 2010). Ever since their wide acknowledgment in the book of Helpman (1998), these technologies are seen as engines of economic development of whole countries (Ott et al, 2009) or industries (Strohmaier and Rainer, 2016). Despite some disagreements on what technologies shall be considered as GPTs, this concept stays relevant up till now and is proved to be important during the first and second industrial revolutions and information age (Bresnahan, 2012, p. 612).

A formal definition of GPT put by Lipsey et al (2005, p. 98) says "A GPT is a single generic technology, recognizable as such over its whole lifetime, that initially has much scope for improvement and eventually comes to be widely used, to have many uses, and to have many spillover effects". The literature claims that in order to be classified as a GPT an innovation has to possess three major characteristics. The first one, *pervasiveness*, implies that a technology or its principle is used in vast amount of products throughout an economy and in various applications (as, e.g., electricity is used from heating and lighting our houses to powering trains). The second, *technological dynamism*, postulates that these technologies experience significant improvement in their efficiency and effectiveness throughout their lifetime (one example is the 'Moore's Law' well-known in the semiconductor industry). Finally, innovation complementarity (also called a 'dual inducement mechanism'), means that these improvements induce innovations in application sectors of this technology (e.g., the evolution of semiconductors has led to the introduction of numerous portable devices) and vice versa (Helpman, 1998).

In the early GPT models, the emphasis was on the attempt to account for a 'residual' in aggregate production functions of mainstream neo-classical models (Bresnahan and Trajtenberg, 1995; Helpman, 1998) and explain the famous 'productivity paradox' (Brynjolfsson, 1993). In these models new transforming technology appears periodically and exogenously and induces changes in economic structures (like in Bresnahan and Trajtenberg (1995), where a switch to a new production regime using a GPT happens after a certain number of new intermediates becoming producible, while agents realize their ability to produce these intermediates at a pre-specified moment). In later models authors followed, among others, the so called 'structuralist-evolutionary approach' (Lipsey et al, 2005), where technologies "evolve under a stream of innovations" and the effect of a newly arrived GPT on the economy is determined endogenously, but the moment of arrival is still exogenous (see also Carlaw and Lipsey (2006)). The work of Lipsey and Carlaw names GPTs as a part of 'structural technologies' with eleven key characteristics incorporating them in a sequential model with simultaneous GPTs (Carlaw and Lipsey, 2011). Similar to others this model uses the concept of aggregate production function which does not allow to reveal the heterogeneity of knowledge stock out of which GPTs emerge. More recent models on GPT focus on a 'dual inducement mechanism' between GPT and its application sectors (Bresnahan, 2012) assuming one in a pair of complementary technologies to have generality of purpose. These works also elaborate on different types of knowledge (Bresnahan, 2012) or 'growth bottlenecks' (Bresnahan and Yin, 2010), but their arguments take introduction of a GPT for granted. The literature, thus, has long been focusing on explaining the effect which GPTs have on economy but none of the models so far tried to address the process of GPT formation, or as pointed by Cantner and Vannuccini (2012, p. 74) in all current models a GPT “arrives from the outside of the system”. Therefore, factors that influence the GPT emergence remain hidden.

The present study attempts to narrow this research gap concentrating solely on the pervasive character of a GPT trying to identify factors fostering inclusion of a technology as an input into newly discovered products. Thus, we consider the emergence of a GPT not as a binary but as a continuous outcome, where certain technologies may exhibit the pervasive property to different extents, and the larger this extent the more likely the technology will be classified as a GPT. For the same reason, this work is not meant to answer the question when a GPT emerges. Instead, we look for forces boosting the process of 'technological convergence' coined by Rosenberg (1976), where economy utilizes the same technologies for different purposes and consumer products become related through similar technologies. We offer a novel perspective on the knowledge discovery process as a network growth, where nodes are single technologies (knowledge pieces), and each new connection (link) represents a new knowledge being discovered (technology combination resulting in value added); each technology allows to produce a certain intermediate input, while fully connected groups of those nodes (cliques) stand for producible final goods.

Our work builds on the literature started by Schumpeter (1934, p. 65) defining innovations as "new combinations" of new or existing knowledge, and continued by theories of architectural innovation (Henderson and Clark, 1990), recombinant growth (Weitzman, 1998), combinatorial technology models (Arthur and Polak, 2006) and works on technological capabilities (Hidalgo and Hausmann, 2009) considering knowledge as a collection of heterogeneous pieces being interconnect-able with each other in one or another way. In other words, technologies are assumed to have a hierarchical structure and be interrelated (Lipsey et al, 2005).[1] Thus, the process of GPT formation transforms into inclusion of a single technology (*potential GPT*) in as many as possible final goods. To consider this, we model the technology to have the *potential* to be included in all final goods, but without *certainty* to do so. This is achieved by allowing multiple alternative ways of producing the same type of good and looking only on the first discovered ways of production. The latter is done as a simplification to concentrate on the process of product discovery and not further competition between substitute goods over production costs,[2] and shall reflect the fact that technologies included in early product discoveries have a lead time advantage over future competitors (Arthur, 1989). The more often the potential GPT enters those early product discoveries (in other words, fulfills its potential), the easier it should become to identify emergence of a GPT.

Hence, the aim of the present work is to reveal factors that may foster or hamper inclusion of the potential GPT as an input to as many final goods as possible. Among the usual suspects we outline the process of knowledge diffusion, the structure of the technological network, the choice over technological trajectories to follow and the pressure from the demand side (and, in particular, its variation over time) in discovering new final goods. The knowledge diffusion is considered because of the famous public good property of knowledge (Arrow, 1962) and the resulting possibility to create “complemen-

[1]While in reality a complex technology can consist of sub-technologies, which in their turn consist of sub-sub-technologies and so on, we simplify this modular structure implying final goods to be producible out of a large group of interconnect-able but single technologies. Note that this is done without loss of generality since those complex technologies can be seen as interconnected groups of intermediates, which in their turn have to be all connected to further technological inputs to invent new final goods.

[2]Introduction of production costs into the model is left for further extensions.

tarities among trajectories" (Dosi, 1982, p. 154). The extent of this effect, however, is contingent on the exact network structure of knowledge considered, since the complex interrelationships between technologies can result in some technological links being present in numerous products (as was the case, e.g., for a steam engine combined with a wheel) or very few only. Another rationale to consider the knowledge network is that the potential GPT is not necessarily the only technology having large scope of applications, but that all technologies have a different potential degree of pervasiveness thus affecting each other chances to become included as an input in final goods. The mechanism behind choosing between technological trajectories, in its turn, is important due to the competition among the aforementioned alternative technological combinations in becoming first to satisfy each consumer need. Since the innovation process is seen a search in complex technology spaces "shrouded in uncertainty" (Silverberg and Verspagen, 2005, p. 226) and characterized by strong path dependence (Nelson and Winter, 1982), it is a key to our model to see how this mechanism affects the GPT adoption. Last not least, the role of the demand side effects is not clear. Is it beneficial for the knowledge discovery process in general and the GPT adoption in particular that society starts favoring a certain product development as it was the case, e.g., for nuclear power plants in the 1950s (Cowan, 1990) or renewable energy generation in the last two decades (Herrmann and Savin, 2016)? In both cases, the policy maker was providing large subsidies to discover a product with certain characteristics, while actual choice among different technological trajectories were left to innovating firms. Clearly enough, none of the four factors shall be considered in isolation from the others, and the rest of the study devotes particular attention to the interplay between those forces on the emergence of GPT.

Our results demonstrate that the knowledge diffusion, both in terms of applying the same technological knowledge (link) to many distinct products but also spreading this knowledge among agents doing R&D, is a key prerequisite for the emergence of a GPT since being discovered once the knowledge spills over benefiting most those technologies having multiple potential applications in combination with other intermediates for production of different final goods. Given the presence of knowledge spillovers, coordination of R&D efforts (concentrating on technological trajectories with more accumulated knowledge) also favors GPT, at least in the short term. However, once the technology network is modeled as a graph growing over time where agents become aware of more complicated technological combinations through inventing simpler products, the aforementioned effect of coordination transforms into an inverted U-shape form illustrating the famous exploitation vs. exploration trade-off. For the same reason, volatility in the rank of expected returns on products has a negative effect on GPT's adoption in the long run: the high pressure from the demand side makes the size of the discovered technology network shrink, limiting the potential knowledge externalities and leading to a technological lock-in. In addition, our model replicates some known stylized facts as S-shaped curve of technology adoption, temporal clustering of innovations in time and some distinct features of networks of the product and technology relatedness discussed by Hidalgo and Hausmann (2009) and Boschma et al (2014).

The rest of the paper is organized as follows. Section 2 describes the basic set up of our model and formulates four propositions on factors triggering the process of GPT adoption. We provide results of the numerical analysis of our baseline model in Section 3 additionally extending it by introducing an increasing in time knowledge base. In Section 4 we outline some stylized facts that our study reproduces, while Section 5 discusses implications of the results and concludes.

2 The Model

2.1 Technology network

In this model we concentrate on the process of knowledge discovery. In particular, it is assumed that to satisfy consumer needs, certain population of product types (P) is necessary to be introduced into the market. Examples of such product type are communication, transportation and clothes, which can be seen as needs. Thus, we share the view on innovation as a problem-solving process (Dosi, 1988, p. 1125). For each product type to be discovered and introduced onto the markets, some intermediates (I) need to be combined, which, in their turn, in reality are typically combinations of other intermediates. We simplify our modelling by considering only two layers (see left panel of Figure 1): the product types (final goods) and the intermediates (technologies used to produce the intermediate input: internet, combustion engine or LED).[3]

From the beginning, the technologies are present in the model as yet not connected nodes of the technology network (mid panel of Figure 1). However, for these technologies to find practical application, they need to become interconnected with other intermediates of the same product type forming a fully connected component,[4] *clique* as it is coined in the network science (demonstrated on the right panel of Figure 1). Thus, the discovery of each link (or edge, as we use those words interchangeably) in the technology network represents the knowledge discovery process, where existing technologies are combined with one another in new ways to produce value for a consumer (i.e. so that connected component of those technologies has a larger value than those technologies taken separately).[5] Hence, a connected clique of any size starting from two (technologies) could represent an invention of one product type (i.e. ones clique is connected, the product becomes producible). Note at this point that such a step-wise discovery of new products ("sub-innovations spread out over time" using the words of Silverberg and Verspagen, 2005, p. 226) is not new for the literature but rather represents a stylized fact.

We make another assumption that each product type has more than one way of production, i.e. there is more than one technology combination satisfying a certain need (compare technology network from Figures 1 and 2 consisting of the same product types and intermediates). The intuition is that there is no consumer need to be satisfied in a unique way. Those alternative technology combinations satisfying the same need can be anything from having very different inputs (e.g., paper towel vs textile one vs electric hand dryer) to fairly similar ones (different types of cheese, all fermented out of milk by yeast).

[3] Henceforth, we use the terms 'technology' and 'intermediate' as synonyms.

[4] In a similar way of reasoning, one could consider a fraction of technological links from the clique also forming a fully connected component to be themselves technologies of a higher complexity (combining more than one technological input) and necessary to be discovered for the respective good to become producible. For simplicity, however, we avoid such a discussion to keep our argument clear and simple.

[5] In fact, the latter definition echoes the definition of complexity by **?**, p. 24: "The whole is greater than the sum of its parts". An illustration of that definition in reality is another quote from **?**: "Take two technological innovations that have revolutionized twentieth-century society, the internal combustion engine and the digital computer. The internal combustion engine combines Volta's sparking device, Venturi's (perfume) sprayer, a water pump's pistons, a mill's gear wheels, and so on. The first digital computers combined Geiger's particle counter, the persistence (slow fade) of cathode ray tube images, the use of wires to direct electrical currents, and so on. In both cases most of the building blocks were already in use, in different contexts, in the nineteenth century. It was the specific combination, among the great number possible, that provided the innovation."

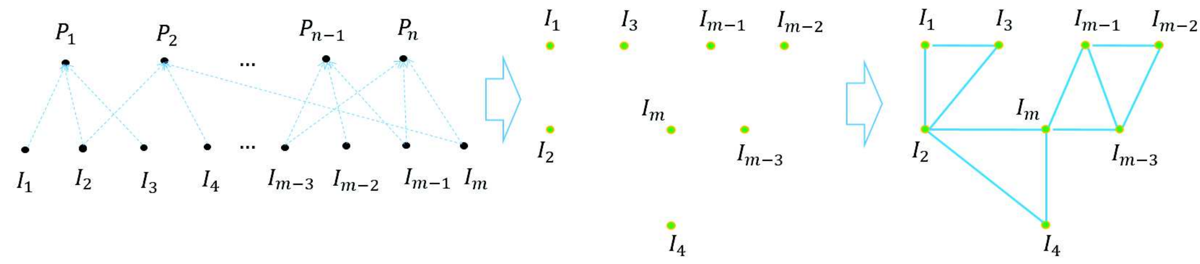


Figure 1: Layers of products and intermediates

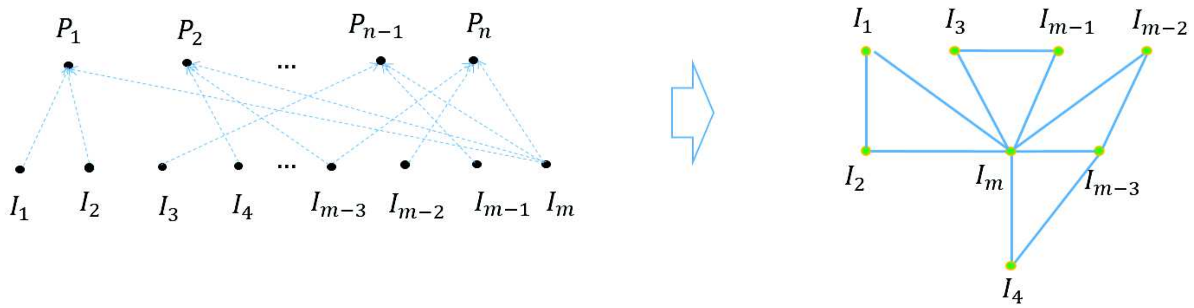


Figure 2: Alternative combinations of intermediates

Important to stress is that combination of two distinct technologies (like $I_{m-1} - I_{m-3}$ on Figure 1 or $I_m - I_{m-3}$ on Figure 2) may enter more than one product both, within one way of technological combination but also between them. This model feature reflects the fact that in real world we may utilize the complementarity arising from combining two technologies together in more than one application.[6] Combining all alternative technology combinations together (constructing a multiplex network) one obtains a '*potential technology network*' - mapping of all possible combinations producing added value, (see left panel of Figure 3). The resulting network can be considered as a *technological paradigm* in accordance with Dosi (1982, p. 148)'s definition: "an 'outlook', a set of procedures, a definition of the 'relevant' problems and of the specific knowledge related to their solution",[7] while each single way of technological combination as a *technological trajectory* – "the direction of advance within a technological paradigm". Clearly, the position of each technology in such a network is different. In accordance with the arguments presented in Section 1, we consider a GPT to be the one with largest generality of purpose, thus, *potentially* entering all product types in at least one technological combination (right panel of Figure 3).[8] However, there is no guarantee that the GPT will eventually be included in any product type discovery. Our aim is to identify factors fostering fulfillment of the GPT's potential application in largest possible number of final goods.[9]

[6]For example, combining tubes and lenses for telescopes, microscopes, photo equipment etc.

[7]In ?, p. 1127, words on ".'pattern' of solution of selected techno-economic problems based on highly selected *principles derived from the natural sciences*" are used.

[8]At the same time we rule out the option that GPT enters all product type within any single way of technological combination to make its inclusion (in all products) a less trivial task.

[9]Henceforth, we refer to GPT as the technology with largest pervasiveness *potential*. While examining to what extent this potential has been fulfilled, we interchangeably call it 'GPT' and 'potential GPT'.

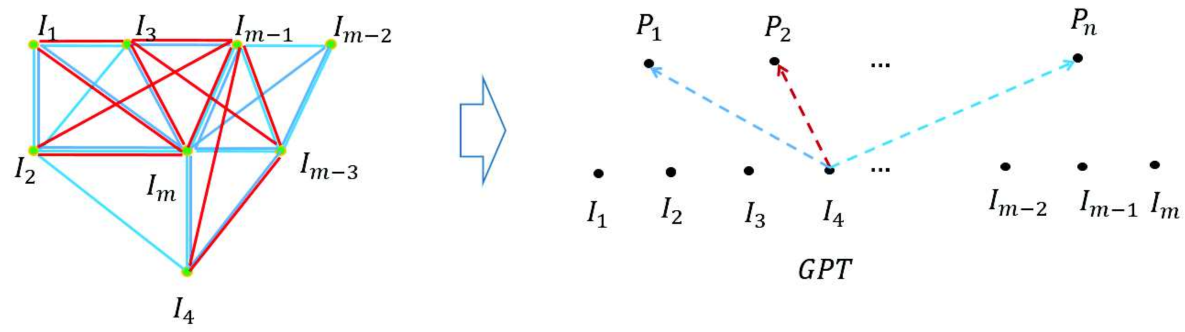


Figure 3: Potential technology network and GPT

2.2 Discovery process

The process of knowledge (and eventual product) discovery is the process of satisfying consumer needs. To keep the demand side simple, we consider each product type having a certain value (V) proxying an expected profit from its discovery. These values are the driver for profit-oriented agents (anyone able to conduct R&D: firms, entrepreneurs, scientists etc.) to conduct the discovery process upon the technology network. In the baseline model agents are considered to be able to see all alternative ways of production (thus, setting the whole potential technological network to be *visible* to all agents), an assumption that is necessary to test the mechanics of our model and which is to be relaxed later (see Section 3.3). Discovery of each technology combination has certain difficulty (d) – resistance of the link to be discovered.[10] This difficulty is not known to agents so that agents can only compare alternative trajectories within one product type in terms of number of links yet to be discovered. The latter introduces *uncertainty* into the model since best strategies are unknown, and agents can at most rank opportunities. In the model, assignment of the values, the difficulties and the technologies are independent. Thus, it may turn out that new knowledge necessary for a very valuable product type can be invented with a small application of effort (e.g., as penicillin discovered accidentally by Alexander Fleming) and the other way around. Also, the GPT is not necessarily attributed with more or less difficult technological links, distinguishing the present model from the existing studies attributing an ex ante advantage to GPT,[11] while the only virtue of a potential GPT we allow is its a priori larger scope of application.

Over time, agents try to discover a certain technological combination from those being visible for each product types, where the order of the products to be considered is random and set anew each cycle. The effort applied is equally distributed among all yet undiscovered links so that once one of the constituent links becomes discovered, the effort is redistributed among the remaining ones creating a cascade effect of product discoveries in time (increasing number of innovations per period over time, see Figure 4).

[10] Note that this does not necessarily introduce a discrete complexity ladder: goods consisting of 3 or 4 technologies would require 3 and 6 technological combinations, respectively, to be discovered. One, however, can smooth the product complexity by randomly assigning zero difficulty values to a certain fraction of edges. We conduct such an exercise as a robustness test.

[11] In perhaps the most related to us studies by Bresnahan (2012, p. 629) combinations of technologies (products) also have values and 'there are two potential ways to create new value': a 'compromise' way does not involve GPT and has lower value than an 'efficient' one including GPT. Thus, the model assumes a higher expected profits to production of a good with GPT pointing out that generality is expensive.

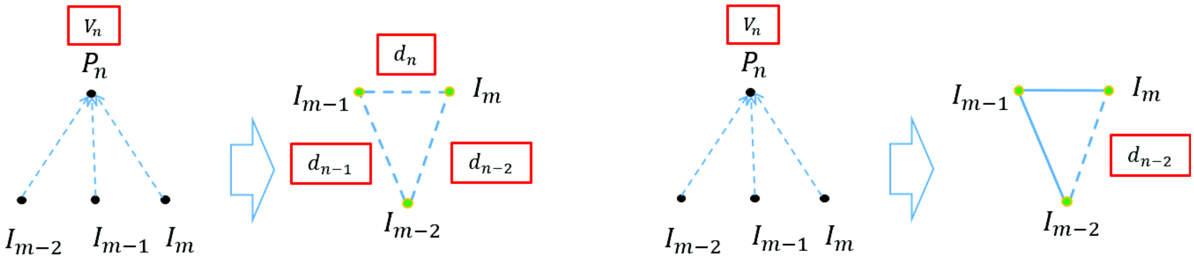


Figure 4: R&D effort redistributed among undiscovered links (dashed lines)

The probability to discover a certain technological link x being a part of the product type y discovery is modeled stochastically as a uniform random number $Pr_x \in U[0,1]$ and turns this link into a discovered one *if*:

$$Pr_x < \frac{V_y}{d_x \times L_x} \tag{1}$$

where L_x is the total number of yet undiscovered links in the clique in which the link x is located. Hence, the higher the product type value or the smaller the resistance of the respective link or the smaller the number of yet undiscovered links in the respective clique, the higher is the chance of that link to be discovered.

The described mechanisms introduce a strong *path dependence* in terms of past decisions and outcomes (which cliques to concentrate effort on and which links become discovered earlier) driving further results (which technology combinations become invented first). Note that given that the present study is a model of discovery, once a certain product type is discovered along one of its technological trajectories, the related pressure from the demand side disappears. We are only interested in first product type discoveries and those are analyzed in terms of GPT adoption. Though the history of innovation has many examples when new products were displacing the existing ones (smart phones against standard mobile phones, alternating current against direct one or VHS against Betamax), this has normally had to do with functional superiority (where it becomes increasingly difficult to compare goods in satisfying exactly the same need) or cost advantage, which are not the focus of the present work. In contrast, we argue that if a technology becomes adopted in as many products as possible at the period of first invention, this does not only give it time and cost advantages but also allows it to become a new GPT.

We model agents in a very simplified way assuming no heterogeneity or interaction among them.[12] Once certain knowledge piece is discovered, it is upon the knowledge property, and not the agents, whether everyone or none of them gets access to this knowledge. Similarly, coordination is made not with respect to which agents shall try to discover which technology link, but in terms of which technology clique one shall try to discover first (see Section 2.3). Thus, one can think of a 'representative' agent having the same incentives (V) and difficulties (d) in R&D process.[13] Also, no budget constraint for the agents is considered.

[12]Similar to production costs, we leave this aspect aside of the model to concentrate on the technology network effect first. In an extension, it will be certainly interesting to explore the issue of heterogeneity and interaction among those agents.

[13]Alternatively, one may think of a number of agents with a perfect information flow that act one at a time and all newly discovered knowledge becomes immediately available for everybody.

The notion of time is also present in our model and is kept simple. In particular, at period $t = 1$ agents start discovering new technological combinations (as aforementioned, none of them is present at the beginning of the simulation) and at each period can apply effort only to *one* way of producing a product. In this way, the model runs until all visible product types become producible (discovered in one of the production ways).

2.3 Factors affecting GPT adoption

Knowledge diffusion

One of the key questions to address in the case of knowledge discovery process is whether and to what extent does this knowledge diffuse to other products. As it has been argued above, some technology combinations can be utilized in more than one good and more than one way of production. A relevant question in such a case is whether the link between the two technologies I_{m-1} and I_m being discovered once (i.e. for one way of producing the respective product type, see Figure 5) opens this link for any other way of technological combination or product type.[14] In the technology network context such a knowledge flow is contingent upon two conditions:

- functional similarity in combining the two technologies is sufficiently high to apply the same knowledge to other contexts: in the example of lenses and optics it means that this knowledge is directly applicable in cameras, telescopes, microscopes etc. This leads us to the discussion on technological standards and dominant design (for an overview, see Abernathy and Clark (1985) and Anderson and Tushman (1990)), where being discovered once certain technology combination becomes *universal* and does not have to be rediscovered for other purposes (e.g., GPS usage from military to civilian applications and from weather forecasting to time synchronization);
- the knowledge discovered flows freely within the population of agents, i.e. there are no firm- or institutional-based barriers preventing the flow of knowledge (so-called knowledge spillovers). This condition addresses the public good property (i.e., not appropriated by the owner) of knowledge coming back in the literature to at least Arrow (1962). The magnitude of this property depends on the extent to which it is codified and the effectiveness of the mechanisms by which knowledge is protected, including the appropriability conditions (Dosi, 1982).

We distinguish between three main regimes of knowledge diffusion (see also Figure 5):

1. *sticky knowledge*. In this regime there is either no functional similarity between products, or no knowledge spillovers preventing the possibility that knowledge discovery for one particular product (one of its production ways) can contribute to a discovery of any other product containing the same link;[15]
2. *partially sticky knowledge*. In this case, while the functional similarity between goods is still limited, the flow of knowledge is not. To distinguish that regime from the previous one, we thus make an assumption that limited functional similarity

[14]Note that in the present study the only mechanism of knowledge spillovers between any two products is thus contingent on the functional (technological) similarity between those two goods sharing those bits of knowledge connecting the respective technologies.

[15]In such a case, inventor literally has to ‘reinvent the wheel’ for every new product.

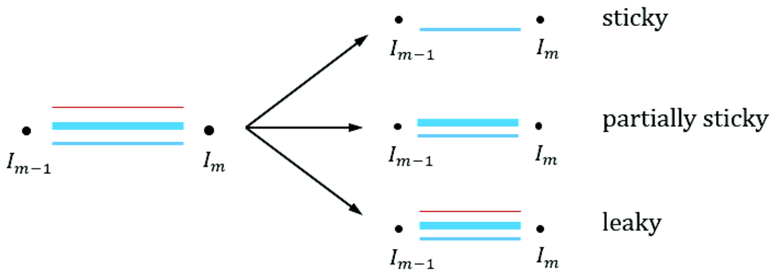


Figure 5: Knowledge diffusion regimes

allows to apply the discovered knowledge to other product types but only within the respective way of technological combination. This should reflect the intermediate status of the regime with imperfect knowledge diffusion;

3. *leaky knowledge.* This is the regime with perfect knowledge diffusion – once certain technological link is discovered in any specific product, it becomes available in all product types across all ways of technology combinations.

It is worth stressing that while we consider the aforementioned regimes of knowledge diffusion to be the result of innovation policy (affecting those through the technological standards and appropriability conditions), we treat those regimes as exogenous in our model, separately considering each of the three scenarios and analyzing implications for the emergence of GPT. In particular, we make the following proposition with respect to the effect the knowledge diffusion has on GPT:

Proposition 1 *The larger the extent of knowledge diffusion, the more likely that potential GPT becomes an input of many different product types at the stage of their discovery.*

Proposition 1 has the intuition that GPT, having in the present study the only distinct property of highest pervasiveness resulting in a large number of links connecting it to many other technologies in the network of intermediates, is also expected to have the largest number of links entering more than one product type in more than one way of production and, thus, must be the major beneficiary (among technologies) of the knowledge diffusion process.

Coordination of R&D efforts

Another mechanism, which may play a considerable role in technology adoption is the decision heuristic of agents on which way of technological combination they shall concentrate while trying to discover any of the product types. As stated earlier, the agents do not know the difficulty of discovering a technological link and, therefore, can choose between the technological trajectories taking only the number of yet to be discovered links into account (i.e. accounting for the accumulation of the problem-solving capabilities). However, the choice in favor of 'smaller' cliques (with least number of links yet to discover) may not always be optimal. First, given the strong uncertainty with respect to the difficulty of links, some cliques being larger in size may still be easier in terms of effort to be applied to discover. Second, agents may prefer *knowledge breadth* over *knowledge depth* because of the interconnectedness between technological problems and the potential to utilize the gained knowledge in other applications.

In the model we introduce the factor of coordination in R&D effort through a logistic function determining the probability of the respective trajectory to be chosen by agents:[16]

$$Pr_i = \frac{e^{10\beta(L-L_i)}}{\sum_j^W e^{10\beta(L-L_j)}} \tag{2}$$

where parameter $\beta \in [0,1]$ varies the scenarios from no (in favor of knowledge breadth) to perfect coordination (knowledge depth) , L stands for the maximum number of links to be discovered across the ways of technological combination W and L_i is the number of yet undiscovered links in the trajectory i. This is illustrated in Table 1. Clearly, with $\beta = 0$ trajectories are chosen randomly without any account for already accumulated knowledge, while for $\beta = 1$ agents always will concentrate on the smallest clique. Intermediate values of β will squeeze probability distributions towards cliques with least number of undiscovered links.

Table 1: An example of how probabilities are distributed for different β

L_j Pr_i	4 edges P_1	2 edges P_2	3 edges P_3	4 edges P_4	5 edges P_5
$\beta = 0$	0.2	0.2	0.2	0.2	0.2
$\beta = 0.2$	0.019	0.85	0.11	0.019	0.002
$\beta = 1$	0	1	0	0	0

We propose the following proposition with respect to the coordination of R&D efforts:

Proposition 2 *Preference for knowledge depth over its breadth fosters adoption of GPT under condition that knowledge spillovers are present between different technological combinations for different product types, and that those spillovers do not change over time.*

Proposition 2 consists of three parts: the first one conjectures that under no knowledge diffusion between different problems agents' coordination on any trajectory is purely random;[17] the second part postulates that in the presence of knowledge diffusion coordination may make agents to switch the trajectory in favor of the one with positive externalities[18] in the form of accumulated knowledge from a different product type. Since our GPT is potentially the most pervasive technology, those positive externalities are expected to be largest for technological trajectories containing it, resulting in a higher adoption of the potential GPT in first product type discoveries; finally, the third part ensures that those spillovers do not change over time: imagine one discovers a technological link which can be utilized in this period in one way of technological combination and one product type only, but many years later people find a different application for this technological

[16]Thus, at one period of time effort can be applied only to one technological trajectory in one of the product types. The fact that agents may not all coordinate in pursuing one technological trajectory is, thus, represented by no coordination.

[17] Remember that earlier we assumed no relation between values, difficulty and technologies involved. Hence, in choice between two ways of production with the same amount of links agents will be indifferent, otherwise they pursue trajectory with smallest amount of edges given R&D coordination.

[18]In the words of Dosi (1982, p. 154) "complementarities among trajectories".

combination in a different product type not considered before. In such a case, time gains in importance in our model and while for invariant knowledge spillovers discovery of the link (or more generally, pursuing the technological trajectory with this link) may look not as attractive originally, this changes if the spillovers alternate over time.[19]

Potential technology network structure

Another core factor affecting the knowledge discovery process is the structure of technological network. Apart from the number of product types (N), intermediates (M) and the ways of technological combination (W), this shall be affected by at least two more parameters: number of intermediate technologies in each technological clique forming a technological combination (clique size, CS) and the pervasiveness of the present intermediates within the product types. To keep the modeling simple, we assume in the baseline model that all product types in all technological combinations consist of the same number of intermediates,[20] while pervasiveness of other technologies is modeled via two opposite views. In particular, while GPT, as discussed earlier, per assumption potentially enters all product types at least once and has the highest *potential* pervasiveness, other technologies (from 2 to M) may either all be very similar or very different in this respect. Based on the latter distinction we formulate the third proposition:

Proposition 3 *The larger the difference between the potential GPT and other technologies in terms of their technological pervasiveness, the more likely that the GPT becomes an input of many different product types at the stage of their first discovery.*

Proposition 3 is based on the intuition that the less potential synergy is concentrated between non-GPT technologies, the easier it must be for the GPT to fulfill its potential. Similar to knowledge diffusion, we consider the technology network structure as an exogenous factor. However, we do not argue that a policy maker may have an impact on technology network structure, as it represents the knowledge space itself; rather this network structure could be indirectly identified in order to adjust policy decisions.

Changes in expected profits

Finally, one may expect some effect on GPT adoption from the demand side. The expected profits for each product type proxy the priority from the side of society (both, consumers and policy makers) on which needs shall be satisfied first. Thus, any change in the rank of priorities can reflect either changes in preferences or institutions.[21] As an example of enforced preference change let us take the one considered by Cowan (1990) on the nuclear power reactors. Because of the Cold War and fierce competition with the Soviet Union for the technological leadership, the U.S. government was heavily subsidizing the nuclear industry in the end of 1950s to foster building of the first commercial prototype and securing the global market. However, to enable such a swift discovery of

[19] In other words, we follow Carlaw and Lipsey (2006, p. 159) in that because of Knightian uncertainty agents do not have a "foresight about an unknowable future" and take decisions based on the externalities "as being constant at the current period level".

[20] In the robustness checks we relax this assumption highlighting that the main results remain valid.

[21] Institutional arrangement change incentives of entrepreneurs and investors to develop new products. For example, policy instruments introduced in the German energy sector made it profitable to concentrate on the renewable energy technologies (above all, wind and solar) (Herrmann and Savin, 2016).

the product type, a critical decision with respect to the preferred technological trajectory had to be taken (in this case between light water, heavy water and gas graphite). Given that "typically [...] when a technology is introduced its future payoffs are not well known" (Cowan, 1990, p. 544), the choice has been made mainly based on knowledge accumulated by the U.S. Navy adopting the light water for submarine propulsion. As history has illustrated, due to that exogenous shock introduced by the policy maker the market eventually became locked into the inferior technology.

To examine such an exogenous effect on the knowledge discovery process and the adoption of a potential GPT, but at the same time keeping the model simple, we allow after a fixed number of periods (throughout the experiments we keep it equal to 100) for a certain fraction of product types to exchange their expected profits, proxied by parameter *Value Dynamics* (VD) between 0 to 100%.[22] Thus, some less 'valuable' needs may instantly gain in priority and the other way around. All other characteristics of the model remain unchanged. Having introduced this mechanism in the model and keeping the example described by Cowan (1990) in mind, we formulate the following proposition:

Proposition 4 *Frequent changes in the rank of product type expected profits negatively affect the adoption of GPT and may lead to a technological lock-in in the long term.*

The intuition behind Proposition 4 is that due to instant change in the product type's expected profit its discovery becomes faster and essentially random with respect to the technological trajectory chosen, leaving not enough time to take an advantage of positive externalities through the knowledge diffusion. Thus, we conjecture that those changes in the rank of priorities diminish the effect of knowledge diffusion combined with coordination of R&D efforts and may lead to a technological lock-in.[23]

3 Numerical Analysis and Model Extensions

In what follows we describe how we set up the numerical experiment and which parameters we use as a default (Section 3.1). Afterwards, results of the simulation exercises (Sections 3.2-3.3) and robustness tests (Section 3.4) are presented.

3.1 Numerical experiment

At the beginning of each simulation restart, a large network of potential technological interconnections has to be generated. For this a subset of technologies (of the size CS) has to be sampled for each product and each of its ways of technological combination. In doing so, three conditions are ensured:

1. The sampling replicates one of the two sampling functions, which are chosen in line with Proposition 3. In particular, both sampling procedures start from ensuring that potential GPT enters first (1 out of W) technological combinations for each product type. Afterwards, one either follows a highly skewed distribution function

[22] This is done to prevent any volatility in the overall amount of effort the agents can apply to discover all product types in at least one production way.

[23] To address the possibility of a technological lock-in, we define as a *lock-in* the situation where the process of knowledge discovery is hampered (e.g., lower number of technological links is discovered), which eventually leads to no or delayed product discovery.

or sets the sampling probability of them equal. Analytically this is achieved by following one of the two probability distributions, respectively:

$$Mprobability_1 = \frac{1}{\sqrt{seq(1, M, M)}} \text{ or } Mprobability_2 = (1, seq(pm, pm, M-1)) \tag{3}$$

where $seq(a, b, l)$ generates a vector of equally distant elements between a and b of size l. As a result, the sampling function to the left in equation (3) creates highly skewed distribution, where the potential GPT still has the largest scope of application and is followed by a small subset of 'competitor' technologies also pretending to become included in many different product types and technological combinations. The sampling function to the right in (3), in contrast, generates a single 'champion' with other technologies having (approximately) equal chance[24] to be included in any technological combination. Needless to say that no technology can enter any technological combination more than once.

2. After all W technological combinations for all N final product types are constructed, they are rearranged randomly to ensure that GPT is equally present in all of them.[25]

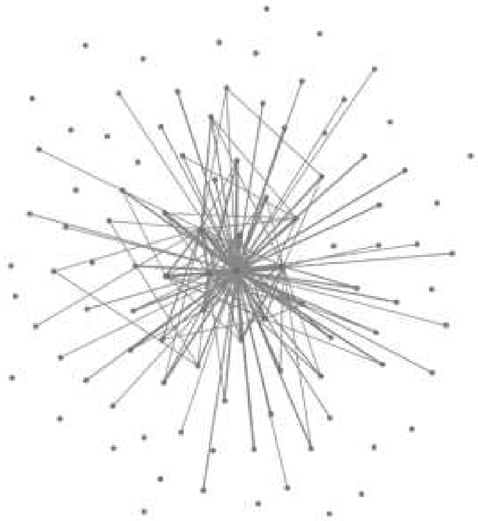
3. While creating the technological combinations, the code ensures no combination is repeated. The motivation behind that is to keep at least moderate technological differences between discovered goods in our model.

The exercise results in a complex weighted network, having both the bipartite (product-technology; presented in Figures 1-2) and multiplex (W alternative technological combinations consisting of the same number of nodes and links but having different link allocation; Figure 3) structure. As default values we consider the number of product types $N = 60$, number of intermediate inputs (technologies) $M = 100$, five ways of technological combination ($W = 5$) and four technologies to be recombined per product ($CS = 4$) so that the resulting network of possible technological links is a highly interconnected graph. To illustrate the difference between the two alternative sampling approaches described above, we examine the two network structures by applying a method similar to the k-core analysis in graph theory filtering edges with a degree below k (in this case smaller than 5).[26] This allows one to concentrate only on those edges which enter several product types. Clearly, in the case of equal pervasiveness potential of other $M - 1$ technologies, a 'star-type' network structure is observed (see right plot in Figure 6), where almost all 'heavy weighted' links lead to GPT, while in the alternative network structure there is a highly interconnected core of five-ten technologies including the potential GPT (left plot in Figure 6), which are also well connected to technologies outside the core (periphery). For this reason, henceforth we denote the two alternative network structures as 'star network' and 'core-periphery network'.

[24] This is proxied by the parameter $pm = \left(\sum\left(\frac{1}{\sqrt{seq(1,M,M)}}\right) - 1\right)/(M-1)$ chosen just to ensure that in both sampling functions GPT has the *same* potential pervasiveness (number of times being sampled for distinct technological combinations. For example, for $M = 100$ $pm \approx 0.178$.)

[25] This is primarily done to avoid any strong assumption that GPT may benefit a lot from limited knowledge diffusion within just one way of technological combination.

[26] The exact value of k is chosen just for visualization convenience. For different parameters of the network, some different value of k may be chosen instead.



a) *core-periphery* type

b) *star* type

Figure 6: Potential technological networks after filtering links with a low weight

For the resulting networks we then randomly distribute values (among final product types) and difficulties (among links). Afterwards, starting from period $t = 1$ agents apply R&D effort in a sequential order to discover final product types. To avoid any effect from specific product type order, each *cycle* the ordering of not yet discovered product types isn rearranged randomly anew. For the basic model described, R&D agents continue inventing new technological links until for each product type at least one way of production is discovered.[27]

To start exploring our basic model with regard to Propositions 1-4, one first has to fix some further parameters we use. We assume expected profits of product types to be exponentially distributed with the parameter rate equaling 10, while the difficulties to discover each of the links are normally distributed with $\mathcal{N}(100, 25)$. These parameters, thus, are chosen to keep the numerical simulation sufficiently fast while avoiding discovery of many technological links within one cycle. Given the stochastic nature of the model and unless specified otherwise, in what follows results are reported for 50 restarts.

In describing the results, we primarily look on the (actual) pervasiveness of the potential GPT (percentage of first product type discoveries where GPT becomes an input). Furthermore, to account for the fact that for different network parameters (like M, W or CS) the potential of GPT to enter all products *relative* to the potential of other technologies varies, we introduce an additional indicator, called $GPT_{Adoption}$, measuring to what extent GPT has fulfilled its potential in comparison to an average other technology in the technological space doing the same:

$$GPT_{Adoption} = \frac{Actual\ Pervasiveness_{GPT}}{\frac{1}{M-1}\sum_{m=2}^{M} Actual\ Pervasiveness_m} \Big/ \frac{Potential\ Pervasiveness_{GPT}}{\frac{1}{M-1}\sum_{m=2}^{M} Potential\ Pervasiveness_m} \quad (4)$$

Thus, $GPT_{Adoption}$ indicates not just how much more pervasive GPT has become in

[27] As stressed before, any subsequent technological combinations, which can be discovered as a byproduct of the R&D process directed on discovery of different product types, are not taken into consideration.

comparison to an average other technology (after the discovery process is finished and one calculates the 'actual pervasiveness'), but compares this ratio with the one using 'potential pervasiveness', i.e. in how many different technological combinations a given technology had a potential to be included.[28] Additionally, we report information on the discovered network size (in terms of number of links discovered) or amount of time spent by agents, which complement the picture on the intuition behind the results we obtain.

3.2 Results of the basic model

On Figure 7 we observe the effect of the extent of knowledge diffusion and coordination of R&D efforts on GPT pervasiveness and adoption. Start from the case of no coordination ($\beta = 0$): the more *leaky* is the flow of knowledge among technological combinations, the more pervasive is GPT and the better is fulfilled its potential. New knowledge embodied in discovered technological edges and applicable in different technological combinations becomes available for agents working on different technological problems and enforces earlier discovery of products containing larger proportion of links with such a multiple application. GPT is the main beneficiary of that 'knowledge propagation' process due to network structure where by definition it potentially has the largest amount of technological links used in more than one product type. Thus, with leaky knowledge and no coordination GPT becomes a part of a much larger number of new products, while in comparison to an average competing technology GPT fulfills its potential 1,3 times better. This result fully supports Proposition 1.

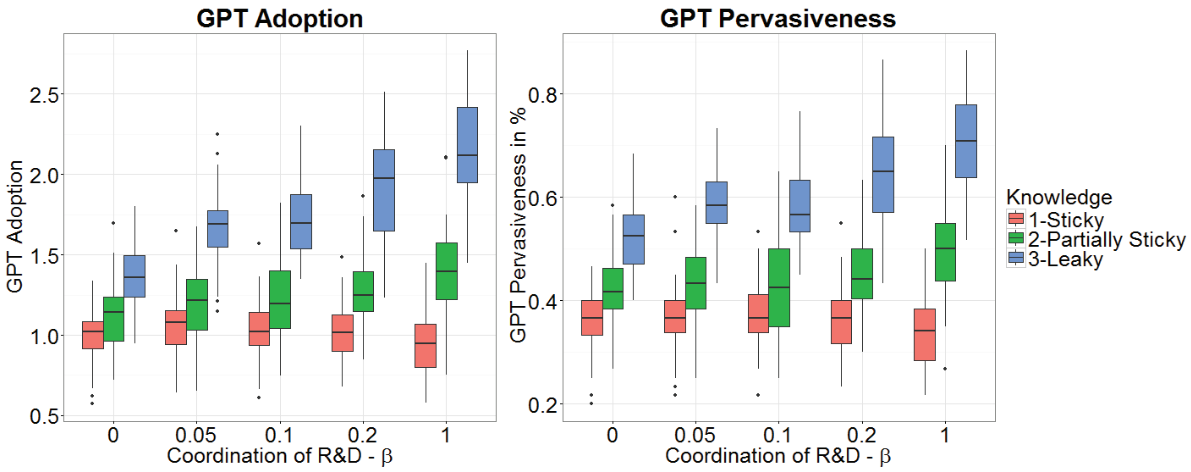


Figure 7: The effect of knowledge diffusion and coordination on GPT adoption

Note: This result is produced for core-periphery type network under no dynamics in product values. The network parameters used: $N = 60, M = 100, CS = 4, W = 5$. A similar result is achieved for the star type network structure (see Figures 22-23 in Appendix).

Furthermore, given that the knowledge diffusion propagates discovered solutions (technological links) to many other applications, it is worth testing whether coordination of R&D can strengthen GPT adoption under leaky knowledge even further and whether this is contingent on the presence of knowledge diffusion (Proposition 2). For this, we

[28] Take an example: GPT had the potential pervasiveness of 60 and other technologies on average only 10, while actual pervasiveness is 30 and 3, respectively. The resulting value of $GPT_{Adoption} \approx 1.67$ implies that in comparison to its 'competitors' GPT has fulfilled its potential 67% better. Note here that the indicator equaling 1 means that technologies have fulfilled their potential equally well.

vary the β parameter between 0 (preference for knowledge breadth) and 1 (preference for knowledge depth) for the three different knowledge diffusion regimes. Clearly, under leaky knowledge, an increase in coordination contributes to a larger GPT pervasiveness and adoption. The more 'sticky' is the knowledge, the smaller this contribution is until it vanishes completely confirming our Proposition 2.[29]

To understand the effect of network structure on GPT pervasiveness, first look at Figure 8 demonstrating how the variation in network parameters affects GPT pervasiveness under *ceteris paribus* principle for *core periphery* and *star* network structures. Increasing the number of technologies M naturally reduces the density of the technological network, thus, lowering the externality effects that favor GPT (top left chart in Figure 8). A similar result with a level-off effect is obtained if we increase number of alternative ways of production (top right chart in Figure 8). Here the explanation is also simple. The more alternative ways of production we have the harder it is for a GPT to become pervasive. A level-off effect appears because we keep ratio of products to technologies constant and at some point GPT starts to enter not one but several ways of producing the same product type in a potential network increasing the variance in the outcomes. An opposite trend is observed if one increases either the number of products (N, bottom left chart in Figure 8) or the number of technologies each product can be made of (CS, bottom right chart in Figure 8): as the network density rises leading to larger externality effects, GPT becomes adopted in a larger number of final goods.

Hence, two conclusions can be made. First, one can observe little difference between two alternative network structures, namely *core-periphery* and *star* types of network, thus, rejecting Proposition 3. Second, the more dense is the network in terms of the amount of weighted links the more likely is the GPT adoption. By 'density' here we mean the amount of links with weight larger than 1. It is clear that a typical definition of network density employed from graph theory will not fit to our type of problem. This definition says that density is a ratio of existing links to all potential links. In our set up we are more interested in which links lead to GPT and which do not. Thus, we construct an index that sums the differences in those occurrences in favour and against GPT adoption, weights it according to the likelihood to encounter in the technological network and normalizes it to the total number of unique links in that network (Ψ):

$$Multiplicity\ Index = \frac{\sum_{\psi=1}^{\Psi} \omega_\psi \left[\max(\omega_\psi^{GPT} - 1, 0) - \max(\omega_\psi^{NoGPT} - 1, 0)\right]}{\Psi}, \qquad (5)$$

where ω_ψ is the number of occurrences of a unique link ψ (the same pair of technologies) in our network of potential technological edges, ω_ψ^{GPT} is the number of times this link leads to cliques containing GPT and ω_ψ^{NoGPT} is the number of times the same link leads to cliques without GPT. In this way, we attempt to capture the effect of knowledge externalities between competing technological trajectories in our model. The larger the resulting *Multiplicity Index* the larger is expected to be the actual GPT pervasiveness.

Figure 9 illustrates how actual GPT pervasiveness depends on the index. Again little difference can be observed regarding two contrast network structures. The dependence is

[29] Note here that varying the extent of knowledge spillovers, we keep those constant in time implying that if a technology combination becomes discovered and the knowledge spillovers regime allows the link to be applied elsewhere, this externality was taking place immediately (without any time lag). In an extension of our model (Section 3.3) we illustrate how such a time delay (in terms of knowledge externalities to be utilized) can be taken into account.

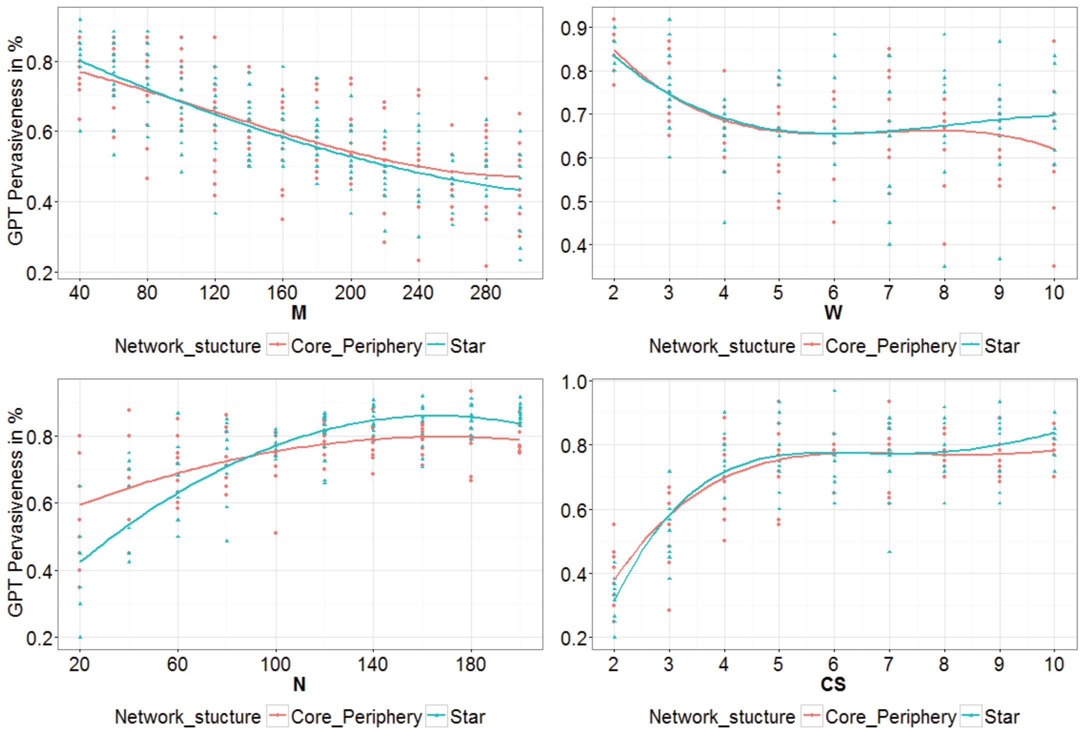


Figure 8: The effect of variation in the number of products (N), ways of production (W), technologies (M), and a cliques size (CS) on the GPT pervasiveness.
Note: We fit two polynomial lines for a better illustration purposes.

not linear and it can be seen that once the index exceeds a value of 1 GPT pervasiveness levels off around 80%. It is important to note that the index reacts to variation in all key network parameters discussed earlier and can serve as a good ex ante estimate of GPT adoption under leaky knowledge and coordination.

Finally, we explore the effect of variation in expected profits by setting the VD parameter equal to values between 0 (no variation) and 1 (all N product types change randomly their rank in expected profits every 100 periods). Results of the exercise are presented in Figure 10. The absence of a clear effect on GPT adoption has an explanation. In our model the demand side is interested in discovery of products (to be precise, first discovery for each product type satisfying a certain need), but puts no difference on which inputs shall be used to do so, leaving this choice to agents doing R&D. The agents, in their turn, pursue trajectories with lowest expected difficulty. As a result, this variation in expected profits has close to no impact on the agent's discovery choices. Hence, as long as we keep the network of potential technological interrelations fixed and knowledge spillovers constant over time, one has to reject Proposition 4.

What the variation in expected profits does affect, however, is the period of time within which all product types are discovered in at least one technological combination (see right chart in Figure 12). Given that a high pressure from the demand side rotates between different product types over time, some more difficult edges become discovered much faster reducing the overall amount of time spent. A similar effect on the time of discovery have the knowledge diffusion and coordination of R&D efforts. Right chart in Figure 11 demonstrates how time reduces with increased coordination, and within each value of coordination increased diffusion reduces time as well. Yet, the nature of those two effects is different. In the former case, present knowledge diffusion stands for the

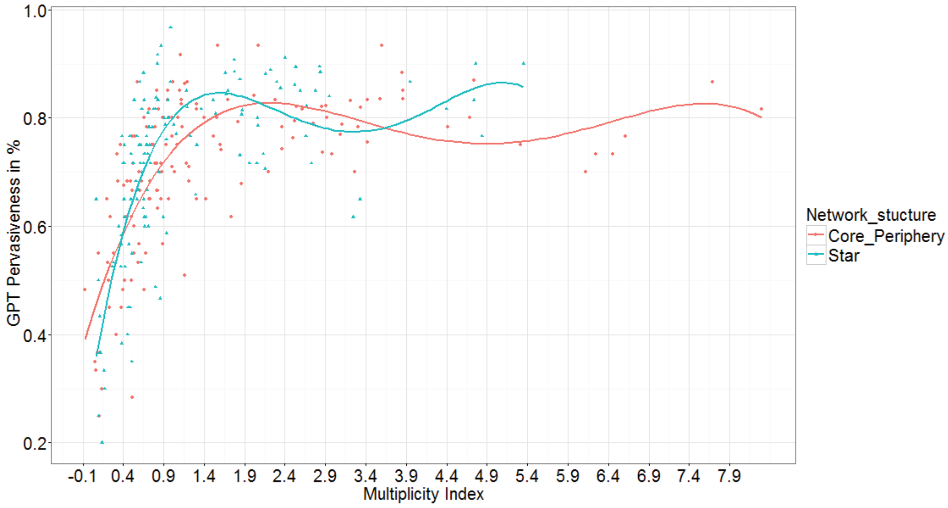


Figure 9: Multiplicity index reflecting resulting GPT pervasiveness under leaky knowledge

Note: We fit two polynomial lines for a better illustration purposes.

possibility of utilizing knowledge discovered elsewhere for a specific technological problem at hand. The latter force leads to focus on technological trajectories where knowledge is already accumulated and results in faster invention.

Another informative result is the size of the discovered graph reflecting the amount of knowledge accumulated. Knowledge diffusion logically increases the knowledge base discovered. Left chart in Figure 11 shows that under all coordination regimes ($\beta = 0...1$) the diffusion has a positive impact on knowledge base. Even though agents were not aiming to discover all possible applications of a unique technological combination, this is done automatically.[30] The same chart demonstrates that coordinating R&D efforts (focus on knowledge *depth*) reduces the discovered base because agents always follow the (seemingly) 'least resistant' clique not trying to discover edges in alternative ways of production of the same product. Finally, left chart in Figure 12 demonstrates the negative effect of variation in expected profits on the amount of accumulated knowledge, which is due to the high pressure from the demand side leading to fast product discovery preventing agents to work more on different technological trajectories. This result is important to understand our findings for the technological network growing over time.[31] Note also that the *core-periphery* type of network has an advantage over *star* type due to higher concentration of edges with large weights: those weighted links being discovered greatly add to the overall knowledge base given *leaky* diffusion regime.

Thus, one could conclude that in order to invent all products in a fastest way and promote adoption of potential GPT, one shall promote knowledge diffusion, stimulate agents to concentrate their innovative efforts on the technological trajectories with largest amount of accumulated knowledge and in parallel stimulate rotation in the demand side pressure towards discovering distinct product types. Yet, as we show in Section 3.3, such a conclusion would be too delusive in the long term perspective.

[30] If in contrast, we would have counted only all unique edges (between unique pairs of technologies), the presence of knowledge spillovers would result in the smallest network discovered.

[31] Note that when coordination of R&D efforts is switched on, variation in profits has no clear effect on the discovered knowledge base since under coordination agents quickly start disregarding alternative trajectories.

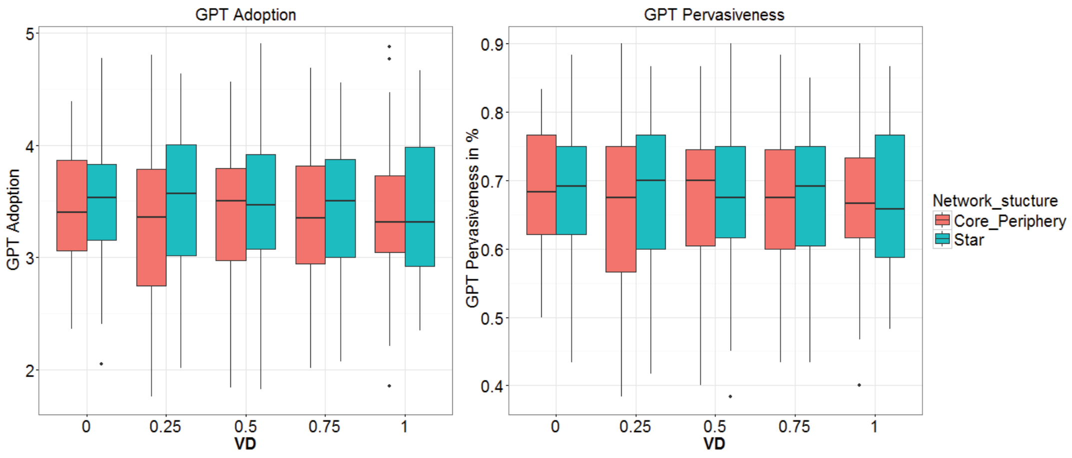


Figure 10: Effect of value dynamics on GPT adoption

Note: The result is obtained under leaky knowledge and full coordination ($\beta = 1$). A similar result with quantitatively smaller values for GPT adoption is observed for no coordination ($\beta = 0$).

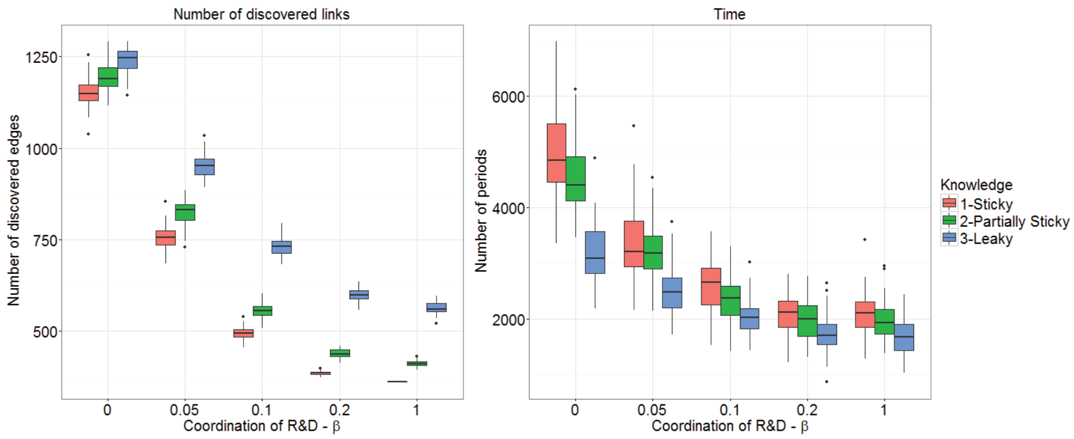


Figure 11: Effect of knowledge diffusion and coordination on time of discovery and discovered graph.

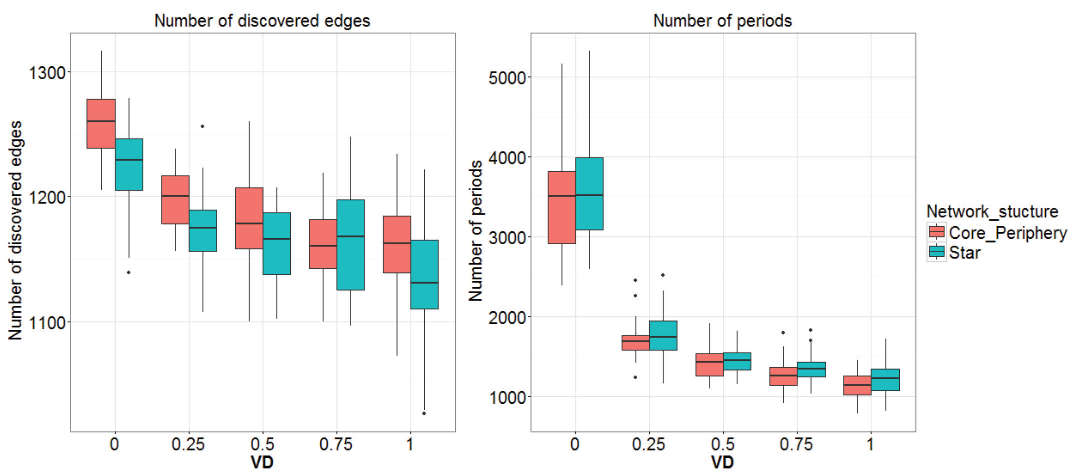


Figure 12: Effect of value dynamics on discovered graph and time of discovery.

3.3 Growing technological network

3.3.1 Modeling the arrival of new ideas

Up till now in our model all ways of technological combinations were visible to agents ex ante and the discovery process was continuing till each product type has been discovered. However, thinking about innovation process in a dynamic context one might say that during the course of technological progress we come up with new ideas of new products and new ways of technological combinations that we were not aware of. Therefore, in the following we relax the assumption of fixed number of technologies (as in Section 2.2) and allow the visible network to grow (both in terms of number of visible technological combinations for a given product type, but also in terms of new product types/needs arising) calling this scenario '*growing technological network*'. This scenario is logically close to the description by **?**, p. 140 of an economy as a complex evolving system, where "structural change is [...] a chain of consequences where the arrangements that form the skeletal structure of the economy continually call forth new arrangements".

We implement this extension into the model by adding a third layer to our multiplex technological network (Figure 13). The model then constitutes *discovered* network (consisting of combinations already discovered by agents), *visible* network (links that agents become aware of, i.e. *realize* those links as we will call henceforth; so far we were considering it to be the entire potential network and fixed over time) and a third *potential* network (all possible technological combinations, including visible ones but also those that agents are not yet aware of).[32] Thus, invisible network contains hidden ideas on new possible technological recombinations of existing but also new product types.

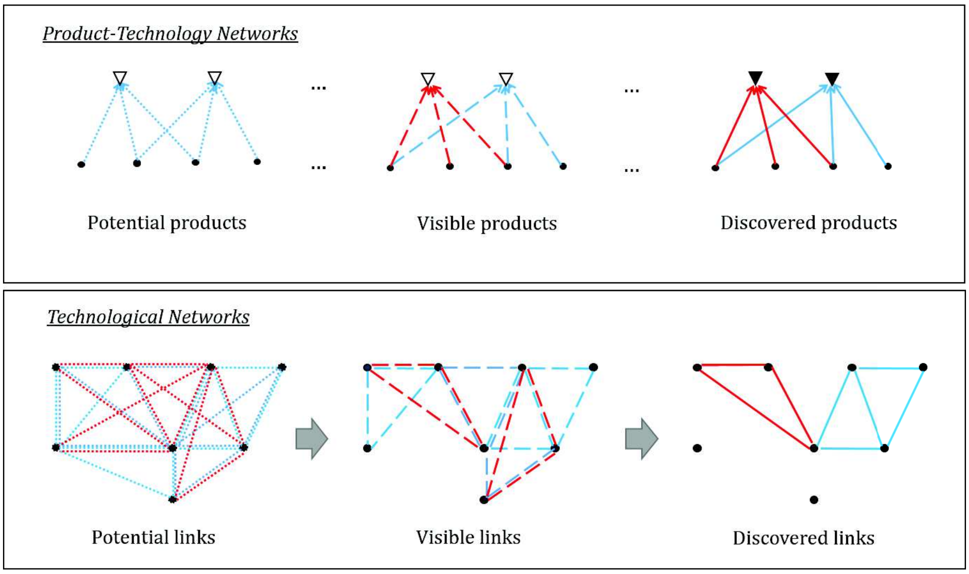


Figure 13: Three states of technological links

While the edge transition from visible to discovered state has been addressed in detail in our baseline model, here we discuss the transition from the invisible to the visible state. In other words, we model the arrival of new ideas to our agents. This process is contingent on the knowledge being already accumulated by them. Thus, the growth of a

[32] Obviously, the latter network is most general one, while the former two represent its fractions (discovered network - part of the visible one, while visible part of invisible potential network).

visible network is highly dependent on the size and structure of the network discovered. In particular, agents tend to learn about new possible product types or new ways of production of known product types depending on the extent they are using constituent technological combinations. One important difference of the mechanism making links visible to the one transforming them into discovered ones is that edges become visible in cliques, while links become discovered individually through practical tests more like an applied knowledge. The second difference comes from our assumption that the process of recognizing new technological combinations requires no R&D effort from the agents.[33] In particular, agents can recognize a new technological combination υ at each period of time if the probability $Pr_\upsilon \in U[0,1]$:

$$Pr_\upsilon < \alpha \exp\left(v(s_1 - 1)\right) + \beta \exp\left(v(s_2 - 1)\right) + \gamma \exp\left(v(s_3 - 1)\right), \tag{6}$$

where s_1, s_2 and s_3 are shares of discovered, visible and invisible links[34] in the technological clique υ, while α, β, γ, v are parameters specifying the function's shape so that it increases exponentially the more links in υ are visible and discovered by agents.[35]

Let us consider an example here. Suppose a product consists of four technologies, which makes a clique size of six (edges). Now suppose that one of these links is also used in another product type which is already discovered by the agents and yet another link is used in a product that is visible but not discovered. In such a case, we calculate a total number of links being equal to 8 and shares of invisible, visible and discovered links as 6/8, 1/8, 1/8 respectively. Thus, if agents are aware of the fact that a pair of technologies has an innovation potential, or they have already discovered that link, it is more likely they will once recognize that there is a new product type containing this link. Note here that even if all edges in a clique are completely unknown to agents this probability is different from zero. We can't deny the fact that there is always a chance of the arrival of new radical idea from different technological paradigm. This chance increases if all respective links are already visible in different products, and by the time 100% of those links are discovered the clique υ becomes visible with certainty. Figure 21 in the Appendix illustrates the equation (6) for different shares of s_1, s_2, s_3. In the words of **?** or Nelson and Winter (1982) agents search locally for new knowledge trying extensions of existing one close to what they already possess and use in some space of technological characteristics. The model now runs until agents discover all product types that they see.

The exercise below takes into account the results of our baseline scenario. We fix knowledge as 'leaky' for the rest of the analysis given that without knowledge diffusion the role of other factors vanishes and discovery process turns random. We also consider only core-periphery network structure as more realistic where a potential GPT is followed by competitors.

3.3.2 Results for growing knowledge base

In this scenario we extend our baseline model adding a second generation of products that is not visible to agents from the beginning. New generation has the same ratio

[33] The process can be better compared with 'Eureka' moments preceding the application of R&D effort.

[34] Once a new link has become either visible or discovered, one automatically updates the probability of yet invisible technological combinations containing this link to become visible.

[35] In particular, we set $v = 5$, $\alpha = 0.01$, $\beta = 0.1$, $\gamma = 1$.

of products to technologies, namely $N = 60$ and $M = 100$ and mainly consists of new technologies that were not present in the first generation (using technologies from 96-195 and the potential GPT itself). Hence, there are only six common technologies between the first and second product generations (see left bottom and upper charts of the Figure 14 respectively). Those two generations are meant to represent two distinct technological paradigms with former of complexity $CS = 4$ and the latter of $CS = 5$ reflecting the fact that consumer products become more complex over time. The value distribution of product types in the second paradigm is taken twice larger than in the former one, primarily to compensate for the complexity and boost the simulation speed.

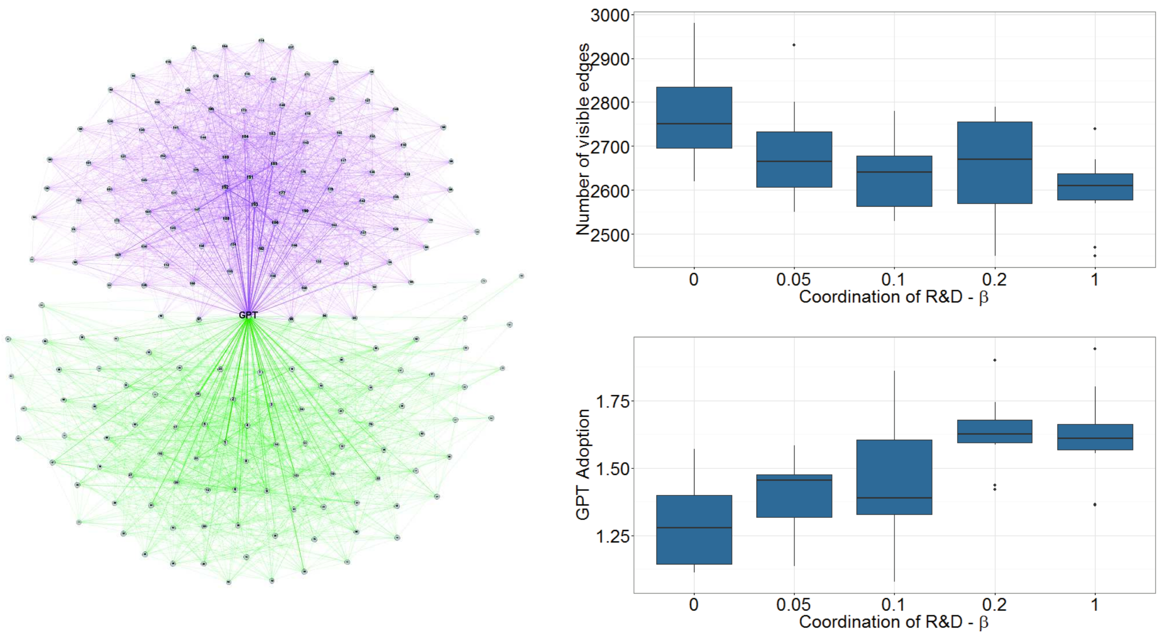


Figure 14: Effect of coordination on GPT adoption for the case with two product generations and a single GPT

The bottom right chart in Figure 14 demonstrates that coordination of efforts instead of a strictly positive (recall Figure 7) exhibits a 'level-off' effect on GPT adoption. A key to understand the nature of this finding is on the upper right chart of the figure: being more focused on knowledge depth strategy (and, as a consequence, discovering less technological links) one reduces the size of the visible network in the second product generation, thus, limiting the externality effect one can exploit. In other words, in the dynamic perspective high coordination hampers agents in realizing technological combinations with more pervasive technologies (including potential GPT). The positive effect of coordination in the first product generation is compensated by the negative effect in the second generation because an agent cannot discover something it has no idea about (yet).

3.3.3 Results for two GPTs with different product generations

The negative effect of coordination becomes more pronounced if we consider the second product generation to have its own potential GPT (for an illustration see left chart in Figure 15). Results of the experiments are demonstrated in the right panel of Figure 15 focusing on the GPT adoption in the second product generation since for the first

GPT the results repeat the pattern described in Section 3.2. Figure 15 demonstrates on the upper right chart that by increasing coordination the size of the visible network also falls. As a result, we observe a pattern similar to an inverted U-shape form illustrating the adoption of the second GPT in coordination. Thus, while moderate coordination is better than no coordination at all, this trend changes its direction once coordination approaches its maximum level demonstrating that neither no nor full coordination is optimal. This trade-off between exploiting externality effects and keeping the size of the visible technological network large enough (a sort of proxy for 'new ideas' in our model) reminds the classical ambidexterity trade-off known in the literature on organization theory (see, e.g., the seminal paper by **?**).[36]

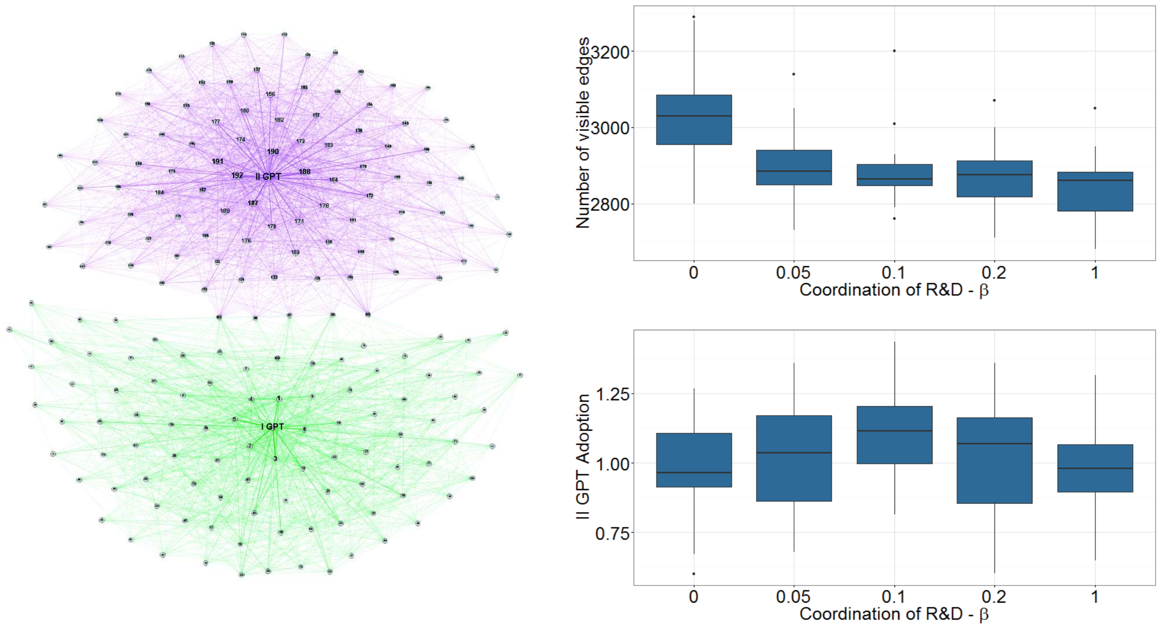


Figure 15: Effect of coordination on GPT adoption for the case with two product generations and two distinct GPTs

As we know from Section 3.2 that variation in expected profits has no clear effect on GPT adoption in the short term while in the long term it reduces the size of the discovered network and the period of time spent on the discovery process (Figure 12), it is easy to foresee that the effect of value dynamics on GPT adoption in the long term is strictly negative.[37]

3.4 Robustness tests

Given the large number of parameters in our model, it becomes infeasible to discuss all of their possible combinations in one paper. It is important to stress, however, that we have conducted the numerical experiments described above for different clique sizes of different product types within one technological paradigm. Though on average the

[36] According to (**?**, p. 72), "choices must be made between gaining new information about alternatives and thus improving future returns (which suggests allocating part of the investment to searching among uncertain alternatives), and using the information currently available to improve present returns (which suggests concentrating the investment on the apparently best alternative)".

[37] For brevity reasons we do not include those results here, but they are available on request.

GPT adoption has reduced over the experiments (because GPT was randomly allocated between combinations of different size), our major results hold. We also conducted the experiments with different distributions of difficulty and expected profits. Among others, we considered the difficulty distribution being exponential reflecting the situation where only few innovations are hard to discover. Additionally, we have considered the case where certain percentage of technological links are given 'for free' implying that their difficulty equals zero. Those modifications affect the speed of discovery process but do not change our findings with respect to Propositions 1-4. Furthermore, we considered alternative parameters for equation (6) and also modified the shape from exponential to logarithmic one. Our main findings do not change as long as our main assumption that arrival of new ideas depending on the visible and accumulated knowledge holds.

4 Stylized Facts of the Model

Apart from theoretical results on GPT emergence, we would like to stress here some of the stylized facts of innovation process that our model replicates and illustrate some steps in empirical verification of our predictions. In Proposition 4 we have already mentioned the *lock-in* effect. This effect is replicated by our model in the scenario with growing technological network, where low amount of knowledge accumulated (either due to co-ordination of R&D efforts or variation in expected profits) leads to many product types remained neither realized nor discovered (Figure 16).

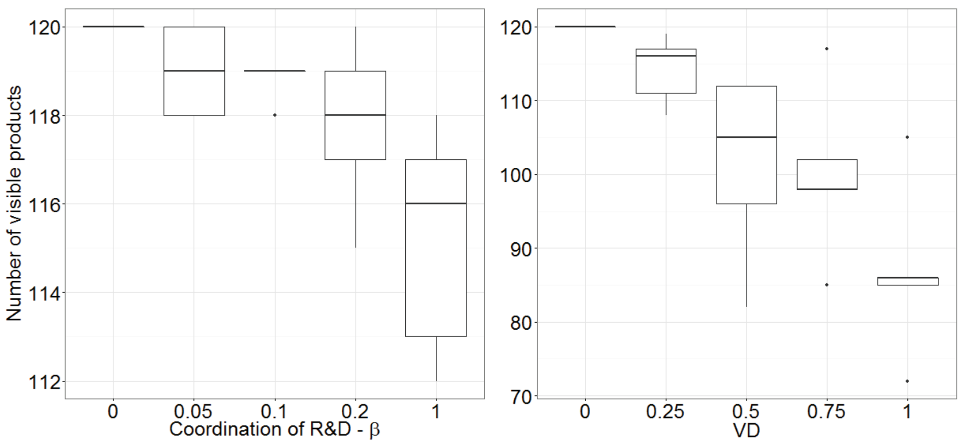
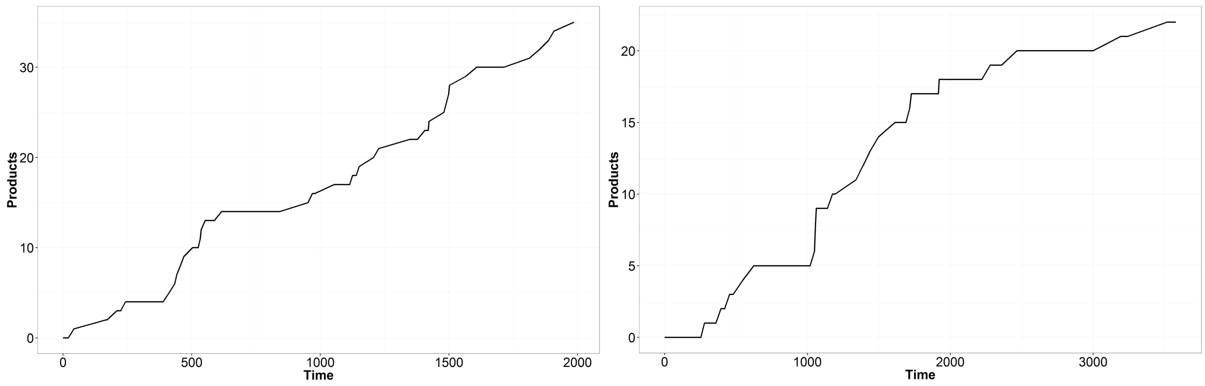


Figure 16: *Lock-in* effect in the case of two product generations and two distinct GPTs.

Next, our model exhibits the well known S-shaped curves of technology adoption if one plots number of products (growing over time) that contain GPT as one of their intermediate inputs (Figure 17).

The model also demonstrates clustering of innovations (discovery of new product) in time (see Silverberg and Lehnert (1993) for a literature review). To ensure that, we replicate the procedure by Silverberg and Verspagen (2003) in generating an innovation time series[38] (periods when a new product type has been discovered, see Figure 18 for an example) and sequentially fit the Poisson and negative binomial models with linear, quadratic and cubic time trends as explanatory variables. The linear and quadratic

[38]Note that by definition of a time period in our model, it is unlikely two innovations to happen at the same period. Therefore, without loss of generality we consider each twenty periods as one time interval.



(a) $N = 60,\ M = 100,$ leaky, $\beta = 1,\ VD = 0$ (b) $N = 60,\ M = 100,$ leaky, $\beta = 0,\ VD = 0$

Figure 17: Two representative cases of S-shaped curves of GPT usage in products.

time coefficients are significant at the 5% level, while the negative binomial model is consistently preferred over the Poisson one for all the three model specifications.[39] The simple explanation of the temporal clustering of innovations by our model is that those innovations share a common knowledge (technological edges), and agents coordinating their R&D activity exploit the knowledge externalities by discovering several product types within a short period of time.[40] This confirms ideas dating back to the concept of 'technological convergence' described by Rosenberg (1976) and shows the power of knowledge diffusion mechanism.

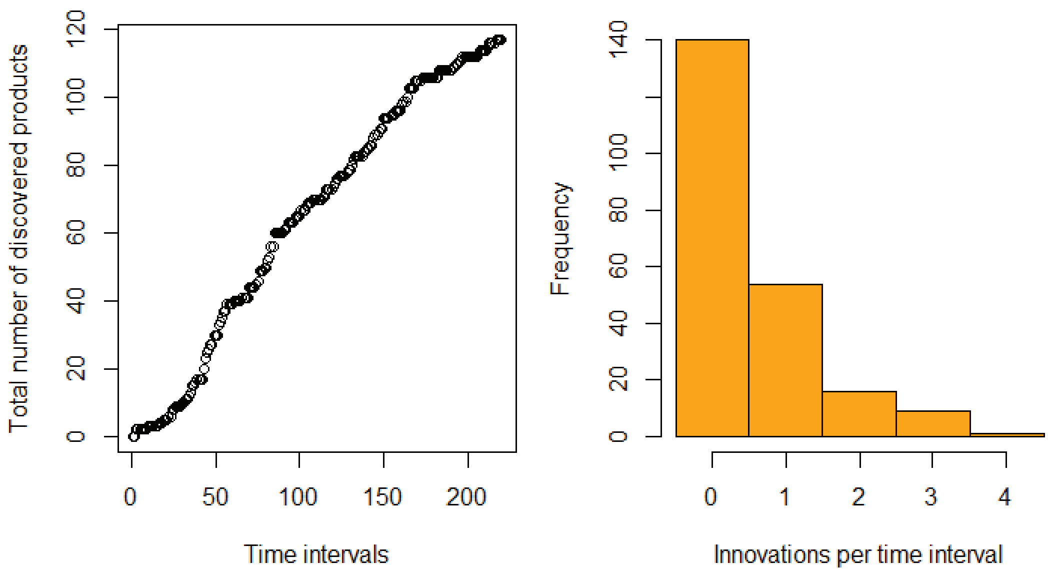


Figure 18: Clustering of innovations in time

Note: The result obtained for the scenario with two distinct GPTs (N=120, M=195, CS=4, $\beta = 1$ and leaky knowledge). The left figure illustrates bursts in the number of innovations (y-axis) in time (x-axis), while the right one is the histogram of innovations, i.e. frequency of periods with 0,1, 2 or more innovations.

In addition to the aforementioned facts, we compare structural similarity of the net-

[39]This finding holds for the majority of parameter values we use. The notable exception is variation in expected profits. If those change often, the process of discovery becomes close to linear in time.

[40]This is particularly true if the technological link had a relatively large difficulty and, thus, likely remaining one of the last barriers to introduce a new product.

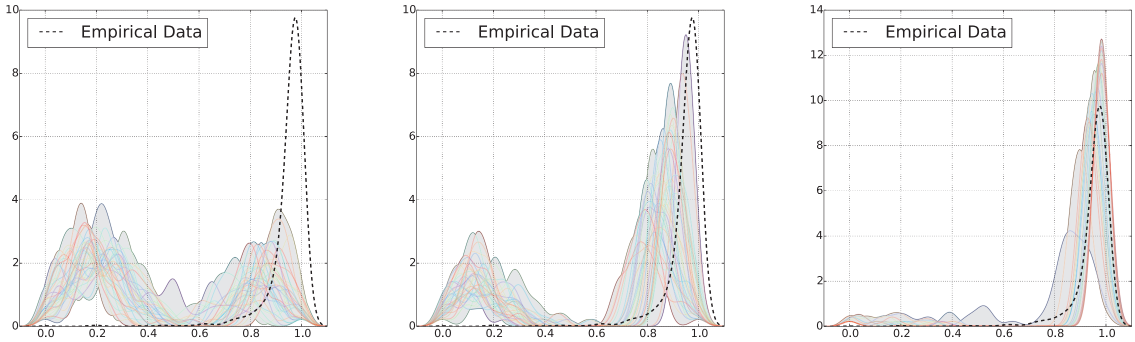
works we generate with those we observe empirically. In particular, networks of technologies and product relatedness are of interest. For many reasons (mainly because of *invisible* and *visible* networks representing an ex-ante state of knowledge we can only hypothesize), we concentrate on the final ex-post *discovered* networks drawing parallels with the works of Hidalgo et al (2007), Hidalgo and Hausmann (2009) and Boschma et al (2014) based on trade and patent data. Hidalgo et al (2007) and Hidalgo and Hausmann (2009) consider a product as a combination of some hidden technological capabilities that economic agents possess. In our model these capabilities are represented by technological combinations (a link between two technologies being discovered). Boschma et al (2014) investigates networks of patent IPC classes and their relatedness providing structural characteristics of those networks constructed by employing similar techniques as Hidalgo et al (2007). We focus here on technological networks where technologies are seen to be related if they share a patent. Thus, we see our model as a mechanism by which these empirical networks of products and technologies are being formed.

There is no consensus in the literature about graph comparison due to the nature of the subject of study. This problem is tackled differently across scientific fields (Mernberger, 2011). In particular, three main strategies are identified: exact graph matching, inexact graph matching, and feature-based approaches. We prefer the latter not least because we observe empirical networks on an aggregate level where each IPC class or product is already a collection of knowledge pieces or smaller products. Hence, we expect our simulated graphs matching some general structural characteristics of empirical graphs to be an indication that forces behind formation of those graphs are similar. We choose four features of graphs to compare: *density*, *degree assortativity*, *average clustering* and *degree distribution*. Density tells us about the interconnectedness (interrelatedness) of technologies in discovered products. Degree assortativity reveals whether more pervasive technologies tend to be connected with less pervasive ones. Average clustering illustrates to what extent do technologies cluster. Finally, degree distribution reveals possible 'hubs' - few technologies dominating the others in terms of their interconnectedness. Note that comparison of those network features does not require to have the same number of nodes in the simulated and empirical networks. We compare our product networks to the 'product space' taken from the atlas of economic complexity (Hidalgo et al, 2007; Hidalgo and Hausmann, 2009). In particular, the data comes from the website of the observatory of economic complexity (Simoes, 2016). Our product networks have similar high density and are disassortative (Table 2). Figure 19 demonstrates how degree distribution changes for simulated graphs with different knowledge diffusion regimes. Only *'leaky'* knowledge ensures products to be technologically highly interconnected as in empirical networks.

Table 2: Comparison of simulated (product) and empirical ('product space') networks

Network parameters	Empirical Network	Simulated Network (Mean)	Simulated Network (Standard Deviation)
N of nodes	773	60	0
N of edges	282402	1720.4	23.8
Density	0.967	0.972	0.013
Degree assortativity	-0.041	-0.042	0.004

To validate the produced technological networks we compare their typical ex-post (discovered) structure (for core-periphery structure, N=60, M=100, CS=4) with empirical networks of patent classes (kindly provided by P.-A. Balland for USPTO data for



(a) *Sticky* (b) *Partially leaky* (c) *Leaky*

Figure 19: Kernel-density estimations of degree distributions for product networks compared to 'product space' (dashed line) under different knowledge diffusion regimes

the years 1976 - 2010). A patent class can be considered as a piece of knowledge needed for production of goods. Characteristics comparison is presented in the Table 3. Both networks have similar density, implying almost the same ratio of existing links to all potential links. They are also similarly disassortative meaning that technologies with high degree centrality tend to be combined with technologies with low degree centrality. This result demonstrates again that technologies become pervasive only when they are combined with many infrequently used ones. We also report a typical for empirical networks heavy-tailed degree distribution (Figure 20). Figure 20 b) also illustrates that empirical technological network has a core-periphery structure: there are important 'gateway' technologies that are connected to the core and peripheral ones.[41]

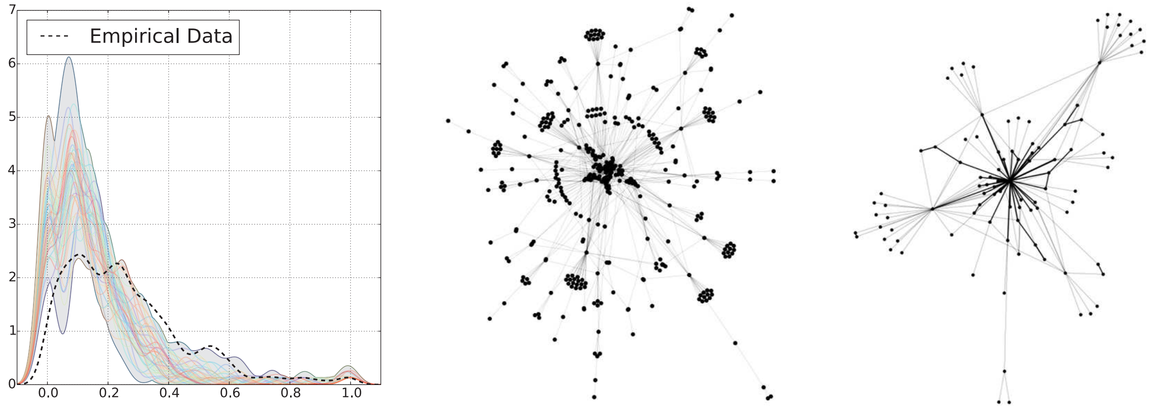
Table 3: Comparison of simulated (technological) and empirical (IPC) networks

Network parameters	Empirical Network	Simulated Network (Mean)	Simulated Network (Standard Deviation)
N of nodes	438	100	0
N of edges	12295	292	12
Density	0.068	0.059	0.009
Degree assortativity	-0.152	-0.150	0.026
Average clustering	0.479	0.501	0.038

5 Conclusion

General Purpose Technologies proved to be important during the past developments supporting the economic growth. Earlier GPT models emphasized their influence on 'productivity paradox' accounting for a 'residual' in aggregate production functions, focused on GPTs' evolution under a stream of innovations as well as explained the 'dual inducement' mechanism between GPT and its application sectors. Despite this extensive body of literature, the emergence of these technologies deserved little attention so far. Our study sheds light on this issue by concentrating on the pervasive nature of GPTs.

[41] If in contrast one looks at the star type network, such gateway technologies are absent.



a) b)

Figure 20: a) - Kernel-density estimations of degree distributions for technological networks compared to empirical network (dashed line) of IPC classes. b) - Visualization of a network of IPC classes among patents (left) and simulated (right) technological networks

Introducing its emergence as a continuous process of technology adoption, we look for mechanisms fostering technological convergence. In particular, we employ methods of network science representing knowledge discovery as a growing technological graph.

Our results demonstrate that knowledge diffusion is absolutely necessary for GPT emergence since being discovered once the knowledge spills over to many other applications benefiting most those technologies having the potential to be used in many distinct products and industries. The structure of our knowledge should have a sufficient density for a GPT to become pervasive, where by structure we mean interconnectedness of our ideas and by density – interchangeability of our knowledge among various applications. With the novel metric (*the Multiplicity Index*) we demonstrate how to measure that density. Given the presence of knowledge spillovers and sufficient density of the network structure, concentrating on technological trajectories where more knowledge is accumulated also favors GPT in the short term. However, once the technology network is modeled as a growing knowledge base where agents become aware of novel possibilities to combine technologies through inventing simpler products, a trade-off between coordinating on existing trajectories and realizing novel technological combinations emerges. This transforms the pure positive effect of coordination into an inverted U-shape form echoing the classical ambidexterity trade-off between exploration and exploitation. Similar to firms in the organizational theory (see, e.g., Sidhu et al (2007)), countries shall apply differentiated technological policy depending on whether the economy is in a more or less dynamic environment. Thus, in the 'path-following' catching-up process (Lee and Lim, 2001) countries aiming to discover certain product types in the knowledge base where most of technological trajectories are known from experience of advanced economies will find exploitative strategy (high coordination on trajectories with most accumulated knowledge) most attractive. In contrast, if the economy is currently at the technological frontier seeking to identify the next GPT, it shall put more focus on exploration of new opportunities and provide incentives for sufficient knowledge breadth. For the same reason, policy maker shall avoid supporting any specific product need before the economic agents accumulate enough information on alternative ways of producing goods to satisfy that need and payoffs to adoption of the respective technologies. Otherwise, the choice of the technological trajectory turns random and due to the increasing returns to adoption described by Arthur (1989) the economy risks to be locked-in to inferior technologies.

Our model reproduces well known stylized facts accompanying innovation processes such as S-shaped curve of technology adoption, temporal clustering of innovations in time and lock in effects. Furthermore, our model replicates many structural features of the empirical product graphs (Hidalgo et al, 2007) and those graphs constructed based on networks of relatedness between technological IPC classes (Boschma et al, 2014).

One shall also stress that the current analysis is limited in a number of ways. First, no production costs are taken into account. This together with explicit budget restriction on the side of agents shall provide a more complete picture of the technological competition, and help to explore 'growth bottlenecks' pointed by Bresnahan and Yin (2010). We preferred to abstract ourselves from those issues here for the sake of clarity. Furthermore, so far we have neglected heterogeneity among agents in terms of their accumulated knowledge and possible cooperation/competition between them. All these aspects provide a natural direction to further develop the present model opening a fruitful trajectory of further extensions in the direction of technological competition/succession.

References

Abernathy W, Clark C (1985) Innovation: mapping the winds of creative destruction. Research Policy 14:3–22

Aghion P, Howitt P (1992) A model of growth through creative destruction. Econometrica 60(2):323–51

Aghion P, Howitt P (1998) On the macroeconomic effects of major technological change. In: Helpman E (ed) General Purpose Technologies and Economic Growth, MIT Press, Cambridge, Massachussets, pp 121–144

Anderson P, Tushman L (1990) Technological discontinuities and dominant designs: a cyclical model of technological change. Administrative Science Quartely 35:604–633

Arrow K (1962) Economic welfare and the allocation of resources for invention. In: Nelson R (ed) The Rate and Direction of Inventive Activity, Princeton University Press, Princeteon, NJ, pp 609–626

Arthur B (2015) Complexity and the Economy. Oxford University Press, NY

Arthur W, Polak W (2006) The evolution of technology within a simple computer model. Complexity 11(5):23–31

Arthur WB (1989) Competing technologies, increasing returns, and lock-in by historical events. The Economic Journal 99(394):116–131

Atkinson AB, Stiglitz JE (1969) A new view of technological change. The Economic Journal 79(315):573–578

Boschma R, Balland P, Kogler D (2014) Relatedness and technological change in cities: the rise and fall of technological knowledge in US metropolitan areas from 1981 to 2010. Industrial and Corporate Change 24(1):223–250

Bresnahan T (2012) Generality, recombination and re-use. In: Lerner J, Stern S (eds) The Rate and Direction of Economic Activity Revised, University of Chicago Press, pp 611–656

Bresnahan T, Trajtenberg M (1995) General purpose technologies 'engines of growth?'. Journal of Econometrics 65(1):pp. 83–108

Bresnahan T, Yin P (2010) Reallocating innovative resources around growth bottlenecks. Industrial and Corporate Change 19(5):1589–1627, DOI 10.1093/icc/dtq048

Brynjolfsson E (1993) The productivity paradox of information technology: Review and assessment. Communications of the ACM

Cantner U, Vannuccini S (2012) A new view of general purpose technologies. In: Heilemann U, Wagner A (eds) Empirische Makoökonomik und mehr, Lucius et Lucius

Carlaw K, Lipsey R (2006) GPT-driven, endogenous growth. The Economic Journal 116(508):155–174

Carlaw K, Lipsey R (2011) Sustained endogenous growth driven by structured and evolving general purpose technologies. Journal of Evolutionary Economics 21:563 – 593

Cowan R (1990) Nuclear power reactors: A study in technological lock-in. Journal of Economic History 50(3):541–567

Dosi G (1982) Technological paradigms and technological trajectories: A suggested interpretation of the determinants and directions of technical change. Research Policy 11(3):147–162

Dosi G (1988a) The nature of the innovative process. In: Dosi G, Freeman C, Nelson R, Silverberg G, Soete L (eds) Technical Change and Economic Theory, Pinter London, pp 221–238

Dosi G (1988b) Sources, procedures, and microeconomic effects of innovation. Journal of Economic Literature 26(3):1120–1171

Helpman E (ed) (1998) General Purpose Technologies and Economic Growth. The MIT Press, Cambridge, MA

Henderson R, Clark K (1990) Architectural innovation. Administrative Science Quarterly 35(1):9–30

Herrmann J, Savin I (2016) Optimal policy identification: Insights from the German electricity market. Tech. Rep. 004, Jena Economic Research Papers

Hidalgo C, Hausmann R (2009) The building blocks of economic complexity. Proceedings of the National Academy of Sciences of the United States of America 106(26):10,570–10,575

Hidalgo CA, Klinger B, Barabási AL, Hausmann R (2007) The product space conditions the development of nations. Science 317(5837):482–487

Holland JH (1995) Hidden Order: How Adaptation Builds Complexity. Addison-Wesley, Michigan

Kauffman S (1995) At Home in the Universe: The Search for the Laws of Self-Organization and Complexity. Oxford University Press, Oxford

Lee K, Lim C (2001) Technological regimes, catching-up and leapfrogging: findings from the Korean industries. Research Policy 30(3):459–483

Lipsey R, Carlaw K, Bekar C (2005) Economic Transformations. Oxford University Press Inc., New York

March JG (1991) Exploration and exploitation in organizational learning. Organization Science 2(1):71–87

Mernberger M (2011) Graph-based approaches to protein structure comparison - from local to global similarity. PhD thesis, Philipps-University Marburg

Nelson R, Winter S (1982) An Evolutionary Theory of Economic Change. Cambridge, MA

Ott I, Papilloud C, Zuelsdorf T (2009) What drives innovation? Causes of and consequences for nanotechnologies. Managing Global Transitions 7(1):5–26

Romer P (1990) Endogenous Technological Change. Journal of Political Economy 98(5):S71–S102

Rosenberg N (1976) Perspectives on technology. Cambridge University Press

Schumpeter J (1934) The Theory of Economic Development. Harvard University Press

Sidhu J, Commandeur H, Volberda H (2007) The multifaceted nature of exploration and exploitation: Value of supply, demand, and spatial search for innovation 18(1):20–38

Silverberg G, Lehnert D (1993) Long waves and 'evolutionary chaos' in a simple Schumpeterian model of embodied technical change. Structural Change and Economic Dynamics 4(1):9–37

Silverberg G, Verspagen B (2003) Breaking the waves: a Poisson regression approach to Schumpeterian clustering of basic innovations. Cambridge Journal of Economics 27(5):671–693

Silverberg G, Verspagen B (2005) A percolation model of innovation in complex technology spaces. Journal of Economic Dynamics & Control 29:225–244

Simoes A (2016) Observatory of economic complexity. Available at `http://atlas.media.mit.edu/en/resources/data/`

Strohmaier R, Rainer A (2016) Studying general purpose technologies in a multi-sector framework: The case of ICT in Denmark. Structural Change and Economic Dynamics 36:34–49

Weitzman ML (1998) Recombinant growth. The Quaterly Journal of Economics 113(2):331–360

Appendix

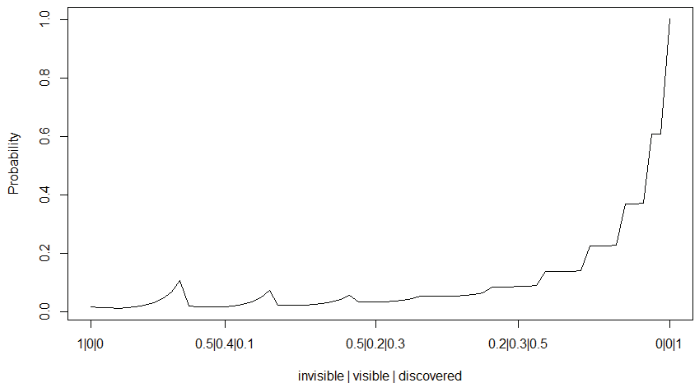


Figure 21: Probability of a technological combination to become visible.

Note: The function illustrates the probability of a technological combination consisting of five technologies and ten edges to become visible. Values on the x-axis denote the distribution of links in the clique being discovered, visible or invisible. For example 0.5|0.3|0.2 means 50% of the links are discovered, 30% – visible and 20% – invisible links.

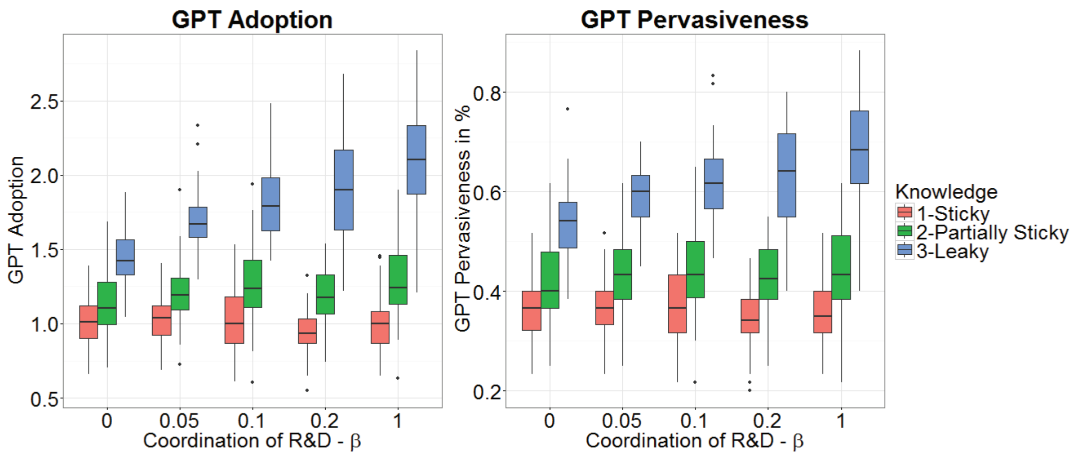


Figure 22: GPT pervasiveness and realization for star type network

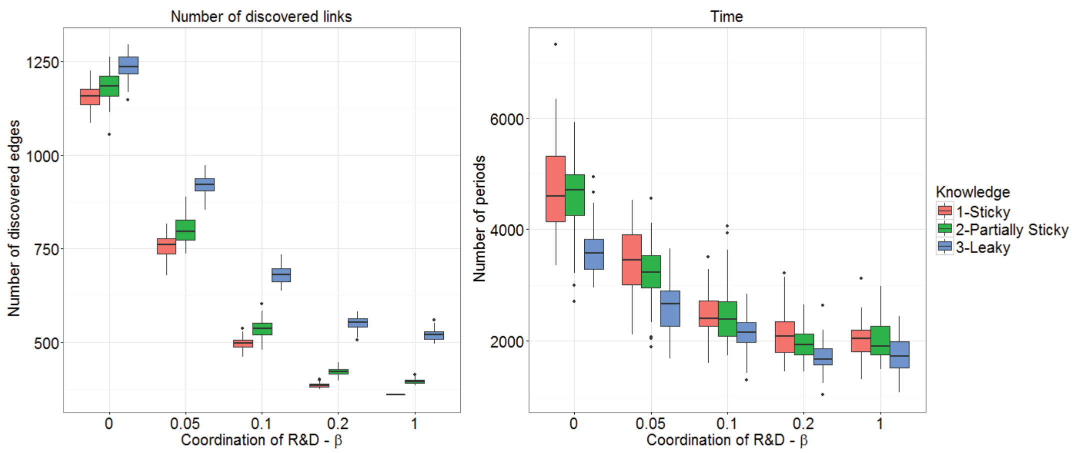


Figure 23: Number of periods and discovered edges for star type network