Do main paths reflect technological trajectories? Applying main path analysis to the semiconductor manufacturing industry

Flavia Filippin[1]

Abstract

It has been proposed that main path analysis can be used to identify technological trajectories in patent-citation networks. In this paper, the method is applied to a network composed of one million US patents and eight million citations in order to trace the backbone of the technological trajectory of the semiconductor manufacturing industry. An in depth discussion of the method is presented, focusing on the many parameters that can be adjusted while applying it and on the consequences of adjusting any of them. Moreover, and differently from other papers on the subject, the result of the algorithm is analysed to determine if it indeed represents the most important technological contributions to the trajectory or if it is merely a collection of relevant and connected patents. This is made easier by the fact that the semiconductor industry has a clear and widely known technological trajectory that spans more than 50 years, Moore's law.

> *"Write it on silicon, not on paper."*
> *Gordon Moore*

Introduction

Since the seminal paper by Dosi ([1982]), the concepts of technological paradigms and trajectories have captured the imagination of researchers and policy makers. More often than not, however, they are used only in the introduction section of scientific papers and reports, revealing that the full meaning of these abstract concepts is difficult to grasp.

✉ Flavia Filippin
   filippin@merit.unu.edu

[1] UNU-MERIT and Maastricht University, Maastricht, The Netherlands

A big step forward was given when Verspagen (2007) proposed that main path analysis could be used to identify technological trajectories in patent-citation networks. The method has since been applied to study the cumulative development of several technologies, from coronary angioplasty to lithium batteries. Yet, the many papers that have been published in the past fifteen years (with the notable exception of Barberá-Tomás et al. (2011), which will be discussed below) skipped one important methodological step, validation, leaving researchers wondering if main paths do indeed reflect technological trajectories.

In this paper, we analyse the method proposed by Hummon and Doreian (1989), main path analysis, as well as the many extensions made to it, discussing, at each step, what options lay before the researcher when applying it. Moreover, and differently from the other papers on the subject, we look at the result of the algorithm, at the specific patents established by the method as the backbone of the technological trajectory in question, and validate if they indeed represent the most important technological contributions to the trajectory or if they are merely a collection of relevant and connected patents, a sample of the population being studied.

To do so, we look at the trajectory that has laid the foundation for much of the economic growth experienced in the last 50 years: Moore's law. This law gave the rhythm to the miniaturization trajectory of semiconductor devices, which has, in turn allowed exponential gains in computing performance with decreasing costs, leading to productivity gains throughout the whole economy.

The paper is structured in six sections besides this introduction. Section 2 discusses the concepts of technological paradigms and trajectories as proposed by Dosi (1982). Section 3 reviews the literature on main path analysis, summarizing what has been done so far. Section 4 discusses the application of main path analysis, step by step, from the selection of the patents of interest, through the computation of transversal weights, to the search of the main path. The patent-citation network used in this paper is composed of more than one million patents and more than eight million citations extracted from the USPTO Patents-View platform, which offers access to the full content of patents granted by the USPTO since 1976. Section 5 evaluates, with the help of a panel of experts, if the patents that compose the main path correspond to the backbone of the main technological trajectory of the semiconductor industry. Finally, Sect. 6 discusses the main results of the paper and raises a few questions about the application of main path analysis and Sect. 7 presents the final remarks.

Technological paradigms and trajectories

Dosi (1982) defines a technological paradigm as "a 'model' and a 'pattern' of solution of *selected* technological problems, based on *selected* principles derived from natural sciences and on *selected* material technologies" (p. 152). A technological trajectory, in turn, is understood as "the pattern of 'normal' problem solving activity (i.e. of 'progress') *on the ground of a technological paradigm* [emphasis added]" (p. 152).

Additionally, "a technological paradigm (or research programme) embodies strong perceptions on the directions of technical change to pursue and those to neglect" (Dosi, 1982, p. 152). Therefore, "technological paradigms have a powerful *exclusion effect*: the efforts and the technological imagination of engineers and of the organizations they are in are focussed on rather precise directions while they are, so to speak, 'blind' with respect to other technological possibilities" (p. 153).

It's worth, nonetheless, recalling Dosi's definition of technology, for the concepts above make more sense when applied to clusters of technologies, rather than to individual pieces of technology, with semiconductor technology being one of the examples mentioned by the author. Technology is "a set of pieces of knowledge, both directly 'practical' (related to concrete problems and devices) and `theoretical' (but practically applicable although not necessarily already applied), know-how, methods, procedures, experience of successes and failures and also, of course, physical devices and equipment" (Dosi, 1982, p. 151–2). Technology is therefore composed both of embodied parts and disembodied parts. Disembodied (i.e. intangible) parts are expertise acquired during past attempts at solving valid technological problems (at the individual level) and the knowledge and achievements of the state of the art (at the social level). Embodied parts (i.e. physical devices), on the other hand, incorporate the achievements in the development of the technological field they belong to. Also, Dosi warns that technological knowledge is less articulated than scientific knowledge, so, in comparison, the definition of a technological paradigm "is bound to be much looser", and "is adequate in some cases but less so in others" (Dosi, 1982, p. 153).

Concerning the identification of technological paradigms, Dosi (1982) argues they have four dimensions: (i) a generic task to which it is applied, (ii), the material technology it selects, (iii) the physical/chemical properties it exploits, and (iv) the technological and economic trade-offs it focusses upon. To illustrate, the author explains that these four dimensions are the following in the case of the semiconductor industry: (i) amplifying and switching electric signals, (ii) semiconductor materials and, more specifically, silicon, (iii) the transistor effect and the field-effect of semiconductor materials, and (iv) the trade-off between circuit-density, speed, unit costs, etc. Progress along a technological trajectory is defined then as improvements on those dimensions, especially on the technological and economic trade-offs. However, it is interesting to note that, in the case of the semiconductor industry, the rate of technological development meant that the trade-off did not materialize until recently: for most of its history, the semiconductor industry delivered increased circuit density (and, therefore, increased performance) with decreased cost per transistor.

The literature on main path analysis

Hummon and Doreian (1989) proposed main path analysis as a method to identify papers that played a central role in the development of a particular theory. Starting from a review of the literature on DNA that had already established the most important papers in the field, these authors compiled a citation network and proposed different ways to measure the extent to which each particular citation was needed for linking any two papers in the network. After computing these transversal weights, the authors applied a local search algorithm: starting from a source node, they looked for the node that had an arc with the highest weight among all sources and added the nodes at the beginning and at the end of the arc to the main path; then, they looked for the arc with the highest weight leaving the node at the end of the previously chosen arc and added the node at the end of this second arc to the main path; and then they repeated the second step until a sink node was reached.

In an attempt to measure and map the concepts developed by Dosi (1982), Verspagen (2007) applied main path analysis to patent documents. According to the author, since technological trajectories can be understood as sequences of interrelated innovations and interrelatedness between innovations can be measured using patent citations, there are, in

a large network of patent-citations, several main streams (or main paths) of knowledge and these streams should summarize the major development in the field.

Verspagen (2007) also proposed two main extensions to the method. The first extension was to take into consideration not only the top main path (either local or global), but all the paths from sources to sinks, thus creating a network of main paths. The second extension was to apply the method to different temporal cuts of the patent-citation network, incorporating, at each turn, more years into the network, resulting in a temporal main path.

Following Verspagen (2007), several other authors applied main path analysis to citation networks. Table 1 summarizes several of these efforts. Columns 3 to 8 explain how the authors of each paper compiled the network used to apply the method and columns 9 to 13 explain how the authors applied the method. From column four, it becomes clear that most authors use either key-words or technology classes to select the papers or patents of interest, i.e. to delimit the field of study.

When authors use key-words as the selection method, the set of nodes selected often needs to be cleaned, since papers or patents that do not actually belong to the field can be included by mistake. Column five presents this information with an arrow indicating a decrease in the number of core nodes, when such information was disclosed by the authors. After collecting the core nodes using the chosen selection method, it is necessary to decide if all the citations made by these nodes will be taken into consideration, thus increasing the final number of nodes in the network, or if only the citations between core nodes will be analysed. The first case is what we call open and the second is what we call closed in column eight.

Column nine shows which transversal weight was used in each paper and columns 10 to 12 show what kind of search was performed to find the main path. Finally, column 13 indicates if any complementary method was used (e.g. temporal paths, hubs and authorities, island, key-route paths, etc.). Concerning the computation of main paths, most authors that mention the use of a software (and most in fact do) state that the software used was *Pajek* (De Nooy et al., 2005).[1] Almost all features of main path analysis are already implemented in *Pajek*, with the notable exception of networks of main paths.

Regarding external validation, although several of the papers mentioned in Table 1 do try to include a historical account of the development of the technology in question, the only work that performs any kind of external validation is the one by Barberá-Tomás et al. (2011). In their paper, they compare the patents in the global main path and in the network of temporal main paths with information from pre-clinical tests (in the case of new inventions) and with products available commercially or in clinical use (in the case of incremental inventions to existing products) and match them using criteria such as design and inventor/owner. As the authors point out themselves, "[their] validation strategy requires in depth knowledge of both the technical aspects and the history of projects" (Barberá-Tomás et al., 2011, p. 478). The authors find that the proportion of patents in the global main path (and to a lesser extent in the network of temporal main paths) associated with products available commercially or in clinical use is higher in comparison to the whole network.

Nonetheless, the external validation we are performing in this paper has a different goal. The objective of Barberá-Tomás et al. (2011) was, in essence, to find out how many inventions had become innovations. What they did not examine was if the whole trajectory of the development of the technology in question was summarized in the main path. And that

---

[1] One exception is Barberá-Tomás et al. (2011), who used *Citpath*.

Table 1 Literature on main path analysis

(1) Author	(2) Case study	(3) Type of node	(4) Selection method	(5) Core nodes	(6) Nodes	(7) Arcs	(8) Open or closed
Hummon and Doreian (1989)	DNA	Papers	Literature review	40	40	61	Closed
Moore et al. (2006)	Social capital	Papers	Keywords	87→65	227	NA	Open[a]
Verspagen (2007)	Fuel cells	Patents	Classes	3371	3371	15,506	Closed
Mina et al. (2007)	Coronary Angioplasty	Papers	Keywords	11,420	94,442	300,000	Open
		Patents		5136	22,095	135,000	
Calero-Medina and Noyons (2008)	Absorptive Capacity	Papers	Original work[b]	1	1213	NA	Open
Lucio-Arias & Leydesdor (2008)	Fulerene	Papers	Keywords	7696	30[c]	NA	Closed
	Carbon nanotubes			9672	30[c]	NA	
Fontana et al. (2009)	LAN and Ethernet	Patents	Classes	581	581	1116	Closed
Barberá-Tomás et al. (2011)	Artificial intervertebral disc	Patents	Keywords	201	1535	6310	Open[d]
Lin et al. (2011)	Automobiles	Patents	Mixed[e]	NA	91,119	224,258	Open
Martinelli (2012)	Telecommunication switches	Patents	Classes	NA	6214	20,848	Open
Liu and Lu (2012)	Hirsch index	Papers	Keywords	1143→421	421	2941	Closed
Ho et al. (2014)	Fuel cells	Papers	Keywords	39,473→7,313[f]	7313	NA	Closed
Hung et al. (2014)	Lithium batteries	Papers	Keywords	1531	1480[g]	17,015	Closed
Xiao et al. (2014)	Data quality	Papers	Keywords	1880	1880	NA	Closed
Triulzi (2015)	Semiconductors	Patents	Classes	114,097	114,097	779,076	Closed
Batagelj et al. (2017)	Peer review research	Papers	Keywords	19,920	45,965	132,604	Open
Kim and Shin (2018)	High voltage direct current transmissions	Patents	Keywords	625→223	1977[h]	NA	Open
Maltseva and Batagelj (2019)	Social network analysis	Papers	Mixed[i]	70,792	222,189	1,521,658	Open

(1) Author	(9) Weight	(10) Local	(11) Global	(12) Network	(13) Additional method
		Type of main path			
Hummon and Doreian (1989)	NPPC, SPLC and SPNP	x			
Moore et al. (2006)	NA	x			
Verspagen (2007)	SPNP		x		Temporal paths
Mina et al. (2007)	NA			x	Island
Calero-Medina and Noyons (2008)	SPC	x			Hubs and authorities

Table 1 (continued)

(1) Author	(9) Weight	(10) Local	(11) Global	(12) Network	(13) Additional method
		Type of main path		Network	
Lucio-Arias & Leydesdor (2008)	SPLC	x			
Fontana et al. (2009)	SPNP		x	x	Temporal paths
Barberá-Tomás et al. (2011)	NA		x		Temporal paths
Lin et al. (2011)	SPC	x		x	Hubs and authorities
Martinelli (2012)	SPLC		x	x	Temporal paths
Liu and Lu (2012)	SPC		x		Backward and key-route paths
Ho et al. (2014)	SPC	x	x	x	
Hung et al. (2014)	SPC				Key-route path
Xiao et al. (2014)	SPC	x	x	x	Key-route path
Triulzi (2015)	SPNP			x	Temporal path
Batagelj et al. (2017)	SPC		x		Key-route path and islands
Kim and Shin (2018)	SPC	x	x		Backward and key-route paths and technology Junctures
Maltseva and Batagelj (2019)	SPC		x		Key-route path and islands

NA: Not Available, *NPPC*: Node Pair Projection Count, *SPC*: Seach Path Count, *SPLC* Search Path Link Count, *SPNP*: Seach Path Node Pair

[a] Only papers cited by more than one core paper were included

[b] All the forward citations generated by the paper that inaugurated a particular strand of the literature were traced

[c] The analysis was restricted to the 30 most frequently cited papers in each dataset

[d] Instead of all citations, only citations to other kinds of medical products were included

[e] Classes, keywords and relevant players were used as selection methods. In regard to the latter, 36 auto-makers and 99 component suppliers were identified and all patents assigned to them were collected

[f] Not only the papers that were outside the field of interest were excluded, but also papers without author information, older papers (published before 2009) with less than 3 citations, and papers that did not contain the keywords of interest in the title

[g] Papers that received no citations and that did not cite other papers in the dataset were excluded

[h] To expand the sample of patents, citations chains of length two and three from core patents to core patents passing through one or two non-core patents were used

[i] A keyword search was complemented by including (i) all papers published in the main journals of the field and (ii) the most frequently cited papers of 100 authors in the field

is what we are proposing to do here, to verify if main path analysis is a valid method for identifying technological trajectories.

Applying main path analysis

Selecting patents

The first step in main path analysis, as in any social network analysis, is to define the population of interest, i.e. select which nodes and edges/arcs will compose the network. Since the objective here is to study semiconductor technology, we started by selecting patents that belong to our object of study, the semiconductor industry. Alternatively, if the focus were on the scientific developments that lay behind this technology, one could build a network using scientific publications.

The most used methods to select patents of interest are key-word search and selecting by technology classes. Key-word search requires (i) carefully choosing the key-words, often with the help of experts, and (ii) cleaning the data afterwards to eliminate unwanted patents. Technology classes, on the other hand, can be more easily chosen by means of concordance tables between classes and economic sectors and by looking at existing literature. Additionally, while the classes used to select patents of interest are often available in papers and reports, keywords are seldom disclosed by the authors. Since using technology classes is more straightforward and allows for more comparability and reproducibility of research, this was the method chosen here.

Our source of data is the United States Patent and Trademark Office (USPTO) PatentsView Platform. It contains patents from 1976 to the present. When data for this paper was retrieved in August 2018, the PatentsView Plataform had last been updated on 28 May 2018 and, therefore, contained patents granted up to that date. The files used to extracted patents from the database by means of technology classes are the "uspc current.tsv" and the "cpc current.tsv". The file used to collect citations is the "uspatentcitation.tsv".

Six different ways of defining the semiconductor industry were found in the literature (see Table 2).[2] Some of them use the United States Patent Classification System (USPC) and others use the International Patent Classification System (IPC).[3] In addition to these six definitions, we included one more ("Van Looy et al. (2015) Plus"), covering all subclasses present in Van Looy et al. (2015) and two other (C23C and G03F), which refer to important processes used in the fabrication of semiconductor devices.

The results presented in this paper were obtained by using the concordance table provided by Van Looy et al. (2015) to select the core patents and by collecting all citations made by these patents (i.e. an open system) to compose the network. We will be referring to this network as our benchmark network; it has 670 thousand core patents, 1.4 million nodes and 8.5 million arcs. In terms of degree distribution, citations received and citations made are both right-skewed, with most nodes receiving and/or making few citations and a few nodes

---

[2] Since in Schmoch (2008) there are two sectors which are related to the semiconductor industry, we use both "Semiconductors" alone (which we refer to as "Schmoch (2008) Narrow") and "Semiconductors" and "Micro-structure and nano-technology" combined (which we refer to as "Schmoch (2008) Broad").

[3] In 2015, the USPTO adopted the Cooperative Patent Classification System (CPC), in partnership with the European Patent Office (EPO). Patents granted until May 2015 are classified using both the USPC and the CPC. Patents granted after May 2015 are classified using only the CPC. The IPC and the CPC are equivalent at the subclass level (four-digit level).

Table 2 Defining the semiconductor industry in terms of technology classes

Author	Classification System	Economic Sector	Technology Class[a]				
Hall et al. (2001)	USPC	Semiconductor devices	257	326	438	505	
Schmoch (2008) Narrow	IPC	Semiconductors	H01L				
Schmoch (2008) Broad	IPC	Semiconductors + Micro-structure and nano-technology	B81B	B81C	B82B	B82Y	H01L
Triulzi (2015)	USPC	Semiconductor industry	257	326	438	505	716
USPTO (2018)	USPC	Semiconductor devices and manufacture	117[b]	257	326	438	716
Van Looy et al. (2015)	IPC	Manufacture of electronic components and boards	B81B	B81C	B82B	B82Y	C30B
			G11C	H01C	H01F	H01G	H01J
			H01L	H05K			
Van Looy et al. (2015) Plus	IPC	Manufacture of electronic components and boards	B81B	B81C	B82B	B82Y	C23C
			C30B	G03F	G11C	H01C	H01F
			H01G	H01J	H01L	H05K	

[a]The titles of the USPC classes and the IPC subclasses are available in Appendix 1

[b]USPTO (2018) includes, besides these five classes, 546 subclasses in its definition of the semiconductor industry
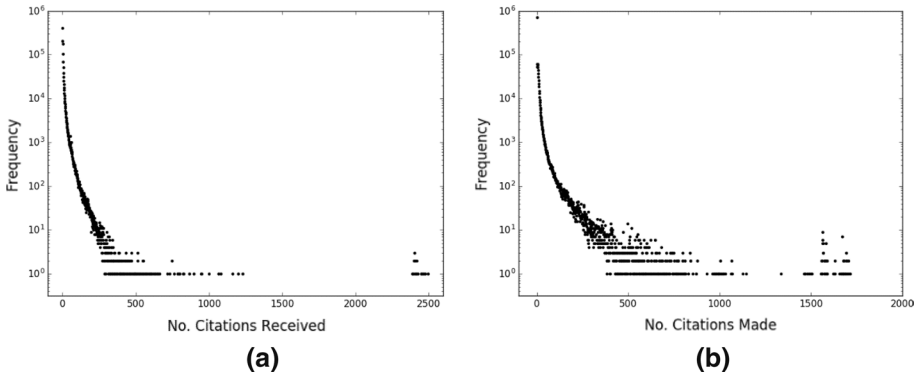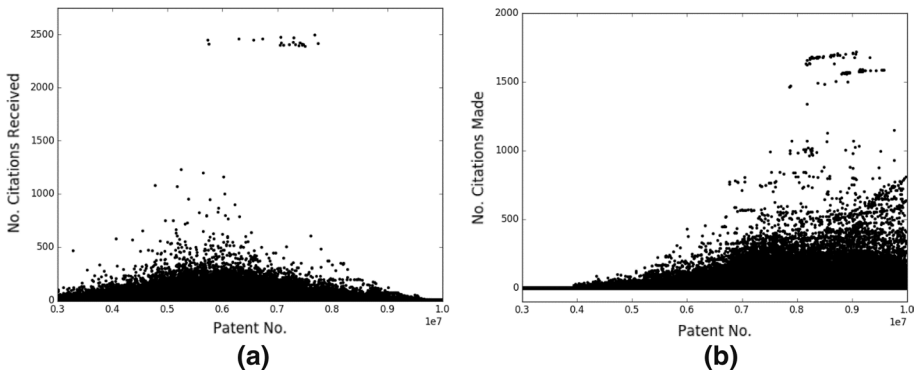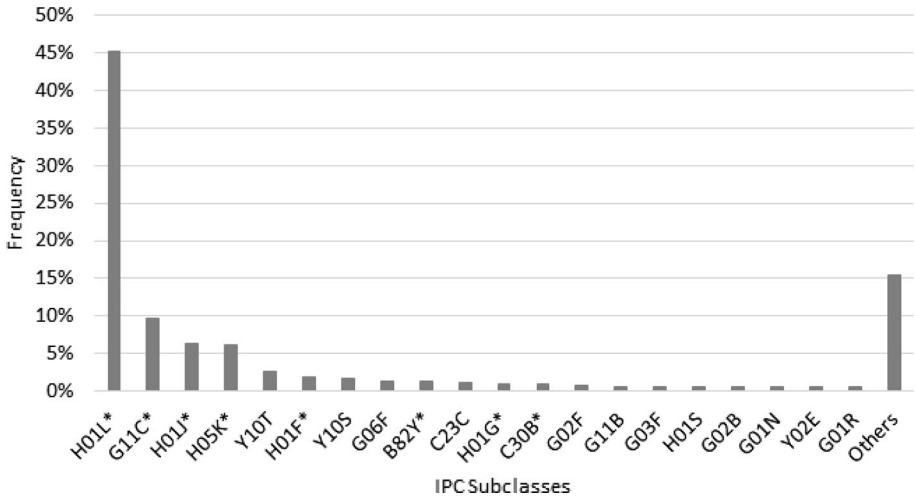
Fig. 1 Frequency Distribution of Citations Received (**a**) and Citations Made (**b**)

Table 3 Descriptive statistics of citations received and made

	Citations received	Citations made
Minimum	0	0
25th Percentile	1	0
50th Percentile	2	0
75th Percentile	5	6
95th Percentile	26	23
99th Percentile	74	86
99.9th Percentile	196	333
Maximum	2495	1716
Mean	6.353471	6.353471
Standard Deviation	19.461416	27.733008

Fig. 2 Citations received (**a**) and citations made (**b**) over time

Fig. 3 Distribution of core patents among IPC subclasses

receiving and/or making lots of citations. Figure 1 shows the frequency distribution of the degree, while Table 3 offers some descriptive statistics of the degree. Figure 2, on the other hand, shows the number of citations received and citations made over time (using the patent number as a proxy for time), and here it can be noted that the number of citations made increases over time, whereas there is no clear trend in the number of citations received.

In terms of technology class distribution, Fig. 3 shows the top 20 most popular IPC subclasses among the core patents of our benchmark network. Subclasses marked with an asterisk belong to our classes of interest, i.e. to the group of classes that were used to select these core patents.[4] It can be clearly seen that subclass H01L is the most important one, accounting for 45% of all patents. Finally, in terms of key players, Table 7 in Appendix 2 shows the top 10 assignees in each of the IPC subclasses mentioned in Table 2.

In order to check the robustness of main path analysis, the other six definitions of the semiconductor industry shown in Table 2 were also tested and yielded similar results.[5] Van Looy et al. (2015) Plus rendered the same main path as our benchmark network and the main paths generated by the other concordances only diverged from the benchmark main path from 2006 onwards in the case of Hall et al. (2001), Triulzi (2015) and USPTO (2018) and from 2008 onwards in the case of Schmoch (2008) Narrow and Schmoch (2008) Broad. Content-wise, while recent patents in the benchmark main path concern tool design and specific process technologies, recent patents in these alternative main paths mainly refer to high-κ materials for the dielectric layer.[6]

---

[4] Fractional counting was used here: if a patent has x technological classes, each class receives a weight of 1/x.

[5] Details regarding all alternative main paths discussed in this paper will be provided by the authors upon request.

[6] Although using different sets of technology classes to define the semiconductor industry produced fairly similar results, it is not possible to state that the method chosen to select the initial nodes is completely unimportant. A similar work to this was done by Epicoco (2013), who also applied main path analysis to the semiconductor industry using a dataset extracted from the USPTO (with patents ranging from 1976 to 2008). Nonetheless, she used keywords to select her initial nodes and the main path she obtained is completely different (i.e. not even a single patent in common) from all the ones showed here.

Other methods of retrieving citations were also tested. If a closed system is used instead, i.e. taking into consideration only citations made by core patents to core patents, a smaller network (with 0.6 million nodes and 5.4 million arcs) is obtained, but only minor changes in the main path are observed, namely in the beginning of the network. There are not significant differences content-wise. Alternatively, if we move back to an open system and consider not only citations made by core patents but also citations received by core patents, we obtain a larger network (with 1.8 million nodes and 10.5 million arcs) and the main path changes drastically: it diverges from the benchmark main path in 1990 and starts to follow a set of patents that refer not to semiconductor technology, but to electronic equipment in general.

Transversal weights

The second step in main path analysis is to attribute weights to the arcs of the network.[7] This weight, called transversal weight, measures "the extent to which a particular citation or article is needed for linking articles", since, "if knowledge flows through citations, a citation that is needed in paths between many articles is more crucial than a citation that is hardly needed for linking articles" (De Nooy et al., 2005, p. 245).

According to De Nooy et al. (2005), there are three ways to compute transversal weights:

- The Search Path Count (SPC) counts all paths[8] starting at a source (a patent that is not citing within the network) and ending at a sink (a patent that is not cited within the network) and determines on how many of these paths each arc, i.e. each citation, is on.
- The Search Path Link Count (SPLC) counts all paths starting at any node and ending at a sink and determines on how many of these paths each arc is on. The authors warn that "citations of early articles receive lower weights because they cannot be part of paths emanating from later articles" (De Nooy et al., 2005, p. 247).
- The Search Path Node Pair (SPNP) counts all paths starting at any node and ending at any node and determines on how many of these paths each arc is on. In this case, arcs in the middle of the network "will receive higher traversal weights" (De Nooy et al., 2005, p. 247).

SPC is the weight preferred by most authors because, besides not introducing biases, it follows Kirchhoff's node law, i.e. the sum of the inflow transversal weights is equal to the sum of the outflow transversal weights (Batagelj, 2003). Nonetheless, using the messenger and tollway analogy,[9] Liu et al. (2019) argue that SPLC is the best weight from a

---

[7] Following the majority of the literature, the software Pajek was used to compute transversal weights and find main paths.

[8] Transversal weights differ from betweenness to the extent that the latter takes into consideration only the shortest paths.

[9] Liu et al. (2019) explain that "knowledge flow in a citation network is carried out by an imaginary messenger who takes knowledge from an origin document and sends it to a destination document through citation chains that connect the documents. For each pair of the specified origin and destination documents, many alternative paths exist running from the origin to the destination. While traversing the chains, the messenger is obliged to pay a toll when passing each citation link. For a citation link situated at a structural position where the messengers are more likely to pass through to complete the mission, it eventually collects more toll than those otherwise" (p. 385).
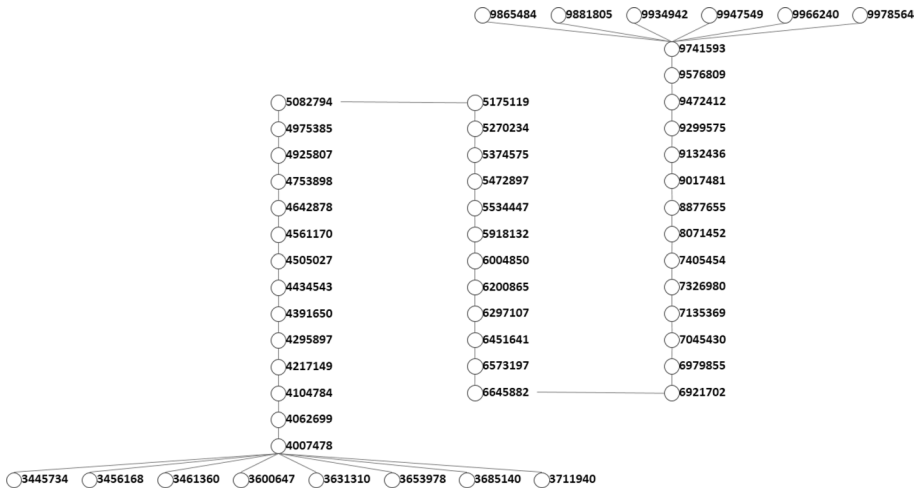
Fig. 4 Main path for the semiconductor manufacturing industry

knowledge flow perspective, because, in this algorithm, the "messenger takes knowledge not only from all sources, but also from all the intermediates and sends them to all the sinks" (p. 385), better representing how knowledge diffusion happens in a science and technology environment.

In order to, again, check the robustness of main path analysis, the three transversal weights were tested using our benchmark network. The same main path was obtained for all of them. Liu et al. (2019) alert, however, that "[…] transversal weights are very sensitive to the network structure. The behaviour in one network structure is not guaranteed in another one" (p. 385).

Search strategies

The third step is to find the main path, which is "the path from a source vertex to a sink vertex with the highest traversal weights on its arcs" (De Nooy et al., 2005, p. 246). There are two main ways to find the main path: local search or global search. The local main path is found by choosing the source node with the outward arc with the highest weight as the starting point, picking the node at the head of that arc, then choosing the outward arc with the highest weight and picking the node at the head of that arc repeatedly until a sink node is reached. The global main path, on the other hand, is found by computing all paths from sources to sinks in the weighted network and choosing the one with the highest accumulated weight.[10]

---

[10] The global search is also known as the critical path method (CPM), which is the name given in project planning to an algorithm that identifies the longest stretch of dependent activities in a project and measures the time required to complete them from start to finish.

Table 4 Top-10 key-routes

Route	Weight		In global path	In local path
	Normalized by flow	Without normalization		
9,132,436 cites 9,017,481	0.22857	3.2704×10^{18}	x	x
9,017,481 cites 8,877,655	0.14500	2.0747×10^{18}	x	x
8,071,452 cites 7,405,454	0.10475	1.4988×10^{18}	x	
7,405,454 cites 7,326,980	0.10260	1.4680×10^{18}	x	
5,082,794 cites 4,975,385	0.08444	1.2082×10^{18}	x	x
9,023,734 cites 8,551,891	0.08218	1.1758×10^{18}		
5,966,597 cites 5,856,225	0.08077	1.1557×10^{18}		
8,877,655 cites 8,071,452	0.07690	1.1003×10^{18}	x	x
9,144,147 cites 8,466,073	0.07439	1.0644×10^{18}		
9,144,147 cites 8,449,942	0.07438	1.0643×10^{18}		

Figure 4[11] shows our benchmark main path, which is the main path obtained when the global search was applied to our benchmark network, i.e. the network which was compiled using the definition of the semiconductor industry provided by Van Looy et al. (2015) and by taking all citations made by core nodes into consideration. The transversal weight used here was the Search Path Count (SPC), although we saw above that the choice of transversal weight would not matter in this case.

If we perform a local search instead of a global one, we obtain a main path that deviates from our benchmark path in the beginning and in several points in the middle. Content-wise, while the initial nodes of the benchmark path concern the fabrication of transistors in general, the initial nodes in the local path refer to specific process technologies involved in the fabrication of transistor, e.g. etching, manufacture of substrates, and preparation of masks, and to the fabrication of memories. The divergences in the middles of the main path are not so significant though, because the patents are fairly similar in content to the patents in the benchmark path.

Another possible search strategy is the key-route search, either global or local, which guarantees that the link(s) with the highest transversal weight(s) are part of the main path. Table 4 shows the 10 arcs with the highest SPC in our benchmark network. If we perform a local key-route search with the top-1 key-route, the main path algorithm starts from the key-route (the link between patents 9,017,481 and 9,132,436) and goes forward in the direction of the sinks and backward in the direction of the sources, always looking for the arc with the highest transversal weight. Even though the top-1 key-route was already present in the local path, the local key-route path differs from the local path in its bottom part—precisely because the algorithm is going from top to bottom instead of going from bottom to top. The local key-route path is now more similar to the benchmark, deviating from it in three short sections and having the same group of nodes as sources.

---

[11] The proper way to draw citation networks is a matter of debate, especially when it comes to the direction of the arrows. Some authors argue that the arrow should point in the direction of the citation, i.e. if i cites j, the arrow starts at i and finishes at j. Others say the arrow should point in the direction of the knowledge flow, i.e. if i cites j, knowledge is flowing from j to i and, accordingly, the arrow starts at j and finishes at i. Here, since all nodes have labels that correspond to the patent number, we decided to use lines instead of arrows to simplify the drawing. When a line connects two nodes, the node with the greater number as label will always be the one citing the other node.

If a local key-route search with the top-5 key-routes is performed instead, a network with more ramifications is obtained, as the algorithm tries to find a path connecting all five key-routes. If, instead of performing a local key-route search, we perform a global key-route search (either with the top-1 or the top-5 key-routes), we obtain a path that is equal to the benchmark path, because all the top-5 key-route are already on it and the global path is the same going forward or backward. Content-wise, patents included in the main path by the local key-route search refer to technologies that were already present in the benchmark main path.

Description of the main path[12]

Our benchmark main path begins with several patents that refer to the fabrication of complementary metal–oxide–semiconductor (CMOS) circuits[13] and to some of the challenges of scaling transistors. Patents 3,461,360 (IBM, 1969) and 3,600,647 (General Electric, 1971), for example, describe the essential process steps to make nMOS and pMOS devices on the same substrate. Patent 3,653,978 (Philips, 1971) refers to ion implantation, which provided better control of the dopant density and profile, and patents 3,685,140 (General Electric, 1972), 3,711,940 (Signetics, 1973), 4,007,478 (Sony, 1977) and 4,062,699 (Western Digital, 1977) refer to engineering the dopant control, with the goal of improving the threshold voltage control, avoiding punch-through and short channel effects, and maximizing the performance. Patent 4,505,027 (Siemens, 1985) deals with silicides, which were important to reduce the contact resistance and overall resistance in CMOS circuits; the technology specifically mentioned on this patent ultimately failed though.

Patent 4,753,898 (Motorola, 1988) described a process for fabricating lightly dopped drain CMOS structures, which was an essential progress in the technology. Patents 4,975,385 (Applied Materials, 1990), 5,082,794 (Motorola, 1992) and 5,175,119 (Fujitsu, 1992) describe progress in the scaling of CMOS gates by engineering the gate structure, spacers and lightly doped drain. Patent 5,270,234 (IBM, 1993) refers to a technique to aggressively scale the gate length of CMOS transistors. Patents 5,472,897 (United Microelectronics, 1995) and 5,534,447 (United Microelectronics, 1996) describe a technology to control punch-through, which was one of the major issues that needed to be solved to allow scaling of devices. Patent 5,918,132 (Intel, 1999) refers to spacer lithography, which allowed scaling beyond what was achievable with conventional lithography methods.

Then, beginning with patent 6,200,865 (AMD, 2001), there is a series of patents referring to materials with high dielectric constants (κ) and to high-κ/metal gate technology. Two patents (6,451,641, AMD, 2002 and 6,573,197, IBM, 2003) describe technologies that combine high-κ dielectrics with poly-Si gates and one patent (6,645,882, AMD, 2003) describes a technology that combines high-κ and standard-κ dielectrics, both of which ultimately failed. Several patents mention materials with high dielectric constants that were never adopted commercially, for example patents 6,979,855 (Micron, 2005) and 7,045,430 (Micron, 2006).

Finally, on the last part of the main path, starting with two patents on the deposition of metal oxides using atomic layer deposition (ALD) (8,071,452, ASM, 2011) and plasma ALD (8,877,655, ASM, 2014), there are several patents from suppliers of tools for the

---

[12] Table 8 in Appendix 3 presents the number, title, assignee and grant date of all patents in the benchmark main path.

[13] Patent 3,445,734 (IBM, 1969) is an exception, as it refers to bipolar transistors.

Table 5 Panel of experts

#	Years of experience	Job title	Type of organization
1	34	Fellow	Leading Research Institute
		Professor	University
2	50	(Formerly) Director of Technology Strategy	Leading Corporation
		(Currently) Chairman	Industry-wide Initiative
		IEEE Life Fellow	
3	45	Professor	University
		IEEE Life Fellow	
4	20	R&D technical manager	Leading Corporation
5	13	Principal Member of Technical Staff	Leading Research Institute

semiconductor industry. Most of the patents that follow are specific tool designs or specific process technologies by Applied Materials.

External validation

The semiconductor industry is a good candidate for the study of technological paradigms and trajectories because it has a very clear trajectory that spans more than 50 years, summarized by Moore's Law and the evolution of CMOS technology. Moore's Law refer to the prediction made by Gordon Moore in 1965 that the number of components in an integrated circuit that resulted in the lowest cost per circuit would continue to double every year in the next ten years. As Moore's prediction began to prove itself remarkably accurate, semiconductor companies started building their production schedule around it. And so, the prediction became a self-fulfilling prophecy, that still holds after more than 50 years, although the exact pace of improvement along the trajectory varied from decade to decade, ranging from 12 to 24 months.

In order to assess if the patents in the main path indeed correspond to this technological trajectory, a panel of experts was contacted. This panel, as it can be seen in Table 5, is composed of five experts from different semiconductor companies, semiconductor research organisations or universities from around the world. All experts were shown files containing information about the patents from the benchmark path (patent number, title, date granted, assignee, inventor, and abstract)[14] and asked to assess whether or not these patents were a good representation or summary of the technological trajectory of the semiconductor industry (i.e. the miniaturization trajectory).

According to the panel, the patents that compose the main path do display a few interesting characteristics and correspond to some of the technologies that were important to the development of semiconductor technology over time and to its miniaturization trajectory in particular. Expert #1, for example, mentioned three interesting aspects of these patents. First, their assignees are more varied in the 1970s and 1980s, reflecting the large pool of players that were active in the semiconductor industry in that period, while more recent patents are assigned to a more limited number of companies and this is consistent with the

---

[14] Some of this information (patent number, the title, the assignee and the date granted) is available in Table 8 in "Appendix 3" section.

overall evolution of the industry, in the opinion of the expert. On this subject, Expert #4 commented that he found it interesting that several of the companies that are present in the main path "are out of the semiconductor business today". Second, the presence of equipment manufacturers can be clearly seen in recent patents, which reflects the increasing role played by these companies in the development of semiconductor technology in recent years (which is now largely embodied in semiconductor equipment and bought from specialized suppliers instead of being developed in house). Third, in regard to the technologies that show up in the main path, notable examples are the lightly doped drain process and high-κ dielectrics, which are technologies that provided solutions for important technological problems that were the focus of research and development in the 1980s and in the late 1990s/early 2000s respectively.

Nonetheless, in the experts' opinions, several factors prevent us from stating that this collection of patents obtained using main path analysis corresponds to the technological trajectory of the semiconductor industry. First, several technologies that were crucial for the miniaturization trajectory are absent. In this regard, Expert #4 provided a list of fourteen technologies that were crucial for the scaling of transistors and counted how many patents on the main path corresponded to each technology (see Table 6). He then noted that several of them are completely absent from the main path. In his words, "especially missing number 10 is odd; FinFETs are the transistor architecture of choice and there are a zillion FinFET-related patents".[15] Moreover, transistor scaling is the main driving force of Moore's law, but is not the only one. Expert #4 noted that other drivers of Moore's law— e.g. improvements in interconnect, lithography, circuit design, and memory—also did not show up in the main path.

Still in regard to important technologies missing from the main path, Expert #1 mentioned lithography and the FinFET as examples, as well as innovations in infrastructure like cleaning techniques and clean room technology in general. Expert #3, in turn, mentioned tungsten plugs, Copper metallization for interconnects, low-κ dielectric materials for multilevel metallization, lithography technologies, FinFET, and Silicon on insulator (SOI) as important technologies that are absent. Finally, Expert #5 stated that "this set of patents appears to be more process and materials focused, and the trajectory of 'scaling' seems to be incomplete" and cited silicidation, strained channels, enhanced electrostatics (FinFET, tri-gate) and UV and immersion lithography as missing technologies.

Second, several important players are completely absent (e.g. ASML, TSMC and Samsung) and other important players are under-represented (e.g. Intel and Toshiba). The fact that there is only one patent assigned to Intel in the main path was particularly striking for Expert #1 because this company is not only deemed as having patents of very good quality, but it is also the company where several of the most important technologies were first developed and adopted (e.g. 300 mm fabs, high-dielectrics and the FinFET). The absence of TSMC is also striking, since this company has been the leader in the dedicated foundry segment since the 1990s and has surpassed Intel as the technological leader of the semiconductor industry a few year ago, being now the main responsible for driving Moore's Law. Expert #5 mentioned Toshiba as an important company that is under-represented in the main path. He also mentioned that another group that is absent in the list is universities and cited the invention of the FinFET, which was patented by UC Berkeley.[16] Additionally, it is important to stress that there are plenty of patents from these players in the network.

---

[15] There are 1.275 patents in our benchmark network and 1.296 patents in the whole USPTO dataset that mention the word "FinFET" (or the variations "finfet", "Finfet", "FinFet", or "FINFET") in the abstract.

[16] Patent 6,413,802 (The Regents of the University of California, 2002).

Table 6 Key Transistor scaling technologies and their presence in the main path

#	Technology	Year	Invented/Developed at	Presence in the Main Path	
				Number of Patents	Patent IDs
1	Ion implantation for doping of silicon	Late 1950s	Bell Labs General Electric	5	3,456,168, 3,461,360, 3,600,647, 3,653,978, 4,434,543
2	Metal–oxide–semiconductor field-effect transistor (MOSFET)	Late 1950s	Bell Labs	0	
3	Complementary MOS (CMOS)	1963	Fairchild	3	4,217,149, 4,295,897, 4,434,543
4	Self-aligned gate process	1969	Fairchild	2	4,062,699, 4,295,897
5	Local oxidation of silicon (LOCOS)	1975	Philips	1	4,561,170
6	Self-aligned silicidation (Salicidation)	1982	Texas Instruments	1	4,505,027
7	Chemical–mechanical polishing (CMP)	1983	IBM	0	
8	Self-aligned spacers and extensions/halo's	Mid-1980s	Motorola	8	4,642,878, 4,753,898, 4,925,807, 4,975,385, 5,082,794, 5,175,119, 5,374,575, 5,534,447
9	Strained silicon	Mid-1990s	Intel	0	
10	FinFET	2000	UC Berkeley	0	
11	Shallow trench isolation (STI)	2000	IBM	1	4,561,170
12	High-k/metal gate by atomic layer deposition (ALD)	2000	Micron Applied Material	12	6,297,107, 6,451,641, 6,573,197, 6,645,882, 6,921,702, 6,979,855, 7,045,430, 7,135,369, 7,326,980, 7,405,454, 8,071,452, 8,877,655
13	Double patterning	Early 2000s	Micron	1	5,918,132
14	Replacement gate process	Early 2000s	Intel	2	6,004,850, 6,200,865

Fig. 5 Assessment of the relevance of main path patents from a technological point of view—Expert #1. Notes: 7–Very important, 6–Important, 5–Important for tool manufacturers/process technology with broad applicability, 4–Specific tool design/Specific process technology, 3–Not important, 2–Failed technological solution, 1–Not CMOS related

Fig. 6 Assessment of the relevance of main path patents from a technological point of view—Expert #4. Notes: 5–Key patent, 4–Patent relevant for digital computation, 3–Patent by equipment vendor, 2–Dead end, 1–Not Moore's law related

As it is shown in Table 7 in Appendix 2, Samsung is the second most important assignee in term of number of patents in subclass H01L, TSMC is the fourth, Toshiba is the fifth and Intel is the nineth.

The fact that there are so many patents assigned to IBM in the main path is also interesting: according to the Expert #1, although IBM is indeed an important and innovative player in the semiconductor industry, the solutions developed by IBM have limited impact throughout the industry because they are often not regarded as cost-effective by other players. Expert #4, on the other hand, believes that "some of the more important inventions by IBM are missing" from the main path.

Third, Experts #1 and #4 noted that, although several of the recent patents in the main path try to solve technological problems that were indeed relevant, they represent technological solutions that ultimately did not work out and could be therefore considered as failed technologies or dead ends. Both these experts submitted a detailed

assessment of the importance of each patent in the main path. To simplify visualization, we created a scale for each expert, trying to remain as truthful as possible to their own words. Figure 5 shows the classification received by patents in the main path according to Expert #1 and Fig. 6 show the classification received by patents according to Expert #4. In both Figures, the patents are organized in a time scale according to the date they were granted.

In Expert's #1 detailed assessment, 20 patents were classified in the top three categories, while 31 were classified in the bottom four categories. It is also interesting to note that from the 10 patents classified as failed technological solutions, nine are concentrated in a short time period from 2002 to 2011 and form a sequence in the main path interrupted by only one patent regarded as important.

Fourth, the experts also alerted that firms' patenting behaviour may play a role and that patents may not be the best indicator to study technological trajectories, especially on their own. In regard to this matter, Expert #1 raised several points. First, he argued that new technologies are not necessarily protected via patents. To illustrate this, the expert explained that, in the early 2000s, everyone in the industry already knew that using materials with a high dielectric constant as gate oxide was the solution to the problem of increasing gate capacitance without the associated leakage effects. Nonetheless, researchers were still looking for the material that would provide the best gate. Intel was the first company to find out which was the best material for applications in logic chips and instead of patenting their discovery (which would in fact have disclosed it to everyone else), they decided to keep it as an industrial secret. To do so, they continued buying precursors for all the possible materials that could be used as gate oxides and managed to successfully protect their discovery for some years. At the same time, other companies were experimenting a lot and that can be seen in the many patents on high-κ materials present in the main path. However, the specific solutions for high-κ dielectrics displayed in some of these patents would eventually not be applied and they can be seen, in retrospect, as failed technological solutions.

There are also other reasons not to patent an innovation. Companies may choose not to patent process technologies because it is difficult to enforce these patents, since it is difficult to determine exactly which process was used to produce a semiconductor device without access to the factory. Companies may also not patent important technologies because they are eco-friendly (e.g. cleaning techniques that save water) or because they want other players to join the effort of further developing that technology.

Second, the expert pointed out that bad results usually yield more publications (either in the form of patents or papers) than good results, because researchers can publish several different incremental results instead of one good result that is final. Additionally, patents that contain a large volume of references probably refer to very incremental technologies and the patenting process of these technologies probably involved several rounds with the patent office in which the company filing the patent had to show additional evidence (in the form of additional references) of the novelty of the invention being patented.
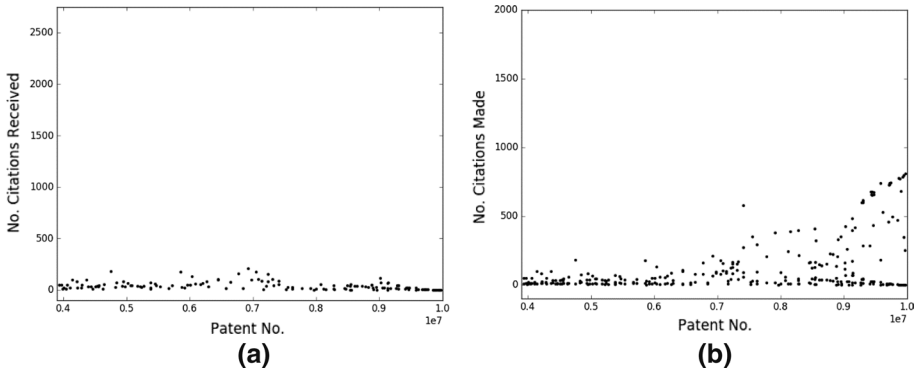
Fig. 7 Citations received (**a**) and citations made (**b**) by main path patents over

Third, patents are nowadays important for reasons other than protecting intellectual property and these reasons influence the propensity to patent of individual companies. Patents serve, for example, as a mechanism for the market to assess the value of companies. They may also be used by companies to make cross-licensing agreements with or to block other companies. On this subject, Expert #2 noted that from the 1950s to the mid-1990s trade secrets and scientific publications marked technological progress more appropriately than patents. Moreover, he reported that Intel actually discouraged employees from filling patents and publishing papers until the mid-1990s. In his words, "[Gordon Moore] was bitter for the Fairchild experience in which good publications and key patents did not lead to the level of economic success that he had in mind. This led to Intel formation with the new guidelines mentioned above [...]". This policy changed in 1995 because Intel became the industry leader and it was important to show why and by how much Intel was leading and because other companies were suing Intel and patents could be used as a defence mechanism. Expert #3, in turn, mentioned the case of a German company, which would either keep new technologies as secrets or publish them in German in local journals, so that it would become prior, but not widespread, knowledge.

In order to cast some light on why so many failed technological solutions and unimportant technologies are present in the main path, we looked at the number of citations received and made by all patents in the main path. Figure 7 shows the number of citations received and the number of citations made by all patents that were mentioned in this paper as being part of a main path or being a top 10 key-route. Again patents are organized in a time scale according to the date they were granted using their ID as a proxy for time. While there is no clear upward or downward trend in the number of citations received, there is a clear upward trend in the number of citations made, notably from the mid-2000s onwards. Additionally, if we look at the average number of citations for this collection of 167 patents, their average number of citations received is only 1.55 standard deviations above the average of the whole network, while their average number of citations made is 6.43 standard deviations above the average of the whole network.

One explanation for this is the integrator effect, which "increases the significance of a document that heavily references others" (Liu et al., 2019, p. 388). The literature on main path analysis has already pointed out that review papers tend do show up in main paths often, since the numerous streams of knowledge that flow into them increase the transversal weight of arcs leading out of them. The presence of review papers in main paths is

usually regarded as a positive aspect though, since review papers are indeed an important source of information (i.e. a building block) for future work. Nonetheless, there is no such thing as a review patent. Alternatively, one could think that a patent that references a large number of other patents refers to a technology that combines several existing technologies to produce a new application. At least here, this is not the case. All patents in our main path with a large number of citations made are related to very specific technologies and represent very incremental innovations. Therefore, the integrator effect appears to produce an artefact in our main path, where patents that cite many other patents are privileged over others.

Discussion

The external validation presented above has indicated that the main path cannot be said to represent the backbone of the technological trajectory of the semiconductor industry. If we rely only on the trajectory produced by main path analysis, we would be led to conclude that the semiconductor industry is limited to a much narrower neighbourhood of the technology space than it really is, because several important technologies are missing. So, instead of a reliable recollection of the history of the semiconductor industry, the main path provided us, at best, with a sample of it.

The presence of so many failed technological solutions in the main path could also point to the inadequacy of the method in portraying the technological trajectory, since it could indicate that the main path represents more a sample of the technology space than of the technological trajectory. Nonetheless, the presence of failed solutions could also indicate that the trajectory builds up more on failed solutions than previously thought. As long as the problems being tackled are relevant from the point of view of the technological trajectory, significant learning could be derived even from failed solutions. Here, it is interesting to note that almost all patents regarded as failed solutions have a high number of citations received (with only two with less than 30 citations). The number of citations made by these patents, on the other hand, progressively increases over time, from around 10–20 earlier to 100–500 later on.

We are left, then, with the question of why main path analysis does not produce the results it arguably would. One possibility is that the method works best with single technologies instead of clusters of technologies. And semiconductor technology is clearly a cluster technology: for the dimension of devices to decrease and the density of devices in integrated circuits to increase (as Moore's Law states), improvements in several sub-technologies must occur. Over the years, the wavelength of lithography decreased from 365 to 193 nm and eventually to extreme ultraviolet (EUV) lithography, Silicon dioxide was replaced by high-κ materials in the gate oxide and Polysilicon was replaced by metals in the gate electrode, the architecture of transistors moved from planar to the FinFET, and wafers increased from 200 to 300 mm, to name a few examples of improvements. Main path analysis, on the other hand, assumes connectivity (i.e. citations) between innovations and it is somewhat not realistic to expect all relevant sub-technologies to show up in the main path, because there is no reason to expect that there is a citation connecting patents from different sub-technologies. Nonetheless, it is important to emphasize that although main path analysis may not be suitable to analyse clusters of technologies, the concepts of technological paradigms and trajectories certainly are, as Dosi (1982) stated himself.

To address the issue above, one could argue that a narrower definition of semiconductors would help. This, however, is not the case. One of the definitions of the semiconductor industry used here, Schmoch (2008) Narrow, includes one subclass only (H01L) and it produced a main path very similar to our benchmark path. And going lower than subclass (into groups, the six-digit level of the CPC) is not very straightforward, since these groups do not each match one sub-technology of semiconductor technology, e.g. there is not one group that corresponds to transistor architecture while another corresponds to gate oxide.

Another option would be to try to find derivate paths, as proposed by Kim and Shin (2018), i.e. to assemble on network for each subclass that makes up our definition of the semiconductor industry, find a main path in each one of them and look for connections between each of these derivative main paths and our benchmark main path. We applied this strategy to the definition of the semiconductor industry labelled as "Van Looy et al. (2015) Plus" in Table 2. This strategy does not greatly improve the results though. Most of these derivate paths are not actually derivate, because they do not cross, at any point, the benchmark main path. Only the main paths generated using subclasses C30B and H01J cross the benchmark main path at some point.

Subclass C30B refers to Single-crystal-growth and its main path contains patents relating mostly to Silicon carbide and, after the late 1990s, to nitrides and gallium-based crystals. Companies like IBM, Micron, Philips and Siemens are present in this path. The path ends on a series of patents by a company that manufactures lighting products called Soraa, Inc. None of the technologies present in this main path is relevant for the main trajectory of the semiconductor industry, since Silicon carbide is used in semiconductor devices that operate at high temperatures or high voltages, or both and Gallium nitride is used in blue light-emitting diodes. Subclass H01J refers to Electric discharge tubes or discharge lamps and its main path contains patents relating to sputtering, plasma deposition and plasma etching. Two thirds of the patents in this main path belong to Applied Materials or Lam Research, which are important suppliers for the semiconductor industry.

Two other derivate paths that are worth mentioning are the ones generated by the networks corresponding to subclasses C23C and G03F, although they are not directly connected with the benchmark path. Subclass C23C refers to coating by sputtering, ion implantation and chemical vapor deposition, which are important processes used in the fabrication of semiconductor devices, and, as expected, its main path is essentially composed of patents related to film deposition. What is interesting here is that the pattern in terms of assignees in similar to the one found in the benchmark path: earlier patents belong to a more varied pool of players who were important semiconductor manufacturers themselves (e.g. Bell Labs, Fairchild, Motorola, Philips and Siemens), while later patents mostly belong to suppliers such as Applied Materials and Lam Research.

Finally, subclass G03F refers to lithography, which is one of the major technologies missing from our benchmark main path, so one would expect this derivative main path to nicely complement it. Nonetheless, the relevance of this derivative main path is questionable: there are only three patents belonging to ASML, which is the undisputed leader on this market, and while most patents refer to the photo masks used in lithography, none of the key suppliers of photo masks (e.g. Dai Nippon Printing, Toppan Photomasks and Photronics) is present here.

Another explanation for the apparent inadequacy of main path analysis in identifying the backbone of the technological trajectory of the semiconductor industry is that the structure of the patent-citation network is being influenced by the patenting behaviour of companies. In our network, the major problem that appears to be related with the patenting behaviour of companies is the fact that most of the more recent patents in our main paths make too

many citations. It could be argued that the high number of citations made is an artefact of method used to compile the network, since all citations made by core nodes were included. Appendix 4 shows the degree distribution (Fig. 8) and the descriptive statistics (Table 9) of the closed network equivalent to our benchmark network, i.e. the network obtained when we include not all citations made by core nodes (i.e. open) but only the citations made by core nodes to other core nodes (i.e. closed). Figure 8b shows that the closed network has almost no patents with more than 700 citations made, which is a remarkable difference when compared to the benchmark network (see Fig. 1b). Nonetheless, the closed network still had a significant number of patents that referenced around 500 other patents, which is still a very large number. Additionally, applying main path analysis to the closed network produced essentially the same result as applying it to the open network.

Final remarks

Our analysis has shown that main path analysis is a robust method. When the method was applied to different networks representing the semiconductor industry (either in terms of core nodes or the method used for collecting citations) and using different specifications (i.e. different transversal weights and different search strategies), similar results were obtained, especially if we analyse them in terms of the content of patents and not of the specific patents present in the main path. The only exception was when both citations made by and received by core nodes were taken into consideration, which created too much noise in the network.

In sum, using a network either with all citations made by core nodes or just with the citations made by core nodes to core nodes, SPC as the transversal weight, and a global search strategy seems to produce good and concise results. Nevertheless, since Liu et al.

([2019](#)) alert that "[main path analysis] results are […] heavily determined on the network structure and are sensitive to citation data" (p. 382), we would recommend that several possibilities be explored every time the method is applied to a new set of data.

In terms of external validity, however, our analysis has indicated that the main path cannot be said to represent the backbone of the technological trajectory of the field in question. It is important to remember that we are not looking for any trajectory, i.e. any sequence of cumulative and selective technological developments. We are looking for the main technological trajectory within a technological paradigm, which corresponds, in this case, to the evolution of CMOS technology.

Our main findings were that the main path is incomplete in terms of the representation of the technological trajectory that we were trying to identify, that there are many failed technological solutions on the main path, and that many patents in the main path have an impressively high number of citations made, and these findings prompted us, therefore, to question the appropriateness of the method for identifying technological trajectories using patent-citation networks. This opens a compelling avenue for future work on the understanding of the main path analysis algorithm, especially regarding the role played by patents that make many citations and the suitability of the method for analysing cluster technologies.

Finally, perhaps it is time to restart the conversation on the use of patents as an indicator of innovation. Researchers usually acknowledge that working with patents has drawbacks, the two most important being that (i) not all patents lead to innovations, and (ii) the propensity to patent an innovation may differ not only between sectors but also from company to company. Nevertheless, researchers also assume that these drawbacks do not compromise the results of the analysis and this is an assumption that we might have to stop making.

On the one hand, we not only cannot assume that all patents lead to innovations, but we also must take into consideration that not all innovations have been patented. There are other mechanisms to protect intellectual property and their use is non negligible, so many inventions do not become patents and are not captured by studies such as the present one, that rely only on patent information. On the other hand, the fact that seeking protection may also not even be the main reason behind patenting may cause the propensity to patent an invention not only vary, but vary greatly between firms in the same sector, making these firms score better than they should in innovation indicators that rely on patent information. This is especially worrisome since many policy recommendations are derived from such works.

Appendix 1:Technology classes' titles

Appendix 1.1: USPC classes

- 117: Single-crystal, oriented-crystal, and epitaxial growth processes; non-coating apparatus therefor
- 257: Active solid-state devices (e.g., transistors, solid-state diodes)
- 326: Electronic digital logic circuitry
- 438: Semiconductor device manufacturing: process
- 505: Superconductor technology: apparatus, material, process
- 716: Computer-aided design and analysis of circuits and semiconductor masks

Appendix 1.2: IPC subclasses

- B81B: Microstructural devices or systems, e.g. Micromechanical devices
- B81C: Processes or apparatus specially adapted for the manufacture or treatment of microstructural devices or systems
- B82B: Nanostructures formed by manipulation of individual atoms, molecules, or limited collections of atoms or molecules as discrete units; manufacture or treatment thereof
- B82Y: Specific uses or applications of nanostructures; measurement or analysis of nanostructures; manufacture or treatment of nanostructures
- C23B: Coating metallic material; coating material with metallic material; Surface treatment of metallic material by diffusion into the surface, by chemical conversion or substitution; coating by vacuum evaporation, by sputtering, by ion implantation or by chemical vapour deposition, in general
- C30B: Single-crystal-growth…
- G03F: Photomechanical production of textured or patterned surface, e.g. for printing, for processing of semiconductor devices; materials there for; originals there for; apparatus specially adapted there for
- G11C: Static Stores
- H01C: Resistors
- H01F: Magnets; inductances; transformers; selection of materials for their magnetic properties
- H01G: Capacitors; capacitors, rectifiers, detectors, switching devices or light-sensitive devices, of the electrolytic type
- H01J: Electric discharge tubes or discharge lamps
- H01L: Semiconductor devices; electric solid state devices not otherwise provided for
- H05K: Printed circuits; casings or constructional details of electric apparatus; manufacture of assemblages of electrical components

Appendix 2: Top assignees by IPC subclass

Table 7 shows the top 10 assignees in each of the IPC subclasses mentioned in this research. The table was produced by first collecting all the patents that belong to each subclass and then obtaining the assignees of each patent using the disambiguated files provided by the USPTO (namely the "patent assignee.tsv" and the"assignee.tsv" files). No further disambiguation was done and the organization's name is presented here in exactly the same way as in the file "assignee.tsv". Simple counts were used to estimate how many patents are owned by each organization in each CPC subclass. The last line of each subclass shows how many patents can be found in that subclass in total.

Appendix 3: Main path patents

See Table 8.

Table 7 Top assignees by IPC subclass

#	Organization	Qty	#	Organization	Qty
B81B			**B81C**		
1	Robert Bosch GmbH	392	1	Robert Bosch GmbH	352
2	Taiwan semiconductor manufacturing company, Ltd	228	2	Taiwan semiconductor manufacturing Company, Ltd	259
3	Texas instruments incorporated	173	3	International business machines corporation	209
4	International business machines corporation	157	4	Texas instruments incorporated	160
5	Infineon technologies AG	153	5	Inneon technologies AG	153
6	Analog devices, Inc	152	6	STMICROELECTRONICS S.R.L	133
7	STMICROELECTRONICS S.R.L.	145	7	Analog devices, Inc	131
8	QUALCOMM MEMS technologies, Inc	135	8	Samsung electronics Co., Ltd	99
9	Honeywell International, Inc	127	9	Honeywell International, Inc	99
10	FREESCALE SEMICONDUCTOR, INC	111	10	QUALCOMM MEMS technologies, Inc	94
Total		6927	Total		6976
B82B			**B82Y**		
1	Nanosphere, Inc	36	1	International business machines corporation	1733
2	NanoProducts corporation	30	2	Kabushiki Kaisha Toshiba	1103
3	William marsh rice University	24	3	Samsung electronics Co., Ltd	1095
4	Northeastern University	18	4	Canon Kabushiki Kaisha	930
5	Tsinghua University	16	5	The regents of the University of California	708
6	Hon Hai precision industry Co., Ltd	16	6	Hitachi, Ltd	625
7	International business machines corporation	14	7	TDK corporation	587
8	The regents of the University of California	12	8	Sony corporation	559
9	PPG Industries, Inc	11	9	Fujitsu limited	545
10	Nanosys, Inc	10	10	Hitachi global storage technologies Japan, Ltd	535
Total		468	Total		48,691
#	Organization	Qty	#	Organization	Qty
C23C			**C30B**		

Table 7 (continued)

#	Organization	Qty	#	Organization	Qty
2	Tokyo Electron Limited	1391	2	Shin-Etsu Handotai Co., Ltd	346
3	General Electric Company	1130	3	Applied Materials, Inc	220
4	Micron Technology, Inc	1096	4	The United States of America as represented by the Secretary of the Navy	188
5	International Business Machines Corporation	728	5	General Electric Company	173
6	Canon Kabushiki Kaisha	508	6	The Regents of the University of California	165
7	HITACHI KOKUSAI ELECTRIC INC	481	7	MEMC Electronic Materials, Inc	159
8	United Technologies Corporation	466	8	Sumco Corporation	156
9	Samsung Electronics Co., Ltd	451	9	International Business Machines Corporation	145
10	SEMICONDUCTOR ENERGY LABORATORY CO., LTD	412	10	Siemens Aktiengesellschaft	140
Total		49,180	Total		13,503
G03F			G11C		
1	ASML Netherlands B.V	2876	1	Micron Technology, Inc	7184
2	Canon Kabushiki Kaisha	2863	2	Samsung Electronics Co., Ltd	6049
3	NOK Corporation	2418	3	Kabushiki Kaisha Toshiba	5520
4	International Business Machines Corporation	1877	4	International Business Machines Corporation	3655
5	Taiwan Semiconductor Manufacturing Company, Ltd	1225	5	Hynix Semiconductor Inc	2691
6	Fuji Photo Film Co., Ltd	1118	6	SK hynix Inc	2173
7	Kabushiki Kaisha Toshiba	1054	7	Intel Corporation	2093
8	Shin-Etsu Chemical Co., Ltd	961	8	Fujitsu Limited	1914
9	Samsung Electronics Co., Ltd	928	9	Mitsubishi Denki Kabushiki Kaisha	1836
10	Carl Zeiss SMT GmbH	837	10	Hitachi, Ltd	1791
Total		50,926	Total		90,412
#	Organization	Qty	#	Organization	Qty
H01C			H01F		

Table 7 (continued)

#	Organization	Qty
1	Murata Manufacturing Co., Ltd	293
2	Matsushita Electric Industrial Co., Ltd	163
3	RCA Corporation	153
4	TDK Corporation	131
5	Rohm Co., Ltd	111
6	General Electric Company	110
7	CMD Corporation	85
8	Siemens Aktiengesellschaft	78
9	Epcos AG	77
10	Alps Electric Co., Ltd	73
Total		5900

H01G

#	Organization	Qty
1	Murata Manufacturing Co., Ltd	951
2	Samsung Electronics Co., Ltd	660
3	TDK Corporation	588
4	Matsushita Electric Industrial Co., Ltd	383
5	Sanyo Electric Co., Ltd	236
6	AVX Corporation	223
7	Taiyo Yuden Co., Ltd	209
8	Kemet Electronics Corporation	185
9	Sprague Electric Company	184
10	Panasonic Corporation	172
Total		15,170

#	Organization	Qty

H01L

#	Organization	Qty
1	TDK Corporation	725
2	General Electric Company	616
3	Murata Manufacturing Co., Ltd	545
4	Kabushiki Kaisha Toshiba	519
5	International Business Machines Corporation	444
6	Samsung Electronics Co., Ltd	444
7	Siemens Aktiengesellschaft	332
8	Hitachi, Ltd	324
9	Hitachi Metals, Ltd	318
10	Matsushita Electric Industrial Co., Ltd	284
Total		27,566

H01J

#	Organization	Qty
1	Applied Materials, Inc	2104
2	Hitachi, Ltd	1992
3	Tokyo Electron Limited	1544
4	U.S. Philips Corporation	1544
5	Canon Kabushiki Kaisha	1317
6	General Electric Company	1240
7	Hitachi High-Technologies Corporation	1162
8	Kabushiki Kaisha Toshiba	1093
9	Lam Research Corporation	951
10	Samsung SDI Co., Ltd	897
Total		59,627

#	Organization	Qty

H05K

Table 7 (continued)

#	Organization	Qty	#	Organization	Qty
1	International Business Machines Corporation	18,793	1	International Business Machines Corporation	4496
2	Samsung Electronics Co., Ltd	14,665	2	Hon Hai Precision Industry Co., Ltd	2006
3	Micron Technology, Inc	14,248	3	Intel Corporation	1876
4	Taiwan Semiconductor Manufacturing Company, Ltd	12,442	4	Samsung Electronics Co., Ltd	1745
5	Kabushiki Kaisha Toshiba	11,869	5	Fujitsu Limited	1387
6	SEMICONDUCTOR ENERGY LABORATORY CO., LTD	9582	6	Matsushita Electric Industrial Co., Ltd	1273
7	Texas Instruments Incorporated	6500	7	Micron Technology, Inc	1168
8	Inneon Technologies AG	5876	8	Hitachi, Ltd	952
9	Intel Corporation	5493	9	Murata Manufacturing Co., Ltd	952
10	Sony Corporation	5181	10	Motorola, Inc	864
Total		397,683	Total		90,459

Table 8 Patents in the main path

Patent No	Title	Assignee	Date granted
3,445,734	Single diffused surface transistor and method of making same	IBM Corporation	20/05/1969
3,456,168	Structure and method for production of doped region semiconductor devices	United Aircraft Corporation	15/07/1969
3,461,360	Semiconductor devices with cup-shaped regions	IBM Corporation	12/08/1969
3,600,647	Field-effect transistor with reduced drain-to-substrate capacitance	General Electric Company	17/08/1971
3,631,310	Insulated gate field affect transistor	U.S. Philips Corporation	28/12/1971
3,653,978	Method of making semiconductor devices	North American Philips Corporation	04/04/1972
3,685,140	Short channel field effect transistor	General Electric Company	22/08/1972
3,711,940	Method for making MOS structure with precisely controlled channel length	Signetics Corporation	23/01/1973
4,007,478	Field effect transistor	Sony Corporation	08/02/1977
4,062,699	Method for fabricating diffusion self-aligned short channel MOS device	Western Digital Corporation	13/12/1977
4,104,784	Manufacturing a low voltage n-channel MOSFET device	National Semiconductor Corp	08/08/1978
4,217,149	Method of manufacturing complementary insulated gate field effect semiconductor device by multiple implantation and diffusion	Sanyo Electric Co. and Tokyo Sanyo Electric Co	12/08/1980
4,295,897	Method of making CMOS integrated circuit device	Texas Instruments Incorporated	20/10/1981
4,391,650	Method for fabricating improved complementary metal oxide semiconductor devices	NCR Corporation	05/07/1983
4,434,543	Process for producing adjacent tubs implanted with dopant ions in the manufacture of LSI complementary MOS field effect transistor	Siemens Aktiengesellschaft	06/03/1984
4,505,027	Method of making MOS device using metal silicides or polysilicon for gates and impurity source for active regions	Siemens Aktiengesellschaft	19/03/1985
4,561,170	Method of making eld-plate isolated CMOS device	Texas Instruments Incorporated	31/12/1985
4,642,878	Method of making MOS device by sequentially depositing an oxidizable layer and a masking second layer over gated device regions	Kabushiki Kaisha Toshiba	17/02/1987
4,753,898	LDD CMOS process	Motorola, Inc	28/06/1988
4,925,807	Method of manufacturing a semiconductor device	Kabushiki Kaisha Toshiba	15/05/1990
4,975,385	Method of constructing lightly doped drain (LDD) integrated circuit structure	Applied Materials, Inc	04/12/1990
5,082,794	Method of fabricating MOS transistors using selective polysilicon deposition	Motorola, Inc	21/01/1992
5,175,119	Method of producing insulated-gate field effect transistor	Fujitsu Limited	29/12/1992
5,270,234	Deep submicron transistor fabrication method	IBM Corporation	14/12/1993

Table 8 (continued)

| Patent No	Title	Assignee	Date granted
5,374,575	Method for fabricating MOS transistor	Goldstar Electron Co., Ltd	20/12/1994
5,472,897	Method for fabricating MOS device with reduced anti-punchthrough region	United Microelectronics Corp	05/12/1995
5,534,447	Process for fabricating MOS LDD transistor with pocket implant	United Microelectronics Corp	09/07/1996
5,918,132	Method for narrow space formation and self-aligned channel implant	Intel Corporation	29/06/1999
6,004,850	Tantalum oxide anti-reflective coating (ARC) integrated with a metallic transistor gate electrode and method of formation	Motorola, Inc	21/12/1999
6,200,865	Use of sacrificial dielectric structure to form semiconductor device with a self-aligned threshold adjust and overlying low-resistance gate	Advanced Micro Devices, Inc	13/03/2001
6,297,107	High dielectric constant materials as gate dielectrics	Advanced Micro Devices, Inc	02/10/2001
6,451,641	Non-reducing process for deposition of polysilicon gate electrode over high-K gate dielectric material	Advanced Micro Devices, Inc	17/09/2002
6,573,197	Thermally stable poly-Si/high dielectric constant material interfaces	IBM Corporation	03/06/2003
6,645,882	Preparation of composite high-K/standard-K dielectrics for semiconductor devices	Advanced Micro Devices, Inc	11/11/2003
6,921,702	Atomic layer deposited nanolaminates of HfO2/ZrO2 lms as gate dielectric	Micron Technology, Inc	26/07/2005
6,979,855	High-quality praseodymium gate dielectrics	Micron Technology, Inc	27/12/2005
7,045,430	Atomic layer-deposited LaAlO3 lms for gate dielectric	Micron Technology, Inc	16/05/2006
7,135,369	Atomic layer deposited ZrAl.sub.xO.sub.y dielectric layers including Zr.sub.4AlO.sub.9	Micron Technology, Inc	14/11/2006
7,326,980	Devices with HfSiON dielectric lms which are Hf–O rich	Micron Technology, Inc	05/02/2008
7,405,454	Electronic apparatus with deposited dielectric layers	Micron Technology, Inc	29/07/2008
8,071,452	Atomic layer deposition of hafnium lanthanum oxides	ASM America, Inc	06/12/2011
8,877,655	Systems and methods for thin-lm deposition of metal oxides using excited nitrogen–oxygen species	ASM America, Inc	04/11/2014
9,017,481	Process feed management for semiconductor substrate processing	ASM America, Inc	28/04/2015
9,132,436	Chemical control features in wafer process equipment	Applied Materials, Inc	15/09/2015
9,299,575	Gas-phase tungsten etch	Applied Materials, Inc	29/03/2016
9,472,412	Procedure for etch rate consistency	Applied Materials, Inc	18/10/2016
9,576,809	Etch suppression with germanium	Applied Materials, Inc	21/02/2017
9,741,593	Thermal management systems and methods for wafer processing systems	Applied Materials, Inc	22/08/2017

Table 8 (continued)

Patent No	Title	Assignee	Date granted
9,865,484	Selective etch using material modi citation and RF pulsing	Applied Materials, Inc	09/01/2018
9,881,805	Silicon selective removal	Applied Materials, Inc	30/01/2018
9,934,942	Chamber with flow-through source	Applied Materials, Inc	03/04/2018
9,947,549	Cobalt-containing material removal	Applied Materials, Inc	17/04/2018
9,966,240	Systems and methods for internal surface conditioning assessment in plasma processing equipment	Applied Materials, Inc	08/05/2018
9,978,564	Chemical control features in wafer process equipment	Applied Materials, Inc	22/05/2018

Appendix 4: Degree distribution in a closed network

Figure 8 shows the degree distribution of the network compiled using the same core patents as in our benchmark network (Van Looy et al., 2015), but including only citations made by core nodes to cores nodes. Table 12 shows the descriptive statistics.
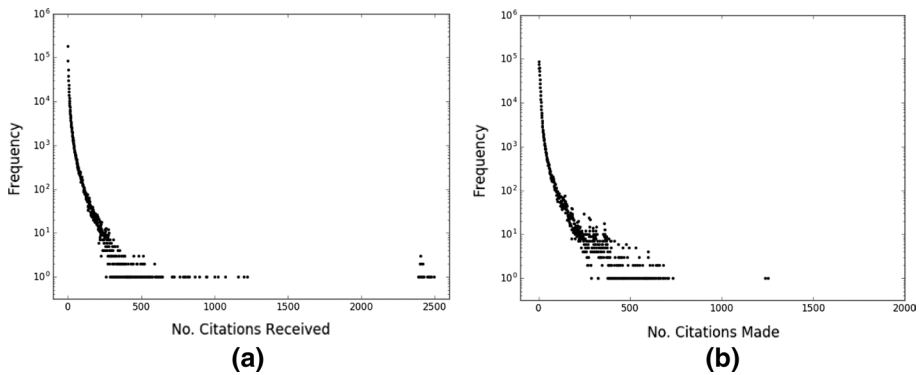
Fig. 8 Frequency distribution of citations received (**a**) and citations made (**b**)

Table 9 Descriptive statistics of citations received and made	Citations received	Citations made
Minimum	0	0
25th Percentile	0	2
50th Percentile	2	4
75th Percentile	8	8
95th Percentile	38	30
99th Percentile	100	96
99.9th Percentile	251	315
Maximum	2495	1254
Mean	8.881323	8.881323
Standard deviation	26.352679	22.498583

Data availability All the data used in this paper was extracted from the US Patent and Trademark Office (USPTO) PatentsView Platform in August 2018. The PatentsView platform can be accessed freely at https://www.patentsview.org/web/.

Code availability Python was used to select the relevant data from the files downloaded from the PatentsView Platform and to produce some of the Figures. Other Figures were generated using Microsof Excel. The network analysis was performed using Pajek, which is available at http://mrvar.fdv.uni-lj.si/pajek/

Declarations

Conflict of interest The author declare that they don't have conflict of interest.

References

Batagelj, V. (2003). Efficient algorithms for citation network analysis. *University of Ljubljana, Institute of Mathematics, Preprint Series, 41*(897), 1–29

Batagelj, V., Ferligoj, A., & Squazzoni, F. (2017). The emergence of a field: A network analysis of research on peer review. *Scientometrics, 113*, 503–532

Barberá-Tomás, D., Jiménez-Sáez, F., & Castelló-Molina, I. (2011). Mapping the importance of the real world: The validity of connectivity analysis of patent citations networks. *Research Policy, 40*(3), 473–486

Calero-Medina, C., & Noyons, Ed. C. M. (2008). Combining mapping and citation network analysis for a better understanding of the scientific development: The case of the absorptive capacity field. *Journal of Informetrics, 2*(4), 272–279

Nooy, De., Wouter, M., & Andrej, & Batagelj, Vladimir. . (2005). *Exploratory social network analysis with Pajek*. Cambridge: Cambridge University Press.

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

Epicoco, M. (2013). Knowledge patterns and sources of leadership: Mapping the semiconductor miniaturization trajectory. *Research Policy, 42*(1), 180–195

Fontana, R., Nuvolari, A., & Verspagen, B. (2009). Mapping technological trajectories as patent citation networks An application to data communication standards. *Economics of Innovation and New Technology, 18*(4), 311–336

Hall, B. H., Jaffe, A. B., & Trajtenberg, M. (2001). The NBER patent citation data file: Lessons, insights and methodological tools. *National Bureau of Economic Research, Working Paper 8498*.

Ho, J. C., Saw, E.-C., Lu, L. Y. Y., & Liu, J. S. (2014). Technological barriers and research trends in fuel cell technologies: A citation network analysis. *Technological Forecasting and Social Change, 82*, 66–79

Hummon, N. P., & Doreian, P. (1989). Connectivity in a citation network: The development of DNA theory. *Social Networks, 11*(1), 39–63

Hung, S.-C., Liu, J. S., Lu, L. Y. Y., & Tseng, Y.-C. (2014). Technological change in lithium iron phosphate battery: The key-route main pathvanalysis. *Scientometrics, 100*(1), 97–120

IC Insights. (2011). Tracking the top 10 semiconductor sales leaders over 26 years. IC Insights.

IC Insights. (2015). The McClean Report. Scottsdale, AZ: IC Insights.

IC Insights. (2018). Nine top-15 2018 semi suppliers forecast to post double digit gains. IC Insights.

Kim, J., & Shin, J. (2018). Mapping extended technological trajectories: integration of main path, derivative paths, and technology junctures. *Scientometrics, 116*(3), 1439–1459

Lammers, D. (2015). Moore's law milestones. *IEEE Spectrum Online*.

Lin, Y., Chen, J., & Chen, Y. (2011). Backbone of technology evolution in the modern era automobile industry: An analysis by the patents citation network. *Journal of Systems Science and Systems Engineering, 20*(4), 416–442

Liu, J. S., & Lu, L. Y. Y. (2012). An integrated approach for main path analysis: Development of the Hirsch index as an example. *Journal of the American Society for Information Science and Technology, 63*(3), 528–542

Liu, J. S., Lu, L. Y. Y., & Ho, M.-C. (2019). A few notes on main path analysis. *Scientometrics, 119*(1), 379–391

Lucio-Arias, D., & Leydesdorff, L. (2008). Main-path analysis and path-dependent transitions in HistCiteTM-based historiograms. *Journal of the American Society for Information Science and Technology, 59*(12), 1948–1962

Maltseva, D., & Batagelj, V. (2019). Social network analysis as a field of invasions: Bibliographic approach to study SNA development. *Scientometrics, 121*, 1085–1128

Martinelli, A. (2012). An emerging paradigm or just another trajectory? Understanding the nature of technological changes using engineering heuristics in the telecommunications switching industry. *Research Policy, 41*(2), 414–429

Mina, A., Ramlogan, R., Tampubolon, G., & Metcalfe, J. S. (2007). Mapping evolutionary trajectories: Applications to the growth and transformation of medical knowledge. *Research Policy, 36*(5), 789–806

Mollick, E. (2006). Establishing Moore's law. *IEEE Annals of the History of Computing, 28*(3), 62–75

Moore, S., Haines, V., Hawe, P., & Shiell, A. (2006). Lost in translation: A genealogy of the "social capital" concept in public health. *Journal of Epidemiology & Community Health, 60*(8), 729–734

Schmoch, U. (2008). Concept of a technology classification for country comparisons. Tech. rept. Fraunhofer Institute for Systems and Innovation Research, Karlsruhe, Germany.

Triulzi, G. (2015). Looking for the right path: Technology dynamics, inventive strategies and catching-up in the semiconductor industry. Ph.D. thesis, Maastricht University.

USPTO. (2018). Semiconductor Devices and Manufacture: All Classified Utility Patents (1976–2015). Tech. rept. US Patent and Trademark Office.

Van Looy, B., Vereyen, C., & Schmoch, U. (2015). Patent statistics: Concordance IPC V8 – NACE Rev.2. Tech. rept. Eurostat.

Verspagen, B. (2007). Mapping technological trajectories as patent citation networks: A study of the history of fuel cell research. *Advances in Complex Systems, 10*(1), 93–115

Xiao, Yu., Lu, L. Y. Y., Liu, J. S., & Zhou, Z. (2014). Knowledge diffusion path analysis of data quality literature: A main path analysis. *Journal of Informetrics, 8*(3), 594–605