



# Rapid improvements with no commercial production: How do the improvements occur?



Jeffrey L. Funk<sup>a,\*</sup>, Christopher L. Magee<sup>b</sup>

<sup>a</sup> National University of Singapore, Division of Engineering and Technology Management (EA-5-34), 9 Engineering Drive 1, Singapore 117576, Singapore

<sup>b</sup> Engineering Systems Division, Massachusetts Institute of Technology, NS2-395, 77 Massachusetts Avenue, Cambridge, MA, United States

## ARTICLE INFO

### Article history:

Received 16 April 2014

Received in revised form 27 October 2014

Accepted 4 November 2014

Available online 20 November 2014

### Keywords:

Performance

Cost

Production

Learning

Technology

R&D

## ABSTRACT

This paper empirically examines 13 technologies in which significant cost and performance improvements occurred even while no commercial production occurred. Since the literature emphasizes cost reductions through increases in cumulative production, this paper explores cost and performance improvements from a new perspective. The results demonstrate that learning in these pre-commercial production cases arises through mechanisms utilized in deliberate R&D efforts. We identify three mechanisms – materials creation, process changes, and reductions in feature scale – that enable these improvements to occur and use them to extend models of learning and invention. These mechanisms can also apply during post-commercial time periods and further research is needed to quantify the relative contributions of these three mechanisms and those of production-based learning in a variety of technologies.

© 2014 Elsevier B.V. All rights reserved.

## 1. Introduction

Rapid improvements in the cost and performance of new technologies enable technological discontinuities (Christensen, 1997) and large improvements in productivity (Solow, 1957), two important issues within the fields of management and economics. But how do these improvements occur during what Dosi (1982) calls a technological trajectory? For cost, most of the literature focuses on the factory floor and links cost reductions with cumulative production. In what has been termed learning by doing (Arrow, 1962), costs fall as firms learn to produce a single design in a single factory more efficiently and thus with lower costs. Workers become better at tasks and firms introduce better work flows (Wright, 1936; Argote and Epple, 1990; Adler and Clark, 1991; Thornton and Thompson, 2001), better process control (Argote, 1999; Lapre et al., 2000), and automated manufacturing equipment (Utterback, 1994), and promote organizational learning (Benkard, 2000).

Some scholars consider cumulative production as a general proxy for effort and thus the driver of new product and process designs and thus improvements in performance and cost (Lieberman, 1984; Dutton and Thomas, 1984; Balasubramanian and Lieberman, 2010). This formulation which is sometimes called

learning by experience suggests that all of the improvements in performance and cost for a technology can be considered endogenous to a model linking cumulative production to the improvements (Dutton and Thomas, 1984; Ayres, 1992; Weiss et al., 2010; Nagy et al., 2013) where the relative contribution of factory floor activities and new product and process designs are unclear. On the other hand, a few scholars have questioned the importance of cumulative production and demand and the possibility that R&D effort or time may be a better independent variable (Koh and Magee, 2006, 2008; Nemet, 2009; Nordhaus, 2009; Thompson, 2012; Funk, 2013a,b).

This paper attempts to better understand the impact of product and process design changes vs. factory floor activities on cost and performance by detailed analysis of improvements in cost and performance in a novel empirical domain. It focuses on new technologies that have experienced rapid improvements in cost and performance before commercial production has been started and it examines the specific mechanisms that enable these improvements to occur. An analysis of these mechanisms enables us to identify more specific modes of learning and to extend models of learning (Argote and Epple, 1990) into the pre-commercialization phase that some define as invention (Arthur, 2007). Our analysis suggests that key aspects of this learning include creating new materials, improving processes, and reducing scale and that this learning occurs in laboratories.

A second contribution of the paper is for theories of invention. Building from others (Fleming, 2001; Fleming and Sorenson,

\* Corresponding author. Tel.: +65 98560674.

E-mail address: [etmfjl@nus.edu.sg](mailto:etmfjl@nus.edu.sg) (J.L. Funk).

2001; Arthur, 2007), the three mechanisms for the improvements in cost and performance suggest that product and process design concepts are improved over time in a recursive process during a transition from invention to commercialization. We view this transition as part of a continuous process of learning in which the technology becomes economically feasible for a growing set of applications both before and after commercial production begins.

This paper first surveys the literature on improvements including learning curves and invention. Second, our methods of finding and analyzing new technologies that are experiencing rapid improvements with little or no commercial production are summarized. Third, time series data on the cost and/or performance of 13 different technologies are analyzed in order to examine the relationship between the rates of improvements and the levels of commercial production. Fourth, a detailed examination of the technical mechanisms that cause these improvements is presented. Fifth, we discuss the extent to which these mechanisms might contribute to improvements after the start of commercial production and the implications of these results for theories of learning and invention, for firm strategy, and for R&D policy.

## 2. Literature review

Since the publication of Wright's (1936) analysis of fighter jet costs in 1936, empirical analyses correlating cost reductions to cumulative production have grown extensively. These analyses plot straight lines of the log of cost vs. the log of cumulative production. As analyzed and named by Arrow, this formalism of learning by doing (Arrow, 1962) has been shown to be an important explanation for improvements in cost (Argote and Epple, 1990). The early work on learning curves was mostly done on single designs in specific factories and thus analyzed the impact of the factory level changes mentioned in the first paragraph of the introduction on factory productivity. Subsequently, learning curves have been applied to technologies that are manufactured with new designs and in new factories where the output variable might be cost or performance, albeit these models are now often called experience curves (Ayres, 1992; Dutton and Thomas, 1984). For example, the costs of ships (Thornton and Thompson, 2001), solar cells (Nemet, 2006), semiconductor memory, chemicals, primary metals, and food have been analyzed using this approach (Ayres, 1992; Nagy et al., 2013), across significant design changes and often throughout all global factories.

Linking cumulative production to reductions in cost or improvements in performance can lead to confusion about how the improvements in cost and performance are occurring. Some believe that such a linkage suggests most of the improvements are occurring on the factory floor while others note that cumulative production *indirectly* leads to improvements in performance. Increases in production are linked with expected future production and lead to increased incentive to perform process-related (Sinclair et al., 2000) and general R&D (Schmookler, 1966) where the findings from the increased R&D spending result in improvements in performance or cost. This argument is also implicit in Christensen's (1997) analysis of hard disk drives, computers and other "disruptive" technologies in that the emergence of a low-end product lead to increases in R&D spending and thus rapid improvements in the product, which leads to replacement of the dominant technology by the low-end innovation. This argument suggests that except for the "invention," all of the improvements in performance can be considered embedded in a model linking cumulative production to the improvements (Dutton and Thomas, 1984; Lieberman, 1984; Balasubramanian and Lieberman, 2010).

Other analyses plot (logs of) performance or costs vs. time for a specific technology and thus do not implicitly argue that cumulative production is a driver of these improvements. This is consistent with mathematical analyses that show cumulative production could simply serve as a surrogate for time (Sahal, 1979; Nordhaus, 2009; Nagy et al., 2013). For performance, this includes the number of transistors on a chip (Moore's Law), the luminosity per Watt of lighting (Azevedo et al., 2009), processing time or speeds of computers, information storage densities and capacities (Koh and Magee, 2006), and energy or power storage densities of batteries and engines (Koh and Magee, 2008). These studies, starting with Moore, have considered different designs and factories over time and typically plot "record setters" or best performers over time. Proposed mechanisms for these improvements in performance (and cost) include changes in product design (Utterback, 1994; Adner and Levinthal, 2001) such as novel combinations of components (Basalla, 1988; Iansiti, 1995) and changes in scale, both increases in production equipment size and reductions in feature size (Gold, 1974; Lipsey et al., 2005; Winter, 2008; Funk, 2013a,b). These mechanisms might occur in response to bottlenecks in a system of materials or components (Hughes, 1983; Rosenberg, 1969; Dosi and Nelson, 2010).

Building from the possibility that cumulative production may be a surrogate and that it cannot be distinguished as a causal variable from time (Sahal, 1979; Nordhaus, 2009; Thompson, 2012), we propose a novel approach to better understand the impact of new product and process designs and thus R&D on improvements in cost and performance. Our approach focuses on technologies that have experienced rapid improvements in time periods of both no and little commercial production. When we can find such technologies, we can exclude factory-floor mechanisms dependent upon commercial production such as better process control, automated equipment, or scale of production equipment and then identify other mechanisms that have caused the improvements **apply** in these cases. Second, we can then examine the extent to which these design changes might continue to impact on improvements in cost and performance after commercial production begins.

Since this analysis focuses on pre-commercialization, it can also help us better understand the process of invention. Most research on invention has focused on developing the concepts that form the basis of new technologies and this research describes a recursive process in which combinatorial search (Basalla, 1988; Fleming, 2001; Fleming and Sorenson, 2001; Arthur, 2007) is done. Recursion occurs in the development of concepts, their translation into working prototypes, and we believe their translation into economically feasible products. For working prototypes, problems and sub-problems often at the system and component levels are recursively solved until a working prototype emerges (Arthur, 2007). This paper's analysis explores the transition from technical to economic feasibility where recursion during the development of a series of working prototypes is found to be an important part of this transition.

## 3. Methodology

We looked for new technologies that have experienced rapid improvements in cost or performance (>10% per year) during time periods in which there was no commercial production. As a point of comparison we note that integrated circuits (ICs) have experienced improvements in the number of transistors per chip of greater than 30% per year – commonly known as Moore's Law. We looked for technologies that are experiencing rapid improvements because technologies with rapid improvements are more likely to have a large impact on improvements in productivity and lead to the

emergence of technological discontinuities and creative destruction than are technologies with less rapid improvements. We also note that annual rates of improvement are very different from rates of improvement for each doubling in cumulative production, which many measure (Weiss et al., 2010). They are only equivalent when cumulative production is doubling each year. Since most technologies experience unit or sales growth rates of much less than 100% a year, annual rates of improvement are often 1/3 to 1/10 those for each doubling.

We first looked for technologies that are *currently* experiencing rapid improvements for two reasons. First, it is less likely that improvements occur with no commercial production for older than for newer technologies. Increases in university R&D over the last 50–100 years (Murray, 2002, 2004) have increased the chances that improvements will occur **for example** before the beginning of commercial production. When there was little university R&D before for example WWII, pre-commercial production improvements depended on the research activities of corporations. Since corporations have much shorter time horizons than do university and government laboratories, it is likely that there is more pre-commercial research done now than 50 years ago. Second, publication in journals is much less likely for corporate work than university work so it is easier to find data for newer than older technologies. Thus, the paper focuses on newer technologies – specifically technologies that have experienced *rapid* improvements in the last 10 years.

We looked for time series data in the leading scientific journals such as *Nature*, *Science*, *Optics*, and *Nanoletters*, annual reports by reputable scientific organizations such as the *International Solid State Circuits Conference*, general technology websites, technology-specific web sites for new technologies such as superconductivity and quantum dots, and in searches on Google. For the latter, we searched for data on technologies that are mentioned in “science-based” journals with terms such as cost, performance, trends, improvements, and specific dimensions of performance.

The search and data analysis led to the creation of a relatively large data set on technologies and their rates of improvement, which are summarized in Table 1. All of these technologies are based on a new concept (Henderson and Clark, 1990) or paradigm (Dosi, 1982) and thus Table 1 does not include new and better versions of memory ICs, computers, liquid crystal displays, or mobile phones, all of which can be defined as new product (and process) designs but probably not new technologies that have been “invented.” The technologies are placed into 7 categories, the first six of which are the transforming, storing, and transporting of energy and information, which is consistent with characterizations of engineering systems (de Weck et al., 2011). In total, there are 30 individual technologies shown in Table 1. Since a variety of performance measures are often relevant for a specific technology, data was collected on multiple dimensions some of which are represented in performance of basic functions per unit cost while others are in performance of functions per mass or per volume totaling 45 metrics in the table.

Second, within the technologies listed in Table 1, we looked for technologies that experienced improvements with no commercial production. To examine this question, we gathered data on the start of commercial production and data on the current level of sales/production. This data was searched for on the Internet. Web searches were done for specific names of technologies along with keywords such as commercial production, start of production, and sales. Commercial production is defined as the production of a technology by firms for a specific application and this does not include production by universities or the development and construction of prototypes by firms. In gathering this data, we looked for data in consulting reports published by organizations such as Dataquest,

BCC Research, Lux Research, ICE (Integrated Circuit Engineering), and DisplaySearch and we were careful to emphasize actual rather than forecasted sales. Consulting reports often focus on forecasts rather than actual sales because most of their readers are more interested in the future than in the past.

We then selected technologies for which our performance data showed rapid improvements before the start of commercial production according to the production/sales data. As a check on the start of commercial production date, we gathered recent sales data on each technology. Technologies with recent start dates for commercial production would not be expected to be currently experiencing large sales. For comparison purposes, we note that the market for semiconductors was \$40 million in 1954 and \$100 million in 1956 (Tilton, 1971) or \$346 million and \$856 million respectively in 2013 dollars. The years 1954 and 1956 were three and five years respectively after the first transistors were introduced in 1951.

Third, for the selected technologies, we looked for information on how the improvements occurred, i.e., the mechanisms. Since each technology experiencing rapid improvements during a time period of no commercial production is based on a single concept (Henderson and Clark, 1990) or paradigm (Dosi, 1982), changes in the concept or paradigm are not the sources of the improvements. Thus, other types of technical changes were the sources of these improvements. We investigated the following questions. Were these changes made to the product and/or process design; what types of product and process design changes occurred? Some argue that changes made to the physical scale of a product design impact on cost and performance (Lipsey et al., 2005; Winter, 2008; Funk, 2013a,b). But how did these changes in scale impact on cost? Were new materials and new material classes involved? Were these product changes completely independent of process design changes (Utterback, 1994; Adner and Levinthal, 2001) or were they inter-related (Linton and Walsh, 2008)? Did the bottlenecks change as the improvements proceeded (Hughes, 1983; Rosenberg, 1969; Dosi and Nelson, 2010)?

Fourth, we looked for information on the relationship between performance and cost and in particular, how improvements in some measures of performance can lead to reductions in cost. While users make tradeoffs between cost and performance at any moment in time (Green and Wind, 1973), improvements over time in a technology can change this tradeoff and thus lead to both improvements in cost and performance (Adner, 2002, 2004; Adner and Zemsky, 2005; de Figueirido and Kyle, 2006). This interaction of improvements in cost and performance can lead to reductions in cost essentially arising from improvements in performance. For example, increases in the efficiency of solar cells (US DoE, 2010), lighting (Azevedo et al., 2009), or displays (e.g., luminosity per Watt), in the densities of transistors or memory cells (e.g., Moore's Law), or in the speeds of electronic devices (Koh and Magee, 2006) can lead to lower costs in the electricity from solar cells, the cost per lumen of lighting or displays, the cost per transistor of microprocessors, and the cost per computation for computers.

To identify the relationship between cost and performance and how these improvements occurred, we analyzed many technical/scientific articles. We looked for discussions of the improvements and the other questions mentioned in the previous paragraph. The sources of the time series data were often good sources for information on the mechanisms improvements. This was particularly true with “science-based” journal papers. In all cases, however, we searched multiple papers, articles, and technical reports in order to obtain a deeper understanding of the improvements and to test any hypotheses that were generated by reading the “Science-based” Journal articles. The mechanisms were only identified by this inductive process and were not pre-selected.

**Table 1**  
Technologies with recent rapid rates of improvement.

Technology domain	Sub-technology	Dimensions of measure	Time period	Improvement rate per year (%)	
Energy transformation	Light emitting diodes (LEDs)	Luminosity per Watt, red	1965–2008	16.8	
		Lumens per Dollar, white	2000–2010	40.5	
	Organic LEDs	Luminosity per Watt, green	1987–2005	29	
		Power density	1987–2007	30	
	GaAs lasers	Cost per Watt	1987–2007	31	
		Square meters per dollar	2001–2011	11.0	
	Liquid crystal displays	External efficiency, red	1998–2009	36.0	
	Quantum dot displays	Peak Watt per price	1977–2013	13.7	
	Solar cells	Efficiency of organic	2001–2013	12.6	
		Efficiency of quantum dot	2010–2013	42.1	
Energy transmission	Superconductors: BSSCO and YBCO	Current-length per cost	2004–2010	115	
		Current × length – BSSCO	1987–2008	32.5	
		Current × length – YBCO	2002–2011	53.3	
Information transformation	Microprocessor ICs	Number of transistors/chip	1971–2011	38	
		Camera chips	Pixels per dollar	1983–2013	48.7
	Power ICs	Light sensitivity	1986–2008	18	
		Current density	1993–2012	16.1	
	MEMS: artificial eye	Number of electrodes	2002–2013	45.6	
		Drops per second	1985–2009	61	
	MEMS: inkjet printers	Mobility	1982–2006	109	
		Organic transistors	1/Purity (% metallic)	1999–2011	32.1
	Single walled carbon nanotube transistors	Density	2006–2011	357	
		Super-conducting Josephson junction-based transistors	1/Clock period	1990–2010	20.3
	Photronics	1/Bit energy	1990–2010	19.8	
		Qubit lifetimes	1999–2012	142	
	Computers	Bits per qubit lifetime	2005–2013	137	
		Data Capacity per Chip	1983–2011	39.0	
	Quantum computers	Instructions per unit time	1979–2009	35.9	
		Instructions per time-cost	1979–2009	52.2	
		Number of qubits	2002–2012	107	
	Information storage	Flash memory	Storage capacity	2001–2013	46.8
			Storage capacity	2006–2013	272
Resistive RAM		Storage capacity	2001–2009	37.8	
		Storage capacity	2002–2011	57.8	
Ferroelectric RAM		Storage capacity	2004–2012	63.1	
		Storage capacity	1991–2011	55.7	
Magneto RAM		Recording density of disks	1993–2011	32.1	
		Recording density of tape	1956–2007	32.7	
Phase change RAM	Cost per bit of disks	1956–2007	32.7		
	Magnetic storage	Recording density of disks	1991–2011	55.7	
		Recording density of tape	1993–2011	32.1	
	Cost per bit of disks	1956–2007	32.7		
Information transmission	Last mile wireline	Bits per second	1982–2010	48.7	
		Bits per second	1996–2013	79.1	
	Wireless, cellular	Bits per second	1995–2010	58.4	
		Bits per second	1996–2008	77.8	
Wireless, WLAN	Wireless, 1 m	Bits per second	1995–2010	58.4	
					Bits per second
Biological transformation	DNA	Sequencing per unit cost	2001–2013	146	
		Synthesizing per unit cost	2002–2010	84.3	
	Cellulosic ethanol	Output per cost	2001–2012	13.9	

Azevedo et al. (2009), Haitz and Tsao (2011), Lee (2005), Sheats et al. (1996), Martinson (2007), Economist (2012a, 2013), Kwak (2010), NREL (2013), Shiohara et al. (2013), Selvamianickam (2011), Wikipedia (2014), Preil (2012), Suzuki (2010), Miller (2012), Chader et al. (2009), Stasiak et al. (2009), Hasegawa and Takeya (2009), Franklin (2013), Fujimaki (2013), Devoret and Schoeldopf (2013), Evans et al. (2011), Koomey et al. (2011), D-Wave (2013), Francis (2011), Yoon (2010), Brown (2011), ISSCC (2013), NHGRI (2013), SingularityHub.com (2013), Service (2013).

#### 4. Results

The 13 technologies listed in Table 2 experienced annual improvements of greater than 10% during a time period (see column 3) in which there was no commercial production; the start of commercial production is shown in the second column of Table 2. In fact, there was commercial production during only 17% of the years (27/156) for which rapid improvements occurred. Table 2 (see column 4) also shows, recent sales figures for each technology and all these figures are smaller than sales of semiconductors in 1956 (\$856 million in 2013 dollars). Except for printed electronics, which probably includes technologies other than organic transistors (see below), all of them had market sizes smaller than did semiconductors in 1954 (\$346 million in 2013 dollars). Thus, high levels of commercial production still do not exist for any of these technologies. The small current market sizes are consistent with

the dates for the start of commercial production and together are strong evidence that these technologies were improved with zero commercial production and still are being improved with small levels of commercial production. While two or three cases might be arguable given uncertainty in the start of commercial production, we were conservative (chose earlier if multiple possibilities exist) about the dates and yet 8 cases are shown in Table 2 where production did not start until 9 or more years after the onset of rapid improvement. In summary, these results clearly indicate that learning by production workers on the factory floor is not a relevant mechanism for these rapid improvements. The following sections address the mechanisms by which these improvements occur and they present more detailed data about each of the cases in Table 2.

Discussion of the technologies listed in Table 2 is covered in five sub-sections (organic materials, quantum dots, new forms of non-volatile memory integrated circuits, carbon nanotubes, and

**Table 2**  
Starts of commercial production and recent sales data.

Technology	Start of commercial production	Time period for rapid improvements	Recent sales data (\$ millions)	Sources of sales data
Organic LEDs	2001	1987–2005	300 in year 2012	Display Search (2013)
Organic transistors	2007	1994–2007	530 (printed electronics) in year 2010	Markets and Markets (2011)
Organic solar cells	2010	2001–2013	4.6 in year 2012	IDTE (2012)
Quantum dot solar cells	2013	2010–2013	Zero until 2013	Investor (2013)
Quantum dot displays	2013	1994–2009	Zero until 2013	Research & Markets (2013)
Resistive RAM	2013	2006–2013	200 (2012)	Yole (2013a,b)
Ferroelectric RAM	2005	2001–2009		
Magneto-resistant RAM	2004	2002–2011		
Phase change RAM	2006	2004–2012		
Single walled carbon nanotubes for transistors	2011	1999–2011	<10 (2011)	BCC (2012)
High temperature superconductor wire (YBaCuO and BiSrCuO)	2006	1987–2008	30 (2011) 30 (2012)	Connectus (2012)
Superconducting Josephson junction-based transistors	2011	1990–2010	First sale of these technologies in 2011	Jones (2013)
Quantum computers	2011	2002–2012		

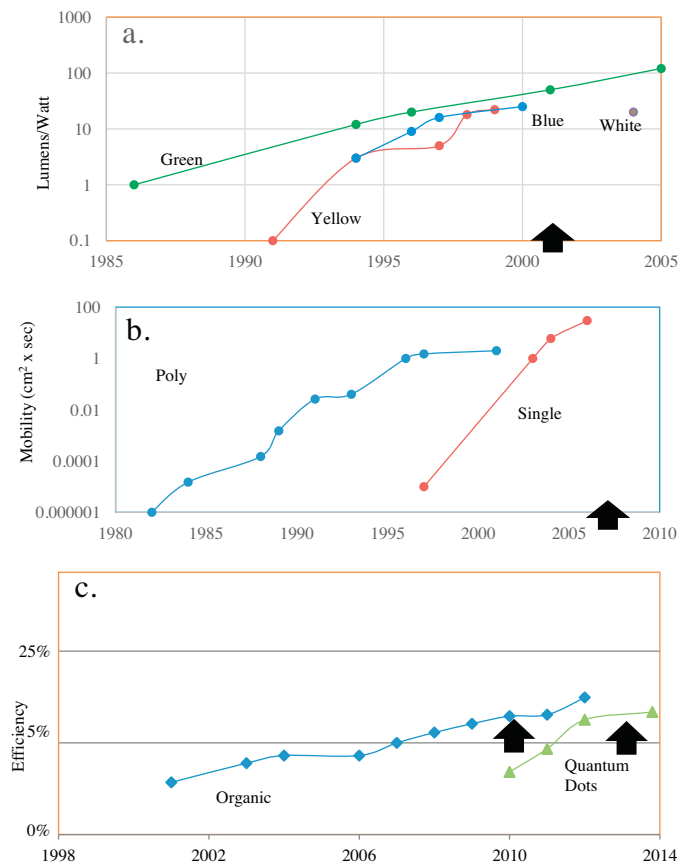
superconducting materials). Several of these sub-sections discuss multiple entries from Table 2 since the mechanisms substantially overlap. Each sub-section describes the relevant dimensions of performance, the impact of these improvements in performance on cost and gives logarithmic plots of the performance measures against time in figures to display the time series data. The starts of commercial production are shown on each graph as a large black arrow and recent sales data is discussed. In addition, each section discusses the technical changes that drove improvements in performance and cost for specific technologies.

4.1. Organic materials

Organic materials are being used in many electronic applications because they are more mechanically flexible than are semiconductor materials and because it is potentially cheaper to fabricate electronic devices with them than with semiconductor materials that require high temperature processes. These electronic devices include lighting, displays, solar cells, and transistors. Luminosity per Watt, shown against time in Fig. 1a is an important dimension of performance for organic light emitting diodes (OLEDs) in lighting and displays and it also impacts on the cost for users. Not only does better luminosity per Watt lead to lower cost per lumen (Azevedo et al., 2009), it also enables smaller devices for a given output and thus lowers material, equipment, and transport costs. This makes improvements in luminosity per Watt an important driver and rough surrogate for reductions in the cost of lighting or displays with OLEDs.

Fig. 1a shows that about 10–100 times improvements were made in luminosity per Watt before documented commercial production was started in 2001 (see black arrow) for mobile phone displays (OLED, 2013). Even after commercial production was started, sales grew slowly with some ups and downs. For example, the leading market research firm, Display Search, claims in two different reports that sales were \$615 million in 2008 (Display Search, 2009) and \$300 million in 2012 (Display Search, 2013). Thus increases in commercial production do not appear to be a strong explanation for improvements in cost and performance after 2001 and are even more questionable before this date where a large fraction of the documented improvements occurred.

Not surprisingly, the literature on OLEDs provides an explanation for the improvements that different than those based upon improvements on the factory floor. Scientists and engineers created new types of organic materials that better exploited the phenomena of electroluminescence. In the literature, one can find graphs similar to Fig. 1a that identify materials changes at many of the data points. For example, according to such plots found in Sheats



**Fig. 1.** (a) Luminosity per watt for organic light emitting diodes (black arrow = start of commercial production). (b) Mobility of single crystal and polycrystalline organic transistors (black arrow = start of commercial production). (c) Efficiency of organic and quantum dot solar cells (black arrows = starts of commercial production)

et al. (1996) and Lee (2005), improvements in the lumens per Watt of green, blue, and yellow OLEDs came from new forms of InGaN, polyfluorenes, phosphorescent materials and molecular solids such as Tris (8-hydroxyquinolinato) aluminum.

The second entry in Table 2 is another technology that is based on organic materials, organic transistors. Mobility is a key dimension of performance for transistors since mobility directly impacts on speed and speed is an important dimension of performance for computers and other electronic products. Mobility also impacts indirectly on cost since computers are often evaluated in terms of

their cost per instruction (Nordhaus, 2007) or cost per processing output (Koh and Magee, 2006) and mobility impacts on the speed at which these computers can perform. Thus, the improvements in mobility shown in Fig. 1b are often a good surrogate for improvements in the cost of processing data with organic transistors.

Organic transistors were first commercially used in 2007 to control the pixel values in flexible electronic paper, which are used in e-books such as the Amazon Kindle (ChemTech, 2008; Mas-Torrent and Rovira, 2008). From Fig. 1b, improvements of more than 1000 times occurred before the start of commercial production in 2007. More recent sales data for organic transistors could only be found under printed electronics, a category that includes any type of electronic circuit that is printed including transistors and displays. Thus, the \$530 million figure in Table 2 is probably a substantial overestimate for the size of the organic transistor market in 2010 or for more recent years and increases in production are not a viable explanation for improvements in cost and performance before 2007 and probably not after that date as well.

Like OLEDs, the literature on organic transistors and even the figures displaying the improvements in mobility in this literature focus on new materials as the sources of improvements. For example, Dong et al. (2010) include the names of organic materials such as polythiophenes, thiophene oligomers, polymers, hthalocyanines, heteroacenes, tetrathiafulvalenes, perylene diimides naphthalene diimides, acenes, and C60 alongside the improvement data. Processes are also important. For example, one scientific paper (Horowitz, 2011) says: “the search for high mobility materials is still very active. However, the mobility is not only dictated by the nature of the organic semiconductor; it also strongly depends on other parameters such as the crystal structure and the quality of the various interfaces that intervene in the device: interfaces between the insulator and the semiconductor and between the electrodes and the semiconductor.” Thus, process optimization and materials creation are inextricably linked for this technology.

The third entry in Table 2 is organic solar cells. The efficiency of solar cells is defined as the percentage of incoming solar energy that is converted into electrical energy and it has a large impact on the cost of electricity from solar cells. While it contributes equally to reductions in cost per peak Watt of physical cells as does reductions in cost per area, it contributes more to reductions in the cost per peak Watt than does reductions in the total cost per area because installation costs depend on the area of the solar cells and better efficiencies enable smaller areas for a given output (US DoE, 2010). Like OLEDs, this makes improvements in efficiency, shown in Fig. 1c, an important driver and useful surrogate for reductions in the cost of electricity from organic and other solar cells.

Organic solar cells were first commercially produced in 2010 by Konarka, a firm that went bankrupt in mid-2012. One market research firms claims that organic solar cells had about \$4.6 million in sales in 2011 (IDTE, 2012), the latest year for which we were able to find sales data. Fig. 1c shows that before commercial production of organic solar cells began in 2010, significant improvements had been achieved in their efficiency, growing from 3% in 2001 to 8% in 2010 and 11.1% by the end of 2012. (For comparison purposes, single crystal silicon solar cells have a best laboratory efficiency of 25%.) Like OLEDs, the literature on organic solar cells focuses on the creation of new types of organic materials and creating these new organic materials often requires new processes. For example, two scientists (Hou and Guo, 2013) summarize these improvements in the following way: “the development of active layer materials is still the key to boost the efficiency. In order to get better photovoltaic properties, many properties, like band gap, molecular energy level, mobility, solubility, etc., should be considered, and how to balance these parameters is the most important part to molecular design.” Continuing with this chapter, Hou and Guo describe the creation of new materials such as

blended films of conjugated polymer (electron donor) and small molecular acceptors. Examples of active materials that enabled improvements in efficiency include polythiophene, polymers with 2,1,3-bezothiadiazole pyrrolo derivatives, and bezo-dithiophene-based polymers.

#### 4.2. Quantum dots

The fourth entry in Table 2 is also a solar cell and thus the efficiencies are also shown against time in Fig. 1c. However, quantum dot solar cells are based on a different physical principle and a different type of material than are organic solar cells. Quantum dots consist of small crystals (usually semiconductor crystals) whose size determines their electronic and optical properties and thus the wavelengths of light that will be absorbed in a quantum dot solar cell. By varying the size of the dots, it is theoretically possible to create solar cells on a single layer of material that have efficiencies greater than 80% or more than three times the maximum theoretical efficiency of conventional solar cells such as those made with organic or semiconductor materials (including silicon). Like other forms of solar cells, improvements in efficiency can be considered a surrogate for reductions in costs (US DoE, 2010).

Commercial production of quantum dot solar cells was reported to have started in late 2013 (Investor, 2013). The market for all types of quantum dots including displays and medical applications were only \$150 million in 2011 (Research & Markets, 2013). Even with significant production, learning from commercial production is not a good explanation for efficiency improvements such as those experienced by quantum dot solar cells shown in Fig. 1c. This figure shows that rapid improvements in the efficiency of quantum dots have been recently achieved, rising from 3% in 2010 to 8.6% in mid-2013. Like organic solar cells, the literature on quantum dot technologies focuses on the creation of new materials and processes. Semiconductor and other crystals are grown with new types of materials or by adding new impurities and/or dopants to these base materials. Types of materials that are mentioned as contributing to the improvements in efficiencies include conventional semiconductors such as silicon or indium arsenide, more complex compositions (i.e., alloys), and selenide or sulfides of metals (e.g., lead sulfide, lead selenide, cadmium selenide). New structures such as quantum dots that are grown within dots are also mentioned where new processes are typically required for these new structures to be effective (Patel, 2011; Chandler, 2013).

Quantum dot displays are the fifth entry in Table 2. Like quantum dot solar cells, the size of the dot determines the relevant wavelength of light although in this case, the important wavelength is of the emitted rather than the absorbed light. Similar to solar cells, this enables a single layer of material to theoretically emit many different colors and thus result in a lower cost display than current display technologies. Also like OLEDs, the efficiency with which electricity is converted to light has an important impact on both the performance and cost of the quantum dot display. For quantum dots, this efficiency is measured in terms of the percentage of available electrons that are converted to photons while for OLEDs, the efficiency is measured in terms of luminosity per Watt. In any case, improvements in the efficiency can be considered a rough surrogate for reductions in the cost of quantum dot displays (Azevedo et al., 2009) and are displayed against time in Fig. 2.

Commercial production of quantum dots for television displays reportedly began in 2013 by Sony. These quantum dots are used in combination with liquid crystal displays (LCDs) to increase the range of colors that can be displayed on a television (Physorg, 2013). As noted above, even the sales for all types of quantum dots in 2011 was \$150 million and these were mostly for biological/medical applications (Anscombe, 2005; Sanderson, 2009). Similarly to the technologies previously discussed, the results indicate that the

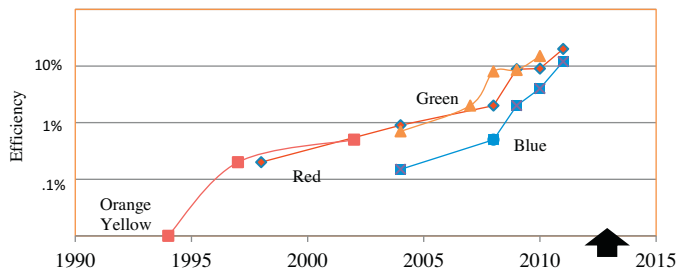


Fig. 2. Efficiency of quantum dot displays for different colors (black arrow = start of commercial production)

significant improvements in the efficiency of quantum dot displays (Fig. 2 shows 1000 times improvement since 1994) are not due to mechanisms inherent to production. Instead, the literature on quantum dot displays focuses on the creation of new materials and typically the same types of new materials, structures and processes that have enabled improvements in the efficiency of quantum dot solar cells (see above). Additional mechanisms stress the intricate linking of new materials and new forms of processes such as layer-by-layer assembly methods (Bae et al., 2010) and the use of ZnO nanoparticles and organic layers in combination with conventional semiconductor-based materials (Kwak et al., 2012).

#### 4.3. New forms of non-volatile memory integrated circuits

The next four entries in Table 2 are for new forms of memory ICs and their improvements are driven by a different set of technical changes than are the previously discussed technologies. These four entries are for different types of RAM (random access memory) ICs that can be defined as non-volatile memory (NVM). NVM refers to memory that retains its value when the power is switched off. The most familiar type of NVM is called flash memory and it is familiar to many of us because it is used in mobile phones. It enables our phones to remember our phone numbers, music, and videos even when the power has been switched off.

Like all forms of chip-based memory, a key dimension of performance for NVM is storage capacity. The performance of semiconductor memory is usually measured in terms of the number of bits per chip and these increases are typically achieved by increasing the number of bits per area. Increases in the number of bits per area usually lead to lower costs per bit because the higher densities lead to lower material and equipment costs per bit, along with enabling faster speeds. This is also the case with Moore's Law, which is typically discussed for microprocessors, but the analysis holds as well for other types of integrated circuits such as RAM and other types of memory ICs. Increases in the numbers of transistors or memory cells per IC chip lead to lower cost and higher speeds (Moore, 2006) and similar improvements will occur with NVM. The four entries on NVM refer to four new types of NVM that are being developed as potential replacements for flash memory. Although these four types of NVM are based on different physical principles, materials, and structures and different ones from flash memory (which uses silicon), patterns on them are fabricated using some of the same processes and equipment that are used to fabricate patterns on ICs including flash memory ICs.

As shown in Fig. 3, substantial improvements have been achieved in these four types of NVM some of which were achieved before commercial production for each alternative (shown by the 4 black arrows) was started. This is particularly true with resistive RAM, for which an improvement of 100 times occurred before commercial production started in 2013. Since the total commercial sales for all four forms of NVM were only \$200 million in 2012 (Yole, 2013a), and since most improvements occurred before the start of

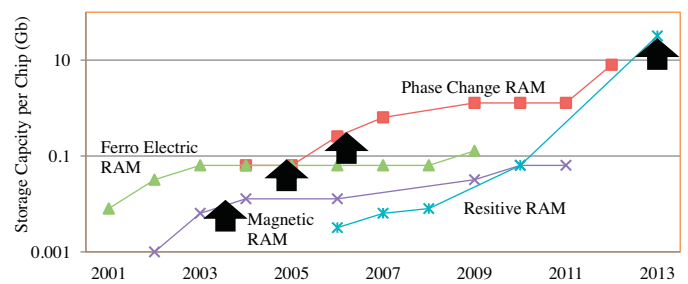


Fig. 3. Number of memory bits (Gb) per RAM (random access memory) chip (black arrows = start of commercial production).

commercial production, it appears again that mechanisms associated with commercial production-based learning have not been a major mechanism for the improvements even after commercial production was started. Instead, most of the improvements were achieved in laboratories where prototypes were fabricated (ISSCC, 2013; Yole, 2013b).

The main mechanism for achieving the improvements in storage capacity has been from reducing the feature sizes associated with the memory cells. Differences in storage capacity between different forms of NVM are largely a function of differences in these feature sizes (Yole, 2013b). This is the same mechanism by which the number of transistors per chip is increased in conventional ICs such as flash memory, DRAM, and microprocessor ICs. Firms such as Intel reduce the size of the features that define a transistor or memory cell and thus are able to increase the number of transistors or memory cells per chip (Kuhn, 2009). In the case of the new forms of NVM, the reductions in feature size are made largely by the appropriate modification of equipment and processes that are borrowed from the manufacture of flash memory or other integrated circuit-related industries.

#### 4.4. Carbon nanotubes for transistors

Carbon nanotubes (CNTs) are another technology that is being developed for ICs largely because of their very high electrical and thermal conductivities. CNTs are composed solely of carbon atoms, just as graphite, diamond, and graphene are. While graphene is a one-atom thick layer of carbon atoms, one can think of CNTs as graphene that is rolled into cylindrical tubes with either open or closed ends. These CNTs can be produced with single, "few," or "multi" walls of which the single wall ones have the highest performance and cost. The market for single walled CNT has reached \$10 million and the market for all three types of carbon nanotubes had reached \$180 million by 2011 (BCC, 2012).

For transistor applications, single-walled CNTs are needed that have both high purity and density (Franklin, 2013). Fig. 4 shows that both of these dimensions of performance have been improved at rapid rates and most of these improvements were achieved before commercial production of single-walled CNTs by firms was begun

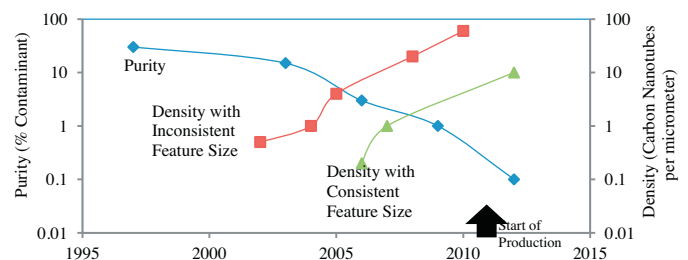
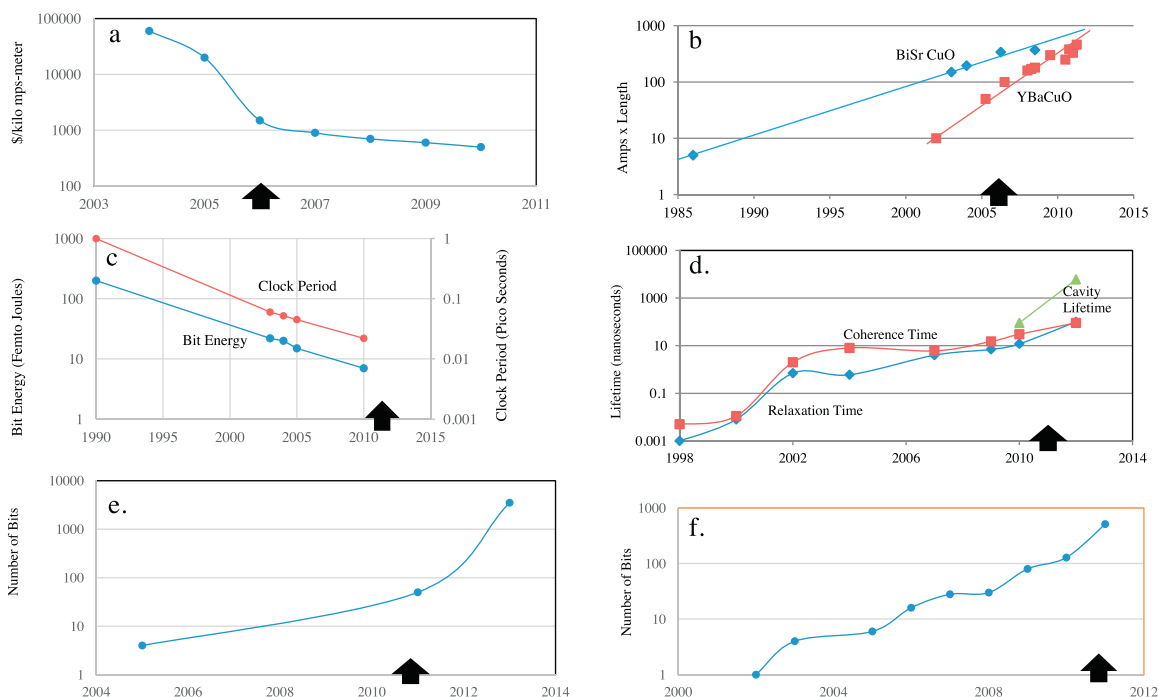


Fig. 4. Purity (left axis) and density (right axis) of carbon nanotubes for transistors. Density is for consistent and inconsistent feature size.



**Fig. 5.** (a) Cost per kiloamps-meter for superconducting cable (black arrow = start of commercial production). (b) Current (Amps)  $\times$  length (km) for two types of superconducting cables (black arrow = start of commercial production). (c) Bit energy (left axis) and clock period (right axis) for superconducting Josephson junctions (black arrow = start of commercial production). (d) Qubit lifetime for several definitions of “Lifetime” (black arrow = start of commercial production). (e) The number of bits per qubit lifetime (black arrow = start of commercial production). (f) Number of qubits in quantum computers (mostly prototypes) (black arrow = start of commercial production).

in 2011 (see black arrow). Improvements in the purity of the CNTs have been achieved by improvements in processes. As described by Liu and Hersam (2010), this includes “post-synthetic efforts to purify and sort carbon nanotubes by their physical and electronic structure” and the “selective growth of carbon nanotubes with predetermined properties.” More recently (Franklin, 2013) writes, “Jin and colleagues have managed to achieve the selective removal of metallic CNTs from an array of such nanotubes on a chip without damaging the semiconducting nanotubes.”

Increases in the density of CNTs have come from improvements in processes and new materials. This includes the growth of CNTs on new types of substrates such as quartz and by coating the CNTs with certain molecules to tune their attraction to different surfaces (Franklin, 2013).

#### 4.5. Superconductors

The final three technologies in Table 2 are superconductor-related. As their name suggests, superconductors conduct electricity with zero resistance and this enables them to be very effective conductors of electricity and to enable the creation of strong magnetic fields. The performance of superconducting materials is typically measured in terms of the highest temperature at which superconducting occurs (the critical temperature), the amount of current and magnetic field they can support before superconducting disappears, and these dimensions in combination with length due to the difficulties of fabricating long superconducting wires. Higher temperatures, currents, and magnetic fields are also related to cost since higher temperature superconductors reduce cooling costs and higher currents and magnetic fields reduce the amount of necessary materials and thus their costs (CCAS, 2014).

More than 33 superconducting materials have been created (CCAS, 2014) of which the highest recorded critical temperature is 153 K. Five of these materials are capable of superconducting at temperatures higher than 77 K and are often called

“high-temperature” superconductors. These higher temperatures enable the replacement of liquid hydrogen with liquid nitrogen as a method of cooling thus reducing the cost of cooling. Furthermore, many of these newly created materials can handle higher currents or magnetic fields at a specific temperature than can previously created materials.

Although the overall production of superconductors has grown steadily over the last ten years reaching \$4.5 billion in 2007, few of these sales are for superconducting wires or rapid single flux Josephson junctions of which the latter are used in quantum computers. Instead, the largest application for superconductors is magnetic resonance imaging in which so-called “low temperature” superconductors are fabricated into magnets. Electric power applications such as cables, transformers, motors, and generators that use high temperature superconductor wires only represent a few percent of the superconductor market (Economist, 2012b). These wires are fabricated from high temperature superconductors such as the ones (BiSrCaCuO and YBaCuO) that have experienced rapid improvements in the cost per kilo-amp-meter (Fig. 5a) or in the current times length (Fig. 5b). The market for high temperature superconductors was only \$30 million in 2011 (Connectus, 2012) with commercial production beginning in 2006 (Selvamanickam, 2011). Thus, mechanisms relating to production experience are probably not important sources of these significant improvements similar to the previously discussed technologies in this paper.

According to the engineering literature, improvements in the current carrying capability and cost of BiSrCaCuO and YBaCuO (Fig. 5a and b) were achieved largely through improvements in processes but also through modifications to the materials that are used to package these materials into wires. For increasing current times length, achieving an in-plane grain alignment of the material’s crystals was important and this was achieved with a new process (ion beam assisted deposition) and a new substrate, a rolling-assisted biaxially textured one (Shiohara et al., 2013). For reducing the cost of YBaCuO, one challenge was to reduce the content of silver due to



its high price while retaining the wire's strength. This was achieved by combining YBCO with nickel and other dopants (Paranthaman and Izumi, 2004).

Rapid single flux quantum (RSFQ) Josephson junctions are another application for superconductors in which the rapid improvements shown in Fig. 5c–e have been achieved prior to commercial production. Named for their discoverer, Brian David Josephson, Josephson junctions consist of a thin non-superconducting material that is sandwiched between two superconducting materials and for which quantum tunneling can occur across the non-superconducting material. These junctions can be used to construct various electronic devices such as single-electron transistors, qubits, superinductors, superconducting quantum interference devices, superconducting tunnel junction detectors, and rapid single flux quantum (RSFQ); RSFQ is the device of interest in this section. RSFQs that are constructed from these junctions are orders of magnitude faster and use orders of magnitude less power than do conventional ICs. Since the cost of computing is often measured in cost per instruction (Koh and Magee, 2006) and cost per energy (Koomey et al., 2011), the improvements in the speed of RSFQ also impact on the cost of computing, as does improvements in power consumption. Thus, improvements in speed and power consumption can be considered surrogates for reductions in cost.

Improvements in speed, i.e., clock period, and power consumption, i.e., bit energy (see Fig. 5c), were being achieved at a rapid rate before commercial production began in 2011 (see black arrow) for quantum computers (Merali, 2011). Like the NVM and conventional ICs that are discussed above, these improvements were achieved by reducing the feature size of the RSFQ Josephson junctions. Reducing the size of the RSFQ Josephson junctions reduced the distance to be traveled by electrons and these reductions in feature size enabled both increases in speed and reductions in power consumption (Fujimaki, 2013).

One application for these RSFQ Josephson junctions is quantum computers. Quantum computers differ from conventional computers in that bits can be in “superposition,” representing 0 and 1 at the same time according to a probability distribution. The bits in a quantum computer are called qubits and by coupling multiple qubits, the performance of a quantum computer rises at a much faster rate than do increases in the number of qubits. While conventional computers operate on a base two system, i.e., 0 or 1, and thus performance rises linearly with increases in the number of bits, the performance of quantum computers rises non-linearly as the number of qubits are increased (Jones, 2013). The problem for quantum computers is that “keeping qubits in superposition long enough to do anything useful with them has proven very hard” (Hardesty, 2011).

Nevertheless, this problem is gradually being solved as improvements in qubit lifetimes and in the number of bits per lifetime have been achieved as shown in Fig. 5d and e. The number of bits per lifetime (Fig. 5d) is equivalent to the number of measurements, each with one bit of precision that would be possible before an error occurs. These improvements have been achieved by creating new types of qubit structures and new processes for making these structures (Devoret and Schoelkopf, 2013). The new structures include different sizes and orientations of tunnel junctions, superinductors, and resonators and new processes include exposure to microwave radiation (Hardesty, 2011). These approaches are given unusual names such as Quantronium, Fluxonium, Transmon, and improved Transmon (Devoret and Schoelkopf, 2013).

The improvements in qubit lifetimes and in the number of bits per lifetime have enabled the number of qubits in a quantum computer to be recently increased more than 100 times as shown in Fig. 5f. The figure demonstrates that much of these improvements were achieved without the commercial production

of quantum computers. The first prototype was constructed in 2002 and commercial production of both quantum computers and RSFQ Josephson junctions started in 2011 with the first sale of a quantum computer (D-Wave, 2013; Jones, 2013).

## 5. Interpretation of results

The 13 technologies listed in Table 2 were shown in section 4 to have achieved rapid improvements in performance and/or cost during periods of zero commercial production and all of them still have low levels of commercial production. Some of the cost reductions are known because cost data was collected while other cost reductions are inferred based on the fact that improvements in some dimensions of performance are equivalent to reductions in cost for a given performance.

The findings detailed in Section 4 demonstrate that the rapid reductions in cost and increases in performance are – as expected since the cases were chosen to avoid production learning – not due to mechanisms associated with factory floor activities and production experience. One predominant mechanism found in these cases was creating materials that better exploit a physical phenomenon where these new materials often required new processes. New materials were important for OLEDs, organic transistors and solar cells, quantum dot solar cells and displays, and superconductors. Scientists and engineers created organic materials that better exploited the phenomenon of electroluminescence for OLEDs, the photovoltaic phenomenon for solar cells, and the semiconducting phenomenon for transistors. They created semiconductor and other materials that better exploited the phenomenon of quantum dots and other materials that better exploited superconductivity. Sometimes, the focus was on a single layer of active material while other times it was for a combination of different materials where each layer in the combination may have been tweaked with impurities and dopants. The multiple ways in which new materials are created for a single technology is consistent with the emergence of bottlenecks in a system of materials or components (Hughes, 1983; Rosenberg, 1969; Dosi and Nelson, 2010).

A second mechanism for improving performance and cost in the thirteen cases was improvements in processes, which as noted in the preceding paragraph, is a subset of the first method since creating new materials often required new processes for the creation or improvement of the material. Nevertheless, improvements in the performance of a single material often involved new processes where these new processes involved slight changes to the material composition, including the addition of impurities or dopants. Our finding of the importance of process research is consistent with other research (Sinclair et al., 2000) that identified new processes arising from research as a major source of cost reductions. For carbon nanotubes, the efforts were aimed at improving their purity and density. This includes post-synthetic efforts to purify and sort carbon nanotubes by their physical and electronic structure, the “selective growth of carbon nanotubes with predetermined properties (Liu and Hersam, 2010), and the growth of CNTs on new types of substrates such as quartz ones and by coating the CNTs with certain molecules to tune their attraction to different surfaces (Franklin, 2013).

A third mechanism (in addition to materials creation and new processes) for improving performance and cost is reducing the scale of features, a mechanism that has perhaps received too little attention in the economics and innovation literature; exceptions are Winter (2008) and Funk (2013a,b). The scale of features was reduced to increase the storage density in new forms of non-volatile memory (RAM) and to increase the speeds of superconducting rapid single flux quantum (RSFQ) Josephson junctions. This involved reducing the size of the memory cells

and RSFQ Josephson junctions, i.e., changes to the product design, and introducing processes that enabled the better control that is necessary to achieve these smaller feature sizes in the product design. Improvements in the performance of RSFQ Josephson junctions in the form of faster speeds and lower power consumption also contributed to improvements in the performance and cost of quantum computers.

## 6. Discussion

Understanding improvements in productivity may be the most important task of economics, since improvements in productivity enable improvements in the standard of living. Although Solow's (1957) Nobel Prize winning research found that most of this growth comes from innovation and subsequent research has illuminated the importance of the computing sector (Jorgenson et al., 2008), more detailed mechanisms have rarely been identified. One broad possibility is that costs and performance are improved by what happens in factory floor production. Since costs do decrease as cumulative production increases (Wright, 1936; Arrow, 1962; Argote and Epple, 1990; Ayres, 1992; Nagy et al., 2013), this has sometimes been taken to mean that this is the primary mechanism for cost reductions. The early work on learning curves focused on single designs that were made in single factories, but such work does not show that production is the key to reductions in cost over multiple factories and designs. By modeling these improvements as a function of cumulative production, it appears to some that cumulative production is necessary for improvements and that factory floor activities are the key to the improvements.

This paper has utilized a different perspective. By focusing on technologies that have experienced rapid improvements during periods of no commercial production, its results show that cumulative production is not needed for significant improvement. These results add significantly to our overall knowledge of the mechanisms for improving performance and reducing cost and in fact raise the potential for a modification of thinking for those who perceive production-based learning as always the primary mode of cost reduction. It showed that most of the improvements in the early stages for our thirteen cases came from changes in product and process designs that are being developed in laboratories. It is only after these technologies are commercialized and volumes increased that learning on the factory floor becomes possible and in most cases important. The production-based learning is particularly important in the implementation of new product and process designs as they are periodically introduced, for example in the form of higher capacity memory chips (Mathews and Cho, 1999) and other ICs (Hatch and Mowery, 1998).

One well known method of reducing costs was not found in our analysis of the thirteen cases, increasing the physical scale of production equipment (Gold, 1974; Lipsey et al., 2005; Winter, 2008; Funk, 2013a,b); this is to be expected since it does not make economic sense to increase the physical scale of the production equipment without large amounts of commercial production. Increases in the scale of production equipment enable economies of scale and some technologies (e.g., chemicals) benefit more from economies of scale than do other technologies (Pavitt, 1984; Chandler, 1994). Geometrical analysis of equipment, particularly with respect to surface area and volume is particularly useful for analyzing the past and future impact of increases in scale on the cost of production (Lipsey et al., 2005; Winter, 2008; Funk, 2013a,b). For example, analyses of chemical products have confirmed the cost advantages of large scale equipment: the capital costs of chemical plants are a function of plant size to the  $n$ th power, where  $n$  is typically between 0.6 and 0.7 (Axelrod et al., 1968; Mannan, 2005). This mechanism for improving performance and cost definitely

requires increases in production volumes but yet still depends on laboratory activities, for example, those of the equipment suppliers as significant geometric increases in production scale are usually not achievable without increased understanding of the underlying phenomena. We expect that the technologies covered in this paper will experience improvements from increases in the scale of production equipment as the levels of commercial production for these technologies are increased and as equipment suppliers and/or manufacturers develop this equipment.

An important question is the extent to which sources of the cost reductions change as cumulative production rises over time. We would expect that as commercial production increases, the importance of larger scale equipment becomes greater. But how much greater? Further research needs to address this question. This paper's results suggest that the three mechanisms identified in this paper may well continue to have a large impact on improvements in performance and cost even after commercial production begins. Indeed, the mechanisms identified in this work are applicable beyond the pre-commercial stage as the analysis of non-volatile memory in Section 4.3 indicates due to the similarities between the improvements in flash memory, other ICs, and the new forms of non-volatile memory. For the six technologies for which we have data, we observe no great changes (either accelerations or decelerations) in improvement rates (see Figs. 1–5) after commercial production begins, which is at least suggestive that combinations of materials creation, process creation and reductions in scale continue to impact on cost and performance in these domains as the levels of commercial production increase. Furthermore, increases in production volumes will also lead to increases in absolute levels of R&D (Schmookler, 1966; Klepper, 1996) that will support the three mechanisms identified in this paper. Thus, it is unclear whether the three mechanisms (or other changes to product and process design) or increases in the scale of production equipment will be important mechanisms for how the improvements occur beyond the early stages of commercialization. Further research on this issue is needed particularly for other technologies.

In summary, we see a continual process of learning that begins in the laboratory during the invention phase and continues into the commercialization and post-commercialization phases with continuing contributions from R&D work. Improvements in cost (and performance) are achieved in a continual manner such that the technology becomes economically feasible over time for a growing number of applications. The ongoing nature of the improvements suggests that it may be possible to model the invention and commercialization processes for many technologies as one continuous process, as opposed to distinct phases. It also suggests that time or research effort might be a superior independent variable to cumulative production since it more effectively explains pre-production activities.

Recursion and recombinant search (Basalla, 1988; Fleming, 2001; Fleming and Sorenson, 2001; Arthur, 2007) are an important part of this continuous process of learning and they include the revision of concepts during invention and beyond. Moreover, we find that achieving economic feasibility begins during the stage of invention, it requires many revisions of concepts, and this recursive process may continue over many years before commercial production begins and probably even after it does. This paper has also identified three types of recursion – choices of materials, modifications to processes and reductions in physical scale – adding more detail to theories of invention.

A final concern is whether the 13 technologies addressed in this paper are representative of all technologies that experience rapid improvements during periods of no commercial production. This issue is not easily dismissed. The paper focused on recent technologies and older ones may reveal different results. Even with recent ones, there are many types of technologies and despite our systemic

attempts to find all recent cases of rapidly cases, this paper may well have missed some rapidly improving technologies that might depend more on commercial production for improvements than do the ones covered in this paper. For example, chemicals and materials that are of fixed chemical formulas might depend more on increases in production volume than do other technologies because the fixed formulas reduce the extent of product design changes that can be productively performed (Stobaugh, 1988). It is also possible that the data for some such cases is not in the literature because they are improved mainly through unpublished corporate development activity. Overall, our search was wide ranging but we cannot claim that our 13 cases – all of which are consistent with each other relative to rapid improvement without commercial production – represent all possible emerging technologies.

This paper has important implications for firm strategies, R&D and governmental research policies. The results indicate that R&D is much more important than sometimes thought for improving cost and performance and that these improvements can sometimes be achieved without commercial production. For firms, quickly ramping up the production of early (and often soon obsolete) designs in order to move down a production learning curve may not be an effective strategy in many new technologies. Instead, developing R&D capabilities, fostering relationships with emerging startups while monitoring and working with universities is likely to be more important in a variety of cases. Furthermore, quantitatively monitoring improvements before the start of commercial production as research goals are pursued could provide an important signal about whether the technology has a large potential for improvements. The results of such monitoring have implications for the types of technologies that should receive research funding from firms and governments.

Further research should look at how improvements occur after commercial production starts and grows. Do improvements occur in ways different from the ways that were identified in this paper? How does this change over time? More specifically, do the contributions from increases in equipment scale and factory floor learning increase and to what extent?

## 7. Conclusions

This paper has shown that some technologies experience rapid improvements in cost and performance with no commercial production and thus mechanisms other than factory floor activities must be the main mechanism for improvements. This suggests that increases in production volume are not as important for achieving cost reductions in newly emerging technologies as the existing literature suggests for later stages of technology development. Creating materials that better exploit physical phenomenon, improving processes, and reducing the scale of features were three mechanisms that were used by scientists and engineers to improve the cost and performance in laboratories before (and even after) the start of commercial production. Although the relative contributions of the various mechanisms at later stages are not known, the three mechanisms identified here can also operate at later stages of development as they result from R&D activities which continue to increase as production and revenue increase. These findings have important implications for theories of learning and invention, for R&D policy and for firm strategies.

## References

- Adler, P., Clark, K., 1991. Behind the learning curve: a sketch of the learning process. *Management Science* 37 (3), 267–281.
- Adner, R., 2002. When are technologies disruptive? A demand-based view of the emergence of competition. *Strategic Management Journal* 23 (8), 667–688.
- Adner, R., 2004. A demand-based perspective on technology lifecycles. *Advances in Strategic Management* 21, 25–43.
- Adner, R., Levinthal, D., 2001. Demand heterogeneity and technology evolution: implications for product and process innovation. *Management Science* 47 (5), 611–628.
- Adner, R., Zemsky, P., 2005. Disruptive technologies and the emergence of competition. *The Rand Journal of Economics* 36 (2), 229–254.
- Anscombe, N., 2005, July. Photonics Spectra. Quantum Dots: Small Structures Poised to Break Big. <http://photonics.com/Article.aspx?AID=22350>
- Argote, L., 1999. *Organizational Learning: Creating, Retaining and Transferring Knowledge*. Springer, NY.
- Argote, L., Epple, D., 1990. Learning curves in manufacturing. *Science* 247 (4945), 920–924.
- Arrow, K., 1962. The economic implications of learning by doing. *The Review of Economic Studies* 29 (3), 155–173.
- Arthur, B., 2007. The structure of invention. *Research Policy* 36 (2), 274–287.
- Axelrod, L., Daze, R., Wickham, H., 1968. The large plant concept. *Chemical Engineering Progress* 64 (7), 17.
- Ayres, R., 1992. Experience and the life cycle. *Technovation* 12 (7), 465–486.
- Azevedo, I., Morgan, G., Morgan, F., 2009. The transition to solid state lighting. *Proceedings of the IEEE* 97, 481–510.
- Bae, W., Kwak, J., Lim, J., Lee, D., Nam, M., Char, K., Lee, C., Lee, S., 2010. Multicolored light-emitting diodes based on all-quantum-dot multilayer films using layer-by-layer assembly method. *Nano Letters* 10, 2368–2373.
- Balasubramanian, N., Lieberman, M., 2010. Industry learning environments and the heterogeneity of firm performance. *Strategic Management Journal* 31, 390–412.
- Basalla, G., 1988. *The Evolution of Technology*. Cambridge University Press.
- BCC, 2012. <http://www.bccresearch.com/market-research/nanotechnology/carbon-nanotubes-markets-technologies-nan024e.html>
- Benkart, L., 2000. Learning and forgetting: the dynamics of aircraft production. *American Economic Review* 90 (4), 1034–1054.
- Brown, R., 2011. Staying with copper in a fiber world. *Rural Telecom* (March–April), 22–26.
- CCAS, 2014. <http://www.ccas-web.org/superconductivity> (last accessed 04.02.14).
- Chader, G., Weiland, J., Humayun, M., 2009. Artificial vision: needs, functioning, and testing of a retinal electronic prosthesis. *Progress in Brain Research* 175, 317–331.
- Chandler, A., 1994. *Scale and Scope: The Dynamics of Industrial Capitalism*. Belknap, Boston.
- Chandler, D., 2013. New Solar-cell Design Based on Dots and Wires. *MIT News*, March 24, 2013.
- ChemTech, 2008. Instant insight: organic field-effect transistors. In: *Highlights in Chemical Technology*, March 18. [http://www.rsc.org/Publishing/ChemTech/Volume/2008/04/organic\\_transistors.asp](http://www.rsc.org/Publishing/ChemTech/Volume/2008/04/organic_transistors.asp)
- Christensen, C., 1997. *The Innovator's Dilemma*. Harvard Business School Press, Boston, MA.
- Connectus, 2012. <http://www.connectus.org/xxmarket.html>
- de Figueirdo, J., Kyle, M., 2006. Surviving the gales of creative destruction: the determinants of product turnover. *Strategic Management Journal* 27 (3), 241–264.
- Devoret, M., Schoelkopf, R., 2013. Superconducting circuits for quantum information. *Science* 339 (March), 1169–1174.
- de Weck, O., Roos, D., Magee, C., 2011. *Engineering Systems*. MIT Press, Cambridge.
- Display Search, 2009. <http://www.ecnmag.com/news/2009/04/worldwide-oled-revenues-forecast-reach-55b-2015>
- Display Search, 2013. <http://www.oled-info.com/displaysearch-most-oled-material-market-growth-will-come-oled-tvs-starting-2014>
- Dong, H., Wang, C., Hu, W., 2010. High performance organic semiconductors for field-effect transistor. *Chemical Communications* 46, 5211–5222.
- Dosi, G., 1982. Technological paradigms and technological trajectories. *Research Policy* 11 (3), 147–162.
- Dosi, G., Nelson, R., 2010. Technical change and industrial dynamics as evolutionary processes. In: Hall, B., Rosenberg, N. (Eds.), *Handbook of The Economics of Innovation*. Academic Press, Burlington.
- D-Wave, 2013. <http://www.dwavesys.com/en/dev-tutorial-hardware.html>
- Dutton, J., Thomas, A., 1984. Treating progress functions as a managerial opportunity. *Academy of Management Review* 9 (2), 235–247.
- Economist, 2012, January. Television Making: Cracking Up, 66 pp.
- Economist, 2012, December. Resistance is Fruiting. <http://www.economist.com/node/21540385>
- Economist, 2013. <http://www.economist.com/news/21566414-alternative-energy-will-no-longer-be-alternative-sunny-uplands>
- Evans, P., et al., 2011. 1.12 Tb/s superchannel coherent PM-QPSK InP transmitter photonic integrated circuit (PIC). *Optics Express* 19 (26), B154–B158.
- Fleming, L., 2001. Recombinant uncertainty in technological search. *Management Science* 47 (1), 117–132.
- Fleming, L., Sorenson, O., 2001. Technology as a complex adaptive system: evidence from patent data. *Research Policy* 30, 1019–1039.
- Francis, G., 2011. Data Storage – Trends and Directions. [http://lib.stanford.edu/files/pasig-jan2012/11B7%20Francis%20PASIG\\_2011\\_Francis\\_final.pdf](http://lib.stanford.edu/files/pasig-jan2012/11B7%20Francis%20PASIG_2011_Francis_final.pdf)
- Franklin, A., 2013. Electronics: the road to carbon nanotube transistors. *Nature* 498, 443–444.
- Fujimaki, A., 2013. Superconductivity web21, January 16, 2012. [www.istec.or.jp/web21/pdf/12\\_Winter/E15.pdf](http://www.istec.or.jp/web21/pdf/12_Winter/E15.pdf)
- Funk, J., 2013a. *Technology Change and the Rise of New Industries*. Stanford University Press.
- Funk, J., 2013b. What drives exponential improvements? *California Management Review* (Spring).

- Gold, B., 1974. Evaluating scale economies: the case of Japanese blast furnaces. *The Journal of Industrial Economics* 23 (1), 1–18.
- Green, P., Wind, Y., 1973. *Multi-attribute Decisions in Marketing: A Measurement Approach*. Dryden Press, Hinsdale, IL.
- Haitz, R., Tsao, J., 2011. Solid state lighting: 'the case' 10 years after and future prospects. *Phys Status Solidi A* 208 (1), 17–29.
- Hardisty, L., 2011. Long Live the Qubit. *MIT News*, 1 June 2011.
- Hasegawa, T., Takeya, J., 2009. Organic field-effect transistors using single crystals. *Science and Technology of Advanced Materials* 10, 024314.
- Hatch, N., Mowery, D., 1998. Process innovations and learning by doing in semiconductor manufacturing. *Management Science* 44 (11), 1461–1477.
- Henderson, R., Clark, K., 1990. Architectural innovation: the reconfiguration of existing product technologies and the failure of established firms. *Administrative Science Quarterly* 35, 9–30.
- Horowitz, 2011. The organic transistor: state-of-the-art and outlook. *European Physics Journal: Applied Physics* 53 (2), 3360.
- Hou, J., Guo, X., 2013. Active layer materials for organic solar cells. In: Choy, W. (Ed.), *Organic Solar Cells*. Springer Verlag, London.
- Hughes, T., 1983. *Networks of Power*. Johns Hopkins Press, Baltimore.
- Iansiti, 1995. Technology integration. *Research Policy* 24 (4), 521–542.
- IDTE, 2012. <http://www.idtechex.com/research/reports/organic-photovoltaics-opv-2012-2022-technologies-markets-players-000319.asp>
- Investor, 2013. <http://investorshub.advn.com/Quantum-Materials-Corporation-QTMM-15185/>
- ISSCC, 2013. International Solid State Circuits Conference. ISSCC Trends.
- Jones, N., 2013. The quantum company. *Nature* 498 (June), 286–288.
- Jorgenson, D., Ho, M., Stiroh, K., 2008. A retrospective look at the U.S. productivity growth resurgence. *Journal of Economic Perspectives* 22 (1), 3–24.
- Klepper, S., 1996. Entry, exit, growth, and innovation over the product life cycle. *American Economic Review* 86 (3), 562–583.
- Koh, H., Magee, C., 2006. A functional approach for studying technological progress: application to information technologies. *Technological Forecasting and Social Change* 73, 1061–1083.
- Koh, H., Magee, C., 2008. A functional approach for studying technological progress: extension to energy technology. *Technological Forecasting and Social Change* 75, 735–758.
- Koomey, J., Berard, S., Sanchez, M., Wong, H., 2011. Implications of historical trends in the electrical efficiency of computing. *IEEE Annals of the History of Computing* 33 (3), 46–54.
- Kuhn, K., 2009. Moore's Law past 32 nm: future challenges in device scaling. In: 13th International Workshop on Computational Electronics, <ftp://download.intel.com/newsroom/bios/pdfs/kkuhn/Kuhn.IWCE.invited.text.pdf>.
- Kwak, J., 2010. PhD thesis, cited in Changhee Lee, Seoul National University. <http://www.andrew.cmu.edu/org/nanotechnology-forum/Forum.7/Presentation/CH.Lee.pdf>
- Kwak, J., Bae, W., Lee, D., Park, I., Lim, J., Park, M., Cho, H., Woo, H., Yoon, D., Char, K., Lee, S., Lee, C., 2012. Bright and efficient full-color colloidal quantum dot light-emitting diodes using an inverted device structure. *Nano Letters* 12, 2362–2366.
- Lapre, M., Mukherjee, A., Wassenhove, L., 2000. Behind the learning curve: linking learning activities to waste reduction. *Management Science* 46 (5), 597–611.
- Lee, C., 2005. OLED 1 – Introduction. <http://wenku.baidu.com/view/783fa93283c4bb4cf7ecd196.html>
- Lieberman, M., 1984. The learning curve and pricing the chemical processing industries. *Rand Journal of Economics* 15, 213–228, R&D expenditures and capital intensity.
- Linton, J., Walsh, S., 2008. *Technological Forecasting & Social Change* 75, 583–594.
- Liu, J., Hersam, M., 2010. Recent developments in carbon nanotube sorting and selective growth. *Materials Research Society Bulletin* 35, 315–321.
- Lipsey, R., Carlaw, K., Bekar, C., 2005. *Economic Transformations*. Oxford University Press, NY.
- Mannan, S., 2005. *Lee's Loss Prevention in the Process Industries*, vol. 1. Elsevier Butterworth-Heinemann, Burlington, MA.
- Markets and Markets, 2011. <http://www.marketsandmarkets.com/Market-Reports/printed-electronics-market-197.html>
- Martinson, R., 2007. Industrial markets beckon for high-power diode lasers. *Optics* (October), 26–27 [www.nlight.net/nlight-files/file/articles/OLE%2010.2007.Industrial%20markets...pdf](http://www.nlight.net/nlight-files/file/articles/OLE%2010.2007.Industrial%20markets...pdf)
- Mas-Torrent, M., Rovira, C., 2008. Novel small molecules for organic field-effect transistors: towards processability and high performance. *Chemical Society Review* 37, 827.
- Mathews, J., Cho, D., 1999. Combinative capabilities and organizational learning in latecomer firms: the case of the Korean semiconductor industry. *Journal of World Business* 34 (2), 139–156.
- Merali, Z., 2011. First sale for quantum computing. *Nature* 474 (18), 194–198.
- Miller, G., 2012. New Power Semiconductor Technologies Challenge Assembly and System Setups. *Embedded.com*, April 16.
- Moore, G.E., 2006. Moore's Law at 40. Chapter 7 in *Understanding Moore's law*. In: Brock, D.C. (Ed.), *Four Decades of Innovation*. Chemical Heritage Foundation, Philadelphia, PA.
- Murray, F., 2002. Innovation as co-evolution of scientific and technological networks: exploring tissue engineering. *Research Policy* 31, 1389–1403.
- Murray, F., 2004. The role of academic inventors in entrepreneurial firms: sharing the laboratory life. *Research Policy* 33, 643–659.
- Nagy, B., Farmer, D., Bui, Q., Trancik, J., 2013. Statistical basis for predicting technological progress. *PLoS ONE* 8 (2), e52669. <http://dx.doi.org/10.1371/journal.pone.0052669NREL2013>.
- Nemet, G., 2006. Beyond the learning curve: factors influencing cost reductions in photovoltaics. *Energy Policy* 34 (17), 3218–3232.
- Nemet, G., 2009. Demand-pull, technology-push, and government-led incentives for non-incremental technical change. *Research Policy* 38 (5), 700–709.
- NHGRI, 2013. National Human Genome Research Institute. [www.genome.gov/sequencingcosts/](http://www.genome.gov/sequencingcosts/) (last accessed 10.09.13).
- Nordhaus, W., 2007. Two centuries of productivity growth in computing. *Journal of Economic History* 67 (1), 128–159.
- Nordhaus, W., 2009. The Perils of the Learning Model for Modeling Endogenous Technological Change, Cowles Foundation Discussion Paper No. 1685.
- NREL, 2013. National Research Energy Lab. [http://en.wikipedia.org/wiki/File:Pveff\(rev131204\).jpg](http://en.wikipedia.org/wiki/File:Pveff(rev131204).jpg)
- OLED, 2013. OLED-info.com. <http://www.oled-info.com/oled-mobile-phones> (last accessed 26.11.13).
- Paranthaman, P., Izumi, T., 2004. High-performance YBCO-coated superconductor wires. *MRS Bulletin* (August), 533–536.
- Patel, P., 2011. Quantum dots as solar cells. *MIT Technology Review*.
- Pavitt, K., 1984. Sectoral patterns of technical change: towards a taxonomy and a theory. *Research Policy* 13, 343–373.
- Physorg, 2013. <http://phys.org/news/2013-01-sony-tvs-high-end-quantum-dot.html>
- Preil, M., 2012. Optimum Dose for EUV: Technical vs. Economic Drivers, April 26, Future Fab 41.
- Research & Markets, 2013. <http://www.prnewswire.com/news-releases/quantum-dot-and-quantum-dot-display-gled-market-shares-strategies-and-forecasts-worldwide-nanotechnology-2013-to-2019-212987981.html>
- Rosenberg, N., 1969. The direction of technological change: inducement mechanisms and focusing devices. *Economic Development and Cultural Change* 18 (1), 1–24.
- Sahal, D., 1979. A theory of progress functions. *AIIE Transactions* 11 (1), 23–29.
- Sanderson, 2009., 2009. Quantum dots go large. *Nature* 459 (June), 760–761.
- Schmookler, J., 1966. *Invention and Economic Growth*. Harvard University Press, Cambridge.
- Selvamanickam, V., 2011. Second-generation HTS wire for wind energy applications. In: *Symposium on Superconducting Devices for Wind Energy*, 25 February, Barcelona, Spain.
- Service, R., 2013. Battle for the barrel. *Science* 339, 1374–1379.
- Sheats, J., Antoniadis, H., Hueschen, M., Leonard, W., Miller, J., Moon, R., Roitman, D., Stocking, A., 1996. Organic electroluminescent devices. *Science* 273, 884–888.
- Shiohara, T., Yoshizumi, M., Takagi, Y., Izumi, T., 2013. Future prospects of high Tc superconductors-coated conductors and their applications. *Physica C: Superconductivity* 484 (January), 1–5.
- Sinclair, G., Klepper, S., Cohen, W., 2000. What's experience got to do with it? Sources of cost reduction in a large specialty chemicals producer. *Management Science* 46, 28–45.
- SingularityHub.com, 2013. <http://singularityhub.com/2012/09/17/new-software-makes-synthesizing-dna-as-easy-as-using-an-ipad/> (last accessed 27.11.13).
- Solow, R., 1957. Technical change and the aggregate production function. *Review of Economics and Statistics* 39, 312–320.
- Stasiak, J., Richards, S., Angelos, S., 2009. Hewlett Packard's inkjet MEMS technology. In: *Proceedings of Society of SPIE*, p. 7318 <http://144.206.159.178/ft/CONF/16431771/16431793.pdf>
- Stobaugh, R., 1988. *Innovation and Competition: The Global Management of Petrochemical Products*. Harvard Business School Press, Boston.
- Suzuki, T., 2010. Challenges of image-sensor development. In: *Int'l Solid State Sensors Conference*.
- Thompson, P., 2012. The relationship between unit cost and cumulative quantity and the evidence for organizational learning-by-doing. *Journal of Economic Perspectives* 26 (3), 203–224.
- Thornton, R., Thompson, P., 2001. Learning from experience and learning from others. *American Economic Review* 91 (5), 1350–1368.
- Tilton, J., 1971. *International Diffusion of Technology: The Case of Semiconductors*. The Brookings Institution, Washington, DC.
- U.S. DoE (Department of Energy), 2010. *\$1/W Photovoltaic Systems: White Paper to Explore a Grand Challenge for Electricity from Solar*.
- Utterback, J., 1994. *Mastering the Dynamics of Innovation*. Harvard Business School Press.
- Weiss, M., Junginger, M., Patel, M., Blok, K., 2010. A review of experience curve analyses for energy demand technologies. *Technology Forecasting and Social Change* 77, 411–428.
- Wikipedia, 2014. [http://en.wikipedia.org/wiki/File:Transistor\\_Count\\_and\\_Moore%27s\\_Law\\_-\\_2011.svg](http://en.wikipedia.org/wiki/File:Transistor_Count_and_Moore%27s_Law_-_2011.svg) (last accessed 04.02.14).
- Winter, S., 2008. Scaling heuristics shape technology! Should economic theory take notice? *Industrial and Corporate Change* 17 (3), 513–531.
- Wright, T.P., 1936. Factors affecting the cost of airplane. *Journal of Aeronautical Sciences* 3 (4), 122–128.
- Yole, 2013a. <http://www.siliconsemiconductor.net/article/76878-Yole-Non-volatile-memories-to-penetrate-many-markets.php>
- Yole, 2013b. Emerging NVM Enter Niche Memory Markets: Expected to Reach \$2B by 2018.
- Yoon, Y., (unpublished PhD dissertation) 2010. *Nano-Tribology of Discrete Track Recording Media*. University of California, San Diego.