Quantitative empirical trends in technical performance

C.L. Magee [a,*], S. Basnet [a], J.L. Funk [b], C.L. Benson [a]

[a] *Massachusetts Institute of Technology (MIT), SUTD-MIT International Design Center, United States*
[b] *National University of Singapore (NUS), Singapore*

ABSTRACT

Technological improvement trends such as Moore's law and experience curves have been widely used to understand how technologies change over time and to forecast the future through extrapolation. Such studies can also potentially provide a deeper understanding of R&D management and strategic issues associated with technical change. However, such uses of technical performance trends require further consideration of the relationships among possible independent variables — in particular between time and possible effort variables such as cumulative production, R&D spending, and patent production. The paper addresses this issue by analyzing performance trends and patent output over time for 28 technological domains. In addition to patent output, production and revenue data are analyzed for the integrated circuits domain. The key findings are:

1. Sahal's equation is verified for additional effort variables (for patents and revenue in addition to cumulative production where it was first developed).
2. Sahal's equation is quite accurate when all three relationships — (a) an exponential between performance and time, (b) an exponential between effort and time, (c) a power law between performance and the effort variable — have good data fits ($r^2 > 0.7$).
3. The power law and effort exponents determined are dependent upon the choice of effort variable but the time dependent exponent is not.
4. *All* 28 domains have high quality fits ($r^2 > 0.7$) between the log of performance and time whereas 9 domains have very low quality ($r^2 < 0.5$) for power law fits with patents as the effort variable.
5. Even with the highest quality fits ($r^2 > 0.9$), the exponential relationship is not perfect and it is thus best to consider these relationships as the foundation upon which more complex (but nearly exponential) relationships are based.

Overall, the results are interpreted as indicating that Moore's law is a better description of longer-term technological change when the performance data come from various designs whereas experience curves may be more relevant when a singular design in a given factory is considered.

© 2015 The Authors. Published by Elsevier Inc. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

1. Introduction

An essential element of many approaches to research on technical change is an understanding of the overall societal impacts of specific technologies. The key methodology for many such studies is essentially historical involving detailed examination of the various interacting social and technical aspects of specific technological changes. Excellent examples of such studies include time keeping (Landes, 1983/2000), electric power (Hughes, 1983), the transistor (Riordan & Hoddeson, 1997), railroad economic impact (Fogel, 1964) and diverse technologies (Rosenberg, 1982). In almost all of these cases, numerous interacting

social changes were identified, but as with all historical studies, the lack of a counterfactual (what happened if a specific technology did not occur) renders precise knowledge unobtainable. The topic of this paper is a complementary way of studying technical change — quantitative empirical performance trends — and the aim of this paper is to improve the utility of this second approach. However, the link between performance trends and overall social impact is not simple.

Even with a narrow focus, for example, on the economic impact of a specific technical change (railroads in America in the late 19th century), there have been significantly different estimates of the actual impact of railroads (vs. a no railroad case) (Fogel, 1964; Fishlow, 1965). This is partly due to the fact that other technologies (for example canals) can be presumed to fulfill very different roles in the counterfactual case and partly due to the fact that the full impact of one technology on

others is highly complex — for example railroads and coal mining (Rosenberg, 1979). More recent work has made progress in decoupling the effects — for example relative to the role of computers in the economy (Brynjolfsson & McAfee, 2014) — but the complications are yet severe for quantitative estimation. Nonetheless there is wide agreement that technical change has enormous impact on society. Improvements in the cost and performance of new technologies enable technological discontinuities (Christenson, 1997) and large improvements in productivity (Solow, 1957) which in turn drive companies out of business, lift the economic level of many and generally transform society in profound ways. While it would be foolish to postulate that quantification will answer all of the important questions about technical change, this paper is based upon the belief that improvement of our theories of technical change will be aided by more dependable quantitative data about improvement of technologies. Indeed, many theories of technical change (Christenson, 1997; Abernathy & Utterback, 1978; Abernathy, 1978; Foster, 1985; Rosenbloom & Christensen, 1994; Tushman & Anderson, 1986; Utterback, 1994; Romanelli & Tushman, 1994) involve assumptions and hypotheses about such trends over the life cycle of a technology.

This paper attempts to make technical performance trends a more reliable part of the empirical arsenal for those studying technical change by clarifying an important issue. In particular, the research question of using an effort variable (such as patent activity, R&D spending, production, or revenue) or time as the independent variable is at the heart of this paper. Section 2 states the research question and analyzes past research concerning effort variables and time as the independent variable while Section 3 presents the data and methods used in our research. Section 4 presents performance trend results for 28 technological domains empirically comparing use of time and patents as effort variables: the section first analytically generalizes study of effort variables. Section 5 interprets the results and discusses their implications in terms of the quantitative technical performance trend of technologies.

2. Multiplicity of independent variables

An issue that must be addressed if one is to improve the utility of quantitative trend description is to determine the most appropriate independent variable. Thus, the first of our two coupled research questions: Is a framework that assumes an exponential relationship between performance and time better, worse or equivalent for quantitative empirical description than a framework that assumes a power–law relationship between performance and an effort-variable? The second research question is how one might empirically answer the first question.

The existing literature has multiple views on the better independent variable. For example, MacDonald and Schrattenholzer (2001) make a strong argument against using time as the independent variable:

> "For most products and services, however, it is not the passage of time that leads to cost reductions, but the accumulation of experience. Unlike a fine wine, a technology design that is left on the shelf does not become better the longer it sits unused."

One counterbalance to this apparent drawback of using time is the fact that measurement of effort introduces more needed data searching. More importantly, measurement of time is unambiguous whereas effort is ambiguous since it can be assessed according to several distinct concepts. The original research by Wright (1936) and further extensions (Alchian, 1963; Arrow, 1962; Argote & Epple, 1990; Benkard, 2000; Thompson, 2012; Dutton & Thomas, 1984) use cumulative production as the independent variable (the equation used will be discussed below). Although Wright treated learning as within a single plant (and for specific airplane designs), the same independent variable is now sometimes used more widely raising significant unit of analysis issues. In particular, researchers often (Argote & Epple, 1990; Dutton & Thomas, 1984; Ayres, 1992) treat cumulative production of an entire

(usually global) industry as the independent variable. However, this requires more careful definition of "industry" than is usually offered. In addition, this broad approach almost always introduces ambiguity about the initial values of output needed for cumulative production and thus also introduces data manipulation issues. To put it simply, determining how many and when unrecorded early units were produced is very problematic.

Another issue involves defining effort since R&D and new designs — not just production — are important in overall technical change. The quotation above (MacDonald & Schrattenholzer, 2001) implies that the unit of analysis is a "technology design" but technical change does not proceed simply by continuing to accumulate experience on existing designs but also through invention and creation of new designs. Recognizing this, some who take the broader view argue that cumulative production is not then "learning by doing" but instead an indirect — more or less total — measure of relevant *effort* (Ayres, 1992). More direct measures of such broader relevant effort include number of patents, R&D spending, and sales revenue: all of these as well as cumulative production have issues in initial values and are more difficult to obtain. For these as well as historical reasons, much of the practice for independent variables for effort remains cumulative production — despite identification of significant issues in interpreting such studies (Benkard, 2000; Thompson, 2012; Dutton & Thomas, 1984).

In addition to its passive nature, time as the independent variable conceptually seems to assume technology development is fully exogenous to what is happening in the economy. Since the consensus is that there are strong endogenous aspects of technology development, a fully exogenous assumption is counter-intuitive to those thinking primarily about causes. However, time indirectly contains the endogenous drivers as well as any exogenous drivers. For example, if the production rate of an artifact is constant, then cumulative production and time are proportional (with the proportionality constant the rate of production) so learning by doing for factory workers is also implicitly contained within the time variable. Similar arguments apply to R&D spending, revenue and numbers of patents with a direct relationship realized if the rates of each are constant over time. The obvious weakness of these indirect entailments for time is that the effort-variable (patent production, revenue or R&D spending) is not necessarily constant over time. A similar issue arises for cumulative production because other suggested effort variables (profits, R&D spending, patents, etc.) are *not* directly proportional to cumulative production. Indeed, cost or revenue per unit is the usual dependent variable so revenue per unit decreases with time: R&D spending and patents are proportional to revenue — not to units. An additional practical and theoretical obstacle to the use of cumulative production as the independent variable is the recent work showing that large performance improvements are often found before any commercial production occurs (Funk & Magee, 2015).

A preliminary conclusion could be that time casts "too wide a net" to give adequate emphasis to the endogenous affects in technological progress but that any specific effort variable "casts too narrow a net" to adequately capture all the endogenous efforts and captures none of the broader effects including "spillover" from efforts outside the implicit unit of analysis.

Perhaps surprisingly, given this qualitative story of differences in the approaches, in a very important way the two approaches are equivalent. Important steps in showing this equivalence have been taken by Sahal (1979), Nordhaus (2009), Nagy et al. (2013). The mathematical relationships (and the inter-relationship among them) specify this equivalence. A generalization of Moore's Law[1] that includes only performance q is

$$q = q_0 \exp\{k(t - t_0)\} \tag{1}$$

---

[1] q in the original or actual Moore's Law is the number of transistors on a wafer.

where $q_0 = q$ at $t = t_0$ and k is a constant; generalizing to an equation that includes cost as well as performance gives

$$q/c = q_0/c_0 \exp\{k(t-t_0)\} \qquad (2)$$

where c is price/cost and $c = c_0$ at $t = t_0$.

Wright's equation is usually formulated as describing only cost and relates it to cumulative production (p) as a power law:

$$c = B \, p^{-w} \qquad (3)$$

where B is the cost for the first unit of production and w is a constant.

A generalization of Wright's Law consistent with Eq. (3) is

$$q/c = (q/c)_0 \, p^{w} \qquad (4)$$

where $(q/c)_0$ is the value of q / c at 1 unit of production.

Sahal (1979) showed that that if the cumulative production, p, (or production) also follows an exponential relationship with time, namely

$$p = p_0 \exp\{g(t-t_0)\} \qquad (5)$$

where g is a constant and $p = p_0$ at $t = t_0$, then eliminating time between Eqs. (2) and (5) yields Eq. (4) with

$$k = w \cdot g. \qquad (6)$$

Thus, Sahal showed that the Wright and Moore formulations were equivalent when production follows an exponential (Eq. (5)): the key parameters are then simply related as shown in Eq. (6). An important issue is why one might expect Eq. (5) to hold. Nordhaus pointed out (Nordhaus, 2009) that as user-based performance increases or cost decreases according to Eq. (2), basic economics (demand elasticity) would result in demand (hence production) increases. Since Eq. (2) is exponential, demand and hence production would "automatically" (if demand elasticity is constant) follow the exponential relationship in Eq. (5). Thus, if either the Moore or Wright equation holds, Nordhaus' research indicates the other is likely to be followed as well.

Beyond these theoretical considerations, Nagy et al. (2013) carried out an important and relatively extensive empirical investigation of these relationships. For 62 cases (but where only price of the artifacts was considered), they found for most cases that production followed exponentials with time and that Eq. (6) showed minimal deviation *in all 62 cases*. The research by Sahal, Nordhaus and Nagy et al. shows that attempting to use fits to Eqs. (1) through (4) to distinguish among the intuitively different interpretations of technological progress is not easily done.

As an answer to the second research question in the first paragraph in this section, we consider other effort variables as a further test of what has been done with production. In particular, we pursue invention as a driver of technical change and utilize the number of patents in a domain as an effort-variable in 28 different domains. Since patents are a relatively direct measure of technologically novel designs, using them as an effort variable will allow a more direct test of whether time or effort is a more appropriate independent variable to use in describing quantitative technological trends. In addition, annual patent output in a technological domain may not continue to follow an exponential with time. If this occurs, we will learn whether Eqs. (2), (5), neither or both is followed, which will significantly illuminate our first research question.

3. Data and methods

3.1. Overview

The research objective in this paper is to compare the reliability of describing quantitative performance trends using a power law as a function of the annual number of patents (the effort-variable as discussed in Section 2) vs. as an exponential function of time. Testing this with a single technological domain is obviously not adequate to address the overall reliability of either approach. Thus, the data and methods described here involve finding patent numbers as a function of time and performance data as a function of time for a substantial number of technological domains. In this case, substantial is the 28 domains for which we have done this. Prior papers have described the methods developed for finding highly relevant and complete sets of patents (Benson & Magee, 2013, 2015) for these 28 domains so the material below is only a summary of that work. However, the performance data for the 28 domains is first reported here so the methods used in gathering that data are described in more detail. The basic tests performed are to look at goodness of fit of performance both for the power-law as a function of the annual output of patents and for an exponential function of time in each domain. The reliability of the two frameworks is then assessed based on all 28 domains.

3.2. Patent data

The Supplementary Information file (see Section 9 for overview and a link) contains annual patent counts from 1976 to 2013 for each domain; the quantity of patents is used as an effort variable for each domain to compare with time dependence in Section 4.4. The patents are all extracted from the PATSNAP database for US patents (Patsnap, 2013). We obtained these highly relevant patent sets by use of a classification overlap method (COM) developed earlier (Benson & Magee, 2013, 2015; Benson, 2014). The COM starts by searching for keywords that are selected as potentially important in the technological domain of interest. Each of the patent sets retrieved with the keyword search are then analyzed by quantitative metrics to assess the *patent classes* containing the patents in each set. The patents that are in *both* the most likely US patent class (UPC) *as well as* the most likely International patent class (IPC) are then taken as the patents in the domain. The basic intuition behind this classification overlap method is that the patent examiners differentially utilize — at least implicitly — the two systems beyond the sub-classifications in each system. Thus, additional confirming evidence of the nature of the technology in a patent is obtained by requiring that the patent be in both the top IPC and top UPC classes. The fact that each patent is classified in several IPC and UPC classes allows this dual classification to not be over-restrictive thus resulting in reasonably good completeness as well as higher relevancy than known alternative techniques (Benson & Magee, 2013, 2015). Each possible set is assessed by reading of patents in the potential set by two different technically-knowledgeable people who independently judge the relevancy of the individual patents to the technological domain of interest.[2] The application of the method for the 28 domains is more fully described in (Benson & Magee, 2015; Benson, 2014).

3.3. Performance data

3.3.1. Overview

In addition to the issue addressed in this paper (the dependent variable chosen), there are two other issues concerning reliable description of quantitative empirical performance trends. The methods used in this research for addressing each of these two issues are covered in the following sub-sections.

---

[2] In the 2 cases where the two raters differed by more than 7% in the relevancy rating, a third rater was used and in both cases, a different overlap was used. Thus, in all cases, the relevancy rating given is the average of the two (closely agreeing) raters. 300 patents are assessed for each set which results in an overall relevancy assessment for the patent set that is +/− 5.7%. This percentage follows from a standard sampling test for very large data sets that states that the uncertainty range at 95% confidence is determined by $1/(N)^{1/2}$ where N is the sampling population size, for N = 300, the uncertainty range is no larger than +/− 5.7%. 5.7% represents the upper limit of the uncertainty range, and the small patent sets have smaller ranges.

3.3.2. Unit of analysis

There are a lot of different approaches to decomposition of technology to specific technologies but the broadest attempts by highly experienced and motivated experts is clearly the US (UPC) and International patent classes (IPC). The UPC has about 400 "top level" classes and about 135,000 subclasses and the IPC (is structured with 628 4-digit classes and 71,437 subgroups at the most granular level of the hierarchy (Patsnap, 2013). Most technological progress researchers find these categories "too detailed" and in some sense not a good match to reality of the technological enterprise (Hall & Jaffe, 2001; Larkey, 1998). The reality of this logical issue is supported by the fact that an average US patent is listed in 4.6 UPCs and 2.4 IPCs indicating impact on multiple streams of technology.

A second way to differentiate among technologies is using Dosi's notions (Dosi, 1982) of "trajectories and paradigms" for technological progress. Dosi uses the idea of a paradigm as normal technology progress (analogous to Kuhn's interpretation of scientific progress) and trajectory as the economic focus of the technological problem solving process inherent in a paradigm. Much more recently, Martinelli (2012) utilized Dosi's concepts in a study of the telecommunication switching industry and in doing so, developed the ideas further.

A third way to differentiate technologies starts with generic functional categories (Koh & Magee, 2006). A refined version of this approach defines a technological domain as within a functional category: specifically, a technological domain is defined here as "artifacts[3] that fulfill a specific generic function utilizing a particular, recognizable body of scientific knowledge". This definition essentially decomposes generic functions along the lines of established bodies of knowledge. This approach is in the same spirit as the Dosi/Martenelli framework and with Arthur's later approach (Arthur, 2007) with the generic function connecting the domains to the economy and the body of scientific knowledge connecting the domain to science and other technical knowledge. Its advantage is that both generic function and domain are less ambiguous than the trajectory and paradigm concepts. The 28 domains studied in this paper are shown in this framework in Fig. 1.

3.3.3. Performance metric selection

Having defined the technological domain, how should one measure performance to quantify improvements in performance? The answer depends on the purpose of the study.

One *purpose* for studying trends of metrics is to indicate the significance of *usage* of a technology to the economy and society over time. The metrics used in such studies (and there is a large body of "diffusion" studies) include the amount consumed (Fisher & Pry, 1971), fraction of potential users who become users (Mansfield, 1961), market penetration of the artifact (Griliches, 1957), and units produced (Grubler, 1991). While this research is very important, it does not clarify trends in technical performance: the metrics used are not measures of technical performance and are not the metrics utilized in this research.

A *second purpose* is to help one anticipate future engineering problems or future design directions. The metrics used in such work (they are numerous and usually as part of more comprehensive design trend studies) include pressure ratio (Alexander & Nelson, 1973), temperature achieved in an artifact (Alexander & Nelson, 1973), energy efficiency (Koff, 1991), mass balance (Koff, 1991) and others. Although such technical metrics are also unquestionably important, effective analysis of technical change and its social and cultural impact requires that technical performance metrics must go beyond these technical metrics. Indeed, our interest in technical performance — the purpose of the current work — resides in its broader impact. As a consequence, the definition for the metrics we utilize is — *technical performance metrics are the properties of artifacts that are coupled to economic usage but are independent of amount of usage, number of possible users, competitive*

offerings, or the scarcity or depletion of resources that are used in building the artifacts.

The "ideal" metric for assessing technical performance is one that would assess the economic value of an artifact independently of purely economic variables such as scarcity and strength of demand. An ideal metric would combine (in the "correct" weight) all performance factors that have a role in a *purchase/use* decision. Thus, these "techno-economic" metrics would measure the performance of an artifact as viewed by a user and not design variables as viewed by an engineer (the technical metrics) and also not the number of users or depletion effects as present in metrics focused more on marketing or economic impact. The desire for such ideal metrics has also been discussed as part of hedonic pricing research (Alexander & Nelson, 1973; Bresnahan, 1986; Willig, 1978).

The metrics we use are thus user focused, avoid incorporation of depletion effects, avoid measuring the amount of use, increase when use is enhanced and are intensive-not extensive or size dependent-metrics. For the 28 domains, we examined trends in 71 metrics and all of these are reported in the supplemental information. We chose the most reliable and meaningful metrics (these are given for all 28 domains in Fig. 1) but none of the conclusions we arrive at are substantially changed if different choices among the 71 are examined.

One can often obtain technical performance data from a variety of sources combining them into metrics that vary with time and other independent variables. Although the entire data set thus obtained can be of interest, for determining the trend, only *non-dominated observations* are typically used. Non-dominated observations are those for which the metric is not surpassed in magnitude by the value achieved by the metric at lower values of effort variables (smaller number of patents, smaller cumulative production, etc.) or earlier time — they are "record setters". Although this reduces the amount of data available for analysis, it is the usual preferred practice because of concern that dominated points may be exceedingly high on a missing variable introducing noise.

4. Results

4.1. Summary of results

Section 4.2 presents the mathematical basis for generalization of Sahal's equation to effort-variables other than production. In Section 4.3, we examine a diversity of effort-variables for the integrated circuits domain attempting a preliminary and wider test of Sahal's relationship than can be done by any singular effort-variable. Section 4.4 presents the results that explicitly focus on the first research question concerning time vs. effort-variables as the independent variable for a technical performance trend. In that section, the goodness of fit according to the two frameworks is compared for the 28 domains using the annual patent output in a domain as the effort-variable.

4.2. Mathematical generalization of effort-variables

In Section 2, we reviewed Sahal's work showing a simple relationship (Eq. (6)) between the power law exponent (w) for production and the exponential (k) with time. The relationship is followed as long as the production (and thus also cumulative production) follows an exponential with time. We demonstrate in this section that Sahal's relationship is expected if *any* effort-variable follows an exponential relationship with time. In particular, if we simply use the chain rule to decompose the derivative of the log of performance[4] vs. time defining E as any effort variable, we obtain:

$$d \log q/dt = d \log q/d \log E \cdot d \log E/dt \qquad (7)$$

The left hand side of Eq. (7) is the familiar slope of the log performance vs. time plot, which is k (Eq. 2 exponent). The first term on the

---

[3] Artifacts include systems, products, subsystems, processes, software and components.

[4] Log as used in this paper represents the natural logarithm.

#	Information	Energy	Material
Storage	**Semiconductor Information storage** *(Solid-state physics, chemistry)* [transistors/die]	**Electrochemical batteries** *(Electro-chemistry)* [W·hr/$]	
	Magnetic information storage *(magnetic materials)* [Mbits/$]	**Capacitors** *(Electrostatics)* [W·hr/$]	
	Optical information storage *(Optical materials)* [Mbits/$]	**Flywheel** *(Mechanics, materials)* [kW.hr/kg]	
Transportation	**Electrical telecommunication** *(Electromagnetism)* [kb/s·$]	**Electrical power transmission** *(electromagnetics)* [W*km/$]	**Aircraft transport** *(Aerodynamics, mechanics)* [passenger*mph]
	Optical telecommunication *(photonics, optics)* [kb/s·km·$]	**Superconductivity** *(solid state physics)* [1/K]	
	Wireless telecommunication *(Electromagnetism))* [bits/s·m^2]		
Transformation	**IC Processors** *(Solid-state physics, chemistry)* [transistors/die]	**Combustion engines** *(Thermodynamics, mechanics)* [W/$]	**Milling Machines** *(Mechanics, dynamics)* [hp/mm]
	Electronic computation *(Solid-state physics, computation)* [1/s]	**Electrical motors** *(Electromagnetism)* [W/kg]	**3D printing** *(Materials, computation)* [mm^3/s·$]
	Camera Sensitivity *(Photonics)* [mV/micron^2]	**Solar PV power** *(Solid-state physics)* [kW·hr/$]	**Photolithography** *(Chemistry, optics)* [1/μm·$]
	MRI *(Nuclear physics)* [1/mm·s·$]	**Wind turbines** *(Aerodynamics, mechanics)* [W/$]	
	CT scan *(Atomic physics, computation)* [1/mm·s]	**Fuel cells** *(Physical Chemistry)* [kW/$]	
	Genome sequencing *(biology, genomics)* [1/$]	**Incandescent Lighting** *(materials)* [1000·lm·hr/$]	
		LED lighting *(Solid-state physics)* [lm/$]	

Fig. 1. The 28 technological domains defined for this study (shown in Bold Type) in the generic functional format used in (Koh & Magee, 2006). The italicized phrase is the scientific knowledge base for the domain. The primary metric reported for each domain (Section 3.3.3) is in normal type after the scientific knowledge base. The generic functional category is the intersection of the operands (across the top) and the operations (down the side). As a specific illustration, the generic functional category energy storage contains three of the 28 domains — namely electrochemical batteries, capacitors and flywheels.

right hand side of Eq. (7) is the power law exponent (w in Eqs. 3 and 4) and the second term is the slope of the exponential fit of the effort-variable with time, g in Eq. (5). Thus, for *any* effort variable, Sahal's relationship (Eq. (6)) holds, k = w · g, where g is now the exponent of Eq. (5) for any effort variable and w is the exponent of a power law (Eq. (4)) or the slope of a log performance vs. log effort plot. Of course, for the relationship to hold the effort variable must be the same for both terms on the right hand side of Eq. (7). As we will see in the next section, each of these quantities can depend upon the effort variable selected. Note that Eq. (7) holds for cumulative or annual versions of the effort variables as long as Eq. (5) holds.

4.3. Comparison of diverse effort-variables for the IC domain

An empirical examination of Eq. (7) using multiple effort variables is possible by examination of one of our 28 domains — namely integrated circuits. In particular, detailed production and revenue data for the IC domain were obtained from (Moore, 2006) to complement the patent output data we have for all domains. The performance data for the IC domain is the Moore's Law dependent variable, transistors/die. Table 1 shows empirical estimates of g and w for ICs for all three effort-variables along with r2 for each estimate. g describes the exponential between the effort-variable and time (Eq. (5)) whereas w is the

Table 1
Empirical values of g and w for IC processors: g from data fit to Eq. (5); w from data fit to Eq. (4); k_s, determined by Sahal's relationship, Eq. (6); k from data fit to Eq. (2)

Independent-effort-variable	g (r^2)	w (r^2)	k_s; {k}
Production/demand	0.59 (.97)	0.6 (.99)	0.35; {0.36}
Revenue	0.095 (.91)	3.4 (.88)	0.32; {0.36}
Number of patents	0.114 (.76)	3.0 (.86)	0.34; {0.36}

power law fit for performance vs. the effort variable (Eq. (4)). These results indicate acceptable power-law (w) fit quality for the performance variable (transistors/die) as a function of each of the three effort-variables — production, revenue and patents for this domain. The results in Table 1 also indicate acceptable fit to the exponential with time (g) for each of these effort-variables.

It is first worth noting that the estimates for w and g are dependent upon the particular effort-variable; g is much higher and w much lower for production as opposed to revenue. This striking result is a natural outcome of the fact that this domain has improved rapidly so that the revenue per transistor has greatly diminished over time. As a result, the exponential increase with time (g) is much lower for revenue than for production. Similarly, the increase in log performance with increase in log effort (w) is understandably much larger for revenue than for production again because of the much more rapid increase in production compared to revenue with the same performance increase. In a given domain, the amount of R&D spending is approximately proportional to revenue and the number of patents is approximately proportional to R&D spending (Margolis & Kammen, 1999); thus, g and w for revenue and patents are expected to be similar. Table 1 confirms empirically that g and w are much more similar for patents and revenue than for production/demand but also shows that patents have increased slightly more rapidly (11.4% per year) compared to revenue (9.5% per year) due to increases in the R&D/revenue ratio in this domain over time (Mowery, 2009).

Despite the systemic change in w and g for the three effort-variables, the last column in Table 1 shows good agreement between direct determination of k and the value of k calculated from Sahal's Equation for all three effort-variables. The estimates are definitely within the confidence interval for k for the IC processors. The agreement with all three effort-variables for IC processors is a strong confirmation of the usefulness of Sahal's relationship and of the generalization derived in Section 4.2. The results also show that patents can potentially be used as an effort variable which is particularly useful since it has been argued (Foster, 1985) that use of an invention-oriented effort-variable is superior to time or production. We now turn to results for using patents as an effort-variable for all 28 of our domains. We first show some plots of actual data to calibrate the reader to different levels of fit found in the data for Eqs. (2), (4) and (5).

4.4. Performance vs. time and patent output 28 technological domains

Fig. 2a shows log-linear plots of performance with time (Eq. (2)) for four domains (optical telecom, LEDs, batteries and 3D printing) using four relevant performance metrics (we call each of these domains
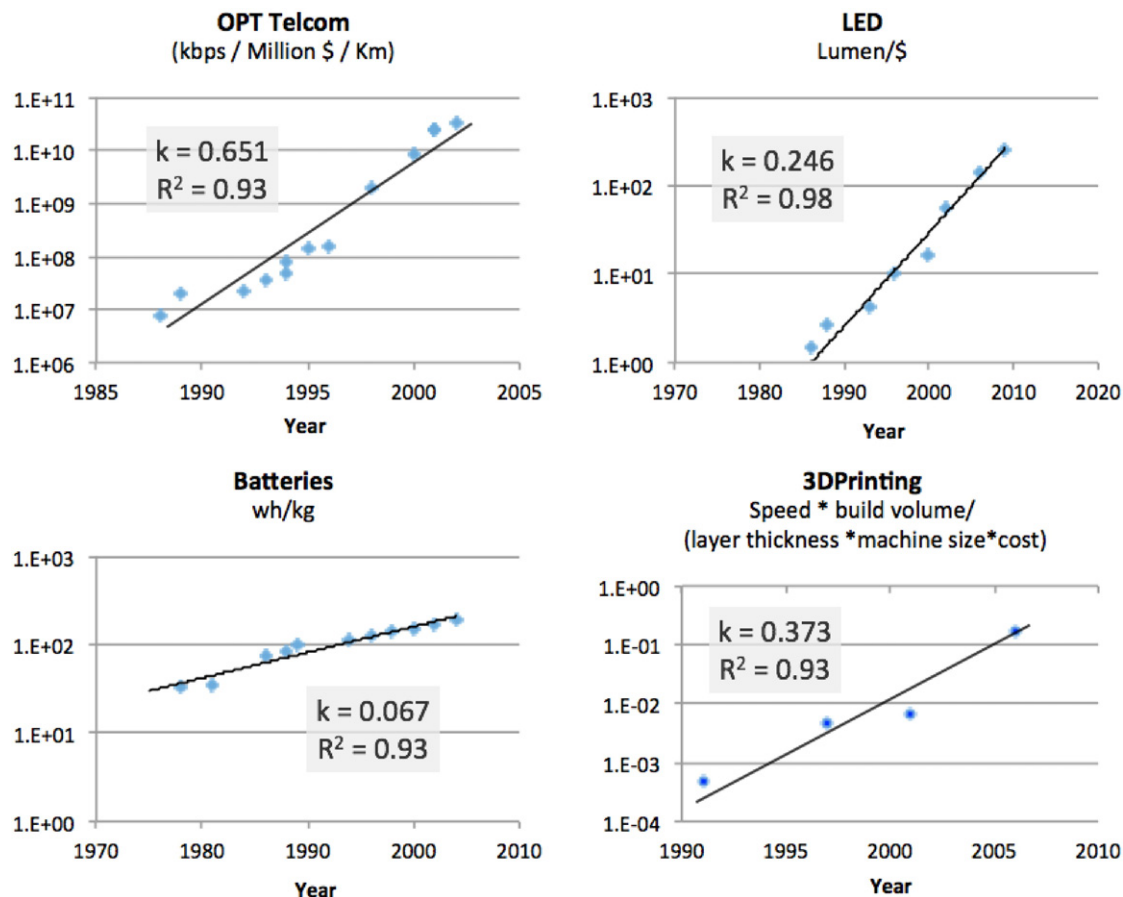

Fig. 2. A: Technological performance (log) against time for four domains (optical telecommunications, LED lighting, electro-chemical batteries, and 3D printing). The performance metric for each domain is shown above the graph. B: Power law fit for four domains (optical telecommunications, LED lighting, electrochemical batteries and 3D SLA printing). The metric for each is above the graph. C: Annual patents against time for four domains (optical telecommunications, LED lighting, electro-chemical batteries, and 3D printing). The performance metric for each domain is above the graph.
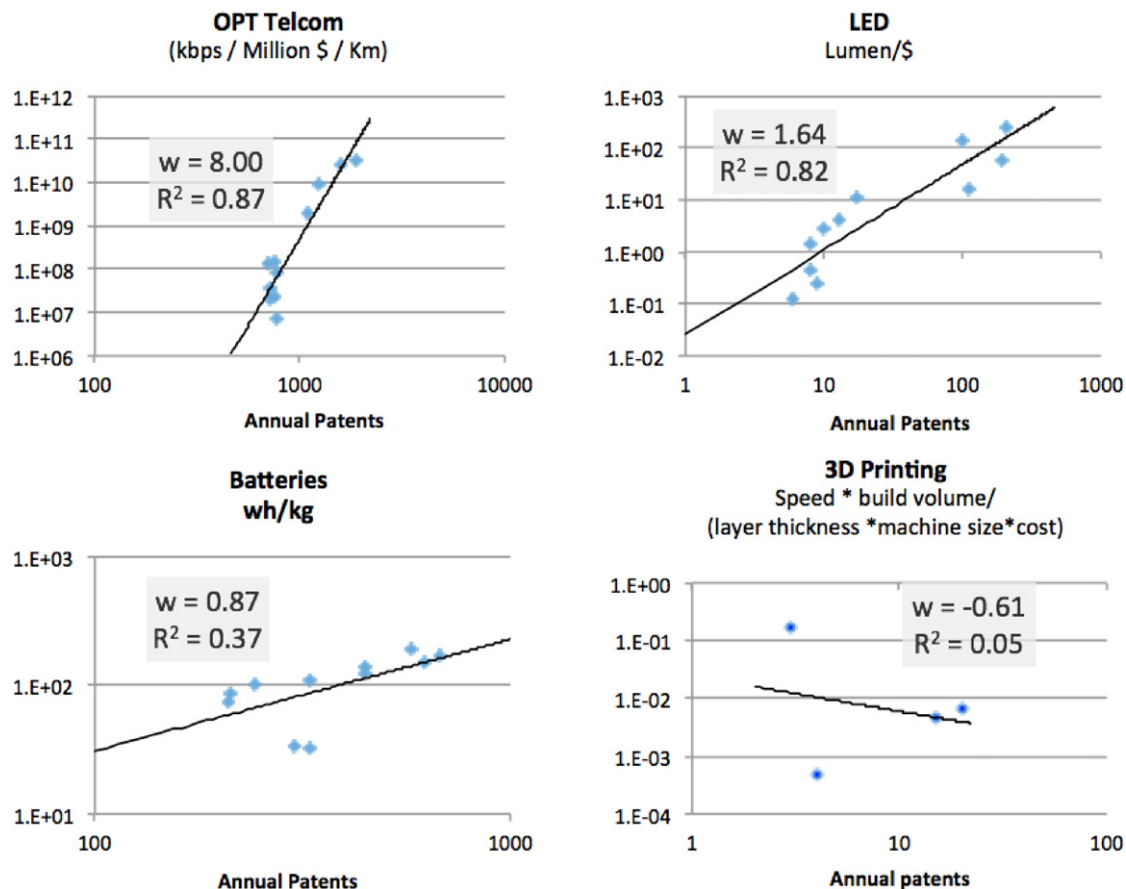
Fig. 2 (*continued*).

with a specific metric, a domain-metric-pair). Fig. 2b shows the log performance vs. log patents[5] (Eq. (4)) for the same four-domain-metric-pairs and Fig. 2c plots log patents vs. time (Eq. (5)) for the same four pairs. These four domain-metric-pairs are chosen because they represent the full range of quality of fits in our larger data set. In particular, the LED and optical telecom plots show good r^2 values and subjectively good fits for all three plots. However, 3D printing and batteries show poorer subjective fits and r^2 values in Fig. 2b and c (but are still fit well in 2a). It is important to note that Sahal's relationship is not expected to be accurate in cases with such poor fit since the parameters on the right hand side of Eq. (7) are not constant. In fact, the k estimated from Eq. (7) for 3D printing and batteries are off from the directly determined value by factors greater than 1.5 (much greater for 3D printing) but are within a factor of 1.2 for optical telecom and LEDs. These results indicate that Sahal's relationship is accurate for cases where good fits ($r^2 > 0.75$) exist for k, w and g. The expected reduction in accuracy of the relationship occurs as the fits deteriorate. This finding does not depend upon the nature of the effort-variable but instead upon whether an exponential describes well the relationship between the effort-variable and time.

Fig. 3 is a distribution of r^2 for all 28 domains for the three key fit parameters (k, g and w). Over all domains, the fits are clearly better for an exponential relationship between performance and time than they are for an exponential relationship of patent output with time or than they are for a power law of performance and annual patent output.

Only 2 of the 28 r^2 values are less than 0.8 for k but that the majority of the r^2 are less than 0.8 for w and for g. This demonstrates that Moore's Law is followed even when a relevant effort-variable does not increase exponentially with time.

This is an important finding because Sahal's equation can be interpreted to say that one needs to have exponential increases with time for effort-variables to get exponential relationships of performance with time. The results in Fig. 3 show that such a conclusion is clearly not true since many cases of very poor exponentials are found for the effort-variable (12 values of r^2 for g are less than 0.5) and yet none are found for exponentials with time. This interpretation of Sahal's equation assumes that a power law between performance and an effort-variable is fundamental and the exponential with time only arises because of a simultaneous exponential of effort with time. However, this suggested interpretation is reversed by the results in Fig. 3. Although the generalization of Moore's Law is followed in all cases, the many instance with low r^2 for w shows that the power law is not followed for this effort-variable when patent numbers do not increase exponentially with time. Thus, Moore's Law (the exponential increase of performance with time) appears fundamental and the power law only applies when a simultaneous exponential of effort as a function of time exists: Eq. (7) then shows that w is also constant (a good power law fit).

5. Discussion

The results that were just presented indicate that the first research question stated in Section 2 about the most effective framework for describing quantitative empirical performance trends is answered in favor of the approach first used by Gordon Moore (1965) fifty years ago. Our second research question was answered by empirical analysis of 28

---

[5] We also tested cumulative patents and got similar results. An additional issue in use of cumulative patents (as with any cumulative variable) is that one does not have actual data for many years (in our patents, we cannot apply COM before 1976) and estimation techniques do not actually add any information. Since Eq. (7) works for cumulative or annual effort variables, and the exponents g and w are the same for annual or cumulative variables, we use the actual data rather than an arbitrary reworking of it.
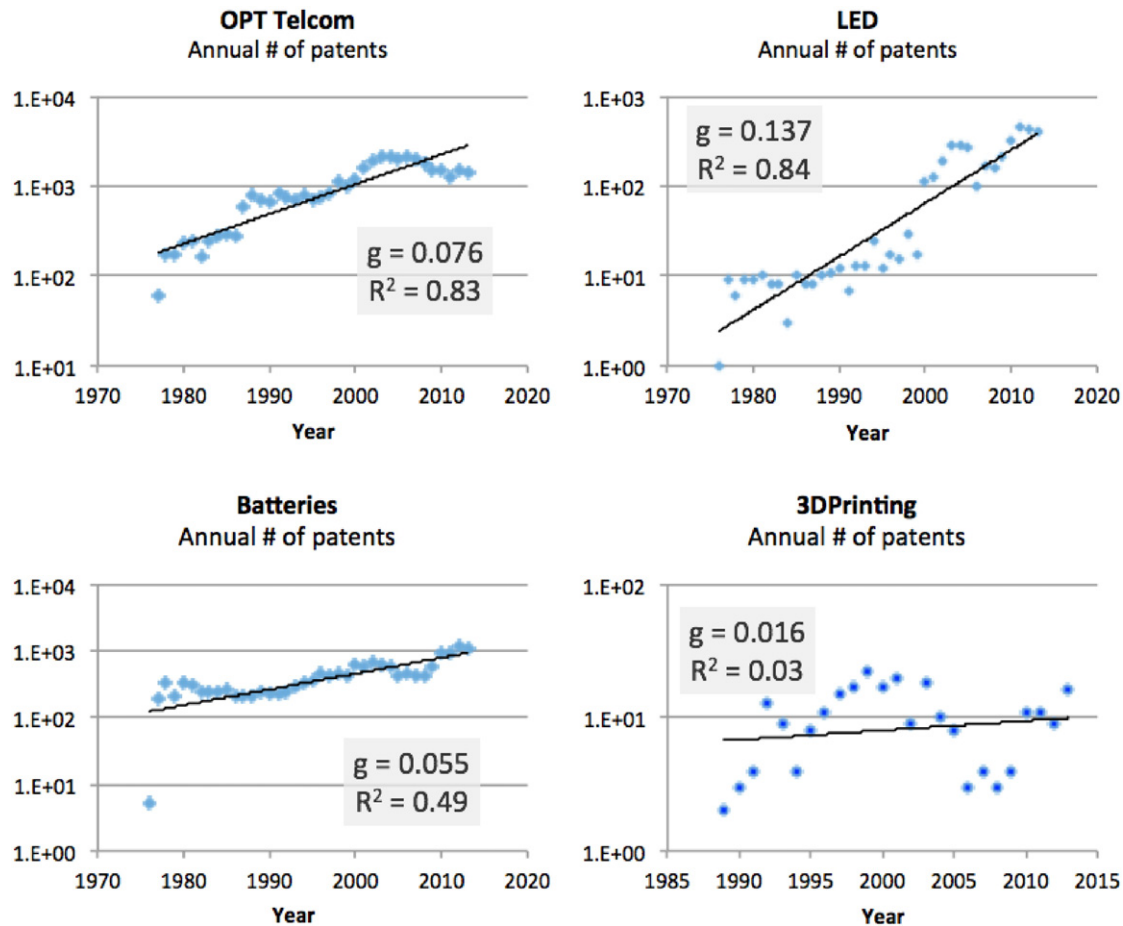
Fig. 2 (continued).

domains where the performance and annual patent output were obtained for the period 1976–2013. Since the patent output as a function of time was often not exponential, these data allowed one to see whether performance then followed a power law function of the patent output or an exponential with time. This decoupling of time and effort made it possible to break away from Sahal's relationship and thus differentiate between two usually coupled approaches. The results show that when patent output does not follow an exponential increase with time, one usually also does not find good fits for power laws between performance and patents (the effort-variable) despite having a good fit for an exponential between performance and time.
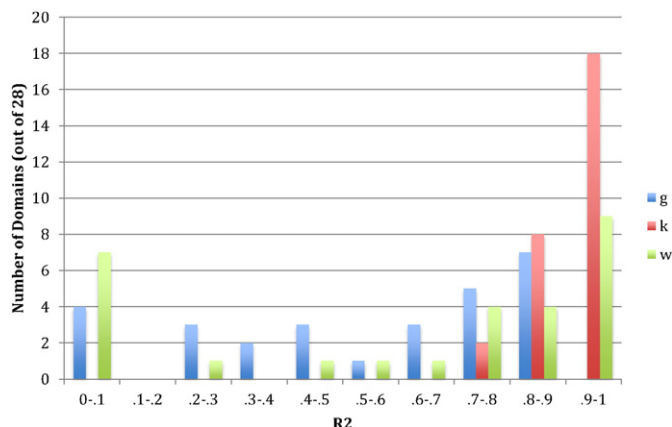
Fig. 3. Distribution of r² for all 28 domains for k, g and w.

This paper also shows theoretically and empirically that Sahal's relationship is followed for diverse effort-variables when the effort-variable increases exponentially with time. In particular, we identify production — the most popular choice in the literature — but also revenue, R&D spending and quantity of patents issued as potentially useful "effort-variables". Our results demonstrate that Sahal's equation is valid for integrated circuits using revenue, patents or production as the effort-variable.

While there is no logical basis for concluding that other effort-variables would lead to a different answer to our first research question (the time dependence appears fundamentally correct and choice of alternative effort-variables would not change this), we do note that our empirical results are only for patent output as a measure of effort in a domain. In the 62 cases studied by Nagy et al. (2013), the most popular effort-variable (cumulative production) was fit adequately by an exponential relation with time and Sahal's relationship (with price as the performance metric) was followed for all 62 cases. From this extensive test, it appears that no differentiation of the two frameworks is possible when one uses only cumulative production as the effort-variable. This could support a conjecture that cumulative production rather than time is the appropriate independent variable but such a conjecture is somewhat weakened by research that has shown rapid improvements in performance before any commercial production occurs (Funk & Magee, 2015). To our knowledge, no-one has obtained effort-variable data for an extensive set of domains beyond these two studies so it is not clear what findings would result from use of effort variables like revenue or R&D spending.

Thus, for examining quantitative trends in performance, time should always be reported since it is always available, requires no more work and appears to be fundamentally important. In cases where new designs

and inventions occur during the trend studied, fitting the exponential with time in addition to the power law with an effort-variable (if sufficient effort-variable data exists) also appears sensible. Our findings show that the generalized Moore's Law formulation of technical trends is the most accurate over a wide range of technological domains where new designs occur. We also recommend explicit discussion of the specific algorithm for estimating missing data when using cumulative effort-variables. This is unfortunately rarely done now and seriously limits the utility of such work since the importance of unknown assumptions cannot be checked. It is preferable to simply use the effort variable (annual or some other fixed period) directly rather than cumulative versions that are undocumented.

One practical implication of the results reported in this paper is for technological forecasting. Our major concern in this paper is to arrive at the best framework for describing the past; the clear value of this is for deeper understanding of what *has* occurred. However, future values of technical performance and cost are critical to such issues as potential diffusion and firm profitability. Thus, accurate projection of future performance is a potentially important element in forecasting potential larger scale change. Even with high r^2, the graphs in Fig. 2 show far from perfect exponentials warning us that extrapolation will not lead to perfect forecasts. Nonetheless, back-casting research (Nagy et al., 2013; Farmer & Lafond, 2015) has demonstrated that extrapolation of past trends is useful in estimating future values and thus overall establish some reality for technological forecasting based upon extrapolation. It is our viewpoint that significantly better forecasting will be enabled by improved quantitative, explanatory theories and the next paragraph argues that the current results and other recent research are important steps towards this goal.

Having the most fundamental framework for describing quantitative technical performance trends for a wide variety of technological domains opens up a number of research questions of significance to understanding technological change and thus improving our foundation for technological forecasting. The 28 domains reported here show variation in improvement rate from 3.1% per year (electric motors) to 65.1% per year (optical telecommunication): such variation is more than sufficient for quantitative empirical and theoretical investigation. Indeed, recent research by two of the authors of this paper (Benson & Magee, 2015) found very strong correlations with patent metadata in a domain and the exponential rate of improvement for that domain. The findings (Benson & Magee, 2015) also support reliable forecasting of rates of improvement for at least 12 years into the future. Moreover, the correlations support a conceptual basis (Benson & Magee, 2014, 2015) for why some domains improve more rapidly than others based upon importance, immediacy and recency of patents in a domain. These findings along with enhanced back-casting research (Farmer & Lafond, 2015) and first-principle modeling (Basnet, 2015) are all enhancing our ability to forecast technological change and further research on hybrid approaches may be of particular utility. This work is enabled by knowledge that performance as an exponential function of time is the best framework for these efforts as established by the research reported here. Nonetheless, even with such results significant further research will be needed to delineate what aspects of technological change can then be forecast: this is not likely to include overall societal change because of our current level of understanding of the complex interaction of technologies and the economy as outlined in the introduction to this paper.

Two other topics involve hypotheses about describing the trends, and we have not fully identified meaningful analytical procedures to address them. S curves are hypothesized to be the usual trend for technical performance when plotted linearly against time or effort (Foster, 1985; Schilling & Esmundo, 2009). Visual inspection of linear plots for all 71 domain metric pairs found that none unequivocally appeared to be S curves as a function of time or effort; however, a desire for a more clearly objective way of determining the reality of S curves is needed. Unfortunately, statistical tools are limited by the fact that logistic (and other equation forms giving S curves) contain additional variables: there are cases (Keyes, 1977) when these curves have been fit to data predicting emerging S curves that have

not yet (even 30 years later) appeared. A second hypothesis about technical performance trends is that they show major discontinuities (Tushman & Anderson, 1986). Testing this hypothesis is not straightforward because increases in technical performance must in reality be discontinuous since advances are typically made by introduction of discretely different designs (inventions and products). Moreover, the level of discontinuity is dependent upon the time between new products and it is not known how many new product observations are missed. One might want to only note discontinuities that in fact are breaks from an existing exponential or power law fit. Objective means for deciding what constitutes a major technological break is also needed to address these questions. Overall, the results reported here give no support to S curves, quantitative discontinuities or life cycle hypotheses in regard to technical change but instead support a generalization of Moore's law as the foundation upon which change occurs. The noise apparent even with good fits (see Fig. 2) does clearly allow room for much variation due to social and economic complexity but such complexity is apparently built upon the regularity of exponential improvement.

Our final important topic for future research (that may well greatly extend work on dependent variable metrics) is the linking of technical performance change with productivity changes with time. Although it is widely agreed that technological change is a major source of economic growth (Solow, 1957; Arthur, 2007; Romer, 1990), there are no economic theories that use quantitative trends in technological performance as input and obtain as output the productivity change over time in an industrial sector. This is at least partly due to the difficulty of the problem of connecting technologies with industrial sectors but the lack of attempts is disappointing. A simpler beginning issue in this regard might be linking technical performance trends with innovation and diffusion. It is widely intuitively understood that the metrics studied here attempt to measure what is "better" and that what is better is generally what diffuses (Griliches, 1957; Mansfield, 1961) but formal treatment has not been attempted. In fact, most diffusion models implicitly consider the relative performance and cost of a diffusing artifact to be constant so a doable first step might be to eliminate this assumption.

6. Conclusions

Twenty-eight technologies (technological domains) are studied in this paper exploring their performance improvement as a function of time and effort: the annual number of patents published in the technological domain is used to measure effort. A total of 71 different performance metrics were studied for these 28 domains.

The major finding is that the results indicate that Moore's exponential law appears to be more fundamental than Wright's power law for these 28 domains (where performance data are record breakers from numerous designs and different factories). This conclusion is supported by:

- The performance metrics in all 28 technological domains have strong exponential correlations with time (Moore's law generalization).
- In contrast, most of these same performance metrics in the 28 domains have much weaker log-log correlations with patents (Wright's law generalization).
- Wright's law is followed only in those domains where published patents in the domain show a strong exponential correlation with time. For these domains, Sahal's relationship is followed: $k = w \cdot g$, where k is the Moore's law exponent, w the Wright power law exponent and g the patent growth exponent. This indicates that the power–law relationship in these cases is not fundamental but instead a shadow of Moore's Law.

Acknowledgments

Appendix A. Supplementary information

In addition to the information presented in this paper, we have compiled key data used into a Microsoft Excel file that can be easily obtained by copying the following link into a web browser where it can be viewed and/or downloaded: http://bit.ly/mageeetalSIMay2014.

This document contains three worksheets which are accessible by clicking the tabs at the bottom of the excel window.

- 28 domains with k, g and w — this worksheet contains the k, g, and w values along with the r^2 values for each of the regressions for the 28 technological domains.
- 71 domain-metric-pairs with Statistical information — this worksheet includes the 71 domain-metric-pairs for the 28 technological domains (sixteen have more than one metric for which trends are determined) along with the relevant statistical information.
- Domain Annual Patenting Rates — this worksheet shows the annual number of patents for each of the 28 domains.

References

Abernathy, W., 1978. The Productivity Dilemma. Johns Hopkins University Press, Baltimore.

Abernathy, W., Utterback, J., 1978. Patterns of industrial innovation. Technol. Rev. 80, 40–47.

Alchian, A., 1963. Reliability of progress curves in airframe production. Econometrica 31 (4), 679–694.

Alexander, A., Nelson, J.R., 1973. Measuring technological change: aircraft turbine engines. Technol. Forecast. Soc. Chang. 4 (2), 189–203.

Argote, L, Epple, D., 1990. Learning curves in manufacturing. Science 247, 920–924.

Arrow, K., 1962. The economic implications of learning by doing. Rev. Econ. Stud. 29 (3), 155–173.

Arthur, B., 2007. The structure of invention. Res. Policy 36 (2), 274–287.

Ayres, R., 1992. Experience and the life cycle. Technovation 12 (7), 465–486.

Basnet, S., 2015. Dept. of Mechanical Engineering, MIT (doctoral thesis).

Benkard, C.L., 2000. Learning and forgetting: the dynamics of aircraft production. Am. Econ. Rev. 90 (4), 1034–1054.

Benson, C.L., 2014. Doctoral Thesis, Department of Mechanical Engineering, MIT.

Benson, C.L., Magee, C.L., 2013. A hybrid keyword and patent class methodology for selecting relevant sets of patents for a technology field. Scientometrics 96, 69–82.

Benson, C.L., Magee, C.L., 2014. On Improvement rates for renewable energy technologies: solar PV, wind turbines, capacitors, and batteries. Renew. Energy 68, 745–751.

Benson, C.L., Magee, C.L., 2015. Quantitative determination of technological improvement from patent data. PLoS ONE 10 (4), e0121635. http://dx.doi.org/10.1371/journal.pone.0121635.

Bresnahan, T.F., 1986. Measuring spillovers from technical advance: mainframe computers in financial services. Am. Econ. Rev. 742–755.

Brynjolfsson, E.,, McAfee, A., 2014. The second machine age: work, progress and prosperity in a time of brilliant technologies. W.W. Norton and Company, New York.

Christensen, C., 1997. The Innovator's Dilemma. Harvard Business School Press, Cambridge.

Dosi, G., 1982. Technological paradigms and technological trajectories: a suggested interpretation of the determinants of technical change. Res. Policy 11, 147–162.

Dutton, J.M., Thomas, A., 1984. Treating progress functions as a managerial opportunity. Acad. Manag. Rev. 9 (2), 235–247.

Farmer, J.D., Lafond, F., 2015. How predictable is technological progress? arXiv:1502.0572 (arXiv:1502.05274)

Fisher, J.C., Pry, R.H., 1971. A simple substitution model of technological change. Technol. Forecast. Soc. Chang. 3, 75–88.

Fishlow, A., 1965. American Railroads and the Transformation of the Ante-Bellum Economy. Harvard University Press, Cambridge.

Fogel, R.W., 1964. Railroads and American Economic Growth: Essays in econometric History. Johns Hopkins Press, Baltimore.

Foster, R.N., 1985. Timing Technological transitions. Technol. Soc. 7, 127–141.

Funk, J.L., Magee, C.L., 2015. Rapid improvements with no commercial production: How do the improvements occur? Res. Policy 44, 777–788.

Griliches, Z., 1957. Hybrid corn: an exploration in the economics of technological change. Econometrica 25 (4), 501–522.

Grubler, A., 1991. Diffusion: long-term patterns and discontinuities. Technol. Forecast. Soc. Chang. 39, 159–180.

Hall, B., Jaffe, A., 2001. The NBER patent citation data file: lessons, insights and methodological tools. NBER Working Paper Series, p. p8498.

Hughes, T.P., 1983. Networks of Power: Electrification in Western Society, 1880-1930. Johns Hopkins University Press, Baltimore.

Keyes, R., 1977. Physical limits in semiconductor electronics. Science 1230–1235.

Koff, B.F., 1991. Spanning the globe with jet propulsion. J. Am. Inst. Aeronaut. Astronaut. http://dx.doi.org/10.2514/6.1991-2987.

Koh, H., Magee, C.L., 2006. A functional approach to technological progress: application to information technology. Technol. Forecast. Soc. Chang. 73, 1061–1083.

D. Landes, Revolution in Time: Clocks and the Making of the Modern World, Belknap Press of Harvard University Press, 1983/2000.

Larkey, L.A., 1998. Patent search and classification system. Proceedings of the fourth ACM conference on Digital Libraries http://dx.doi.org/10.1145/313238.313304.

MacDonald, A., Schrattenholzer, L., 2001. Learning rates for energy technologies. Energ Policy 29, 255–261.

Mansfield, E., 1961. Technical change and the rate of imitation. Econometrica 29 (4), 741–766.

Margolis, R.M., Kammen, D.M., 1999. Underinvestment: the energy technology and R&D policy challenge. Science 285, 690–692.

Martenelli, A., 2012. An emerging paradigm or just another trajectory? Understanding the nature of technological changes using engineering heuristics in the telecommunications switching industry. Res. Policy 41, 416–429.

Moore, G.E., 1965. Cramming more components onto integrated circuits. Electron. Mag. 38 (8).

Moore, G.E., 2006. In: Brock, D.C. (Ed.), Moore's law at forty. Understanding Moore's law: four decades of innovation. Chemical Heritage Foundation (chapter 7).

Mowery, D.C., 2009. Plus ca change: industrial R&D in the third industrial revolution. Ind. Corp. Chang. 18 (1), 1–50.

Nagy, B., Farmer, J.D., Bui, Q.M., Trancik, J.E., 2013. Statistical basis for predicting technological progress. PLoS ONE 8, 1–7.

Nordhaus, W.D., 2009. The perils of the learning model for modeling endogenous technological change. Cowles Foundation Discussion Paper No. 1685.

Patsnap, 2013. Patsnap patent search and analysis. Retrieved July 1, 2013, from http://www.patsnap.com.

Riordan, M., Hoddeson, L., 1997. Crystal Fire: The Birth of the Information Age. W. W. Norton & Company, New York.

Romanelli, E., Tushman, M., 1994. Organizational transformation as punctuated equilibrium. Acad. Manag. J. 34, 1141–1166.

Romer, P.M., 1990. Endogenous technological change. J. Polit. Econ. 98 (5), S71–S102.

Rosenberg, N., 1979. Technological interdependence in the american economy. Technol. Cult. 1, 25–50.

Rosenberg, N., 1982. Inside the Black Box: Technology and Economics. Cambridge University Press, Cambridge.

Rosenbloom, D., Christensen, C., 1994. Technological discontinuities, organizational capabilities and strategic commitments. Ind. Corp. Chang. 3, 655–686.

Sahal, D., 1979. A theory of progress functions. AIIE Trans. 11 (1), 23–29.

Schilling, M.A., Esmundo, M., 2009. Technology S curves in renewable energy alternatives: analysis and implications for industry and government. Energ Policy http://dx.doi.org/10.1016/j.enpol.2009.01.004.

Solow, R., 1957. Technical change and the aggregate production function. Rev. Econ. Stat. 39, 312–320.

Thompson, P., 2012. The relationship between unit cost and cumulative quantity and the evidence for organizational learning-by-doing. J. Econ. Perspect. 26 (3), 203–224.

Tushman, M., Anderson, P., 1986. Technological discontinuities and organization environments. Adm. Sci. Q. 31, 439.

Utterback, J., 1994. Mastering the dynamics of innovation. Harvard Business School Press, Cambridge.

Willig, R.D., 1978. Incremental consumer surplus and hedonic price adjustment. J. Econ. Theory 17, 227–253.

Wright, T.P., 1936. Factors affecting the cost of airplanes. J. Aeronaut. Sci. 3 (4), 122–128.

Christopher L. Magee has been with MIT since January 2002 as a Professor of the Practice in the Engineering Systems Division (IDSS) and Mechanical Engineering. He also co-directs a multidisciplinary research center (SUTD/MIT International Design Center). Dr. Magee is currently engaged in research and teaching relative to technological progress and complex system design research. Dr. Magee is a member of the National Academy of Engineering, a fellow of ASM and SAE and a participant on major National Research Council Studies. Dr. Magee is a native of Pittsburgh, PA and received his B.S. and Ph.D. from Carnegie-Mellon University in that city. He later received an MBA from Michigan State University.

Subarna Basnet received his PhD in the Mechanical Engineering department at MIT in September 2015. He is a native of Nepal who received his previous degrees from the Indian Institute of Technology and Manhattan College.

Jeffrey L. Funk is Associate Professor at the National University of Singapore. He has published widely in the field of technical change including two recent books. He has previously held faculty positions at the University of Tokyo and Pennsylvania State University. He received his PhD in Engineering and Public Policy from Carnegie Mellon University.

Christopher L. Benson received his PhD at MIT in Mechanical Engineering in May 2014. His thesis involved linking patents and quantitative performance technical performance trends and he has much experience in teaching innovation and founded Startup Labs while receiving his undergraduate degree at MIT. He is currently an officer in the U. S. Air Force.