Technological Forecasting & Social Change 104 (2016) 237–246

Contents lists available at ScienceDirect

Technological Forecasting & Social Change

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

ARTICLE INFO

Article history:
Received 3 May 2014
Received in revised form 18 December 2015
Accepted 22 December 2015
Available online 29 January 2016

Keywords:
Moore's Law
Power laws and experience curves
Forecasting
Performance trends
Quantitative
Empirical trends

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.

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

* Corresponding author.
E-mail address: cmagee@mit.edu (C.L. Magee).

http://dx.doi.org/10.1016/j.techfore.2015.12.011
0040-1625/

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.