ARTICLE IN PRESS

Int. J. Production Economics ∎ (∎∎∎∎) ∎∎∎–∎∎∎

Contents lists available at ScienceDirect

Int. J. Production Economics

journal homepage: www.elsevier.com/locate/ijpe

Production economics and the learning curve: A meta-analysis

Eric H. Grosse*, Christoph H. Glock, Sebastian Müller

Institute of Production and Supply Chain Management, Technische Universität Darmstadt, Hochschulstr. 1, 64289 Darmstadt, Germany

ARTICLE INFO

Article history:
Received 13 April 2014
Accepted 18 June 2015

Keywords:
Learning
Learning curve
Experience curve
Curve fitting
Literature survey
Literature review
Meta-analysis

ABSTRACT

For almost a century, researchers and practitioners have studied learning curves in production economics. Learning, in this context, refers to performance improvements of individuals, groups or organizations over time as a result of accumulated experience. Various learning curves, which model this phenomenon, have been developed and applied in the area of production economics in the past. When developing planning models in production economics, the question arises which learning curve should be used to best describe the learning process. In the past, the focus of the literature has been on empirical studies that investigated learning processes in laboratory settings or in practice, but no effort has been undertaken so far to compare existing learning curves on a large set of empirical data to assess which learning curve should be used for which application. This study systematically collected empirical data on learning curves, which resulted in a large database of empirical data on learning. First, the data contained in the database is categorized with the help of meta-tags along different characteristics of the studies the data was taken from. Second, a selection of well-known learning curves is fitted to the empirical datasets and analyzed with regard to goodness of fit and data characteristics. We identify a set of data/task characteristics that are important for selecting an appropriate learning curve. The results of the paper may be used in production economics to assist researchers to select the right learning curve for their modeling efforts.

© 2015 Elsevier B.V. All rights reserved.

1. Introduction

Since Wright's (1936) seminal work on the functional relationship between the time required to perform a task and task repetition, a plethora of works has been published that investigated this functional relationship that is also termed the learning curve. Learning (or experience) curves assume that performance (output) improves as a task is repetitively performed, which is attributed to experience that is accumulated by the individual or group performing the task. Learning curves have frequently been the subject of research. Empirical studies focused on measuring learning by collecting empirical data, either in laboratory settings or in field studies. Learning effects were observed in various areas, such as assembly production (Shafer et al., 2001; Smunt and Watts, 2003), online ordering in supply chains (Kull et al., 2007), manual order picking (Grosse and Glock, 2013), or construction (Hinze and Olbina, 2009), to name just a few examples. Although the concept of learning curves in the field of production economics has been introduced almost a century ago, it is still of importance for manufacturing firms, for example as a performance measure, an aid in setting labor standards, a forecasting tool, or an application in decision support tools. Recent examples, for instance the market launch of Boeing's Dreamliner, confirm the practical importance of learning curves (Nolan, 2012).

Learning curves can be of multivariate or univariate type, where log-linear, exponential and hyperbolic models have most often been used (Anzanello and Fogliatto, 2011). Besides studying learning empirically, many authors have modeled the effects of learning on industrial and logistics processes by including learning curves in decision support models. Examples are inventory models that consider learning in the production rate, in setups or in fuzziness (e.g., Jaber et al., 2008, 2009; Kazemi et al., 2015), supplier selection models (e.g., Glock, 2012), models of manual order picking that consider picker learning (e.g., Grosse et al., 2013; Grosse and Glock, 2014), or vehicle routing models that involve driver learning (e.g., Zhong et al., 2007). Learning curves and their applications have been surveyed in a number of literature reviews, such as in Yelle (1979), Anzanello and Fogliatto (2011), or Fogliatto and Anzanello (2011).

* Corresponding author. Tel.: +49 6151 165281.
E-mail addresses: grosse@pscm.tu-darmstadt.de (E.H. Grosse), glock@pscm.tu-darmstadt.de (C.H. Glock), sebastianmueller-lm@gmx.de (S. Müller).

http://dx.doi.org/10.1016/j.ijpe.2015.06.021
0925-5273/© 2015 Elsevier B.V. All rights reserved.

Surprisingly, the question of how different learning curves perform and which learning curve to use in which application has not yet been addressed in a comprehensive study. Researchers and practitioners alike face the problem of selecting an appropriate learning curve each time learning effects are modeled, which can be challenging and time-consuming given the large number of learning curves that have been developed in the past. To assist researchers and practitioners in their efforts to model human learning, this paper provides a comprehensive study of learning curves and their applicability. Based on an extensive review of the literature, empirical data on learning is collected, which is then used to evaluate a selection of popular learning curves. With the help of meta-tags (see Section 3.4 for a detailed description and definition of meta-tags) on the general setup and purpose of the datasets contained in our sample, we compare the performance of different learning curves and derive propositions as to which learning curves perform best in which application. The results of this paper may assist researchers and practitioners to select learning curves for future studies.

The remainder of this paper is structured as follows. Section 2 first discusses popular learning curve models. The results of a comprehensive literature review on empirical studies of learning are presented in Section 3. Section 4 analyses the goodness of fit of the learning curves presented in Section 2 on the empirical datasets obtained in Section 3. Section 5 summarizes the findings of the study and concludes the paper.

2. Learning curve models

This section presents a selection of learning curves that have frequently been studied in the past. Learning curves presented below have been selected based on their popularity, which was evaluated with the help of the reviews cited above, and to make sure that a broad range of learning curves is used for data fitting. We note that the literature contains many more models of learning that are not discussed in this paper, and refer the reader to the reviews that were cited above.

2.1. Log-linear models

2.1.1. Wright's model (WLC)

A seminal paper on learning curves is the one of Wright (1936), who showed that the average unit production costs in airplane assembly reduced as a function of the number of airplanes produced. He suggested that this phenomenon is caused by increasing worker skill levels, fewer setups and a decreasing number of errors. Wright's learning curve has the following form:

$$y_x = y_1 \cdot x^{-b}, \tag{1}$$

where y_x is the time needed for the xth repetition of the task, y_1 is the time required for the first repetition, x the number of repetitions, and b the slope of the learning curve (learning exponent), with $0 < b < 1$. Note that Wright's learning curve (and other log-linear models) can be used to model both reductions in time or in cost as a result of learning.

2.1.2. Plateau model (PM)

The Plateau model is similar to the one proposed by Wright, with the difference that a constant C is added to the model to take into account that a minimum time exists for performing a task that is independent of the learning effect (Baloff, 1971). The plateau learning curve is formulated as

$$y_x = C + y_1 \cdot x^{-b} \tag{2}$$

2.1.3. Stanford B model (SBM)

The Stanford B learning curve extends Wright's learning curve by considering prior experience (Carlson, 1973). The model assumes that an equivalent of $B > 0$ cycles has been processed earlier, either because the same or a similar task has been performed, which led to the acquisition of knowledge. The Stanford B model is formulated as follows:

$$y_x = y_1 \cdot (x + B)^{-b} \tag{3}$$

2.1.4. De Jong's model (DJM)

De Jong (1957) assumed that there is an incompressible component in each process where no learning and thus no productivity improvement occurs, and thus extended Wright's (1936) learning curve by adding a factor of incompressibility to the model. De Jong's learning curve has the following form:

$$y_x = y_1 \cdot (M + (1 - M) \cdot x^{-b}) \tag{4}$$

The factor M ($1 \geq M \geq 0$) depends, for example, on the degree of automatization of the production process. If a production process is partially automatized, we may assume that no learning takes place in automated tasks. Thus, the fewer manual tasks a production process contains, the earlier learning may be assumed to plateau, which is expressed by a higher value for M.

2.1.5. S-curve model (SCM)

The S-curve model combines the characteristics of the Stanford B model and De Jong's model. The name derives from the fact that this learning curve is s-shaped when plotted in logarithmic scale. It can be expressed as follows (Nembhard and Uzumeri, 2000):

$$y_x = y_1 \cdot (M + (1 - M) \cdot (x + B)^{-b}) \tag{5}$$

2.1.6. Jaber–Glock learning curve model (JGLCM)

The *JGLCM* extends the dual-phase learning curve introduced by Dar-El et al. (1995) and accounts for the fact that in most industrial tasks, both cognitive and motor learning occur (Jaber and Glock, 2013). The *JGLCM* consists of two components, cognitive and motor, where p represents the share of both types of learning. It is modeled as follows:

$$y_x = p \cdot y_1 \cdot x^{-b_c} + (1 - p) \cdot y_1 \cdot x^{-b_m}, \tag{6}$$

where b_c is the learning exponent for cognitive learning and b_m the one for motor learning.

2.2. Exponential models

Exponential learning curve models contain more parameters than log-linear models to account for empirically observed characteristics (such as worker's prior experience) and to include more information on the learning process. Exponential models that are fitted to empirical data in this paper are discussed briefly in this section.

2.2.1. 2-Parameter exponential model (2PE)

The 2-parameter exponential model of Mazur and Hastie (1978) is formulated as

$$y = k \cdot \left(1 - e^{-(t/R)}\right), \tag{7}$$

where y represents the number of units produced since the start of production, t the time that has elapsed since the start of production (or the time that has elapsed during training), k the prediction of maximum performance after an infinite amount of training ($k \geq 0$), and R the learning rate parameter which measures how fast an individual learns.

2.2.2. 3-Parameter exponential model (3PE)

In the 3-parameter exponential model, a parameter p is added to the 2-parameter exponential model ($p \geq 0$), which accounts for the worker's prior experience. It is measured in the same units as t in the 2-parameter exponential model, such as time or amount of training (Anzanello and Fogliatto, 2011):

$$y = k \cdot \left(1 - e^{-((t+p)/R)}\right) \tag{8}$$

2.2.3. Group learning curve (GLC)

While most learning curves in the literature have been applied on the individual level, only a few studies proposed special learning curves for groups. In this paper, we use the *GLC* developed by Glock and Jaber (2014) and fit this curve to empirically observed group learning data. In the *GLC*, $Z(T)$ is the total number of units produced in the group in T units of time, and it is of the following form:

$$Z(T) = \sum_{i=1}^{n} Y_i(T) + \sum_{i=1}^{n} \sum_{j=1}^{n} X_{i,j}(T), \tag{9}$$

with $Y_i(T)$ representing the number of units individual i would produce by time T, $Y_i(T) = \frac{y_{1,i}}{1+b_i} T^{1+b_i}$ with b_i representing the learning rate of individual i, and $X_{i,j}(T)$ being the number of units individual i produces in T units of time due to the knowledge received from individual j in a knowledge transfer.

2.3. Hyperbolic models

Similar to exponential models, learning curves can also be expressed in hyperbolic form. Two popular hyperbolic learning curves that have often been discussed in the literature will be described in the following (Mazur and Hastie, 1978) sections.

2.3.1. 2-Parameter hyperbolic model (2PH)

The 2-parameter hyperbolic model is described as

$$y = k \cdot \left(\frac{t}{t+R}\right), \tag{10}$$

where y is the number of items produced in t units of time (or the amount of training), R the learning rate and k the maximum output level (i.e. the asymptote for learning).

2.3.2. 3-Parameter hyperbolic model (3PH)

A parameter p is added to the 2-parameter hyperbolic model to account for prior experience of the workforce ($p \geq 0$). The model is formulated as:

$$y = k \cdot \left(\frac{t+p}{t+p+R}\right) \tag{11}$$

Hyperbolic learning curves can be applied to a wide range of different scenarios, such as the reduction of defective items as a result of learning, for example. In this case, t would represent the number of conforming items and R the number of defective ones. The fraction of defective items, y, decreases as the output quantity increases (Anzanello and Fogliatto, 2011). Hyperbolic models can express both increasing ($R > 0$) and decreasing ($R < 0$) productivity (Nembhard and Uzumeri, 2000).

2.4. Summary of learning curve models

The learning curve models considered in this paper are summarized in Table 1 (cf. Anzanello and Fogliatto, 2011). Table 1 also shows which learning curve expresses learning as a reduction in time/cost or as an increase in output/productivity, and illustrates their typical profile.

3. The literature review approach

Literature reviews are conducted to structure a certain research area, to identify popular research streams within this area and/or to synthesize research findings (Rhoades, 2011; Hochrein and Glock, 2012). Literature reviews, in general, can be differentiated into narrative reviews, systematic reviews and meta-analyses (Rhoades, 2011). While narrative literature reviews typically summarize the literature relevant to a certain topic without using a systematic evaluation procedure, systematic literature reviews require a clearly defined, rigorous and reliable approach (Tranfield et al., 2003; Rhoades, 2011). Meta-analyses, in turn, have a primary focus on empirical studies and extract, summarize, consolidate and synthesize data from several works in a replicable way in order to create consensus and reliability on a certain research field and to promote future research opportunities (Cooper, 1998, 2010). A meta-analysis usually requires a systematic search of the literature, wherefore a methodology for systematically reviewing and presenting the literature is required (Rhoades, 2011; Tranfield et al., 2003).

This paper presents the results of a meta-analysis of empirical learning curve data. First, a systematic search of the literature on empirical learning curve studies was performed. Subsequently, datasets were extracted from the relevant literature, and meta-tags were assigned to each of the datasets. Meta-tags represent characteristics of the studies the data was taken from, e.g. individual learning observed in a field study (see Section 3.4). In a third step, the datasets were analyzed to gain insights into the comparative performance of different learning curves and to evaluate which learning curve should be used in which application. This section explains the methodology used in this paper, which is based on the works of Cooper (1998, 2010), Tranfield et al. (2003), Rhoades (2011), Glock and Hochrein (2011), Hochrein and Glock (2012), Glock et al. (2014), and Hochrein et al. (In Press).

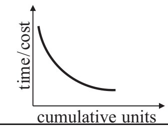
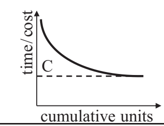
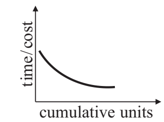
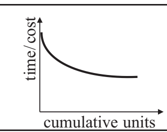
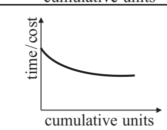
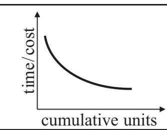
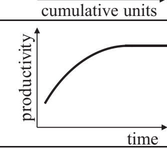
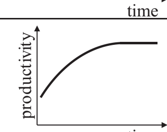
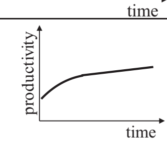
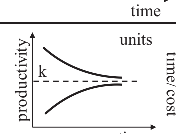
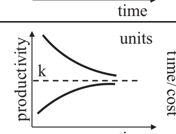
3.1. Problem formulation

This paper conducts a meta-analysis on empirical learning curve data. First, a systematic search of the literature is performed to identify articles that studied human learning in industrial settings (or tasks related to production) and that contain empirically observed data. Subsequently, data on learning contained in the sample is extracted and categorized with the help of meta-tags, which describe the setting where the empirical data was collected as well as the type of learning considered in the study. Learning curve models that have often been used in the literature are then fitted to the empirical data and evaluated with regard to goodness of fit. Goodness of fit and meta-tags are evaluated to study which learning curve model is most suitable to describe which data best, and in which setting (or for which task, type of learning) which learning curve should be used for modelling learning. In the next section, the conceptual taxonomy of this paper is explained.

3.2. Taxonomy

The purpose and content of this meta-review can be classified according to the following taxonomy that is based on Cooper (2010) and Hochrein and Glock (2012):

Table 1
Summary of learning curve models.

Nr	Learning-Curve	Model	Usually measures learning as	Typical learning curve
1	*WLC*	$y_x = y_1 \cdot x^{-b}$	*reduction in time/cost*	
2	*PM*	$y_x = C + y_1 \cdot x^{-b}$	*reduction in time/cost*	
3	*SBM*	$y_x = y_1 \cdot (x + B)^{-b}$	*reduction in time/cost*	
4	*DJM*	$y_x = y_1 \cdot (M + (1 - \mathrm{M}) \cdot x^{-b})$	*reduction in time/cost*	
5	*SCM*	$y_x = y_1 \cdot (M + (1 - \mathrm{M}) \cdot (x + B)^{-b})$	*reduction in time/cost*	
6	*JGLCM*	$y_x = p \cdot y_1 \cdot x^{-b_c} + (1 - p) \cdot y_1 \cdot x^{-b_m}$	*reduction in time/cost*	
7	*2PE*	$y = k \cdot \left(1 - e^{-\left(\frac{t}{R}\right)}\right)$	*increase in output/productivity*	
8	*3PE*	$y = k \cdot \left(1 - e^{-\left(\frac{(t+p)}{R}\right)}\right)$	*increase in output/productivity*	
9	*GLC*	$Z(T) = \sum_{i=1}^{n} Y_i(T) + \sum_{i=1}^{n} \sum_{j=1}^{n} X_{i,j}(T)$	*increase in output/productivity*	
10	*2PH*	$y = k \cdot \left(\frac{t}{t + R}\right)$	*reduction in time/cost increase in output/productivity*	
11	*3PH*	$y = k \cdot \left(\frac{t + p}{t + p + R}\right)$	*reduction in time/cost increase in output/productivity*	

1. Focus: The focus is on works that evaluate empirical data on learning in industrial production processes.
2. Goal: The goal is to consolidate research findings on learning and to gain insights into the comparative performance of learning curves and to evaluate which learning curve should be used in which application.
3. Perspective: The study adopts a neutral perspective and tries to analyze datasets in a balanced way.
4. Coverage: The study aims to provide an exhaustive overview of the literature by using several established methodologies to search the literature. Its aim is to include all existing works that meet the selection criteria defined in this study.
5. Organization: The study adopts a conceptual organization and groups works that are found in the literature search into a set of initially defined content categories. These content categories are based on a typology of empirical learning curve studies (see Section 3.4).
6. Audience: The audience of the study are general and specialized scholars as well as practitioners.

Please cite this article as: Grosse, E.H., et al., Production economics and the learning curve: A meta-analysis. International Journal of Production Economics (2015), http://dx.doi.org/10.1016/j.ijpe.2015.06.021

3.3. Methodology

Providing a review of empirical learning curve data makes it necessary to consider research from several disciplines. Although we focus on human learning in production-related settings (such as assembly line work), works from other disciplines, such as psychology or ergonomics, may be relevant as well. In a first step, we identified two databases that contain research from the disciplines management, engineering, psychology and ergonomics, namely Business Source Premier and Scopus. Both databases were used to search for relevant articles.

In a second step, to identify relevant articles, we defined two lists of keywords, where one list (A) contained keywords related to learning curves, and the second one (B) keywords related to empirical data collection methods (see Table 2). Subsequently, each keyword from list A was combined with each keyword from list B to generate the final keyword list. These keywords were then used to search the two databases for relevant works.

Works were considered relevant if they contained at least one keyword from the final keyword list either in their title, abstract or list of keywords. The language of the papers was limited to English for any year of publication. In addition, we focused only on works that appeared in peer-reviewed academic journals. Our search led to 777 initial hits in Scopus and 380 initial hits in Business Source Premier. After deleting duplicates, 938 papers were added to our initial sample. In the next step, the abstracts of the papers contained in the initial sample were read by all authors of this paper to verify their relevance. In this step, 653 papers were excluded from further analysis, which led to 285 relevant papers.

Subsequently, the reference lists of the papers contained in our sample were checked to find additional works that could be relevant to our study and that had not been selected before (this procedure is often referred to as a ‘snowball approach’, see Hochrein and Glock, 2012). The snowball approach led to 35 additional papers, which led to a working sample of 320 papers.

In the last step, all pre-selected papers were read completely, and papers that did not meet the selection criteria were excluded. To be included in the final sample, works had to show the following characteristics:

- The focus of the paper had to be on human learning, and only works that analyze empirically collected data on learning were considered.
- The empirical data presented in the papers had to be collected in industrial settings, or it had to be on tasks that are closely related to manual tasks that occur in production, such as assembly or disassembly tasks.
- Papers that investigated only one of the two keyword domains and extraneous works, such as experiments with children or animals, were excluded.

After this step, we arrived at a final sample of 44 works and 115 datasets. The results of this search procedure were presented at an international conference to experts in the field of production economics. The ensuing discussion confirmed that all substantial articles had been considered in the sample.

Table 2
Keywords used in the systematic search of two databases.

Learning curve keywords (A)	Empirical data keywords (B)
Experience curve	Data
Learning curve	Study
Progress curve	Empirical
Improvement curve	Survey
Progress function	Test
Forgetting	Experiment
	Investigation

3.4. Data extraction process and data meta-tags

Table 3 suggest a typology of empirical learning curve studies, which differentiates works according to the following meta-tags: type of study, task duration, individual(s) that learn(s), and type of learning. The attributes shown in Table 3 were deduced from the works contained in the sample, and they can be explained as follows:

- Type of study: field study (data was collected in a practical setting) vs. laboratory study (data was collected in a test setting or in class).
- Task duration (interruption): Continuous learning (no interruption) vs. interrupted learning (interruption, which could imply that forgetting occurs).
- Individual(s) that learn(s): individual learning (one person learns) vs. group learning (several persons learn) vs. organizational learning (an entire institution learns).
- Task type: Motor learning vs. cognitive learning vs. motor and cognitive learning (there are tasks that are either motor or cognitive and tasks that require/induce both types of learning).

All works contained in our sample were categorized according to the attributes shown in Table 3. Afterwards, data on learning was extracted from the papers contained in our sample. If a paper contained numerical data, it was directly transferred to our database. If a paper displayed data in graphs, in turn, it was necessary to transform these graphs into numerical values. For this purpose, we used the *MatLab 2013* plugin *Grabit*. Note that transferring graphically represented data into numerical values is likely to produce minor errors in the data. To make sure that the transformation does not bias the results of our study, the data was extracted multiple times and by two researchers, and the average value obtained for each data point was used. In case papers were considered relevant, but did not display the original data points but rather modified, consolidated or fitted results, the respective corresponding authors were contacted by email and asked to provide the original data. However, in most cases, data could not be shared due to data protection and non-disclosure agreements or author retirement.

3.5. Descriptive analysis

In this section, the papers contained in our final sample are grouped according to the meta-tag “individual(s) that learn(s)”.

Table 3
Typology and related meta-tags of empirical learning curve studies.

Type of study	Field study		Laboratory study	
Task duration	Continuous learning		Interrupted learning	
Individual(s) that learn(s)	Individual learning	Group learning		Organizational learning
Type of learning	Motor learning	Cognitive learning		Motor/cognitive learning

ARTICLE IN PRESS

Table 4
Papers that contain data for individual learning.

Study type	Field						Laboratory						Datasets
Task duration	Continued			Interrupted			Continued			Interrupted			
Learning type	m	c	m/c	m	c	m/c	m	c	m/c	m	c	m/c	
Almgren (1999)		x											1
Anderson et al. (2009)									x				3
Bailey (1989)										x			1
Barlow (1928)								x					1
Bevis et al. (1970)			x						x				4
Braden (1924)							x						1
Davies (1945)							x						2
Easley (1933)							x						1
Ehrlich (1943)							x						2
Eyring et al. (1993)								x					2
Gray (1918)							x						12
Grosse and Glock (2013)			x										3
Hamade et al. (2005)									x				1
Hamade et al. (2009)									x				1
Kellogg (1946)									x				1
Leslie and Adams (1973)									x				3
Levy (1965)	x												3
Nakamura et al. (1996)								x					1
Nembhard and Osothsilp (2001)						x							1
Nembhard and Uzumeri (2000)			x										4
Perrin (1919)							x						2
Reid and Mirka (2007)							x						2
Rodrigue et al. (2005)							x						3
Rohmert and Schlaich (1966)							x						4
Towill (1977)	x												2
Towill (1990)		x											1
Towill et al. (1989)				x									1
Uzumeri and Nembhard (1998)			x										7

Table 5
Papers that contain data for group learning.

Study type	Field						Laboratory						Datasets
Task duration	Continued			Interrupted			Continued			Interrupted			
Learning type	m	c	m/c	m	c	m/c	m	c	m/c	m	c	m/c	
Argote et al. (1995)									x				1
Baloff and Becker (1968)								x					11
Guetzkow and Simon (1955)									x				1
Leavitt (1951)								x					1
Shure et al. (1962)								x					1

Table 6
Papers that contain data for organizational learning.

Study type	Field						Laboratory						Datasets
Task duration	Continued			Interrupted			Continued			Interrupted			
Learning type	m	c	m/c	m	c	m/c	m	c	m/c	m	c	m/c	
Adler and Clark (1991)	x												2
Chambers and Johnston (2000)			x										1
Fessia et al. (2007)	x												4
Foster and Adam (1996)			x										1
Franceschini and Galetto (2004)	x												2
Hinze and Olbina (2009)				x									1
Huntley (2003)		x											2
Jarkas and Horner (2011)	x												1
Junginger et al. (2006)			x										1
Lapré (2011)			x										8
Macher and Mowery (2003)			x										7

The results for individual learning are summarized in Table 4, for group learning in Table 5, and for organizational learning in Table 6. The column "datasets" shows how many different datasets on learning could be extracted from each paper.

In the case of field studies, the type of industry, which formalizes the sector the data was collected in, was extracted as well (note that this is not possible for laboratory studies). Table 7 categorizes field studies according to industry type.

4. Curve fitting

4.1. Methodology

The empirical data extracted from our sample was categorized with the help of meta-tags (see Section 3.5), and then the learning curves described in Section 2 were fitted to the data with *MS Excel Solver* and a regression add-in of *MatLab 2013*. Subsequently, the quality of the fit was analyzed. To fit learning curves to the empirical data, the parameters of the learning curves were estimated within their respective ranges by using the least squares method, where the relevant parameter ranges where taken from the original works that proposed the learning curves considered in this paper. This method has frequently been used in the past to fit learning curves to empirical data (e.g., Bevis et al., 1970; Hackett, 1983; Hinze and Olbina, 2009; Grosse and Glock, 2013; Glock and Jaber, 2014). We formulate the objective function for fitting learning curves to the empirical data as follows:

$$min \sum_{i=1}^{n} (\hat{y}_i - y_i)^2 \tag{12}$$

where y_i is the data point observed in period i taken from the empirical data (e.g., cost or production time), n is the number of

ARTICLE IN PRESS

Table 7
Field studies industry type.

Adler and Clark (1991)	Electronic equipment manufacturing
Bevis et al. (1970)	
Fessia et al. (2007)	
Macher and Mowery (2003)	
Nembhard and Osothsilp (2001)	
Towill (1977)	
Almgren (1999)	Automotive parts manufacturing
Foster and Adam (1996)	
Franceschini and Galetto (2004)	
Hinze and Olbina (2009)	Construction
Jarkas and Horner (2011)	
Bevis et al. (1970)	Tobacco manufacture
Chambers and Johnston (2000)	Aviation service
Lapré (2011)	
Grosse and Glock (2013)	Household products manufacturer warehouse
Huntley (2003)	Software service
Junginger et al. (2006)	Energy generation
Levy (1965)	Commercial printing
Nembhard and Uzumeri (2000)	Textile manufacturing
Towill (1990)	Electrical inspection
Towill et al. (1989)	Machine-building
Uzumeri and Nembhard (1998)	Manufacturing

data points/periods, and $\hat{y}_i$ the ith estimated value which is calculated with the help of the learning curve under study. When fitting learning curves to empirical data, we made sure that each learning curve was only fitted to data it was developed for. Thus, the *GLC*, for example, was only fitted to group learning data, which is the application this learning curve has been developed for.

To assess the quality of the fit, we calculated the coefficient of determination, R^2, as follows (cf. Dodge, 2008):

$$R^2 = \sum_{i=1}^{n} \frac{(\hat{y}_i - \bar{y})^2}{\sum_{i=1}^{n} (y_i - \bar{y})^2} \tag{13}$$

with $\bar{y}$ representing the mean value of the observed data points. Based on the R^2-value of each fit, a ranking was established to evaluate which learning curve performed best (the higher the R^2-value, the better the performance and thus the rank). For this purpose, we first derived the average R^2 value for each learning curve over all fitted datasets as an approximate benchmark. Second, as the average R^2 could be influenced by outliers, we tracked how often a learning curve led to a particular result, for example the best fit, the second-best fit and so on, to evaluate the performance of each learning curve in relation to the number of datasets. To establish a detailed performance ranking based on how often each learning curve led to the best fit, we used the following index σ_l (cf. for a similar approach in dynamic lot sizing Zoller and Robrade, 1988):

$$\sigma_l = \frac{\sum_{i=1}^{I} n_{r_{i,l}} . r_{i,l}}{N} \tag{14}$$

where $r_{1,l}$ represents the best fit ($r_{1,l}=1$), $r_{2,l}$ represents the second best fit ($r_{2,l}=2$) and so on with $r_{I,l}$ representing the worst fit ($r_{I,l}=I$; $r_I=8$ for group A and $r_I=4$ group B) for learning curve l. N is the number of fitted datasets and $n_{r_{i,l}}$ the number of times learning curve l yielded fit r_i ($i=1, 2,\ldots, I$). Thus, the smaller σ_l, the better the rank of the learning curve.

4.2. Results

This section presents the results of the curve fitting procedure, where the goodness of fit (R^2) is used as evaluation criterion first. Note that the datasets were split up into two groups: Group A contains data that expresses learning as a reduction in time or cost (Table 8), while group B expresses learning as an increase in output/productivity (Table 9). To avoid that transforming the original data leads to biases, we fitted log-linear ($0<b<1$) and hyperbolic models ($R<0$) to datasets contained in group A (decreasing time or cost) and exponential and hyperbolic models ($R>0$) to datasets contained in group B (increasing output). The last rows in Tables 8 and 9 (shaded) contain the average R^2 for each learning curve over all datasets under study.

Data in group B shown in Table 9 contains group learning data (for group A data, no group learning data could be found in the systematic evaluation of the literature). To make sure that the average R^2 value can be compared for each learning curve, group B datasets were split up into two sets, namely ‘all data’ and ‘group learning data’ (note that ‘group learning data’ is displayed in italics). The first average R^2 value in the second last row (shaded) in Table 9 considers all datasets in group B, while the average R^2 value in the last row (shaded) considers only group learning data. The *GLC* was only fitted to group learning data.

To evaluate the performance of the learning curves under study, we established rankings for each learning curve as described in Eqs. (13) and (14). First, the learning curves were ranked according to their average R^2 value over all datasets in each group. Table 10 summarizes the average performance of each learning curve according to their average R^2.

To develop the second ranking based on Eq. (14), we tracked the performance of each learning curve with respect to goodness of fit, i.e. we evaluated how often (#) each learning curve led to a certain rank of fit, e.g. the best fit, the second best fit etc. (for a similar approach applied to the dynamic lot sizing problem, see, for instance, Berry, 1972; Blackburn and Millen, 1985; Zoller and Robrade, 1988). This approach makes it possible to evaluate the performance of each learning curve in relation to the number of datasets. Tables 11 (for group A) and 12 (for group B) present the count of how often a learning curve achieved a specific rank and the related percentage share per learning curve. For example, fitting the *S*-curve (*SCM*) to the empirical data obtained the best coefficient of determination (rank 1) for 31% (17 times) of the tested datasets in group A (see Table 11). In turn, the 2-parameter hyperbolic model (*2PH*) obtained the last rank (rank 8) for 78% (43 times) of the datasets (see Table 11). For group B (Table 12), the 3-parameter exponential model (*3PE*) obtained the best coefficient of determination (rank 1) for 42% (25 times) of the tested datasets in group B, for example.

Please cite this article as: Grosse, E.H., et al., Production economics and the learning curve: A meta-analysis. International Journal of Production Economics (2015), http://dx.doi.org/10.1016/j.ijpe.2015.06.021

ARTICLE IN PRESS

Table 8
Results (R^2 values) of the curve fitting procedure (Group A).

Dataset	*WLC*	*PM*	*SBM*	*DJM*	*SCM*	*2PH*	*3PH*	*JGLCM*
Almgren (1999)	0.7261	0.7345	0.8713	0.7485	0.8429	0.3491	0.8537	0.7485
Anderson et al. (2009) A	0.9879	0.9886	0.9879	0.9940	0.9940	0.5313	0.9935	0.9879
Anderson et al. (2009) B	0.9672	0.9733	0.9672	0.9733	0.9725	0.8182	0.9655	0.9757
Anderson et al. (2009) C	0.9662	0.9774	0.9662	0.9774	0.9770	0.4616	0.9718	0.9774
Bailey (1989)	0.8262	0.8487	0.8262	0.8475	0.8474	0.7781	0.8475	0.8262
Barlow (1928)	0.8532	0.8067	0.8067	0.8067	0.8067	0.5888	0.8025	0.9310
Chambers and Johnston (2000)	0.7595	0.7595	0.7661	0.7595	0.7704	0.4906	0.7626	0.7595
Ehrlich (1943) B	0.9972	0.9970	0.9691	0.9970	0.8567	0.8439	0.9978	0.9972
Fessia et al. (2007) A	0.9795	0.9796	0.9781	0.9838	0.9827	0.9920	0.9932	0.9289
Fessia et al. (2007) B	0.9779	0.8994	0.9927	0.9075	0.9972	0.2527	0.9970	0.9779
Fessia et al. (2007) C	0.9760	0.9754	0.9733	0.9754	0.9754	0.4875	0.9399	0.9733
Fessia et al. (2007) D	0.9600	0.9974	0.9601	0.9974	0.9974	0.7562	0.9943	0.9976
Franceschini and Galetto (2004) A	0.8305	0.8305	0.8625	0.8305	0.8690	0.4144	0.8360	0.8305
Franceschini and Galetto (2004) B	0.8426	0.8426	0.8426	0.8426	0.8426	0.7646	0.8957	0.8426
Grosse and Glock (2013) A	0.0249	0.0249	0.0485	0.0249	0.0273	0.0038	0.0037	0.0249
Grosse and Glock (2013) B	0.1324	0.1324	0.1555	0.1324	0.1422	0.0301	0.0841	0.1324
Grosse and Glock (2013) C	0.0044	0.0044	0.0019	0.0048	0.0074	0.0383	0.0692	0.0044
Hamade et al. (2005)	0.9839	0.9863	0.9863	0.9863	0.9863	0.9807	0.9977	0.9839
Hamade et al. (2009)	0.9995	0.9995	0.9996	0.9995	0.9995	0.9168	0.9994	0.9995
Hinze and Olbina (2009)	0.4864	0.7679	0.4864	0.7679	0.7679	0.8656	0.8687	0.7679
Jarkas and Horner (2011)	0.6770	0.6766	0.6911	0.4256	0.7746	0.5018	0.1922	0.6770
Junginger et al. (2006)	0.9335	0.9335	0.9834	0.9335	0.9565	0.3079	0.9777	0.9335
Kellogg (1946)	0.9204	0.9402	0.9204	0.9422	0.9444	0.6843	0.9423	0.9422
Lapré (2011) A	0.4070	0.3904	0.5178	0.3904	0.4034	0.1171	0.4922	0.3904
Lapré (2011) B	0.5735	0.5623	0.5679	0.5623	0.5623	0.3385	0.5696	0.5623
Lapré (2011) C	0.5741	0.5564	0.6641	0.5564	0.5635	0.2177	0.6381	0.5564
Lapré (2011) D	0.5573	0.5411	0.5937	0.5411	0.5411	0.2746	0.5881	0.5411
Lapré (2011) E	0.2451	0.2538	0.2672	0.2538	0.2538	0.1371	0.2729	0.2538
Lapré (2011) F	0.7368	0.7381	0.7380	0.7381	0.7381	0.4597	0.7304	0.7332
Lapré (2011) G	0.1253	0.3015	0.1763	0.3015	0.3015	0.3230	0.3207	0.3015
Lapré (2011) H	0.3292	0.3377	0.4770	0.3377	0.3697	0.0757	0.4562	0.3377
Leslie and Adams (1973) A	0.3943	0.3943	0.4766	0.3943	0.4399	0.2069	0.4399	0.3943
Leslie and Adams (1973) B	0.4042	0.4042	0.4410	0.4042	0.5529	0.2154	0.4434	0.4042
Leslie and Adams (1973) C	0.8844	0.8844	0.8993	0.8993	0.9124	0.7989	0.9031	0.8993
Macher and Mowery (2003) A	0.7281	0.7281	0.8246	0.7281	0.8665	0.4107	0.8184	0.7281
Macher and Mowery (2003) B	0.7618	0.7618	0.8459	0.7618	0.8664	0.4472	0.8462	0.7618
Macher and Mowery (2003) C	0.1336	0.1336	0.2664	0.1336	0.8561	0.0149	0.2392	0.1336
Macher and Mowery (2003) D	0.4058	0.5490	0.2323	0.5490	0.5490	0.5628	0.5759	0.5490
Macher and Mowery (2003) E	0.8106	0.8106	0.8452	0.8106	0.9011	0.5751	0.8462	0.8106
Macher and Mowery (2003) F	0.7470	0.7470	0.8992	0.7470	0.8625	0.3751	0.8867	0.7470
Macher and Mowery (2003) G	0.8675	0.8675	0.8914	0.8675	0.8853	0.5627	0.8927	0.8675
Nakamura et al. (1996)	0.9400	0.9399	0.9809	0.9400	0.9863	0.6540	0.8075	0.9400
Nembhard and Osothsilp (2001)	0.2913	0.3078	0.2929	0.2784	0.3013	0.1191	0.0003	0.2989
Nembhard and Uzumeri (2000) A	0.4559	0.4561	0.4557	0.4942	0.4756	0.5211	0.5261	0.4559
Perrin (1919) A	0.5014	0.6028	0.5014	0.6028	0.6028	0.6424	0.6509	0.6028
Perrin (1919) B	0.5788	0.6685	0.5794	0.6685	0.6685	0.7906	0.7895	0.6685
Reid and Mirka (2007) A	0.9896	0.9943	0.9896	0.9943	0.9943	0.8973	0.9150	0.9897
Reid and Mirka (2007) B	0.9611	0.9611	0.9735	0.9611	0.9633	0.7560	0.7624	0.9611
Rodrigue et al. (2005) A	0.9658	0.9657	0.9680	0.9658	0.9803	0.6383	0.9646	0.9658
Rodrigue et al. (2005) B	0.9409	0.9411	0.9409	0.9409	0.9636	0.6885	0.9285	0.9409
Rodrigue et al. (2005) C	0.8930	0.8930	0.9117	0.8930	0.9558	0.4863	0.9227	0.8930
Rohmert and Schlaich (1966) A	0.9627	0.9622	0.9783	0.9622	0.9866	0.5133	0.1502	0.9622
Rohmert and Schlaich (1966) B	0.8976	0.9156	0.9685	0.9156	0.9321	0.4756	0.9597	0.9415
Rohmert and Schlaich (1966) C	0.9203	0.9203	0.9836	0.9203	0.9826	0.5992	0.9830	0.9203
Rohmert and Schlaich (1966) D	0.9274	0.9274	0.9274	0.9274	0.9274	0.6275	0.9061	0.9274
Average R^2	**0.7041**	**0.7181**	**0.7294**	**0.7145**	**0.7543**	**0.4978**	**0.7129**	**0.7211**

Based on Tables 11 and 12, the results for the second ranking according to Eq. (14) can be derived. The results are summarized in Table 13.

4.3. Discussion

As can be seen in Tables 10 and 13 for group A, the *S*-curve generated the best results on average in both rankings σ_l and R^2. The three-parameter hyperbolic model is ranked second in σ_l, although it led to poor results for 7% of the datasets, which is why it is only ranked sixth in R^2. It is worth noting that Wright's simple learning curve performed reasonably well (although leading to a lower R^2), obtaining rank 1 to 3 in σ_l for 22% of the datasets studied. This result confirms prior studies that claimed that the Wright learning curve, although being of simple structure, is able to approximate empirically observed learning quite well (Nembhard and Osothsilp, 2001; Jaber, 2013).

The Stanford B model led to good results in both rankings, with rank 2 in σ_l and rank 3 in R^2, although it led to the poorest fit for 9% of the data. The Plateau and the De Jong models obtained rank 4 and 5 in both rankings, leading to close σ_l and R^2 values.

The Jaber–Glock learning curve was ranked third in R^2, but only seventh in σ_l, which indicates that this learning curve led to few very good fits and predominantly poorer results for the tested datasets (see Table 11).

When fitting the De Jong, the Plateau, the Stanford B and the Jaber–Glock learning curve model to some datasets, we observed

Please cite this article as: Grosse, E.H., et al., Production economics and the learning curve: A meta-analysis. International Journal of Production Economics (2015), http://dx.doi.org/10.1016/j.ijpe.2015.06.021

Table 9
Results (R^2 values) of the curve fitting procedure (Group B).

Dataset	*2PE*	*3PE*	*2PH*	*3PH*	*GLC*
Adler and Clark (1991) A	0.3229	0.7490	0.9986	0.9524	
Adler and Clark (1991) B	0.2506	0.6775	0.2512	0.6749	
Argote et al. (1995)	*0.9513*	*0.4816*	*0.4958*	*0.9927*	*0.6965*
Baloff and Becker (1968) A	*0.8590*	*0.8797*	*0.8590*	*0.8773*	*0.8620*
Baloff and Becker (1968) B	*0.8188*	*0.8100*	*0.8100*	*0.8102*	*0.8277*
Baloff and Becker (1968) C	*0.7878*	*0.7672*	*0.7878*	*0.7641*	*0.7438*
Baloff and Becker (1968) D	*0.0053*	*0.1119*	*0.0007*	*0.0036*	*0.0099*
Baloff and Becker (1968) E	*0.5528*	*0.5176*	*0.5407*	*0.5176*	*0.4576*
Baloff and Becker (1968) F	*0.4413*	*0.4482*	*0.4435*	*0.4435*	*0.4513*
Baloff and Becker (1968) G	*0.7361*	*0.7342*	*0.7349*	*0.7328*	*0.7202*
Baloff and Becker (1968) H	*0.7252*	*0.7265*	*0.7252*	*0.7804*	*0.8171*
Baloff and Becker (1968) I	*0.8304*	*0.8334*	*0.8303*	*0.8327*	*0.8367*
Baloff and Becker (1968) J	*0.8763*	*0.8705*	*0.8763*	*0.8648*	*0.9076*
Baloff and Becker (1968) K	*0.5993*	*0.6035*	*0.5916*	*0.6085*	*0.6821*
Bevis et al. (1970) A	0.9497	0.9859	0.9594	0.9792	
Bevis et al. (1970) B	0.9755	0.9956	0.9897	0.9943	
Bevis et al. (1970) C	0.6982	0.3527	0.7285	0.0000	
Bevis et al. (1970) D	0.7414	0.8786	0.4833	0.8246	
Braden (1924)	0.8568	0.9474	0.4935	0.0696	
Davies (1945) A	0.9520	0.4743	0.4912	0.9601	
Davies (1945) B	0.9844	0.0000	0.5014	0.0000	
Easley (1933)	0.9838	0.9983	0.4253	0.0000	
Ehrlich (1943) A	0.8388	0.4964	0.5055	0.9969	
Eyring et al. (1993) A	0.9160	0.0000	0.4996	0.2015	
Eyring et al. (1993) B	0.7471	0.9480	0.5005	0.9228	
Foster and Adam (1996)	0.3902	0.4901	0.3706	0.3235	
Gray (1918) A	0.0000	0.0000	0.0139	0.1606	
Gray (1918) B	0.3362	0.4903	0.4837	0.0000	
Gray (1918) C	0.8855	0.4841	0.4803	0.9372	
Gray (1918) D	0.6898	0.4800	0.4532	0.0000	
Gray (1918) E	0.9040	0.4874	0.5194	0.7779	
Gray (1918) F	0.8134	0.0000	0.4924	0.0000	
Gray (1918) G	0.6721	0.9083	0.4801	0.9089	
Gray (1918) H	0.7555	0.0000	0.5056	0.7461	
Gray (1918) I	0.6125	0.0000	0.5078	0.0000	
Gray (1918) J	0.8698	0.0000	0.4896	0.000	
Gray (1918) K	0.5170	0.5498	0.5359	0.5870	
Gray (1918) L	0.8646	0.4798	0.4831	0.000	
Guetzkow and Simon (1955)	*0.9684*	*0.9776*	*0.9685*	*0.9780*	*0.9855*
Huntley (2003) A	0.0642	0.8017	0.5815	0.7398	
Huntley (2003) B	0.7876	0.9864	0.8354	0.9816	
Leavitt (1951)	*0.8663*	*0.8694*	*0.8653*	*0.8734*	*0.7888*
Levy (1965) A	0.4445	0.9522	0.7798	0.3851	
Levy (1965) B	0.8961	0.9390	0.8962	0.0000	
Levy (1965) C	0.2382	0.9561	0.5759	0.0000	
Nembhard and Uzumeri (2000) B	0.9473	0.0000	0.9169	0.9526	
Nembhard and Uzumeri (2000) C	0.5566	0.8363	0.6679	0.8359	
Nembhard and Uzumeri (2000) D	0.0450	0.0000	0.1039	0.1215	
Shure et al. (1962)	*0.9150*	*0.3541*	*0.9148*	*0.8485*	*0.9863*
Towill (1977) A	0.1952	0.8480	0.5770	0.8475	
Towill (1977) B	0.8548	0.5976	0.4972	0.1064	
Towill (1990)	0.7212	0.9738	0.3856	0.0000	
Towill et al. (1989)	0.5854	0.6057	0.3692	0.0235	
Uzumeri and Nembhard (1998) A	0.9350	0.9400	0.9243	0.9430	
Uzumeri and Nembhard (1998) B	0.9128	0.9420	0.9365	0.9427	
Uzumeri and Nembhard (1998) C	0.9461	0.9919	0.9572	0.9900	
Uzumeri and Nembhard (1998) D	0.6791	0.6494	0.6759	0.6567	
Uzumeri and Nembhard (1998) E	0.6150	0.7939	0.7747	0.8273	
Uzumeri and Nembhard (1998) F	0.6913	0.8095	0.7560	0.8029	
Uzumeri and Nembhard (1998) G	0.2704	0.2703	0.4041	0.5608	
Average R^2	**0.6808**	**0.6125**	**0.6117**	**0.5610**	
Average R^2 (group learning data)	***0.7289***	***0.6657***	***0.6963***	***0.7285***	***0.7182***

convergence of the model parameters such that the results of Wright's model were obtained. Worst results were, on average, obtained when fitting the 2-parameter hyperbolic model to the data, which resulted in the poorest fit for 78% of the datasets and an average R^2 of 0.4978.

For group B, we observed a clear ranking for non-group learning data, as the exponential models outperformed the hyperbolic models in both rankings (see Tables 11 and 13). The 3-parameter exponential model obtained the best fit for 42% of the datasets, followed by the 2-parameter exponential model in the σ_l ranking. However, the 2-parameter exponential model led to better R^2 values on average.

The three-parameter hyperbolic model is ranked third with 27% best fits, and the 2-parameter hyperbolic model led to the poorest fit for 20% of the datasets in the σ_l ranking. As in the case of the exponential models, the 2-parameter hyperbolic model obtained better R^2 values than the 3-parameter hyperbolic model.

Table 10
Ranking of the fitted learning curves according to average R^2.

Group A			Group B			Group B (*group learning data*)		
Rank	***Learning curve***	**Average R^2**	**Rank**	***Learning curve***	**Average R^2**	**Rank**	***Learning curve***	**Average R^2**
1	*SCM*	0.7543	1	*2PE*	0.6808	1	*2PE*	0.7289
2	*SBM*	0.7294	2	*3PE*	0.6125	2	*3PH*	0.7285
3	*JGLCM*	0.7211	3	*2PH*	0.6117	3	*GLC*	0.7182
4	*PM*	0.7181	4	*3PH*	0.5610	4	*2PH*	0.6963
5	*DJM*	0.7145				5	*3PE*	0.6657
6	*3PH*	0.7129						
7	*WLC*	0.7041						
8	*2PH*	0.4978						

Table 11
Count of the ranking positions achieved by the fitted learning curves (Group A).

Learning curve/rank of fit	***WLC***		***PM***		***SBM***		***DJM***		***SCM***		***JGLCM***		***2PH***		***3PH***	
	#	%	#	%	#	%	#	%	#	%	#	%	#	%	#	%
1	2	*4*	3	*5*	14	*25*	2	*4*	17	*31*	4	*7*	2	*4*	11	*20*
2	1	*2*	5	*9*	11	*20*	1	*2*	8	*15*	2	*4*	6	*11*	21	*38*
3	9	*16*	8	*15*	9	*16*	11	*20*	8	*15*	5	*9*	0	*0*	5	*9*
4	10	*18*	10	*18*	3	*5*	17	*31*	10	*18*	5	*9*	0	*0*	0	*0*
5	10	*18*	15	*27*	2	*4*	12	*22*	7	*13*	7	*13*	0	*0*	2	*4*
6	7	*13*	11	*20*	6	*11*	8	*15*	4	*7*	18	*33*	1	*2*	0	*0*
7	14	*25*	3	*5*	5	*9*	4	*7*	1	*2*	13	*24*	3	*5*	12	*22*
8	2	*4*	0	*0*	5	*9*	0	*0*	0	*0*	1	*2*	43	*78*	4	*7*

Table 12
Count of the ranking positions achieved by the fitted learning curves (Group B).

Learning curve/rank of fit	***2PE***		***3PE***		***2PH***		***3PH***		***GLC*** (fitted only to group learning data)	
	#	%	#	%	#	%	#	%	#	%
1	17	*28*	25	*42*	2	*3*	16	*27*	8	*53*
2	12	*20*	10	*17*	21	*35*	17	*28*	1	*7*
3	17	*28*	12	*20*	25	*42*	6	*10*	2	*13*
4	14	*23*	13	*22*	12	*20*	21	*35*	0	*0*
5									4	*27*

Table 13
Ranking of fitted learning curves according to σ_l.

Group A			Group B			Group B (group learning data)		
Rank	***l***	σ_l	**Rank**	***l***	σ_l	**Rank**	***l***	σ_l
1	*SCM*	3.0	1	*3PE*	2.22	1	GLC	1.43
2	*3PH*	3.5	2	*2PE*	2.47	2	*3PE*	2.67
3	*SBM*	3.6	3	*3PH*	2.53	3	*2PE*	3.07
4	*PM*	4.3	4	*2PH*	2.78	4	*3PH*	3.13
5	*DJM*	4.4				5	*2PH*	3.73
6	*WLC*	5.0						
7	*JGLCM*	5.2						
8	*2PH*	7.0						

As for group learning data, the group learning curve obtained the best results for 53% of the group learning data in the σ_l ranking, while the 2-parameter hyperbolic model led to the poorest fit for 40% of the datasets. For the group learning curve, it is interesting to note that it was outperformed by the *2PE* and the *3PH* if the average R^2 is used as evaluation criterion. Although the group learning curve led to the best fits for more than half of the datasets, it seems that it performed poorly for some datasets, which led to a relative low average R^2 value.

Future research could concentrate on investigating whether there are specific reasons why the group learning curve underperforms in certain scenarios. The results on group learning, in general, confirm that learning curve models that were developed for special data characteristic – group learning in this case – have a tendency to outperform general learning curves when this type of data is studied.

In a final step, we studied patterns in the performance of the learning curves with respect to the meta-tags that were assigned to each dataset (see Tables 4–6). This permits us to derive propositions about which learning curve should be used for which application. The results can be summarized as follows:

- For *individual continuous learning*, we could observe a dominance of the 3-parameter exponential model for field data and motor learning, and a dominance of the 3-parameter hyperbolic model for motor/cognitive learning.
- For *individual learning under laboratory settings*, the Jaber–Glock learning curve and the *S*-curve performed well for motor/cognitive learning, and the *S*-curve, the Stanford B and the 3-parameter hyperbolic model for motor learning.
- For *group learning in laboratory settings*, the group learning curve approximated both motor/cognitive and cognitive data quite well.
- As far as *organizational learning and field data* are concerned, the 3-parameter hyperbolic model yielded good results for motor tasks, and the *S*- and Stanford-*B* curve fitted motor/cognitive data best.
- The sample contained only a few datasets on *interrupted learning*, so no clear proposition could be derived; however, for the datasets under study, the plateau learning curve obtained good results for individual interrupted learning, both motor and motor/cognitive.
- In addition, the results were ambiguous for *cognitive learning* and no clear tendency could be observed; however, for all learning types, the 3-parameter exponential model seemed to perform quite well with cognitive learning data.

Table 14 summarizes the propositions that could be derived from our analysis.

5. Conclusion

The intention of this paper was to survey works that study learning in production processes and related tasks, to extract empirically observed learning data from the literature and to analyze how well popular learning curves describe the datasets. In addition, the paper aimed to present a comparative performance analysis of different learning curves and to derive recommendations about which learning curve should be used to

Please cite this article as: Grosse, E.H., et al., Production economics and the learning curve: A meta-analysis. International Journal of Production Economics (2015), http://dx.doi.org/10.1016/j.ijpe.2015.06.021

Table 14
Propositions on which learning curve obtained good results according to meta-tags.

Learning curve	Meta-tags
SBM	• Organizational, field, continuous, motor/cognitive • Individual, laboratory, continuous, motor
SCM	• Organizational, field, continuous, motor/cognitive • Individual, laboratory, continuous, motor, motor/cognitive
3PH	• Organizational, field, continuous, motor, motor/cognitive • Individual, field, continuous, motor, motor/cognitive • Individual, laboratory, continuous, motor
JGLCM	• Individual, laboratory, continuous, motor/cognitive
3PE	• Individual, field, continuous, motor
GLC	• Group, laboratory, continuous, motor, motor/cognitive

approximate learning in different practical settings. For this purpose, we conducted a meta-analysis of empirical learning curve data by systematically searching for relevant literature from different research disciplines. First, we presented a typology for categorizing data on human learning. A selection of well-known learning curve models was then fitted to the empirical datasets and analyzed with regard to goodness of fit and data characteristics. Our analysis showed that some learning curves performed well on average, such as the *S*-curve, the 3-parameter hyperbolic as well as the 3-parameter exponential models. Wright's basic learning model also led to acceptable results in predicting performance and thus can be recommended as a basic and simple tool for modelling learning. Other learning curve models, especially the 2-parameter hyperbolic model, led to poor results. We also observed a relative performance advantage of learning curves that were developed for specific tasks and data characteristics, such as the group learning curve for group learning data and the Jaber–Glock learning curve for motor/cognitive task data.

This paper extended the existing literature on learning curve models by giving a broad overview of works that contain empirically observed learning data and by comprehensively evaluating the performance of different learning curve models to predict empirically observed learning data. The paper supports researchers in their effort to find relevant papers that contain empirical data, and it facilitates selecting an appropriate learning curve for their modeling efforts.

This paper also has limitations. First, only articles published in peer-reviewed journals were considered in this study. Including other works, such as book chapters or working papers, in the sample as well could result in more available datasets and more insights on learning curve performances. In addition, only a few studies that provided data on interrupted or cognitive learning were contained in our sample, which is why no clear recommendations on the use of learning curves could be derived for this type of data. Furthermore, assigning the selected papers to the developed typology of empirical learning curve studies involved some amount of judgment, as some papers did not state clearly under which conditions the data had been collected. Future studies on empirical learning should, as far as possible, present more relevant data in tables or figures to make reusing the data possible.

Future research could build on this study and extend the search methodology to find more relevant datasets. In addition, this study showed that there are many works that empirically study learning at the individual level, but that there are only few studies on group or organizational learning. Future research could concentrate on the transfer of knowledge within groups and organizations to gain further insights into how groups and organizations learn, and how this can be modelled mathematically. This also implies a need for a stronger focus on learning in specific industries.

Acknowledgments

This paper is an extended version of the conference paper "Grosse, Glock, Müller: Production Economics and the Learning Curve: A Meta-Analysis" that was presented at the Eighteenth International Working Seminar on Production Economics in Innsbruck in 2014. The authors are grateful to the special issue editor and the anonymous reviewers for their constructive comments that helped to improve an earlier version of this paper. They also appreciate the assistance of Pia Behnisch in collecting the empirical data analyzed in this paper.

References

Adler, P.S., Clark, K.B., 1991. Behind the learning curve: a sketch of the learning process. Manage. Sci. 37 (3), 267–281.
Almgren, H., 1999. Towards a framework for analyzing efficiency during start-up: an empirical investigation of a Swedish auto manufacturer. Int. J. Prod. Econ. 60-61 (1), 79–86.
Anderson, A.M., Mirka, G.A., Joines, S.M., Kaber, D.B., 2009. Analysis of alternative keyboards using learning curves. Hum. Factors: J. Hum. Factors Ergon. Soc. 51 (1), 35–45.
Anzanello, M.J., Fogliatto, F.S., 2011. Learning curve models and applications: literature review and research directions. Int. J. Ind. Ergon. 41 (5), 573–583.
Argote, L., Insko, C.A., Yovetich, N., Romero, A.A., 1995. Group learning curves: the effects of turnover and task complexity on group performance. J. Appl. Soc. Psychol. 25 (6), 512–529.
Bailey, C.D., 1989. Forgetting and the learning curve: a laboratory study. Manage. Sci. 35 (3), 340–352.
Baloff, N., Becker, S.W., 1968. A model of group adaptation to problem-solving tasks. Organ. Behav. Hum. Perform. 3 (3), 217–238.
Baloff, N., 1971. Extensions of the learning curve—some empirical results. Oper. Res. Q. 22 (4), 329–340.
Barlow, M.C., 1928. A learning curve equation as fitted to learning records. Psychol. Rev. 35 (2), 142–160.
Berry, W.L., 1972. Lot sizing procedures for requirements planning systems: a framework for analysis. Prod. Inventory Manage. 13 (1), 19–34.
Bevis, F.W., Finniear, C., Towill, D.R., 1970. Prediction of operator performance during learning of repetitive tasks. Int. J. Prod. Res. 8 (4), 293–305.
Blackburn, J.D., Millen, R.A., 1985. A methodology for predicting single-stage lot-sizing performance: analysis and experiments. J. Oper. Manage. 5 (4), 433–448.
Braden, S.R., 1924. An extensive experiment in motor learning and re-learning. J. Educ. Psychol. 15 (5), 313–315.
Carlson, J.G., 1973. Cubic learning curve: precision tool for labor estimating. Manuf. Eng. Manage. 71 (5), 22–25.
Chambers, S., Johnston, R., 2000. Experience curves in services: macro and micro level approaches. Int. J. Oper. Prod. Manage. 20 (7), 842–859.
Cooper, H.M., 1998. Synthesizing Research: A Guide for Literature Reviews, third ed.. Sage Publications, Thousand Oaks.
Cooper, H.M., 2010. Research Synthesis and Meta-Analysis: A Step-By-Step Approach, fourth ed.. Sage Publications, Thousand Oaks.
Dar-El, E.M., Ayas, K., Gilad, I., 1995. A dual-phase model for the individual learning process in industrial tasks. IIE Trans. 27 (3), 265–271.
Davies, D.R., 1945. The effect of tuition upon the process of learning a complex motor skill. J. Educ. Psychol. 36 (6), 352–365.
De Jong, J.R., 1957. The effects of increasing skill on cycle time and its consequences for time standards. Ergonomics 1 (1), 51–60.
Dodge, Y., 2008. The Concise Encyclopedia of Statistics. Springer Science & Business Media, LLC.
Easley, H., 1933. An analysis of the curves of learning and forgetting code material. J. Educ. Psychol. 24 (8), 634–640.
Ehrlich, G., 1943. A method of constructing learning curves for a motor skill involving total body speed and accuracy. J. Appl. Psychol. 27 (6), 494–503.
Eyring, J.D., Johnson, D.S., Francis, D.J., 1993. A cross-level units-of-analysis approach to individual differences in skill acquisition. J. Appl. Psychol. 78 (5), 805–814.
Fessia, P., Regis, F., Rossi, L., 2007. Industrial learning curves: series production of the LHC main superconducting dipoles. IEEE Trans. Appl. Supercond. 17 (2), 1101–1104.
Fogliatto, F.S., Anzanello, M.J., 2011. Learning curves: the state of the art and research directions. In: Jaber, M.Y. (Ed.), Learning Curves: Theory, Models, and Applications. CRC Press, Boca Raton, FL, pp. 3–22.
Foster, S.T., Adam, E.E., 1996. Examining the impact of speed of quality improvement on quality-related costs. Decis. Sci. 27 (4), 623–646.

Franceschini, F., Galetto, M., 2004. An empirical investigation of learning curve composition laws for quality improvement in complex manufacturing plants. J. Manuf. Technol. Manage. 15 (7), 687–699.
Glock, C.H., Hochrein, S., 2011. Purchasing organization and design: a literature review. BuR—Bus. Res. 4 (2), 149–191.
Glock, C.H., 2012. Single sourcing versus dual sourcing under conditions of learning. Comput. Ind. Eng. 62 (1), 318–328.
Glock, C.H., Grosse, E.H., Ries, J.M., 2014. The lot sizing problem: a tertiary study. Int. J. Prod. Econ. 155, 39–51.
Glock, C.H., Jaber, M.Y., 2014. A group learning curve model with and without worker turnover. J. Modell. Manage. 9 (2), 179–199.
Gray, C.T., 1918. A comparison of two types of learning by means of a substitution test. J. Educ. Psychol. 9 (3), 143–158.
Grosse, E.H., Glock, C.H., 2013. An experimental investigation of learning effects in order picking systems. J. Manuf. Technol. Manage. 24 (6), 850–872.
Grosse, E.H., Glock, C.H., 2014. The effect of worker learning on manual order picking processes. Int. J. Prod. Econ. , http://dx.doi.org/10.1016/j.ijpe.2014.12.018.
Grosse, E.H., Glock, C.H., Jaber, M.Y., 2013. The effect of worker learning and forgetting on storage reassignment decisions in order picking systems. Comput. Ind. Eng. 66 (4), 653–662.
Guetzkow, H., Simon, H.A., 1955. The impact of certain communication nets upon organization and performance in task-oriented groups. Manage. Sci. 1 (3-4), 233–250.
Hackett, E.A., 1983. Application of a set of learning curve models to repetitive tasks. Radio Electron. Eng. 53 (1), 25–32.
Hamade, R.F., Artail, H.A., Jaber, M.Y., 2005. Learning theory as applied to mechanical CAD training of novices. Int. J. Hum. Comput. Stud. 19 (3), 305–322.
Hamade, R.F., Jaber, M.Y., Sikström, S., 2009. Analyzing CAD competence with univariate and multivariate learning curve models. Comput. Ind. Eng. 56 (4), 1510–1518.
Hinze, J., Olbina, S., 2009. Empirical analysis of the learning curve principle in prestressed concrete piles. J. Constr. Eng. Manage. 135 (5), 425–431.
Hochrein, S., Glock, C.H., 2012. Systematic literature reviews in purchasing and supply management research: a tertiary study. Int. J. Integr. Supply Manage. 7 (4), 215–245.
Hochrein, S., Glock, C.H., Bogaschewsky, R., Heider, M., In Press. Literature reviews in supply chain management: a tertiary study. Manage. Rev. Q., http://dx.doi.org/10.1007/s11301-015-0113-4.
Huntley, C.L., 2003. Organizational learning in open-source software projects: an analysis of debugging data. IEEE Trans. Eng. Manage. 50 (4), 485–493.
Jaber, M.Y., Goyal, S.K., Imran, M., 2008. Economic production quantity model for items with imperfect quality subject to learning effects. Int. J. Prod. Econ. 115 (1), 143–150.
Jaber, M.Y., Bonney, M., Moualek, I., 2009. Lot sizing with learning, forgetting and entropy cost. Int. J. Prod. Econ. 118 (1), 19–25.
Jaber, M.Y., 2013. Learning and forgetting models and their applications. In: Badiru, A.B. (Ed.), Handbook of Industrial and Systems Engineering, second ed. CRC Press, Boca Raton, FL.
Jaber, M.Y., Glock, C.H., 2013. A learning curve for tasks with cognitive and motor elements. Comput. Ind. Eng. 64 (3), 866–871.
Jarkas, A., Horner, M., 2011. Revisiting the applicability of learning curve theory to formwork labour productivity. Constr. Manage. Econ. 29 (5), 483–493.
Junginger, M., de Visser, E., Hjort-Gregersen, K., Koornneef, J., Raven, R., Faaij, A., Turkenburg, W., 2006. Technological learning in bioenergy systems. Energy Policy 34 (18), 4024–4041.
Kazemi, N., Shekarian, E., CárdenasBarrón, L.E., Olugu, E.U., 2015. Incorporating human learning into a fuzzy EOQ inventory model with backorders. Comp. Ind. Eng. 87, 540–542.
Kellogg, W.N., 1946. The learning curve for flying an airplane. J. Appl. Psychol. 30 (5), 435–441.
Kull, T.J., Boyer, K., Calantone, R., 2007. Last-mile supply chain efficiency: an analysis of learning curves in online ordering. Int. J. Oper. Prod. Manage. 27 (4), 409–434.
Lapré, M.A., 2011. Reducing customer dissatisfaction: how important is learning to reduce service failure? Prod. Oper. Manage. 20 (4), 491–507.
Leavitt, H.J., 1951. Some effects of certain communication patterns on group performance. J. Abnorm. Soc. Psychol. 46 (1), 38–50.
Leslie, J.H., Adams, S.K., 1973. Programmed safety through programmed learning. Hum. Factors: J. Hum. Factors Ergon. Soc. 15 (3), 223–236.
Levy, F.K., 1965. Adaptation in the production process. Manage. Sci. 11 (6), B136–B154.
Macher, J.T., Mowery, D.C., 2003. "Managing" learning by doing: an empirical study in semiconductor manufacturing. J. Prod. Innovat. Manage. 20 (5), 391–410.
Mazur, J.E., Hastie, R., 1978. Learning as accumulation: a reexamination of the learning curve. Psychol. Bull. 85 (6), 1256–1274.
Nakamura, Y., Yasuyuki, K., Mitsunaga, Y., 1996. Proposal of a skill level index based on user's thinking time. Electron. Commun. Jpn. 79 (8), 416–423.
Nembhard, D.A., Osothsilp, N., 2001. An empirical comparison of forgetting models. IEEE Trans. Eng. Manage. 48 (3), 283–291.
Nembhard, D.A., Uzumeri, M.V., 2000. An individual-based description of learning within an organization. IEEE Trans. Eng. Manage. 47 (3), 370–378.
Nolan, 2012. Ubiquitous IT: the case of the Boeing 787 and implications for strategic IT research. J. Strategic Inf. Syst. 21 (2), 91–102.
Perrin, F.A.C., 1919. The learning curves of the analogies and the mirror reading test. Psychol. Rev. 26 (1), 42–62.
Reid, S.A., Mirka, G.A., 2007. Learning curve analysis of a patient lift-assist device. Appl. Ergon. 38 (6), 765–771.
Rhoades, E.A., 2011. Literature reviews. Volta Rev. 111 (3), 353–368.
Rodrigue, K.M., Kennedy, K.M., Naftali, R., 2005. Aging and longitudinal change in perceptual-motor skill acquisition in healthy adults. J. Gerontol.: Psychol. Sci. 60B (4), P174–P181.
Rohmert, W., Schlaich, K., 1966. Learning of complex manual tasks. Int. J. Prod. Res. 5 (2), 137–145.
Shafer, S.M., Nembhard, D.A., Uzumeri, M.V., 2001. The effects of worker learning, forgetting, and heterogeneity on assembly line productivity. Manage. Sci. 47 (12), 1639–1653.
Shure, G.H., Rogers, M.S., Larsen, I.M., Tassone, J., 1962. Group planning and task effectiveness. Sociometry 25 (3), 263–282.
Smunt, T.L., Watts, C.A., 2003. Improving operations planning with learning curves: overcoming the pitfalls of 'messy' shop floor data. J. Oper. Manage. 21 (1), 93–107.
Towill, D.R., 1977. Exponential smoothing of learning curve data. Int. J. Prod. Res. 15 (1), 1–15.
Towill, D.R., Davies, A., Naim, M.M., 1989. The dynamics of capacity planning for flexible manufacturing system startup. Eng. Costs Prod. Econ. 17 (1), 55–64.
Towill, D.R., 1990. Forecasting learning curves. Int. J. Forecasting 6 (1), 25–38.
Tranfield, D., Denyer, D., Smart, P., 2003. Towards a methodology for developing evidence-informed management knowledge by means of systematic review. Br. J. Manage. 14 (3), 207–222.
Uzumeri, M.V., Nembhard, D.A., 1998. population of learners: a new way to measure organizational learning. J. Oper. Manage. 16 (5), 515–528.
Wright, T.P., 1936. Factors affecting the cost of airplanes. J. Aeronaut. Sci. 3 (4), 122–128.
Yelle, L.E., 1979. The learning curve: historical review and comprehensive survey. Decis. Sci. 10 (2), 302–328.
Zhong, H., Hall, R.W., Dessouky, M., 2007. Territory planning and vehicle dispatching with driver learning. Transp. Sci. 41 (1), 74–89.
Zoller, K., Robrade, A., 1988. Dynamic lot sizing techniques: survey and comparison. J. Oper. Manage. 7 (4), 125–148.

Please cite this article as: Grosse, E.H., et al., Production economics and the learning curve: A meta-analysis. International Journal of Production Economics (2015), http://dx.doi.org/10.1016/j.ijpe.2015.06.021