COVER SHEET

This is the author-version of article published as:

Falta, Michael and Wolff, Rodney C (2004) Recent developments of statistical approaches in cost accounting: a review. *International Statistical Review* **72.**

Accessed from [**http://eprints.qut.edu.au**](http://eprints.qut.edu.au)

Recent Developments of Statistical Approaches in Cost Accounting: a Review

Michael Falta and Rodney Wolff

Department of Mathematics, The University of Queensland,
Qld 4072 Brisbane, Australia

Summary

We review and simultaneously introduce a convenient statistical concept for the mathematical representation of the Statistical Activity Cost Theory (SACT) introduced by Willett (1987 and 1988). Further, we discuss, and present a critique of, a variety of statistical models with respect to long debated accounting problems, such as the allocation of joint costs and depreciation. We finally propose that taking the effort to combine those models results in a novel *statistical* accounting system and this is discussed by means of the so-called virtual firm. As it has been shown that any statistical model discussed here outperforms associated deterministic counterparts, this review presents promising outcomes and useful perspectives for the accounting profession.

Key words: Accounting systems; Conditioning; Depreciation; Markov chains; Statistical Activity Cost Theory (SACT); Virtual firm;

1 Introduction

Surprisingly, statistical methods and approaches in (cost) accounting can be only found in the academic accounting literature whereas deterministic models prevail in accounting practice. The purpose of research, as we understand it, is to apply its results in practice, where appropriate. There then follows the legitimate question why this has not been done in this particular area of research since most of the articles reviewed in the following sections report about a superiority of statistical over comparable deterministic methods or results which arise out of them.

At first, we want to outline the contrast taken here between an accounting system and a model in accounting and then to answer the above question.

An *accounting system* supports management in decision-making by predicting results in the planning stage and reporting on the performance upon realized transactions in the control stage of a decision cycle (Datar, Foster & Horngren, 2000). The same authors report that Activity Based Costing (ABC) is the system predominantly used by companies nowadays. This has been confirmed in personal communication with Value Focused Consulting, a leading representative of the accounting business based in Brisbane, Australia. ABC was introduced by Cooper (1988) and Cooper & Kaplan (1988). In ABC, costs of individual activities are calculated based on cause and effect relations and then assigned to cost objects (products or services) on the basis of the underlying activities and cost drivers. In contrast to the deterministic ABC systems, we review the Statistical Activity Cost Theory (SACT) introduced by Willett (1987, 1988 and 1991a), which in fact, is the only existing approach having the potential to be used as the underlying framework for a statistical accounting system due to its axiomatic basis. Note that the definition of a SACT-activity differs from the interpretation of activities found in ABC systems. In SACT, activities are input and ouput functions which are the basic construct of separated production and cost structures according clearly formulated mathematical axioms. The relevance of the separation of accounting entities into cost and production structures with respect to accounting is thoroughly discussed in Willett (1998). The strength of SACT applied to accounting measurement is best summarized in the work of Hillier (1998):

> ... In it [the framework of SACT], indirect accounting measurements such as earnings are viewed as statistical constructs. Adjustments such as depreciation and the like are interpreted as devices potentially varying the statistical properties of earnings. On this view, the properties of earnings in representing underlying economic events may be evaluated according to statistical criteria. Questions regarding forms of distributions, time series behaviour, and efficacy as estimators and predictors of economic parameters and realisations respectively arise naturally in this context ...

and later more specifically:

> ... the statistical interpretation means interpreting the elementary events and central terms of the theory [SACT], transaction costs and production relations, as outcomes of stochastic processes.

So far, the stochastic processes applied to activities have included the Poisson process for activity starting points, exponentially distributed activity durations and Normally distributed activity costs. Using these processes, most of the SACT literature reports on the smoothing of a series of accounting earnings numbers. To do so, it is necessary to determine a decision parameter, e.g., long-term profitability ({revenues-expenses}/time), and then to define a criterion, e.g., least-squares error criterion, with respect to which the decision parameter is analysed. With this respect, the SACT literature divides into three areas: firstly, pure statistical analysis (Lane & Willett, 1997 and 1999), both of which are further developments of Willett (1991b); secondly, results from simulations (Willet, 1991b; Hillier, 1998; and Hillier & Willett, 2001); and third, various papers which conceptualize the idea of a virtual firm (Gibbins & Willett, 1997; Gibbins, Hillier & McCrae, 1998), a model of real companies. We will critically report about the statistical contents of the first and third area and postpone the discussion of the second area to a future timepoint because the results presented there are of limited information content, resulting from simplistic assumptions.

By (statistical) *models in accounting* we refer to approaches which deal with a particular accounting problem such as the allocation of joint costs (Brief, 1967; Brief & Owen, 1968, 1969, 1970 and 1973), the depreciation problem or the allocation of costs of intangible assets (e.g. Hodgson, Okunev & Willett, 1993). We review the Brief and Owen paper series, and in the case of depreciation the following approaches undertaken. Ijiri & Kaplan (1969 and 1970) derive some properties by conditioning and Zannetos (1963), Pollard & Tippett (1994) and Pollard, Rhys & Tippett (1994) use Markov chains in connection with depreciation. Based on the latter approach, two recent articles by Rhys (2000) and Rhys & Tippett (2002) present further advances in statistical depreciation.

About depreciation, there has been, and still is, a considerable debate over the methodology for determining the depreciation adjustment, among others due to its close relation to profit and loss numbers, and accounting earnings numbers in financial statements. Therefore, many SACT articles use depreciation to demonstrate, e.g., its smoothing capabilities of earnings time series (e.g. Hillier & McCrae, 1998; Lane & Willett, 1999) or to make some statements about its non-arbitrary character (Lane & Willett, 1997). The review article by Butler, Rhys & Tippett (1994) surveys the theory of depreciation and reports, besides the above Markovian and SACT approach, on depreciation based on discounted cash flows. The content of that article includes numerical examples which accompany the different depreciation models and clearly targets the accounting readership.

We return to the question why there has been no implementation in practice of either a statistical model nor a statistical accounting system and argue in the following way.

Firstly, the statistical content of accounting study course books such as Datar, Foster & Horngren (2000) is rather limited or non-existent. Accounting is still regarded as a pure economics discipline although articles like Mattessich (1984) and Brief (1990) suggest a more scientific (statistical) approach to the matter. Secondly, SACT is a 'young' theory and as we will show below, results following from the fundamental propositions are still scarce. There is also no evidence in the literature of an attempt to implement an accounting systems based on SACT, whereas in the case of deterministic accounting systems (in which we count all ABC systems), there are many in both the literature (e.g. Bassett, 1987; Cokins, 1996) and, obviously, in practice.

Our review reports about statistical models for the generation or allocation of accounting data but does not include articles where time series analysis is applied to reported accounting earnings numbers (see, e.g., Manegold (1981); Ball & Watts (1972)).

The remainder of this article is structured as follows. In Section 2 we review SACT and implicitly introduce a convenient and uniform statistical concept of its formulae and terms. That is, this document presents a correct representation of deterministic and random variables and in the latter case, of their realisations. We also introduce a uniform notation of variables and indices because throughout the SACT literature already exist different notations of identical terms. This happened in other disciplines like thermodynamics and unnecessarily confuses readers and students. We also want to point the reader to Section 2.4 where we comment on the reviewed results and also suggest future research topics in the SACT framework. In Section 3 we present a critical review of statistical models in accounting and include again some modifications in their statistical representation. This section consists of three paper series: firstly by Brief and Owen; secondly by Ijiri and Kaplan; and third by Pollard, Rhys, Tippett and Zannetos. Section 4 then goes on to discuss how SACT and the models presented in Section 3 might be combined and build virtual firms. The final Section 5 includes the summary discussion.

2 Statistical Activity Cost Theory (SACT)

2.1 The SACT Framework

In general, SACT accounting earnings numbers Y_i are calculated with respect to accounting periods i, $(T_{i-1}, T_i]$, $i = 1, 2, ...$, with a constant duration $T_i - T_{i-1}$ for all i. For an accounting point of view, it is a matter of taste whether or not to include or exclude T_{i-1} or T_i in the interval, as long as consecutive intervals disjointly cover the entire time line. This definition also equates to the general statistical convention of defining events.

An accounting earnings number Y_i for period i equals the sum over k accounting earnings numbers Y_{ik}, k finite, each of which denotes the earnings figure of a particular activity class and thus,

$$Y_i = \sum_{all\ k} Y_{ik}. \tag{2.1}$$

An *activity class* groups activities into equivalence classes defined by identical distributions for starting times, duration times and costs, that is, identical technologies and cost structures.

In the SACT literature, there are two different models for Y_{ik}.

In the first model, the accounting earnings function Y_{ik} for period i of a particular activity class k is represented as the sum of an undepreciated contribution U_{ik} and the depreciation function L_{ik}, thus,

$$Y_{ik} = U_{ik} + L_{ik}. \tag{2.2}$$

The most popular examples in accounting practice for a depreciation function α are the linear or 'straight-line' depreciation ($\alpha = (C - Sel)/D$) and the reducing balance depreciation ($\alpha = 1 - (Sel/C)^{1/D}$), where Sel is the selling or residual value of an asset, C the purchase price and D the duration of use. However, this might not be true in real situations and Sel can additionally vary over time. Friberg (1973) discussed deterministic *versus* probabilistic depreciation and accounted in the latter approach for a varying selling value Sel. For further depreciation methods and a detailed analysis and overview the interested reader is referred to Reynolds (1962).

In detail, Y_{ik} consists of direct measurements U_{ik}, which are the contribution margin (Willett, 1991b) or realised revenues and expenses (Lane & Willett, 1997) of finished activities, and derived or indirect measurements L_{ik}, which appear only in fixed asset expense accounts and represent the total change in depreciation (Lane & Willett, 1997).

The decomposition shown in (2.2) can be expanded to

$$Y_i = \sum_{all\ k} \sum_{r=0}^{i} \sum_{j=1}^{Q_k(r,i)} C_{jrk}(i) + \sum_{all\ k} \sum_{r=0}^{i} \alpha_k(r) \Big\{ M_k(r,i) - M_k(r,i-1) \Big\}, \quad (2.3)$$

where the left term of the sum is the undepreciated part summed over $j = 1, ..., Q_k(r,i)$ completed random activity costs $C_{jrk}(i)$ in period i, $r = 0, ..., i$. The quantity $Q_k(r,i)$ is the random number of finishing activities in period i and starting point r periods before i for a particular activity class k. In the right term, which represents the depreciation adjustment, the quantity $M_k(r,i)$ is the random number of activities purchased r periods before i and still held at i, such that the difference $M_k(r,i) - M_k(r,i-1)$ is the change in the number of activities whose accumulated depreciation charge on activities finishing in the period is $\alpha_k(r)$.

Note that we have replaced the infinite sum in Equation 1 (Lane & Willett, 1999), which implicitly assumes an infinite life time of the firm, by a sum over $r = 1, ..., i$ periods.

Note that in (2.3) a clear distinction between short-term and long-term activities is difficult to assess. In the second, and more intuitive model to specify the representation of Y_i, the application of any term is based on the activity classification discussed in the introduction. Thus, three different classifications are applicable corresponding to the single terms in (2.4). They are in order: short-term or undepreciable long-term activities, or their respective activity costs; unfinished long-term activities, or the depreciable amount; and finished long-term activities, or the accumulated write-off. Note that the separation between short-term and undepreciable long-term activities in the first term below is made by assigning either activity category to particular activity classes. The accounting period (AP) is usually assumed to have a constant duration $AP = T_i - T_{i-1}$, as mentioned above. Thus,

$$\begin{aligned} Y_i \;\; = \;\; & \sum_{all\ k} \sum_{j=1}^{N_k(i)} C_{jk}(i) + \sum_{all\ k} \sum_{j=1}^{N_k^a(i)} \alpha_k \Big\{ C_{jk}(i), D_{jk}(i) \Big\} \\ & - \sum_{all\ k} \sum_{j=1}^{N_k^e(i)} \sum_{h=1}^{\lceil M_{jk}(i) \rceil} \alpha_k \Big\{ C_{jk}(h), D_{jk}(h) \Big\}, \end{aligned} \quad (2.4)$$

where, in general, the random variable C stands for activity costs, the random variable D for activity durations and α for the depreciation function.

In the first term, activity costs $C_{jk}(i)$ are only included if they are either short-term ($D_{jk} < AP$) and finish within the relevant accounting period i, that is, for $T_{i-1} < S_{jk} + D_{jk} \leq T_i$, where S_{jk} is a random variable representing the starting point of a single activity j of class k. The summation goes from $j = 1, ..., N_k(i)$, where $N_k(i)$ is a random variable. The other activity type included in the first term are undepreciable long-term activities, that is, activities for which $D_{jk} > AP$ and which also finish within the relevant accounting period i.

In the last two terms, the depreciable long-term activities, $D_{jk} > AP$, are split into long-term assets and long-term equities which are indicated by the respective superscripts a and e in the random numbers $N_k^{a,e}(i)$ of associated activities. The second term accounts for unfinished long-term activities which are characterized by $S_{jk} < T_i < S_{jk} + D_{jk}$ and $T_i - S_{jk} \leq D_{jk}$, and the depreciation α_k is added. This depreciable amount depends on the random variables $C_{jk}(i)$ and $D_{jk}(i)$ at period i. The last term accounts for finished long-term activities characterized by $T_{i-1} < S_{jk} + D_{jk} \leq T_i$ and writes off the amount of accumulated depreciation amounts prior to period i. Thus, h goes from 1 to $M_{jk}(i)$, where $M_{jk}(i)$ is the rounded up integer value of the difference $T_{i-1} - S_{jk}$ and is, because a function of i, j and k, a random variable. The terms in the depreciation function are no longer estimates, because the summation is over the past history of $M_{jk}(i)$ periods.

The model described by (2.4) includes all the basic constructs of the activity cost that provides the building blocks for the virtual firm.

2.2 Analytical Results

Lane & Willett (1997 and 1999) based their analysis on (2.3) and only dealt with one activity class. We therefore omit the index k below and write Y_i for Y_{ik}. Then, the mean and the variance of Y_i are $\mathrm{E}(Y_i) = \mathrm{E}(U_i) + \mathrm{E}(L_i)$ and $\mathrm{Var}(Y_i) = \mathrm{Var}(U_i) + \mathrm{Var}(L_i) + 2\mathrm{Cov}(U_i, L_i)$, respectively. To achieve a reduction in the variance of the earnings function Y_i with respect to the variance of its undepreciated part U_i, Lane & Willett (1997) noted that $\mathrm{Var}(Y_i) < \mathrm{Var}(U_i)$ if and only if $\mathrm{Cov}(U_i, L_i) < 0$ and $\mathrm{Var}(L_i) < 2 \cdot |\mathrm{Cov}(U_i, L_i)|$.

The resulting optimal variance-reducing depreciation function for Y_i over undepreciated earnings U_i is given by

$$\alpha = \mu_C S(\delta)/R(\delta), \tag{2.5}$$

where δ is a fixed depreciation period, $R(\delta) \propto \mathrm{Var}(L_i)$ and $S(\delta) \propto \mathrm{Cov}(U_i, L_i)$ (see Lane & Willett (1997), p.189, for an explicit form for $R(\delta)$ and $S(\delta)$). Obviously, α differs from the traditional and simple linear or reducing balance depreciation methods used in practice. Hence, the reduction of variance for accounting earnings numbers is given by $\mathrm{Var}(Y_i) = \mathrm{Var}(U_i) - \mu_C^2 S^2(\delta)/R(\delta)$.

Later, Lane & Willett (1999) assumed that the firm operates under the condition of stationarity, such that $M(r, i)$ has a constant mean. It follows, that Y_i is an unbiased estimator of $\mathrm{E}(U_i)$, since the expected value $\mathrm{E}(L_i)$ equals zero. Because the minimization of the mean square error of Y_i depends on an optimal depreciation function, they derived the following second-order difference equation to be satisfied by $\alpha(r)$,

$$\sigma^2_{M(r+1,i)} \Delta\alpha(r+1) = \sigma^2_{M(r,i)} \Delta\alpha(r) - \{\sigma^2_{M(r,i)} - \sigma^2_{M(r+1,i)}\}\{\mu_C - \alpha(r)\}, \tag{2.6}$$

where $\Delta\alpha(r) = \alpha(r) - \alpha(r-1)$, the initial condition $\alpha(-1) = 0$, and $M(r, i)$, $Q(r, i)$ and $C_{jr}(i)$ mutually independently and identically distributed, for fixed r, with variances $\sigma^2_{M(r,i)}$, $\sigma^2_{Q(r,i)}$ and σ^2_C, respectively.

Thus, the optimal $\alpha(r)$ depends on $\sigma^2_{M(r,i)}$ of the process generating the activity starting points and its durations, and μ_C, the mean of the activity costs. Note that (2.6) and the following analysis is only valid for the special assumptions that $M(r, i)$ has a constant mean with respect to i and therefore $\alpha(r)$ does not need to be specified for particular periods i.

In personal correspondence, J. Lane explained that the independence of $M(r, i)$ and $Q(r, i)$, for fixed r, follows from the assumption of random starting times generated by the Poisson process. This is a common simplification strategy in queuing type problems. Further, for different duration times r, independence seemed less controversial as in the presented model $M(r, i)$ and $Q(r, i)$ refer to different streams of activities, starting in different intervals in the past. Lane & Willett (1999, p.9) extend the basic model and introduced dependency, that is, costs C and R are allowed to have means and variances which depend on the duration r. This is discussed in detail below.

Reparametrizing (2.6) and defining $\beta(r) = \mu_C - \alpha(r)$ yields

$$\sigma^2_{M(r+1,i)}\beta(r+1) - 2\sigma^2_{M(r,i)}\beta(r) + \sigma^2_{M(r,i)}\beta(r-1) = 0, \tag{2.7}$$

and the initial condition is $\beta(-1) = \mu_C$. On the interpretation of (2.6) Lane & Willett (1999, p.5) stated that "There is a natural interpretation of $\beta(r)$ as the *balance* between expected cost and current depreciation, and so the solution will be a variant of *declining-balance* depreciation."

The minimum variance achieved by using the optimal $\alpha(r)$ is given by

$$\text{Var}_{min}(Y_i) = \sigma^2_{Q(0,i)}\mu^2_C + \sigma^2_{M(0,i)}\mu_C\alpha(0) + \lambda_p\sigma^2_C, \tag{2.8}$$

where λ_p is the expected number of activity starting points and equal to $\mu_{all\ Q(r,i)}$, the sum of all $\mu_{Q(r,i)}$, $r = 0, ..., i$, of a particular period i. The equality of the above quantities is due to the stationarity assumption.

The proportional reduction in variance is given by

$$\frac{\text{Var}(U_i) - \text{Var}_{min}(Y_i)}{\text{Var}(U_i)} = \frac{\{\sigma^2_{all\ Q(r,i)} - \sigma^2_{Q(0,i)}\}\mu^2_C - \sigma^2_{M(0,i)}\mu_C\alpha(0)}{\lambda_p\sigma^2_C + \sigma^2_{all\ Q(r,i)}\mu^2_C}, \tag{2.9}$$

where $\sigma^2_{all\ Q(r,i)}$ equals the sum of all $\sigma^2_{Q(r,i)}$, $r = 0, ..., i$. It might be useful for the interested reader to compare Lane & Willett (1999, Ch. 4) where a worked example clarifies the analytical results and shows that the proportional reduction in variance (2.9) can be significant in size.

Furthermore, Lane & Willett investigated the impact of non-capitalized activity costs, which are written off due to the uncertainty about the completion of activities, on the optimal variance-reducing depreciation function. Hence, they defined (besides C_{rj}) activity costs R_{rj}, $j = 1, ..., M(r, i)$; $r = 0, 1, ..., i$, for those activities which continue beyond period i but nevertheless contribute to expenses in that period. Both C_{jr} and R_{jr} are assumed to depend on the duration r and to be independently distributed with means

$\mu_{C(r,i)}$ and $\mu_{R(r,i)}$ and variances $\sigma^2_{C(r,i)}$ and $\sigma^2_{R(r,i)}$, respectively. Thus, the undepreciated earnings are given by

$$U_i = \sum_{r=0}^{i} \Big(\sum_{j=1}^{M(r,i)} R_{rj} + \sum_{j=1}^{Q(r,i)} C_{rj} \Big). \tag{2.10}$$

The optimum variance reducing depreciation function $\alpha(r)$ then satisfies

$$\begin{aligned} \sigma^2_{M(r+1,i)} \Delta\alpha(r+1) \;&=\; \sigma^2_{M(r,i)} \Delta\alpha(r) - \\ &\Big([\sigma^2_{M(r,i)} - \sigma^2_{M(r+1,i)}][\mu_{C(r+1,i)} - \alpha(r)] - \\ &\sigma^2_{M(r,i)} \mu_{R(r,i)} + \sigma^2_{M(r+1,i)} \mu_{R(r+1,i)} \Big), \end{aligned} \tag{2.11}$$

where again for convenience $\alpha(-1) = 0$.

Letting the means $\mu_{C(r,i)}$ and $\mu_{R(r,i)}$ have constant values μ_C and μ_R, respectively, (2.11) simplifies to

$$\sigma^2_{M(r+1,i)} \Delta\alpha(r+1) = \sigma^2_{M(r,i)} \Delta\alpha(r) - [\sigma^2_{M(r,i)} - \sigma^2_{M(r+1,i)}][\mu_C - \mu_R - \alpha(r)], \tag{2.12}$$

and allows one to investigate the impact of ongoing costs on the optimal depreciation function $\alpha(r)$. Further, replacing μ_C by the difference of expected costs on completion and expected ongoing costs, $\mu_C - \mu_R$, yields the same form as presented in (2.6). Hence, Lane & Willett concluded that depreciation should be calculated proportional to $\mu_C - \mu_R$. In practice "μ_R could represent recurrent costs over a project's duration, while μ_C represents revenue paid on completion." It follows that μ_R and μ_C are likely to have opposite signs.

2.3 Results from Simulations

The simulation model which is applied on the results in the previous section is discussed in Lane & Willett (1997, 1999), and is based on the assumption that activity starting points, hence all the numbers $Q(r,i)$ and $M(r,i)$, are mutually independently Poisson distributed with rate λ. Moreover, $Q(r,i)$ and $M(r,i)$ have the form of a generalized Poisson shot-noise process (e.g., Parzen, 1962) whose characteristic function is known (e.g., Takacs, 1954; Lane, 1984). The activity durations are independently and exponentially distributed with mean $\gamma = 1/\theta$, $\theta > 0$.

Lane & Willett (1997) used Campbell's theorem (Parzen, 1962) to derive the mean and variance of a shot-noise process, such that for $M(r,i)$,

$$\mu_{M(r,i)} = \sigma^2_{M(r,i)} = \frac{\lambda}{\theta}(1 - e^{-\theta})e^{-r\theta}, \tag{2.13}$$

and for $Q(r,i)$,

$$\mu_{Q(r,i)} = \sigma^2_{Q(r,i)} = \frac{\lambda}{\theta} e^{-(r-1)\theta}(1 - e^{-\theta})^2. \tag{2.14}$$

Then, substituting (2.13) and (2.14) into (2.5) yields the optimal variance-reducing α for a given δ,

$$\alpha_{opt}(\delta_{fixed},\theta)=\mu_C\frac{(1-e^{-\delta\theta})(1-e^{-\theta})}{2\{1-(\delta+1)e^{-\delta\theta}+\delta e^{-(\delta+1)\theta}\}}. \qquad (2.15)$$

Note that,

$$\lim_{\delta\to\infty}(\alpha_{opt})=\frac{\mu_C}{2}(1-\exp^{-\theta}), \qquad (2.16)$$

which is the expected activity cost multiplied with the probability that the duration D is less than one period.

To optimize α and δ jointly they inserted the values of $S(\delta)$ and $R(\delta)$ into the variance-reducing term $f(\delta)=\mu_C^2 S^2(\delta)/R(\delta)$ and derived a table for values $\gamma=2,3,4,...$ and $\delta=1,2,3,...$, from which they read off the optimal δ for a given γ. The derivation of a table is necessary, since $\partial f(\delta)/\partial\delta=0$ can not be solved analytically.

In the final result of this simulation Lane & Willett compared the popular depreciation rate, $1/\gamma$, with the derived optimal depreciation rate, $h(\gamma)=\alpha_{opt}(\delta_{opt})/\mu_C$. They observed that if the expected life γ is small, the optimal depreciation rate $h(\gamma)$ underestimates the ordinary depreciation rate $1/\gamma$ by over 50% for $\gamma=2$ periods. In the case of long expected service lives, $h(\gamma)$ overestimates $1/\gamma$ by over 50% for $\gamma>15$ periods.

In Lane & Willett (1999, p.9) the above model was extended and allowed for the activity starting points to have a periodic Poisson rate $\lambda(t)=\lambda(t+j)$, $0<t\leq 1$; j integer. The activity durations were defined as independently distributed exponential random variables with mean $1/\theta$. Two cases appear to be of special interest. One assumes a constant rate $\lambda(t)=\lambda$ while the other assumes a periodic rate $\lambda(t)=\lambda(t+j)$ for any integer j, such that $M(r,i)$ does not depend on i and the difference equation (2.7) simplifies to

$$e^{-\theta}\beta(r+1)-2\beta(r)+\beta(r-1)=0, \qquad (2.17)$$

such that the solution $\beta(r)=\mu_C(1-d^{r+1})$, $d=(1-(1-e^{-\theta})^{1/2}/e^{-\theta}$ only depends on θ. This solution is a geometrically declining balance function because $\alpha(r+1)-\alpha(r)=d(\mu_C-\alpha(r))$ and $0<d<1$.

Willett (1991b, p.123) discussed the theoretical limiting behaviour of the undepreciated earnings function U_{ik} in (2.2), where U_{ik} consists of only short-term activities. He assumed that the costs $C_{jk}(i)$ and the number of activities $N_k(i)$ are mutually independently and identically distributed random variables with finite variances. For the costs we define the distribution $F(c)=\Pr\{C_{jk}(i)\leq c\}$. Furthermore, the considered short-term activities are supposed to produce large numbers of individual instances of an identical product or service (see Hillier 1998, p.75-77 for a thorough discussion).

Here, we follow Robbins (1948, p.1151) and assume that the distribution of $N_k(i)$ depends on a parameter $\zeta(i)$. Thus, the distribution of U_{ik} becomes a function of $\zeta(i)$ which may have an asymptotic expression as $\zeta(i)\to\infty$. Robbins then specified that the distribution of $N_k(i)$ for any $\zeta(i)$ is determined by the values $\omega_j=\Pr\{N_k(i)=j\}$, $j=0,1,...$, where the ω_j are functions of $\zeta(i)$ such that $\omega_j \geq 0$ and $\sum_{j=0}^{\infty} \omega_j = 1$, for all $\zeta(i)$. In this case, the mean and variance of U_{ik} are given by

$$\mathrm{E}\{U_{ik}\} = \mathrm{E}\{N_k(i)\}\mathrm{E}\{C_{jk}(i)\} \tag{2.18}$$

$$\mathrm{Var}\{U_{ik}\} = \mathrm{E}\{N_k(i)\}\mathrm{Var}\{C_{jk}(i)\} + \mathrm{Var}\{N_k(i)\}[\mathrm{E}\{C_{jk}(i)\}]^2. \tag{2.19}$$

The normalized random variable $V_{ik} = [U_{ik} - \mathrm{E}\{U_{ik}\}]/[\mathrm{VarU_{ik}}]^{1/2}$ has the limiting distribution $H(c)$ if $\lim_{\zeta(i)\to\infty} \Pr\{V_{ik} \leq c\} = H(c)$ whenever c is a continuity point of $H(c)$. Thus, V_{ik} has the limiting distribution $H(c)$, if for every t,

$$\lim_{\zeta(i)\to\infty} \phi(t) = h(t) = \int e^{ict} dH(c), \tag{2.20}$$

where $\phi(t) = \mathrm{E}(e^{iV_{ik}t})$ is the characteristic function of V_{ik}. If (2.20) holds for $h(t) = e^{-t^2/2}$, then for every c,

$$\lim_{\zeta(i)\to\infty} \Pr\{V_{ik} \leq c\} = \frac{1}{(2\pi)^{1/2}} \int_{-\infty}^{c} e^{-u^2/2} du = H_0(c), \tag{2.21}$$

and hence U_{ik} is asymptotically Normal.

A useful result presented in Corollary 4 (Robbins, 1948, p. 1159) states that if $N_k(i)$ is asymptotically Normal then U_{ik} is asymptotically Normal. We therefore might focus on the distributional properties of $N_k(i)$ to gain information about the limiting distribution of U_{ik}.

Hillier (1998) performed an exhaustive investigation to assess distributional properties of the accounting earnings function Y by simulating single activity parts by different statistical processes and using several levels of assumptions. It goes far beyond the scope of this report to review all the experimental results. Nonetheless, Hillier & Willett (2001) summarised the limiting and non-limiting distributional properties of accounting earnings numbers under independence and stationary assumptions investigated by Hillier.

2.4 A Comment on the Single Activity Class Models

In the SACT literature, investigations to date have considered only the behaviour of a single activity class. Of course, investigations in any field start with simple assumptions. However, no real firm can be represented by just one activity class or equivalently by single assets because results based on one-tentacle octopuses (Johnston, 1960) have only limited explanatory power. This is also the reason why we put off the many results on the distributional properties of earnings from simulation reported in Hillier (1998) for comparison with more sophisticated simulations to a future point in time. This also includes the results reported about estimation and prediction properties of earnings by the same author.

As we understand, the main objective of SACT is to model financial statement entries, where financial statements stands for, for example, ledgers, balance sheets or cash flow statements, based on the two constructs activity and activity class. Because most of the single numbers in financial statements are likely to result from complicated production and cost structures, it is straightforward to investigate the impact on Y_i of a class mix. The next step would be to introduce stochastic dependence or correlation between single activity components (start, duration and cost), activities and activity classes. It also seems natural, in a statistical context, to assume the parameter for different activity classes in (2.3) or (2.4) to be interpreted as a random variable K. However, in dependence on practice, K might either be fixed or assume a (random) value in a suitable interval.

Further, we suggest equation (2.4) rather than (2.3) for any future investigation in this field because the concept of an activity class is represented far more intuitively by the prior equation. For instance, Lane & Willett (1999, p.3) assumed a firm "carrying out a large number of similar activities with random starts and completion dates". To achieve 'similar' activities, they then had to assume the activity durations to be uniformly distributed in a particular interval. However, if we assume that the activity duration is governed by, for example, the exponential distribution and values between zero and infinity can occur, it becomes difficult to explain what 'similar' refers to. That is, although by definition activities where their parts have identical distributional properties are grouped into one activity class, it is hard to think of an example in practice, where an activity with duration of a couple of days should be related to an activity with duration of one or several accounting periods (or years). We further argue in favour of (2.4) because the clear separation between short-term and long-term activities solves the above problem.

3 Statistical Models in Accounting

3.1 Introduction to the Allocation Problem

In the Brief & Owen paper series a slow development towards more realistic models passed off, introducing statistical analysis simultaneously to describe the models appropriately. In other words, the fluent combination of the accounting measurement problem of joint cost with financial decision and estimation theory made it possible, under already relatively realistic assumptions, to answer the question "what is the optimal (with respect to the least squares criterion and a target ratio of a series of rates of return) allocation method of the known cost of an asset over its known useful life?".

The earlier Brief and Owen papers use very basic statistics such as simply plugging in an estimator (Brief ,1967; Brief & Owen, 1968a) or simple assumptions about the character and quantity of its deterministic variables (Brief & Owen, 1969). Therefore, we limit our review on those papers which present reasonable models and do not make constrained and artificial assumptions.

Unfortunately, we have to note that the quite numerous 'typographical' mistakes found in the original papers revealed an inaccurate mathematical administration, which might induce confusion, especially among the targeted readership. A list of errata can be obtained via correspondence.

In general, an allocation pattern $Q(p_1, ..., p_n)$, n a fixed number of periods, is sought by minimizing Q with respect to the proportions p_i of a known overhead cost C, with $p_i \geq 0$ and $\sum_i p_i = 1$, such that

$$\frac{\partial Q(p_1, ..., p_n)}{\partial p_i} = \frac{\partial}{\partial p_i} \sum_{j=1}^{n} \frac{(X_j - p_j C)^2}{W_j} = 0. \tag{3.1}$$

The quantity X_i denotes some known numerical characteristic describing or assigned to the i^{th} product and the quantity W_i another known numerical characteristic chosen to standardize the differences $(X_i - p_i C)^2$. Not until Brief & Owen (1970) were the measures X_i and W_i interpreted as random variables. Then, it was possible to formulate the total expected measurement error from using an allocation pattern $p_1, ..., p_n$ prior to observing the variables X_i and W_i, given by

$$Q(p_1, ..., p_n) = \mathrm{E}\Big(\sum W_i (r_i - k)^2 \Big), \tag{3.2}$$

where $(r_i - k)$ is the difference between rate of returns in single periods and the overall rate of return on sales. Simply to avoid sign problems, the square of this difference is considered.

Whenever independent standard Normal random variables are squared and summed they result in the χ^2-distribution. It would be therefore possible to make inferences about the Normal character of the measurement error defined by Brief & Owen using a goodness-of-fit χ^2-procedure.

Minimizing this measurement error yields, in approximation by using Taylor expansions up to the quadratic order, a bias corrected allocation pattern

$$p_i = \{\mathrm{E}(a_i)\}^{-1}\mathrm{E}(b_i) - \frac{\{\mathrm{E}(a_i)\}^{-1}}{\sum_j \Big(\{\mathrm{E}(a_j)\}^{-1}\Big)} \cdot \Big(\sum_j \Big(\{\mathrm{E}(a_j)\}^{-1}\mathrm{E}(b_j)\Big) - 1 \Big), \tag{3.3}$$

where $a_i = C^2/W_i$, $b_i = X_i C/W_i$ and $j = 1, ..., n$. Results so far are applicable for a single assets. Brief & Owen (1973) discussed the multi-asset firm and introduced an algorithm which updates the distribution of the allocation pattern after every period. However, although the following model was presented under the title of a *multi-asset* firm, there is no obvious explanation and introduction of how to incorporate several assets. In fact, the example given also considers only one asset.

To do so, Brief & Owen assumed that the life of the firm is a random variable H where its density g_H is known. They described the uncertain cash flows $\boldsymbol{x} = (X_1, ..., X_H)$ by the joint density $f(\boldsymbol{x}|H)g_H$, which summarizes subjective opinion about future cash flows. To allow for a periodical updating of both the time horizon H and the cost allocation pattern $\boldsymbol{p} = (p_1, ..., p_n)$, which depends on the estimated horizon $n = \mathrm{E}(H)$, they defined the following entities: the set of the accountant's estimates, $a = (\boldsymbol{p}, n)$, and the expected future cash flows and the life of the firm, $\alpha = (\boldsymbol{x}, H)$. The density of α, $f(\alpha) = f(\boldsymbol{x}|H)g_H$ is assumed to be known. The estimating horizon n and the associated allocation table $\boldsymbol{p}$ are then updated in terms of their dependence on the expected future cash flows x_i and the life of the firm h. The loss function L, which accounts for the accountant's estimating error of choosing non-optimal values p_i, $i = 1, .., n$, and n, is given by

$$\begin{aligned} L(a, \alpha) &= L_1(\boldsymbol{p}, \boldsymbol{x}) + L_2(n, H) & (3.4) \\ &= (\boldsymbol{x} - \boldsymbol{p}C)\,\boldsymbol{w}(\boldsymbol{x} - \boldsymbol{p}C)^T + t(n - H)^2, & (3.5) \end{aligned}$$

where t is a constant and $\boldsymbol{w}$ is the $n \times n$-matrix with diagonal elements $(1/W_i)$. Brief & Owen used the loss function to measure the quality of different allocation tables.

The smallest expected loss a^* must satisfy $\mathrm{E}\{L(a^*, \alpha)\} = \min_a \mathrm{E}\{L(a, \alpha)\}$, which is equal to the value a which minimizes

$$\int L(a, \alpha) f(\alpha) d\alpha = \int \int L(a, \alpha) f(\boldsymbol{x}|H) g_H \boldsymbol{x} dH. \quad (3.6)$$

Constraining $\sum_i p_i = 1$ leads to the minimization problem with Lagrangian function

$$F(\boldsymbol{p}, n, \lambda) = \mathrm{E}\{(\boldsymbol{x} - \boldsymbol{p}C)^T \boldsymbol{w}(\boldsymbol{x} - \boldsymbol{p}C) + t(n - H)^2\} + \lambda(\boldsymbol{l}\boldsymbol{p} - 1), \quad (3.7)$$

with respect to $\boldsymbol{p}$, n and λ, where $\boldsymbol{l}$ is the $1 \times n$-vector with unit entries.

The *ex ante* problem is solved by determining $n^* = \mathrm{E}(h)$ and

$$\boldsymbol{p}^* = \frac{\{\mathrm{E}(\boldsymbol{w})\}^{-1}\mathrm{E}(\boldsymbol{w}\boldsymbol{x})}{C} + \frac{C - \boldsymbol{l}^T\{\mathrm{E}(\boldsymbol{w})\}^{-1}\mathrm{E}(\boldsymbol{w}\boldsymbol{x})}{C\boldsymbol{l}^T\{\mathrm{E}(\boldsymbol{w})\}^{-1}\boldsymbol{l}}\{\mathrm{E}(\boldsymbol{w})\}^{-1}\boldsymbol{l}, \quad (3.8)$$

where the * again denotes values for the smallest expected loss. Here, the solution for $\boldsymbol{p}^*$ is presented in vector form but is identical to (3.3).

Conditioned on the j observed cash flows $\boldsymbol{x}_{k,j} = (x_1, ..., x_j)$, where the index k stands for known quantities, the task to solve with the updating problem is to find the future states $a_j = (\boldsymbol{p}_j, n)$ including the possible selection of the remaining allocation probabilities $\boldsymbol{p}_j = (p_j, p_{j+1}, ..., p_n)$, which minimize

$$\int L(a_j, \alpha_{f,j}) f(\alpha_{f,j}|\boldsymbol{x}_j) d\alpha_{f,j}, \quad (3.9)$$

where $\alpha_{f,j} = (\boldsymbol{x}_{f,j}, h)$, the possible selection of an allocation pattern for the remaining future, denoted by the index f. The function $f(\alpha_{f,j}|\boldsymbol{x}_{k,j}) = f_1(\boldsymbol{x}_{f,j}|\boldsymbol{x}_{k,j}) \cdot f_2(h|\boldsymbol{x}_{k,j})$ is the conditional density function of $\alpha_{f,j}$ including the future cash flows $\boldsymbol{x}_{f,j} = (X_{j+1}, ..., X_H)$, given $\boldsymbol{x}_{k,j}$, and equals the product of two respective conditional density functions f_1 and f_2 as indicated.

The algorithm then starts again with (3.5), such that the updating problem is solved iteratively.

3.2 Depreciation by Conditioning

Ijiri & Kaplan (1969) showed that deterministic depreciation (DD) is only an approximation to probabilistic depreciation (PD) and underestimates PD in early periods. In their model, PD accounts for the uncertainty of the expected life of the asset by estimating its probability distribution at the outset, that is, before the asset is put into service. For DD, the expected life of the asset,

$$\mathrm{E}(L) = \sum_{l=1}^{n} p_l l, \tag{3.10}$$

is a weighted average, where L is the life of the asset and p_l the probability that the asset will be retired at the end of the l^{th} period with a known maximal service life $n = \max(l)$. We write $\mathrm{E}(L) = \mu_L$.

They further defined the depreciation vector $\boldsymbol{h}_l = (h_{1l}, h_{2l}, ..., h_{ll})$, where the proportion of the overhead costs h_{il}, with $h_{il} \geq 0$ and $\sum_i h_{il} = 1$, is to be depreciated at the end of the i^{th} period if the actual service life is $L = l$ periods $(i = 1, ..., l; l = 1, ..., n)$. It follows that for the deterministic approach the depreciation rate for period i, $d_i^{(DD)}$, equals $h_{i\mu_L}$ and for the probabilistic approach, the depreciation rate for period i, $d_i^{(PD)}$, is given by

$$d_i^{(PD)} = \sum_{l=i}^{n} p_l h_{il}, \tag{3.11}$$

such that the depreciation rate is a weighted average, too. It is computed by summing over all possible asset lives $l = i, ..., n$ and weighted by the respective probability p_l that they will occur.

Ijiri & Kaplan (1970) extended their early work by introducing sequential probabilistic depreciation (SPD), which accounts for the new information about the future expected life of an asset gained at the end of each period.

For the single asset case, Ijiri & Kaplan assumed that the distribution of the service life of the asset is known and given by $\boldsymbol{p} = (p_1, p_2, ..., p_l)$, where $p_l < \infty$ is the probability that the asset retires at the end of j periods in service. (In the last section but one of their article, they use Bayesian analysis to derive, period by period, the entries of $\boldsymbol{p}$, assuming that the service life still follows any parametric distribution but where its parameters have to be estimated.) Considering period i, the conditional probability that the random service life of the asset is $L = l$ periods, given that it is still in use at the end of the i^{th} period, $i < l$, is given by

$$\Pr\{l|i\} = p_l S_i^{-1}, \qquad \text{with } S_i = \sum_{j=i+1}^{\infty} p_j, \tag{3.12}$$

where we write for short $\Pr\{l|i\} = \Pr\{L = l|\text{period } i \text{ survived}\}$.

The sequential probabilistic depreciation rate $d_i^{(SPD)}$ is a random variable, since we do not know at the beginning of period i about the future expected life of the asset. Ijiri & Kaplan defined three conditional depreciation rates relative to a particular period i which describe $d_i^{(SPD)}$. They are, in order, d_i^- with probability $\Pr\{l < i|i\}$, that is, in case that the asset is retired before period i, and therefore $d_i^- = 0$. In case the asset is retired at the end of the current period, the depreciation rate d_i^0 with probability $\Pr\{l = i|i\}$ and is therefore $d_i^0 = h_{ii}$. The last depreciation rate is d_i^+ with probability

$\Pr\{l > i|i\}$, that is, in case that the asset has survived the i^{th} period, and where the conditional expected value of d_i^+ is given by

$$\mathrm{E}(d_i^+|i) = \sum_{all\ d_i^+} d_i^+ \cdot \Pr\{d_i^+|i\} = S_i^{-1} \sum_{l=i+1}^{\infty} p_l h_{il}, \tag{3.13}$$

using (3.12), and denotes a weighted average over the remaining random $(l-i)$ periods. This means that that for a fixed i, d_i^- and d_i^0 are constants and d_i^+ a random variable.

The expected value for $d_i^{(SPD)}$ for a particular period i is given by

$$\begin{aligned} \mathrm{E}(d_i^{(SPD)}) &= \sum_{l=1}^{i-1} p_l d_i^- + p_i d_i^0 + \sum_{l=i+1}^{\infty} p_l \mathrm{E}(d_i^+|i) \\ &= p_i h_{ii} + \sum_{l=i+1}^{\infty} p_l \mathrm{E}(d_i^+|i). \end{aligned} \tag{3.14}$$

Inserting (3.13) into (3.14) and using (3.12) yields

$$\begin{aligned} \mathrm{E}(d_i^{(SPD)}) &= p_i h_{ii} + \sum_{j=i+1}^{\infty} p_j h_{ij} \\ &= \sum_{l=i}^{\infty} p_l h_{il}, \end{aligned} \tag{3.15}$$

which is equal to $d_i^{(PD)}$ in (3.11), where a finite service life L was assumed.

They further showed that the average squared error over all possible service lives l given that the asset has survived the i^{th} period, between the proper depreciation rate h_{il} and the equally weighted actual depreciation rates d_i^*, is given by

$$S_i^{-1} \sum_{l=i+1}^{\infty} p_l (h_{il} - d_i^+)^2 + (d_i^+ - d_i^*)^2. \tag{3.16}$$

The d_i^*'s are interpreted as the accountant's guesses, and only the choice $d_i^* = d_i^+$ minimizes (3.16).

Using the same approach for the expected accumulated sequential depreciation rate after period i, $\mathrm{E}(a_i^{SPD})$, as they did for $\mathrm{E}(d_i^{SPD})$, yielded

$$\begin{aligned} \mathrm{E}(a_i^{(SPD)}) &= \sum_{all\ a_i^{(SPD)}} a_i^{(SPD)} \Pr\{a_i^{(SPD)}|i\} \\ &= \sum_{k=1}^{i} \sum_{l=k}^{\infty} p_l h_{kl}, \end{aligned} \tag{3.17}$$

which equals the accumulated depreciation rate for PD under the static probabilistic depreciation assumptions.

In the multi-asset case, where N assets are put into service at the beginning of the first period, Ijiri & Kaplan assumed it to be just an extension of the single asset case discussed above. The assets are grouped into three categories which describe relative to a period i, $i = 1, ..., l$, whether the asset was retired before, within or will be retired after that period and the same depreciation rates d_i^-, d_i^0 and $\mathrm{E}(d_i^+|i)$ apply, respectively.

The expected depreciation rate for the i^{th} period, based on the total depreciable cost for the N assets, is given by

$$\begin{aligned} \mathrm{E}(d_i^{(N-SPD)}) &= \frac{n_{i-1}}{N} d_i^- + \frac{m_i}{N} d_i^0 + \frac{(N-n_i)}{N} \mathrm{E}(d_i^+|i) \\ &= \frac{m_i}{N} h_{ii} + \frac{(N-n_i)}{N} S_i^{-1} \sum_{j=i+1}^{\infty} h_{ij} p_j, \end{aligned} \tag{3.18}$$

where n_i is the number of assets that have been retired at the end of the i^{th} period and m_i the number of assets that were retired just at the end of the i^{th} period, such that $n_i = n_{i-1} + m_i$. In order to get $\mathrm{E}(d_i^{(N-SPD)})$ for period i, only the expected values for m_i/N and $(N - n_i)/N$ have to be computed.

Ijiri & Kaplan used the binomial distribution as the mass function of the random variables m_i and $N-n_i$ such that the expected values are $\mathrm{E}(m_i/N) = p_i$ and $\mathrm{E}\Big((N - n_i)/N\Big) = S_i$. Then, inserting the last result into (3.18), the expected value for $d_i^{(N-SPD)}$ is therefore

$$\mathrm{E}(d_i^{(N-SPD)}) = p_i h_{ii} + S_i S_i^{-1} \sum_{l=i+1}^{\infty} p_l h_{il} = \sum_{l=i}^{\infty} p_l h_{il}, \tag{3.19}$$

which is equal to the single asset sequential probabilistic depreciation in (3.15) and to the static probabilistic rate in (3.11).

Similar analysis yields for the conditional expected value of the accumulated group depreciation rate, given that the asset survived period i, $\mathrm{E}(a_i^{+(N-SPD)}|i)$, the same result as it did for the single asset case. Substituting for the expected value of the accumulated group depreciation rate after i periods, $\mathrm{E}(a_i^{(N-SPD)})$, yields the same result as in (3.17).

3.3 Depreciation with Markov Chains

From the discussion in the previous section, it follows that the decision whether an asset is still in service at the beginning of a period i depends only on the events during the preceding period $i-1$. It arises naturally that this can be modelled by discrete time (stochastic) Markovian processes. This approach has been pursued by different authors. In the following we review the work of Zannetos (1963), which is based upon two articles Zannetos (1962) and Bissinger (1961). The articles of Pollard & Tippett (1994) and Pollard, Rhys & Tippett (1994) present further development of statistical results in the context of Markovian depreciation techniques.

Besides presenting a basic introduction on the applicability of Markov chains to group depreciation, Zannetos (1962) was concerned to justify that, in fact, a probabilistic treatment of depreciation and grouping assets with similar economical properties yields reliable information on the economic life of the assets, which would then support managerial decisions. Zannetos' aim was by other means to convince the reader that the presented approach does not complicate the accountant's work in the sense of complex and unmanageable amounts of data. In particular, a finite Markov chain was used to describe the behaviour of a group of assets through time and the transition probability matrix P was derived empirically. A single entry p_{jk} in P was defined as the probability that the assets were new at the beginning of period j and had k years of life remaining. The canonical representation of P is given by

$$\boldsymbol{P} = \begin{pmatrix} 1 & 0_{1\times n} \\ R_{n\times 1} & Q_{n\times n} \end{pmatrix}, \tag{3.20}$$

where R includes the transition probabilities from transient states to the absorbing state, Q the transition from transient states to other transient states and where n is a fixed number for the maximal life time of the assets. Zannetos then derived the depreciation expense D_k for the k^{th} year, which is given by

$$D_k = \frac{1}{n}\sum_{j=0}^{n} A_j(\pi_{k-1} - \pi_k)N^T, \tag{3.21}$$

where fixed A_j's represent the total dollar amount of those assets having j years of life remaining at the end of a time period and where N^T denotes $(0, 1, ..., n)^T$, the vector of the years. The input vectors π_k, $k = 0, ..., n-1$, can be derived by multiplying the initial input vector π_0 with the k^{th} power of P, where the entries of π_0 are derived by classifying the assets in terms of their remaining life, that is,

$$A_j / \sum_{j=0}^{n} A_j. \tag{3.22}$$

Further, under the assumption that new assets of the same price magnitude are acquired every period, the age distribution of the total assets will assymptotically converge towards a steady state after c years, such that for any year $k > c$ the depreciation amount is given by

$$D_k = \frac{1}{n}\sum_{j=0}^{n} A_j \pi_c (1 - P) N^T. \tag{3.23}$$

Zannetos (1963) further briefly mentioned that the two basic methods for testing the validity of the entries in P are classical (frequentist) hypothesis testing and the Bayesian approach. However no further efforts have been made in that direction. Finally, some comments were made with respect to the expected life of assets in a group, based on a matrix $N = (1 - Q)^{-1}$, where 1 is the $n \times n$ identity matrix and Q is defined above. The entries of N show the mean number of economic years the assets spend in a transient state j, given that they originated in state i. From this, the number $M_i(t)$, which represents the mean number of steps t to reach from a transient state i to the absorbing state, can be derived.

Zannetos (1963) demonstrated that from a same amount of initial information about a group of assets, that is, a fixed maximal life-time and the total initial value, the Markovian approach described above provides more information than the survival curve depreciation method presented in Bissinger (1961). Bissinger combined given mortality curves with a function which describes the trend or rate of the assets' behaviour. The reason why the Markovian approach turned out to be more informative, and thus superior, is that the survivor rates for different time periods for the survival curve method are conditional probabilities of survival for one additional period and they "smooth out non-uniformities in life expiration from year to year", as Zannetos (1963, p.162) noted.

Pollard & Tippett (1994) and Pollard, Rhys & Tippett (1994) also used homogeneous Markov chains to investigate some of the properties of probabilistic depreciation. However, their assumptions were less restrictive than those in the model presented above. Both articles assumed that n machines are simultaneously put into service at time $t = 0$ and that the sequence $0 \leq S(t) \leq n$, $t = 0, 1, 2, ...$, denotes the number of machines remaining in service at time t. The probability of moving from one state to another is described by the transition probabilities $p_{jk} = \Pr\{S(t+1) = k | S(t) = j\}$, where both $k, j = 0, 1, 2, ..., n$ and $j \geq k$.

The corresponding transition matrix is given by

$$\boldsymbol{P} = \begin{pmatrix} 1 & 0 & 0 & \cdots & 0 \\ p_{21} & p_{22} & 0 & \cdots & 0 \\ p_{31} & p_{32} & p_{33} & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ p_{n+1,1} & p_{n+1,2} & p_{n+1,3} & \cdots & p_{n+1,n+1} \end{pmatrix}, \qquad (3.24)$$

where the probabilities p_{jk} in every row, that is, for $j = 0, 1, ..., n$, sum to one. The first state of $\boldsymbol{P}$ is an absorbing state because $p_{11} = 1$. This was justified in both articles by the fact that "... once an asset is taken out of service it remains permanently out of service ..." . All other states are non-recurrent, that is, once an asset has left a given state, it never returns to it.

Unlike Zannetos' empirically derived transition mechanism, Pollard & Tippett discussed three different transition mechanisms for fixed j's. The first assumed that all conditional probabilities are equal (which throughout the rows is similar to a reducing balance depreciation method), the second and third are known as the 'sum-of-years' digits' and the 'sum-of-squared digits' techniques (see, e.g., Reynolds, 1962), and all are well known methods in practice. Further, they computed for every method $\mathrm{E}\{S(t)|S(0)\}$ and $\mathrm{Var}\{S(t)|S(0)\}$, the expected number and variance of machines in service at time t given that a certain number of machines was put into service at time zero. Because in all cases the mean and the variance approached zero with increasing time, Pollard & Tippett defined depreciation "in terms of the rate at which machines are taken out of service", $d_{0,t} = \{S(t) - S(0)\}/S(0)$, which is the proportional reduction of the number of machines still in service during the interval $[0, t]$.

A clear separation can be drawn between the results of the first two transition mechanisms (sort of reducing balance and 'sum-of-years' digits) and the 'sum-of-squared digits' mechanism: the previous two are special cases of the transition mechanism defined by the probabilities

$$p_{jk} = \frac{\rho\Gamma(k+\rho)\Gamma(j+1)}{\Gamma(k+1)\Gamma(j+\rho+1)}, \tag{3.25}$$

where Γ is the gamma function and $\rho \in \mathbb{R}_0^+$ a 'continuity' index. Rhys & Tippett (2002) give an interpretation of the parameter ρ in a similar model: ρ can be identified as the weighted average age of items, that is, "higher values of ρ imply that the asset deteriorates more slowly".

Pollard & Tippett (1994) showed that the expected depreciation rate $\mathrm{E}(d_{0,t})$ is independent of $S(0)$ for the linear and the 'sum-of-years' digits' depreciation. However, the 'sum-of-squared digits' mechanism revealed that this independence of the number of machines put into service at time zero does not generally apply. Furthermore, they mentioned that any transition mechanism has to satisfy the additivity property, which was illustrated by a simple example (see Pollard & Tippett, 1994, p.71) and, suggested to design Markovian transition mechanisms based on Bose-Einstein statistics as a possible future research area. This was followed up by Rhys (2000) and is reviewed below.

Pollard, Rhys & Tippett (1994) extended the analysis based on (3.25) and introduced a more flexible Markovian model. They determined the probability of having k machines in service at time $t+1$ by using the Chapman-Kolmogorov equation $\boldsymbol{p}^T(t+1) = \boldsymbol{p}^T(t)\boldsymbol{P}$, where $\boldsymbol{p}(t)$ is the vector containing the state probabilities at time t.

Using (3.25) and the above Chapman-Kolmogorov equation yields after some analysis the recursion formula for the i^{th} factorial moment

$$m^{(i)}(t+1) = \frac{\rho}{\rho+i} m^{(i)}(t) = \mathrm{E}_t\Big[\prod_{j=0}^{i-1}\{S(t+1)-j\}\Big], \tag{3.26}$$

where $E_t(.)$ is the expectation operator taken at time t. Computing the mean depreciation rate for the first period, thus $\mathrm{E}_0(d_{0,1}) = \{\mu_1 - S(0)\}/S(0) = -1/(\rho+1)$ was used to calculate ρ. Based on that constant one-period depreciation rate all higher moments in (3.26) for the particular transition mechanism can be derived.

This methodology is commonly used in epidemic models where the starting settings determine the properties of the system during its future trend (see, e.g., Daley & Gani, 1999).

Rhys (2000) extended the analysis based on a slight variant of the model above and represented an asset by n discrete items which are put randomly into r cell during a time period. Further, that article concentrated on two different interpretations of the n items: distinguishable and indistinguishable.

It is clear to physicists that distributional properties of distinguishable items can be described by Maxwell-Boltzmann statistics and for the case of indistinguishable items by Bose-Einstein statistics.

Rhys noted that the transition probabilities given in (3.25) represent Bose-Einstein probabilities and rewrote the expected depreciation rate, which follows from (3.26), in terms of r, thus, $1/(\rho + 1) = 1/r$. He further noted the important difference of the existence of discrete and continuous assets. Obviously, (3.26) suggests a discrete distribution for discrete assets and for continuous assets, as Pollard, Rhys & Tippett (1994) showed, a continuous distribution with a log-gamma probability density function.

Rhys then went on to show that if Maxwell-Boltzmann statistics are used, the transition probabilities turn out to have a binomial distribution. It follows that for discrete assets, the expected depreciation rate is independent of the initial state, as we have already seen above for the model based on Bose-Einstein statistics. However, for continuous assets, that is when $n \to \infty$, the variance becomes zero such that the limiting distribution is deterministic and the mean with probability one resembles the widely used reducing balance method in accounting.

Rhys further expressed the useful life of an asset by a random variable Y_n, the waiting time to absorption and derived the probability generating function under distinguishability and indistinguishability assumptions. He showed that infinite asset life is impossible and that the expected value of the useful life increases without bound as n increases. Higher order moments were not derived because it is straightforward to do so and would not improve the information content of an analytical paper. However, for simulation purposes, higher order moments could be used to assess the rate of convergence towards a particular limiting distribution.

An interesting article was recently published by Rhys & Tippett (2002) where stochastic depreciation based on Bose-Einstein statistics is not only investigated on its own but within a model of capital accumulation (Merton, 1973). In the model, an agent seeks to optimize units of capital through time by dividing it between consumption and investments in productive facilities. Rhys and Tippett used for this optimization a variant of an algorithm suggested by Dreyfus (1965, p.218) which accounts for both sources of uncertainty (consumption and depreciation). At the end of this very technical paper, they concluded that "consumption will be higher when capital deteriorates stochastically in comparison to the case where capital deteriorates deterministically". By now, the many models discussed in this chapter should have convinced the reader that depreciation is of stochastic nature. Therefore, the above result shows that there is a large systematic bias if deterministic depreciation based on Maxwell-Boltzmann statistics (Rhys, 2000) is used.

4 The Virtual Firm

A *virtual firm* is a model of a real firm in which it is intended to assess simulation results from the combination of the SACT framework with particular statistical models for simple business structures. The timeline is divided into discrete periods representing accounting periods where the activities are embedded. Activities are further classified by their positioning, that is, there are *equities* (finishing activities in a period of concern) and *assets* (unfinished activities in a period of concern) (for example, Willet (1991b), p.120). Furthermore, activities are classified with respect to their duration into short-term and long-term activities, where the latter ones can be depreciated.

In an advanced virtual firm, there will be activity classes which separately account for production or cost relations for short-term, non-depreciable and depreciable long-term activities. Such a virtual firm also accounts for a periodic updating of realised amounts, similar to the algorithm suggested by Brief and Owen (equations (3.5) to (3.9)). It further accounts for SACT depreciable long-term activities which make use of conditional expected life times suggested by Ijiri and Kaplan (equations (3.15) and (3.17)). Another approach to interpret SACT production relations for depreciable long-term activities is to apply Markovian transition mechanisms suggested by different authors and reviewed in Section 3.3 and to account in this way for their depreciable life. Further improvements which a sophisticated virtual firm should account for is the dependence between single activities, single activities and activity classes or even between activity classes, properties which were already suggested in association with SACT.

Of course, this idea of simulating what is observed in reality is not new and similar procedures are used to, for example, optimize assembly plant processes or to assess characteristics of proteins by simulation. One common and very important part of all these simulations, including the virtual firm, is the quality of the boundary conditions.

Unfortunately, in this field of research, additional problems which one might be confronted with are the following. It is, in general, not a simple task to obtain information about the entries found in companies' balance sheets, cash flow statements or profit and loss statements and therefore the comparison of simulation output and real data is complicated. Furthermore, the assessment of estimation and prediction properties, two basic functionalities of an accounting system, via time series analysis faces the problem that most real data series are very short. Companies merge, companies go bankrupt, companies change their policies. To get better statistics out of real data, we suggest to use the bootstrap method introduced by Efron (1979) or modified versions to meet the specific requirements. To account for changes in companies' structures, change point methods might be applicable (see, e.g., Ahsanullah, Rukhin & Sinha, 1995).

4 Summary

We have reported on recent statistical developments in cost accounting. At first glance, SACT might be seen closely related to ABC. The important difference between the two is the definition of what an activity is: ABC pools activities, SACT separates them into cost and production structures. In detail, SACT presents a completely different philosophy of how to interpret accounting numbers, namely, that financial statement and even ledger entries are random variables. In other words, each entry in financial statements is not viewed as a fixed number anymore but as a statistic with associated distributional properties and thus much more information content. This in turn suggests, for example, an introduction of risk to put emphasize on those numbers. It seems that an introduction of any reliable risk measures would be highly appreciated by the private sector because there, derived risk numbers are solely based on sensitivity analysis and subjective opinion of future events.

We have further shown that, so far, limited analytical results and results from simulation were derived using the SACT framework. It is up to future research to investigate more sophisticated models, which basically include an introduction of stochastic dependence and correlation among the SACT input and output relations and additionally account for several activity classes, which are the second basic construct in SACT besides single activities. The objective of SACT to represent accounting earnings numbers faithfully necessarily includes further investigations in that direction because in practice, dependence is likely to occur.

Further, we reviewed statistical models which are mainly concerned with the famous depreciation problem in accounting and suggested to combine them with the SACT framework. In this article, this combination is referred to as a statistical accounting system and the so-called virtual firm is used to yield results from simulating such accounting or business structures.

As a last remark, we conclude that statistical analysis is inevitable in modern cost accounting and that more attention should be paid to it, foremost in accounting courses.

Acknowledgements

We are thankful to John Lane for his useful comments which clarified some issues in the SACT review. Michael Falta holds an International Postgraduate Research Scholarship (IPRS) at The University of Queensland and gratefully acknowledges this financial support.

References

Ahsanullah, M., Rukhin, A. & Sinha, B. (1995). *Applied Change Point Problems in Statistics*. New York: Nova Science Publishers.

Ball, R. & Ross, W. (1972). Some time series properties of accounting income. *J. Fin.* **27**, 663-681.

Basset, P.H. (1987). *Computerised Accounting*. Manchester: NCC Publications.

Bissinger, B.H. (1961). An application of general plant account theory to the fitting of a survivor curve to a physical group property account stub life table. *Metroecon.* **13**, 156-163.

Brief, R.P. (1967). A late nineteenth century contribution to the theory of depreciation. *J. Acct. Res.* **5**, 27-38.

Brief, R.P. (1990). Accounting error as a factor in business history. *Acct., Bus. Finan. Hist.* **1**, 7-21.

Brief, R.P. & Owen, J. (1968). A least squares allocation model. *J. Acct. Res.* **6**, 193-199.

Brief, R.P. & Owen, J. (1969). On the bias in accounting allocations under uncertainty. *J. Acct. Res.* **7**, 12-16.

Brief, R.P. & Owen, J. (1970). The estimation problem in financial accounting. *J. Acct. Res.* **8**, 167-177.

Brief, R.P. & Owen, J. (1973). A reformulation of the estimation problem. *J. Acct. Res.* **11**, 1-15.

Butler, D., Rhys, H. & Tippett, M. (1994). On the wasting of assets - a review article. *Pacific Accounting Review* **6**, 70-93.

Cokins, G. (1996). *Activity-based costing management making it work: a manager's guide to implementing and sustaining an effective ABC system.* Boston: McGraw Hill.

Cooper, R. (1988). The rise of activity-based costing - part one: what is an activity-based cost system? *Journal of Cost Management* **2**, 45-54.

Daley, D.J. & Gani, J. (1999). *Epidemic Modelling: an Introduction.* New York: Cambridge University Press.

Datar, S.M., Foster, G. & Horngren, C.T. (2000). *Cost Accounting - A Managerial Emphasis.* 10th ed. Upper Saddly River: Prentice Hall International, Inc.

Dreyfus, S. (1965). *Dynamic Programming and the Calculus of Variations.* New York: Academic Press.

Efron, B. (1979). Bootstrap methods: another look at the jackknife. *Ann. Statist.* **7**, 1-26.

Friberg, R.A. (1973). Probabilistic depreciation with a varying salvage value. *The Acctg. Rev.* **48**, 50-60.

Gibbins, M., Hillier, J.R. & Willett, R.J. (1998). The foundations of statistical activity cost theory with applications to some old accounting measurement issues: a simulation approach. *J. Pol. Econ.* **64**, 57-78.

Gibbins, M. & Willett, R.J. (1997). New light on accrual, aggregation and allocation, using an axiomatic analysis of accounting numbers' fundamental and statistical character. *ABACUS* **33**, 137-167.

Hillier, J.R. (1998). *Properties of accounting earnings statistics and time series cost averaging.* University of Otago. Dunedin. New Zealand.

Hillier, J.R. & McCrea, M. (1998). The earnings smoothing potential of systematic depreciation. *ABACUS* **34**, 75-91.

Hillier, J.R. & Willett, R.J. (2001). The impact of depreciation-type adjustments on the distribution of accounting earnings. working paper.

Hodgson, A., Okunev, J. & Willett, R.J. (1993). Accounting for intangibles: a theoretical perspective. *Acc. Bus. Res.* **23**, 138-150.

Ijiri, Y. & Kaplan, R.S. (1969). Probabilistic depreciation and its implications for group depreciation. *The Acctg. Rev.* **44**, 743-756.

Ijiri, Y. & Kaplan, R.S. (1970). Sequential models in probabilistic depreciation. *J. Acct. Res.* **8**, 34-46.

Johnson, J. (1960). *Statistical Cost Analysis.* New York: McGraw-Hill.

Lane, J. & Willett, R.J. (1997). Depreciation need not be arbitrary. *Acc. Bus. Res.* **27**, 179-194.

Lane, J. & Willett, R.J. (1999). Optimal smoothing of accounting earnings. *IMA Journal of Mathematics Applied in Business & Industry* **10**, 1-14.

Lane, J.A. (1984). The central limit theorem for the poisson shot-noise process. *J. Appl. Prob.* **21**, 287-301.

Manegold, J.G. (1981). Time-series properties of earnings: a comparison of extrapolative and component models. *J. Acct. Res.* **19**, 360-373.

Mattessich, R. (1984). The scientific approach to accounting. In *Modern Accounting Research: History, Survey and Guide.* Ed. R. Mattessich, pp. 1-19. Vancouver: Canadian Certified General Accountants' Research Foundation.

Merton, R. (1973), An intertemporal capital asset pricing model. *Econometrica* **41**, 867-887.

Parzen, E. (1962). *Stochastic Processes.* San Francisco: Holden-Day.

Pollard, B., Rhys, H. & Tippett, M. (1994). Markovian depreciation. *IMA Journal of Mathematics Applied in Business & Industry* **5**, 163-169.

Reynolds, I.N. (1962). *An Analysis of Depreciation Methods and Bases.* University of North Carolina. Chapel Hill.

Rhys, H. (2000). Depreciation: a stochastic approach. *Journal of Business Finance & Accounting* **25**, 777-784.

Rhys, H. & Tippett, M. (2002). Stochastic depreciation and optimal consumption-investment decisions. *Journal of Business Finance & Accounting* **29**, 273-286.

Robbins, H. (1948). The assymptotic distribution of the sum of a random number of random variables. *B. Am. Math. S.* **54**, 1151-1161.

Takacs, L. (1954). On secondary processes generated by a poisson process and their applications in physics. *Acad. Sci. Hungary* **5**, 203-236.

Willett, R.J. (1987). An axiomatic theory of accounting measurement. *Acc. Bus. Res.* **17**, 155-171.

Willett, R.J. (1988). An axiomatic theory of accounting measurement - part II. *Acc. Bus. Res.* **19**, 79-91.

Willett, R.J. (1991a). An axiomatic theory of accounting measurement structures. *IMA Journal of Mathematics Applied in Business & Industry* **3**, 45-59.

Willett, R.J. (1991b). Transaction theory, stochastic processes and derived accounting measurement. *ABACUS* **27**, 117-134.

Willett, R.J. (1998). Production and cost functions: an accountant's viewpoint. In *Production and Cost Functions: Specification, Measurement and Applications*, Ed. E.I. Bairam, pp. 17-41. Aldershot, UK: Ashgate.

Zannetos, Z.S. (1962). Statistical attributes of group depreciation. *The Acctg. Rev.* **37**, 713-720.

Zannetos, Z.S. (1963). Markov chain depreciation model and survivor curve depreciation. *Metroecon.* **15**, 155-166.

Résumé

Nous avons revu les développements récents des modèles statistiques dans la comptabilité des coûts et simultanément, présenté un concept statistique convénient. On a appliqué ceci sur la théorie statistique des coûts des activités (Statistical Activity Cost Theory, SACT) introduit par Willett (1987 et 1988). Cette théorie interprete les numéros rencontrés dans les rapports

financiers comme des statistiques et puis, avec plus d'information que des numéros fixés et déterministes.

Nous proposons la combinaison de SACT avec plusieurs modèles statistique d'amortissement, qui montrent des avantages sur des modèles comparables et déterministes, et concluons que le résultat est en effet le premier système de la comptabilité *statistique*. Cette proposition est signifiante parce que en pratique, des entreprises utilisent seulement des modèles déterministes connus comme des "Activity Based Costing (ABC)"-systèmes.

De plus, nous avons introduit l'entreprise virtuel, l'instrument avec lequel on simule des saisies dans les rapports financiers basé sur des résultats de la théorie proposé précédemment. Nous discutons aussi des problèmes que l'on peut rencontrer pendant cette implémentation.