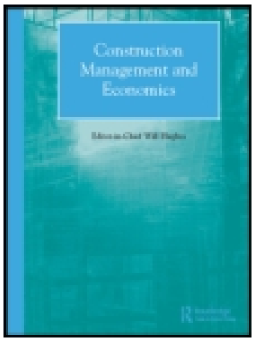

Construction Management and Economics

ISSN: 0144-6193 (Print) 1466-433X (Online) Journal homepage: https://www.tandfonline.com/loi/rcme20

A construction project cash flow model—an idiographic approach

Russell Kenley & Owen D. Wilson

To cite this article: Russell Kenley & Owen D. Wilson (1986) A construction project cash flow model—an idiographic approach, Construction Management and Economics, 4:3, 213-232, DOI: 10.1080/01446198600000017

To link to this article: https://doi.org/10.1080/01446198600000017

Published online: 28 Jun 2007.

Construction Management and Economics, 1986, **4**, 213–232

A construction project cash flow model – an idiographic approach

RUSSELL KENLEY and OWEN D. WILSON

Department of Architecture and Building, University of Melbourne, Parkville, Victoria, Australia 3052

The paper introduces and supports the contention that an idiographic methodology is appropriate to the post-hoc *study and interpretation of individual construction project cash flows. A cash flow model based on the logit transformation is proposed to be consistent with this methodology. The model is based on historical data, and yields two parameters to describe each individual project. The model is tested using two samples totalling* 72 *projects. Goodness of fit for the model, using a measure of standard deviation from* 1.0% *to* 4.6%, *with a median of* 2.5%, *is found for individual projects.*

The experimental hypothesis (that there is substantial variation between projects) is supported by the graphical and statistical evidence of deviation, which is argued to be the result of the individual ontology of each project – systematic error – rather than random error from an ideal. The paper concludes that forecasts of individual cash flows are invalid when derived from analysis of grouped data.

Keywords: Construction management, cash flow, estimating, logit transformation, idiographic methodology

Introduction

General overview

The study of construction project cash flows has become increasingly popular over the last 20 years. Several approaches to the analysis have been used and they may all be characterized as nomothetic, in that they attempted to discover general laws and principles across categorized or non-categorized groups of construction projects, with the purpose of *a-priori* prediction of cash flows. In contrast, the present study adopts an idiographic methodology; the search for specific laws pertaining to individual projects. This paper aims firstly at examining the underlying theories implicit in the study of building project cash flows, and secondly at developing an idiographic model which allows the *post-hoc* study and interpretation of individual cash flows.

The idiographic–nomothetic debate flourished within the social sciences, from the 1950s through to the early 1960s (Runyan, 1983). The contention arose, according to De Groot (1969), from the inability to classify the social sciences as either cultural or natural sciences. The social sciences, to which construction management must belong, have components of both cultural sciences (for example history) and natural sciences (for example physics), and have aspects which are 'individual and unique; they are own – (character) – describing: idio-graphic' (De Groot,

0144–6193/86 $03.00+.12

1969). De Groot claimed that 'if one seeks to conduct a scientific investigation into an individual, unique phenomenon . . . the regular methodology of (natural) science provides no help'. As construction projects are unique it would seem logical that their cash flows should be considered as individual and unique. The evidence in this paper supports this view, and the view that an idiographic methodology has a place in the study of individual cash flows.

Individual variation between projects is caused by a multiplicity of factors, the great majority of which can neither be isolated in sample data, nor predicted in future projects. Some existing cash flow models hold that generally two factors, date and project type, are sufficient to derive an ideal construction project cash flow curve. Such convenient divisions ignore the complex interaction between such influences as economic and political climate, managerial structure and actions, union relations and personality conflicts. Many of these factors have been perceived to be important in related studies such as cost, time and quality performance of building projects (Ireland, 1983), and therefore models which ignore all these factors in cash flow research must be questioned.

As with the majority of previous studies, the present study uses historical data. Standard curve models, based on historic data, have been extensively used in cash flow research (for example: Balkau, 1975; Bromilow, 1978; Bromilow and Henderson, 1977; Drake, 1978; Hudson, 1978; Hudson and Maunick, 1974; Kerr, 1973; McCaffer, 1979; Singh and Woon, 1984; Tucker and Rahilly, 1982). Although these approaches have gained general acceptance, they have not been without criticism. Hardy (1970) found that there was no close similarity between the ogives for 25 projects considered, even when the projects were within one category. This implicit support for an idiographic methodology was subsequently ignored, despite the problems which some researchers found in supporting their models. Hudson observed that 'difficulties are to be expected when trying to apply a simple mathematical equation to a real life situation, particularly one as complex as the erection of a building' (Hudson and Maunick, 1974; Hudson, 1978). It is interesting to contrast the size of Hardy's sample of 25 projects, with the relatively small samples used by many of the researchers finding nomothetic, ideal curves. For example Bromilow and Henderson (1977) used four projects, while Berdicevsky (1978) used three and Peer (1982), four.

One group of authors (for example: Kennedy *et al.*, 1970; Peterman, 1973; Reinschmidt and Frank, 1976; Ashley and Tiecholz, 1977; Berdicevsky, 1978; Peer, 1982) have modelled cash flows, prior to construction, through the use of forecast work schedules. Although this approach is valid for modelling predicted cash flow schedules, comparisons between the forecasts and the data for the actual project have not been published to the authors' knowledge. This method of using forecast work schedules is likely to be predominantly a measure of an estimator's consistency, rather than of central trends within project groups. Historical data, on the other hand, is derived from existing projects and therefore allows direct analysis of actual, as opposed to imagined, cash flow curves.

There has been an implied trend over time towards an idiographic construction project cash flow model. The early models, which may be termed 'industry' models, searched for generally applicable patterns across the entire industry. When it was recognized that this was unlikely to be achieved, greater flexibility was introduced by searching for patterns within groups or categories of project (the division usually being made according to project type and/or dollar value – e.g. Hudson and Maunick, 1974). This still wholly nomothetic approach was modified by Berny and Howes (1982) who adapted the Hudson and Maunick (1974) category model to a form which could reflect the specific form of individual projects. Even within categories, it had been found that there were occasional projects which did not fit the forecast expenditure well (Hudson and Maunick, 1974; Hudson, 1978).

Berny and Howes (1982) designed methods for calculating the specific curve for a given project, based on their general equation. In doing so they pointed the way for future research in this field. Their model made a very important cognitive step. By proposing an equation for the general case of an individual project curve, as distinct from the curve of the general (standard) function, it moved from a nomothetic to an idiographic approach. Unfortunately, for reasons which do not require elaboration in this paper, the Berny and Howes model has been found to be unsuitable for the aim of this research (to allow for a net cash flow model), and thus another model is sought.

As stated above, one aim of this research is to produce a model which, while recognizing that variable influences exist, does not need to predict them in order to model cash flows. This is a model which can take on the individual shape of a project cash flow curve regardless of influences and allow the variation between projects to be examined and quantified. Having formulated the model the idiographic approach can then be contrasted with the nomothetic approach. The search for, and testing of, such a model forms a significant portion of this paper.

The proposed model utilizes the Lorenz curve, which illustrates a relationship between two variables expressed as percentages and is common to most cash flow ogive models. It has been widely shown that the Lorenz curve expressing the relationship between percentage of project time completed (as the independent variable on the abscissa) and cumulative value at any time as a percentage of final value (the dependent variable on the ordinate), takes on a sigmoid form. Fig. 1 illustrates an example project data set for cash flow from a client to a contractor with a suggested fitted curve.

A second aim of this research is to allow for the subsequent modification of the model to a net cash flow model. Nazem (1968) suggested that cash balance curves could be obtained indirectly from the inward and outward cash flow curves. With two such curves, generating each net cash flow curve, the problems inherent in a nomothetic methodology are accentuated. The net cash flow model will be presented in a later paper.

Within this paper the term cash flow is used to describe the flow of cash or commitment from the client to the contractor. The model is equally applicable to flows from the contractor to his various subcontractors and suppliers. However an important clarifying distinction must be made between cash flow and net cash flow (cash balance), the latter being the residual after both income and outgoings are assessed. The model is equally capable of handling the value of work in place, the value of progress payments claims, or the value of progress payments actually paid to the contractor, despite the differences in time and monetary value which each of these may have for a given project.

The principles of regression analysis and independent samples

A significant rationale for the introduction of an idiographic methodology is based upon a consideration of the principles of regression analysis. Researchers modelling cash flows under a nomothetic methodology have utilized polynomial regression as their curve-fitting technique.

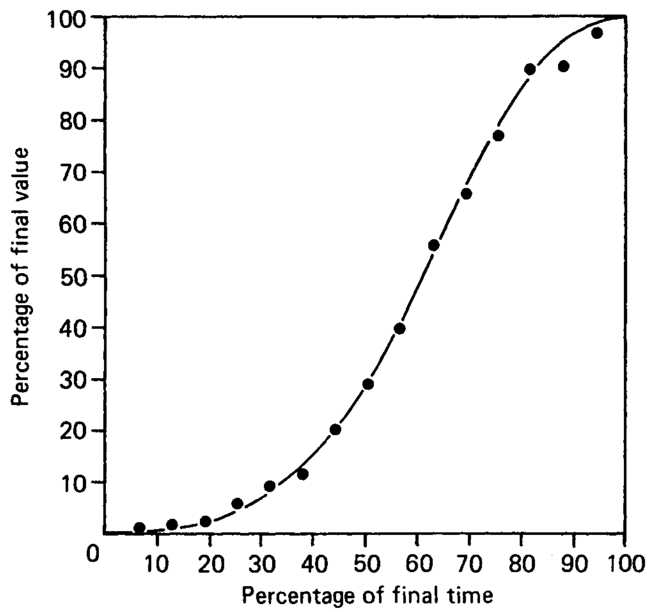

Fig. 1. Example of project data set with fitted curve.

Given sufficient constants and terms it seems possible to fit a polynomial to any required sigmoid form. However, the nomothetic regression operates upon grouped data, a procedure which involves overlaying the data from many projects. Given the nature of the data involved and the principles of regression analysis, the validity of this procedure is questionable.

The principles of regression analysis assume that dispersion of points about a central tendency is normally distributed and caused by random error. Spirer (1975, p. 329) states the assumptions applicable to a relationship in a bivariate population to be:

(i) The conditional mean, $\mu_{Y.X}$ is related to the independent variable *Y*.

(ii) The conditional standard deviation, $\sigma_{Y.X}$ is the same for all values of the independent variable *X*.

(iii) The conditional distribution of *Y* for any *X* is modelled by the normal probability distribution.

The regression model therefore pre-supposes no relationship between *Y* (dependent) values except for their dependency upon *X* (independent) values.

In reality, discrepancy from the average trend line is caused by systematic, idiosyncratic variation in the profile for each project, rather than random error. The *Y* value for each point is specifically linked to the preceding or succeeding points for that project, whilst maintaining *no relationship whatsoever* to any other data points from other projects in the analysis. Thus each project has an individual line of central tendency, with an associated error scatter about the line. A group regression is only applicable if it can be shown that there is no significant difference between the individual conditional means $\mu_{Y.X}$, and therefore no significant difference between the lines of best fit. This has not been demonstrated in any work known to the authors, and evidence will be put forward in this paper to show that it is fundamentally incorrect.

The only situation in which regression may be appropriate is for the analysis of the trend line of a single project, where the hypothesis that all deviation from the central mean is caused by random error may reasonably be accepted. It follows then, that if regression analysis is to be used it cannot correctly analyse more than one project at a time.

There are other problems stemming from the use of polynomial regression. Firstly a large number of constants are generally required. Peer (1982) used a bi-quadratic equation with five constants; Berdicevsky (1978) used a cubic with four; so also did Bromilow and Henderson (1977), whose inverted polynomial expressed time as a function of value. Hudson (Hudson and Maunick, 1974; Hudson, 1978) used a cubic polynomial, re-arranged to contain only two variable constants. In addition a full polynomial regression is necessary for each fit, and the lines of best fit derived do not always intersect the origin (0,0) and end (100,100) (which would be an advantage for a cash flow model intended for modification to a net cash flow model).

There is an alternative to the use of polynomial regression analysis for curve fitting. This is linear transformation, which involves altering the values of either the independent, the dependent, or both, data variables by a mathematical function in order to render a linear relationship between the transformed data variables. The methods for choosing transformations require specific knowledge about the relationship between variables, or use diagnostics to suggest possible transformations (Weisberg, 1980). The establishment of a suitable linear transformation for construction project cash flow sigmoids would be a valuable contribution.

The logit model

The cash flow ogive is a sigmoid and similar curves have been found in connection with growth patterns in economics and biological assay. Investigations in these areas have found that specific transformations of sigmoid curves can produce linear functions, the parameters of the sigmoid function are provided by the parameters of the linear equation, which in turn are found through linear regression.

The equation of the curve of best fit for data which approximates one of the sigmoid growth functions, may be found by a linear regression, of suitably transformed data, and then substitution of the linear parameters into the sigmoid function. The result may then be tested for goodness of fit against the data.

The selection of an appropriate sigmoid function in bio-assay was examined by Ashton (1972), who outlined four of the best known sigmoids and their transformations; the integrated normal curve, the logistic curve, the sine curve and Urban's curve. Ashton found the above curves to be very similar in shape, almost all variation being seen only at the extremes. He therefore concluded that the selection of an appropriate sigmoid was more a matter of application than anything else. If it could be shown that the ontological form of the cash flow, described by Hudson and Maunick (1974) as the *underlying relationship* between expenditure and time, approximates the form of the sigmoids listed above, then it would follow that any of these sigmoids could be used. This has been found to be so, although it is necessary to use a double transformation in cash flow analysis.

The Logit (given in equation 1) is the simplest of the Ashton sigmoid transformations and most easily allows the change to double transformation. The linear equation is found by a logit transformation of both the independent and dependent variables:

$$\text{Logit} = \ln \frac{z}{1-z} \tag{1}$$

where z is the variable to be transformed, and Logit is the transformation. The logistic equation for cash flows can be expressed using value (v) as the dependent variable and time (t) as the independent variable

$$\ln \frac{v}{1-v} = \alpha + \beta X$$

where

$$X = \ln \frac{t}{1-t}$$

therefore

$$\ln \frac{v}{1-v} = \alpha + \beta\left(\ln \frac{t}{1-t}\right) \tag{2}$$

This then forms the equation of the sigmoid curve which describes the flow of cash on a specific building project. It may also be expressed as follows

If

$$\ln \frac{v}{1-v} = \alpha + \beta\left(\ln \frac{t}{1-t}\right)$$

which can be shown to be equivalent to

$$v = e^{\alpha}\left(\frac{t}{1-t}\right)^{\beta} \Big/ \left[1 + e^{\alpha}\left(\frac{t}{1-t}\right)^{\beta}\right] \tag{3}$$

or

$$v = \frac{F}{1+F} \qquad \text{where } F = e^{\alpha}\left(\frac{t}{1-t}\right)^{\beta} \tag{4}$$

The logit cash flow model given above uses scales from 0.0 to 1.0, where the ratio (on the abscissa or ordinate) 1.0 is equivalent to 100%. As percentage scales are to be used in accordance with convention, the equations should be expressed as follows

If

$$\ln \frac{v}{100-v} = \alpha + \beta\left(\ln \frac{t}{100-t}\right) \tag{5}$$

which can be shown to yield the following equations

$$v = 100.e^{\alpha}\left(\frac{t}{100-t}\right)^{\beta} \Big/ \left[1 + e^{\alpha}\left(\frac{t}{100-t}\right)^{\beta}\right] \tag{6}$$

or

$$v = \frac{100 \times F}{1+F} \qquad \text{where } F = e^{\alpha}\left(\frac{t}{100-t}\right)^{\beta} \tag{7}$$

The practical application of the logit transformation cash flow model implies that construction project cash flow curves approximate the sigmoid form yielded by equation 7. This being so, a transformation of the data should approximate a line described by equation 8, and with parameters α and β

$$Y = \alpha + \beta X \tag{8}$$

where

$$Y = \ln \frac{v}{100-v}$$

and

$$X = \ln \frac{t}{100-t}$$

The logit transformation model is represented by equations 5, 7 or 8.

In Fig. 2 the data for a sample project is illustrated firstly in the Lorenz format, and secondly

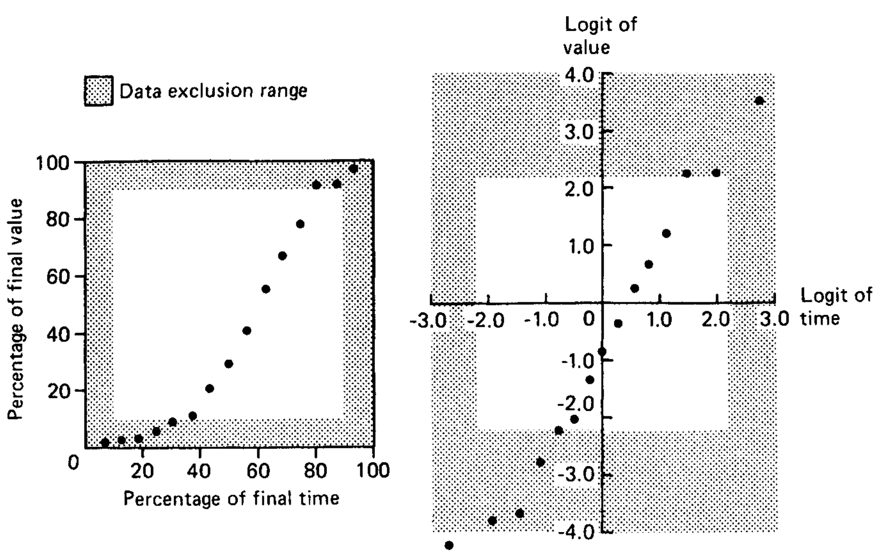

Fig. 2. Linear transformation for a sample project.

as transformed. From this it can be seen that the transformed data can indeed be fitted by a straight line.

In order to transform the data, X and Y must be calculated for each value of t and v respectively. Deriving the constants α and β is then simply a matter of linear regression of the transformed data

where

$$\beta = \frac{\Sigma(X-\bar{X})(Y-\bar{Y})}{\Sigma(X-\bar{X})^2} \tag{9}$$

and

$$\alpha = \bar{Y} - \beta\bar{X} \tag{10}$$

Aims and predictions

In summary the following rationale for an idiographic methodology for construction project cash flows have been presented:

(a) Consideration of the idiographic–nomothetic debate in the literature led to the conclusion that the natural science methodology was inappropriate for unique phenomena such as construction projects.

(b) A multiplicity of factors and influences effect project cash flows, many of which are unquantifiable and have differential impact.

(c) Regression analysis for grouped data, associated with a nomothetic methodology, is of questionable validity for construction projects.

It is therefore contended that an idiographic methodology is more appropriate to the study of construction project cash flows, than is a nomothetic methodology, and a nomothetic methodology can only be supported if a significant similarity can be shown to exist within groups. The experimental hypothesis is that there is substantial variation between projects.

The principal aim of this paper is to develop an idiographic construction project cash flow model. Secondary aims are:

(i) to determine an optimum exclusion range for data points;

(ii) to examine the model for goodness of fit and to identify projects where the model has failed to adequately fit the data (referred to as outliers);

(iii) to use the model to support the hypothesis above, by demonstrating variation between projects;

(iv) to support the contention above by contrasting the two methodologies quantitatively; and

(v) to allow for the subsequent development of a net cash flow model.

Method

Data

The model was tested on data from two separate, and different, sources.

The first sample (S1) comprised data for 32 medium to large scale commercial and industrial projects, provided by a Sydney-based construction group. The requirements of anonymity prevented the provision of a detailed breakdown into subcategories, as the only data provided were the dollar amounts of monthly claims from the client and the dates of approval. The projects were all constructed in the mid to late 1970s, in and around the Sydney and New South Wales area.

The second sample (S2) comprised data for projects from throughout Australia, provided by the Melbourne office of a quantity surveying firm. The sample included many projects, of which 40 had sufficient data for analysis. S2 differed from S1 in that it covered a wide range of contractors, using differing contractual arrangements. The data included the type of construction, which allowed for a more detailed breakdown into subcategories. The projects were constructed in the period from the late-1960s to the mid-1970s.

Two further samples were examined during the course of the analysis. Sample 3 (S3) was derived from S2 by removing outliers. Sample 4 (S4) is S1 adjusted for inflation – through the use of a building cost index.

Goodness of fit

In order to draw comparisons between this and other models, and in order to examine its performance, it is necessary to develop a measure of the goodness of fit. The measure chosen, put forward as a risk index by Jepson (1969) and given the acronym 'SDY' by Berny and Howes (1982), is the standard deviation about the estimate of Y. SDY adopts the common measure of dispersion

$$\sigma = \sqrt{(\text{var})} = \sqrt{\{\Sigma(X-\bar{X})^2/N\}} \tag{11}$$

The fitted model is then declared to be the measure of central tendency or conditional mean of the Y values of the data. Therefore $\bar{X}$ is replaced by the estimated (or fitted) value of Y (Y_E) for a given X co-ordinate. The variance being measured is of Y about Y_E.

The rationale underlying this step is that for each point X there is a true mean value of Y and that any deviance can be explained by random, normally distributed error. A systematic error would imply a limitation of the fitted model.

The measure of dispersion for the model is therefore

$$SDY = \sqrt{\{\Sigma(Y-Y_E)^2/N\}} \tag{12}$$

This measure permits models to be compared: where the model with the lowest SDY value demonstrates the best fit and is therefore the most desirable.

Exclusion of data at the extremes

The nature of the logit transformation is such that as the data approaches either 0% or 100%, then the logit will approach $-\infty$ or $+\infty$ respectively. One of the limitations of linear regression is that any extreme values will dominate the analysis, to the extent that α and β values might be unduly influenced by a small number of extreme values rather than by the bulk of the data.

Within cash flow analysis, the extreme data points are arguably the least significant whereas they are the most dominant in the regression analysis. A simple method for countering the problem is to exclude the data points outside an acceptable range from the analysis. An appropriate cut off percentage for end point data (to be referred to as the exclusion range) will be found by trial and error, using the *SDY* as an indicator.

Design and procedure

The conventions used when collecting and analysing the data in the current analysis are set out below so that comparisons may be made, and the tests repeated. Time was measured in calendar days. It was assumed that the Lorenz value ogive for working days would approximate that for calendar days. This approach ignores the effect of the four week holiday taken in January by the Australian construction industry, which may result in discontinuites in the ogives for specific projects. As project data for days worked was not available, this crude assumption was considered necessary and the results indicate it was adequate.

The base figure for 100% value was taken to be the final amount certified; the equivalent figure for time was taken to be the time at which 100% value was certified, this point was considered preferable to the date of practical completion sometimes proposed (for example by Balkau, 1975; Bromilow, 1978; Bromilow and Henderson, 1977; Tucker and Rahilly, 1982) because the latter is irrelevant to a flow of cash and would be meaningless in a net cash flow model. The origin was taken to be the commencement of work on site, a convenient and easily identified point in time. All projects have an origin, and an end, and thus a model which incorporates these points would be advantageous.

The projects were built during a period of economic instability, with associated high inflation. For this reason the impact of inflation on the model was used as a measure of the model's ability to adapt to varying data conventions. By comparing results for projects before and after adjustment for the Building Cost Index (BCI), a direct comparison could be made to see if the goodness of fit was affected.

The procedure for testing the model involved the calculation of α and β values. The *SDY* measure was used as an indicator of comparative accuracy.

A systematic trial and error process was used to locate the optimum exclusion range. The analysis was run 31 times for each project in each sample (S1 and S2), excluding 0% to the 30% ranges in steps of 1%. The lowest mean sample *SDY* value indicated the optimum exclusion range. The mean sample *SDY* could then be used to compare the two sample groups using a non-parametric test – the W test (Wonnacott and Wonnacott, 1972) – and to compare the fit of the model for sample one, before and after adjustment for inflation (S1 and S4).

The distribution of the *SDY* values for the projects, with extreme data excluded, formed the basis for a further non-parametric test (the Boxplot test available on MINITAB – see Hoaglin, Mosteller and Tukey, 1983; Velleman and Hoaglin, 1981) designed to identify outliers in the population, which are projects for which the model does not fit the data with a statistically acceptable *SDY*.

A random sample of profiles was selected from each of S1 and S2 to demonstrate the diversity and extent of variation attained from the data, and thus the variation between projects.

Finally the results of a traditional nomothetic standard curve analysis were compared with the results of the idiographic analysis. In this experiment an average curve was calculated for each of S1 and S3. Then the *SDY* for each project, based on the average curve as the line of central tendency, was calculated and contrasted with the individual *SDY*s already achieved.

In summary the following experiments are reported in the following order:

(i) testing for optimum cut off point using a trial and error basis and mean sample SDY as an indicator;

(ii) examining the samples for significant differences which would indicate they were from separate based populations;

(iii) using the model derived from (i) to identify and remove projects which may be considered outliers according to a non-parametric test;

(iv) examining the influence of inflation on project *SDY* for the model through the use of a Building Cost Index which tests the prediction that the model is equally good for differing data conventions;

(v) assembling a random selection of profiles from S1 and S2 and inspecting for variation in the fitted models of individual projects; and

(vi) comparison of deviation from a calculated average curve with deviation from a project specific curve using the *SDY* measure, to demonstrate the failure of a group average model.

Results and discussion

The results of experiment (i) proved to be very significant. Fig. 3 illustrates the mean sample *SDY* values for the trial and error analysis, and the graph shows that a lower mean sample *SDY* is achieved if the upper and lower 13% (S1 and S4) or 11% (S2 or S3) extremes of data are

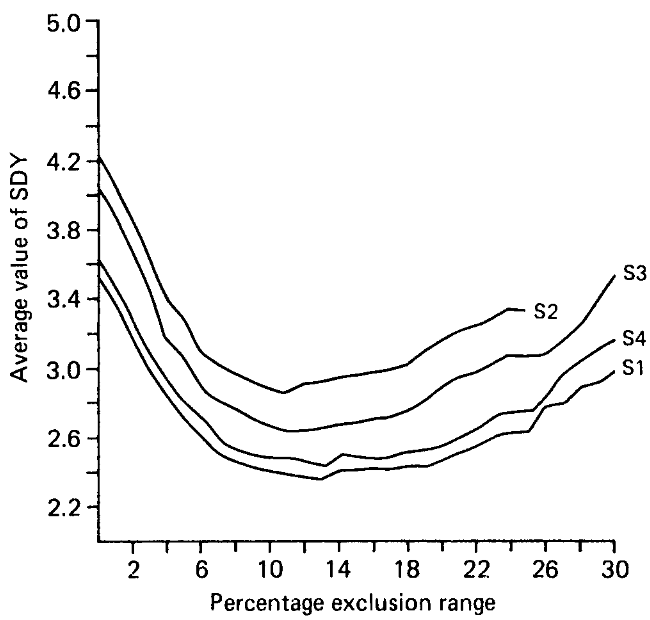

Fig. 3. Mean sample *SDY* against exclusion range.

excluded from the analysis. It was decided to choose 10% as the optimum exclusion range, because the curves reach a minimum from about 10% to 16% and it is desirable to select the percentage which utilizes as much of the data as possible, without affecting results detrimentally. Table 1 compares the mean *SDY* and standard deviation values for the model for all data, and for the 10% exclusion range. These figures show clearly that not only are the mean sample *SDY* values lower, but the standard deviation of the mean sample *SDY* also decreases markedly, at the 10% exclusion range. This means that at the 10% exclusion range the model is not only a better fit but has a smaller range of *SDY* values. Thus fewer projects are likely to be excluded as outliers by the *SDY* measure.

Table 1. Mean sample *SDY*

Sample	Exclusion range (%)	Mean sample *SDY*	σ
S1	0	3.52	1.61
	10	2.40	0.81
S2	0	4.23	2.35
	10	2.88	1.30
S3	0	4.05	2.25
	10	2.66	0.88
S4	0	3.62	1.66
	10	2.48	0.88

The lower mean sample *SDY* and standard deviation for S1 as compared with S2 prompts the question, is the model achieving significantly different results for each sample? This was tested in experiment (ii) by the use of a non-parametric W-test (Wonnacott and Wonnacott, 1972) chosen for its ability to cope with suspected outliers. The test examined whether the two underlying populations were centered differently, or were one and the same. The results of the W-test indicated that with a 95% confidence the samples were from one and the same population. This result was supported by a *t*-test. Thus there is no support for the treatment of the results from the model as coming from independent samples.

Given that the samples could be treated as one, the samples were examined in experiment (iii) for the existence of outliers. Another non-parametric test, the Boxplot (MINITAB) shown in Fig. 4, found no outliers in S1, but two possible outliers in S2. It can therefore be concluded that of the 72 projects tested, only two failed to be fitted adequately. Fig. 4 illustrates the Boxplot output for S1 and S2 combined; the asterisks identify outliers. Effectively an SDY of 6% can be deemed the determinant SDY value for outliers, and thus these are projects for which the model is unreliable.

The prediction that the model can adequately handle differing data conventions is supported by the results of experiment (iv). There was no significant difference in the goodness of fit for S4 as compared with S1. The results of experiment (iv) shown in Table 1 and Fig. 3 suggest a slight decrease in the average goodness of fit, but this can be rationalized. The reduced performance may be a function of the construction of the building cost index, which is an independent

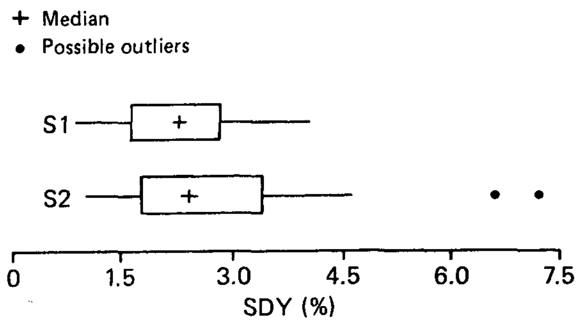

Fig. 4. Boxplot of project *SDY* values.

measure, divorced from the specific needs of a project, and applied at discrete intervals which do not always correspond to payments.

The visual inspection required for experiment (v) supports the experimental hypothesis. Fig. 5 illustrates a random sample of actual project curves produced by the model in this analysis, and demonstrates the wide range of profiles possible. The 'ideal sigmoid' is an illusion, cash flows may be concave or convex as well as the more common 'S'-shape. The projects vary in slope and lag, and some demonstrate a prolonged start or conclusion; one curve was an inverted sigmoid; whilst others do approximate to the ideal curve reasonably well. It is unfortunate that space does not allow for the publication of all of the graphs modelled within this work. However, as the sample demonstrated a wide range of profiles, the α and β constants for each project have been included in Table 2, so that the graphs may be reproduced by the interested

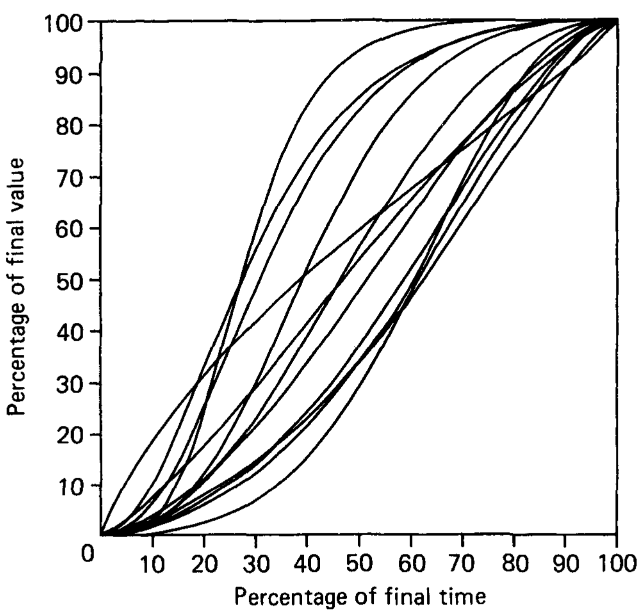

Fig. 5. Random sample of project curves.

Table 2. α and β for S1 and S2 with associated *SDY*

S1			S2		
α	β	*SDY*	α	β	*SDY*
0.94	2.13	1.65	−0.82	1.83	1.81
−0.20	1.98	2.89	−0.57	1.27	1.60
0.07	1.24	1.57	−0.69	1.51	3.23
0.37	1.46	3.65	0.36	2.30	3.42
−0.15	1.43	2.72	0.23	3.86	4.09
0.13	1.38	3.71	0.98	1.45	2.72
0.38	0.84	3.69	1.10	2.12	2.26
−0.31	1.34	1.42	2.83	2.86	2.63
0.36	1.26	1.03	0.77	1.88	1.88
−0.23	1.25	2.77	−0.05	1.67	1.55
0.42	1.72	1.92	0.69	2.09	4.10
0.25	1.30	2.41	0.65	1.80	3.38
0.23	1.71	3.16	1.58	1.94	4.66
−0.16	1.65	1.85	0.67	1.85	7.33*
−0.25	1.18	2.37	−0.38	1.52	1.91
0.52	1.78	1.28	0.44	1.81	1.97
0.68	2.35	3.15	−0.82	1.13	2.18
−0.54	1.50	2.47	0.86	2.15	2.27
−0.69	1.63	4.11	1.76	1.80	3.52
1.00	1.75	2.26	−0.53	1.34	2.56
−0.19	1.46	2.86	0.65	1.34	2.64
0.50	1.35	2.00	−1.05	0.96	3.69
−0.76	1.21	2.66	1.81	2.40	4.29
−0.68	1.30	1.86	0.66	1.53	2.68
0.00	1.26	1.32	−0.10	1.44	2.26
0.29	1.44	1.74	−0.29	1.58	2.24
0.59	1.74	1.90	1.27	1.78	1.64
−0.51	1.39	1.85	−0.27	1.43	1.42
−0.24	1.24	2.42	−0.83	1.51	1.87
−0.90	2.01	1.72	0.06	2.40	1.92
0.76	1.31	3.51	0.14	1.22	2.34
0.73	1.71	3.05	0.69	1.77	3.50
			0.12	1.16	3.55
			1.41	1.91	6.80*
			−0.19	1.06	3.76
			0.81	1.72	2.70
			−0.66	1.31	1.22
			−0.04	1.81	2.19
			0.83	2.44	2.73
			0.71	1.46	2.70

**Possible outlier.*

reader. These findings support inferences which may be drawn from the work of Singh and Woon (1984) whose graphs illustrate a high degree of inter-project variability.

A powerful argument in support of the idiographic methodology is provided by experiment (vi). Figs 6 and 7 illustrate the scatter plots for all projects in S1 and S3 (outliers were excluded from this analysis). The lines shown are the lines of best fit for each sample. A casual inspection would find the results to be in support of a nomothetic approach as the fitted lines appear as a

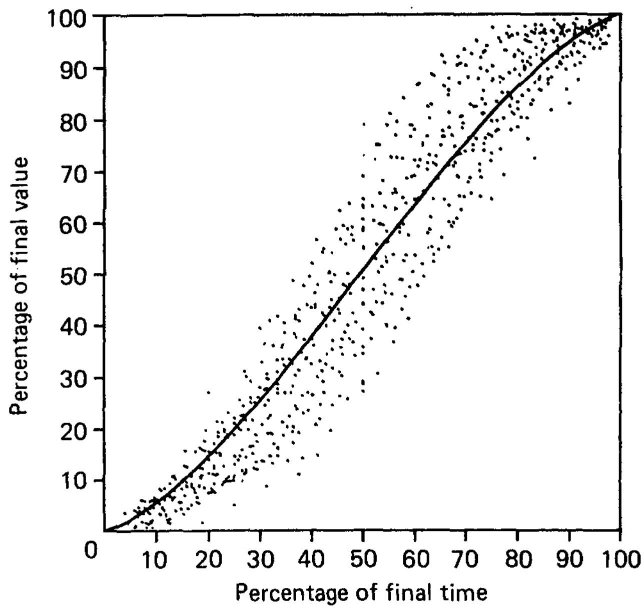

Fig. 6. Standard curve analysis for S1.

good average of all the points. However it can be seen from Fig. 5 that the lines which connect the points, while passing through the overall envelope of possible points, can do so in significantly different ways, and thus the average curve is misleading. It should be noted that the best fit lines for this experiment excluded data in the exclusion range, a technique designed to optimize fit for individual projects. The resultant fit is poor in this range but this does not detract from the argument above.

The contrast between the average curve and the individual curve is demonstrated by the results shown in Table 3. This shows the *SDY* values for each project based on, firstly the average curve as the central tendency, and secondly the individual curve as the central tendency. In all but one case the *SDY* value for the general model is higher than the *SDY* value for the specific model. In 80% of projects (69% for S1 and 89% for S3) the fit for the general model exceeds the SDY value set as the determinant for outliers for the specific model. In other words the fit of the general model is only unacceptable for 20% of projects. These results support the contention made, in the discussion of the principles of regression analysis, that project data deviates from a central mean by a combination of both systematic (resulting from the individual ontology of the project) and random error.

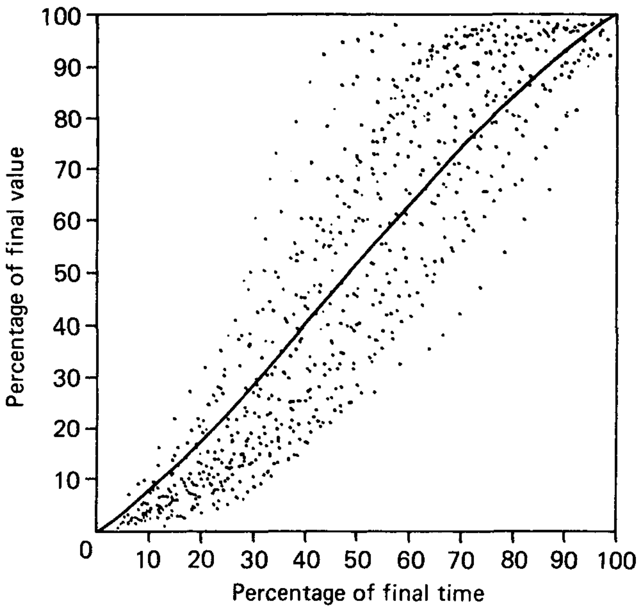

Fig. 7. Standard curve analysis for S3.

General discussion

An idiographic construction project cash flow model has been developed and successfully tested within this paper. The model has achieved excellent results over large samples, and only two of 72 projects were rejected by the model. The variation from the model was restricted to an *SDY* ranging from 1.0% to 4.6% (excluding outliers) for individual projects, with a median of 2.5% approximately (including outliers), which suggests that there is little systematic error left in the model. In contrast the results of experiment (vi) suggest that random error is insufficient to explain variation in a nomothetic model, which must therefore contain substantial systematic error. It is unfortunate that it is not possible to compare the goodness of fit for the model presented with other models, as *SDY* results have not hitherto been published.

Several rationales for the introduction of the idiographic methodology have been presented, the first of which is philosophically based. The nomothetic approach assumes there are consistent similarities between projects and the proponents of this approach then produce what are viewed as non-transient industry averages for groups of projects. This rationale ignores idiosyncratic differences between projects, discounting their significance by treating such variation as random (hence implying unimportant) error. If there exist for groups of projects discoverable, consistent, non-transient industry averages, the error component of which is truly random, then the prevailing useage of the nomothetic models, for predictive purposes, is justified. Conversely, if the above assumptions are violated, nomothetic prediction is invalid and probably meaningless.

It is contended that the conditions required (for predictive purposes) are not fulfilled. This

Table 3. Comparison of *SDY* for individual and average models for each sample

S1		S3	
Average	Individual	Average	Individual
13.60	1.65	14.38	1.81
7.99	2.89	11.24	1.60
1.97	1.57	12.42	3.23
6.38	3.65	12.77	3.42
3.82	2.72	17.40	4.09
5.29	3.71	15.64	2.72
11.55	3.69	17.30	2.26
5.46	1.42	29.82	2.63
6.41	1.03	12.71	1.88
4.10	2.77	6.27	1.55
6.83	1.92	13.65	4.10
4.55	2.41	12.19	3.38
6.41	3.16	21.90	4.66
5.08	1.85	8.14	1.91
5.17	2.37	8.71	1.97
8.43	1.28	15.35	2.18
12.01	3.15	14.58	2.27
9.33	2.47	24.45	3.52
12.23	4.11	10.16	2.56
14.36	2.26	11.04	2.64
4.44	2.86	20.93	3.69
9.03	2.00	24.69	4.29
13.19	2.66	10.34	2.68
11.12	1.86	4.80	2.26
1.43	1.32	7.59	2.24
4.75	1.74	18.06	1.64
9.44	1.90	6.77	1.42
8.00	1.85	15.41	1.87
4.62	2.42	10.90	1.92
12.94	1.72	2.91	2.34
13.12	3.51	12.21	3.50
11.86	3.05	3.48	3.55
		4.43	3.76
		13.23	2.70
		12.10	1.22
		7.87	2.19
		14.66	2.73
		12.14	2.70

paper demonstrates that variation between projects is a product of their individuality rather than a random error about an established ideal. Therefore there is no such ideal curve for groups. Furthermore the process of updating ideal curves, by the recalculation of parameters at intervals, is simply the selection of a new sample – and not an allowance for time.

The contention that an idiographic methodology is more appropriate than a nomothetic methodology for the study of construction project cash flows, has therefore been supported. The projects examined have yielded individual profiles, which support Hardy's (1970) contention that no close similarities exist between projects. It is the authors' belief that group models are both functionally as well as conceptually in error.

The above conclusion may have severe repercussions within the industry. Nomothetic models have achieved widespread support, especially where forecasting is concerned. However, consistent non-transient industry averages, with only a random error component of deviation, have not been found in the present work. Therefore standard or average curves do not reflect the individual projects, and cannot be used for forecasting.

The industry has expressed a need to predict construction project cash flows, and will continue to do so, despite any arguments that present methods are invalid. The desire of the industry to forecast cash flows is entirely understandable. Fortunately the construction industry already deals with such uncertainty in its everyday operations. Elaborate methods have been established to deal with such 'fuzzy' areas as time and cost estimating, and bidding. The role of the estimator is essentially that of predicting the unpredictable, and therefore encapsulates cash flow estimating.

To suggest that estimating differs in some way from forecasting and prediction may appear pedantic, but is fundamental to the debate at hand. The idiographic model provides a tool for the estimator, as indeed does the nomothetic model, both of which can be combined with personal judgement to arrive at an estimate of future cash flows. There can be no objection to an estimator predicting cash flows based on the evidence available, providing that the limitations of this subjective decision process are recognized (for example that extrapolation from a nomothetic model is uncertain). There will always be a place for personal judgement within estimating in the building industry.

The logit cash flow model is extremely flexible mathematically. It allows for the development of a net cash flow model (which will be the subject of a future paper). It is equally functional under nomothetic and idiographic methodologies, and therefore can be used by the estimator in whatever form required.

The model presented has some shortcomings. Exclusion of data below 10% or above 90% may gain the optimum goodness of fit for the model, but further study of the data within these bounds may be advantageous. Indeed the development of a net cash flow model demands that the model be more sensitive in the early stages of a project. It is proposed that an allowance for late starts of payments be included in such a model. At this stage the model is purely descriptive, and does not yet explain the variation between projects, although it may assist in interpretation. Knowledge of the effects which various factors might have on the profiles for individual projects would be of utmost benefit to estimators, and would merit further enquiry.

Another possible development of the model would be for in-progress cash flow forecasting. This would be a tendency for the project cash flow to adopt its final sigmoid form at an early stage in its development, which would imply that the final profile could be known before completion (providing final completion could be adequately estimated). The time at which the α and β parameters could be reliably derived, and the accuracy of the result, would be of significant value to the manager.

In conclusion, the logit transformation construction project cash flow model has been found suitable and accurate in idiographic *post-hoc* analysis of project cash flows.

Acknowledgements

The authors wish to acknowledge the support of the Department of Architecture and Building, and in particular the encouragement of Professor J.C. Scrivener. Thanks go to Ken Sharpe for mathematical and statistical advice, and to Catherine Dossetor for her constructive criticism, and acknowledgement is also due to the Commonwealth Post Graduate Awards scheme for support of the PhD candidature of Russell Kenley.

References

Ashley, D.B. and Tiecholz, P.M. (1977) Pre-estimate cash flow analysis. *Journal of the Construction Division*, ASCE, Proc. Paper 13213, 103 (CO3): 369–379. See also discussion by Gates, M. and Scarpa, A. (1978), 1 (CO1): 111–113, and closure by Ashley *et al.* (1978), 104 (CO4): 554.

Ashton, W.D. (1972) *The Logit Transformation*. Griffins statistical monographs and courses, No. 32. Griffin, London.

Balkau, B.J. (1975) A financial model for public works programmes. Paper to national ASOR conference, Sydney, August 25–27.

Berdicevsky, S. (1978) Erection cost flow analysis. MSc Thesis, Technion, Israel Institute of Technology, Haifa, Israel.

Berny, J. and Howes, R. (1982) Project management control using real time budgeting and forecasting models. *Construction Papers* **2**, 19–40.

Bromilow, F.J. (1978) Multi-project planning and control in construction authorities. *Building Economist*, March, 208–13, also September, 67–70.

Bromilow, F.J. and Henderson, J.A. (1977) *Procedures for Reckoning the Performance of Building Contracts*, 2nd edn. CSIRO, Division of Building Research, Highett, Australia.

De Groot, A.D. (1969) *Methodology – Foundation of Inference and Research in the Behavioral Sciences*. Mouton & Co., Belgium.

Drake, B.E. (1978) A mathematical model for expenditure forecasting post contract. In *Proceedings of the Second International Symposium on Organisation and Management of Construction*. Vol. 2. Technion, Israel Institute of Technology, Haifa, Israel, pp. 163–83.

Hardy, J.V. (1970) Cash flow forecasting for the construction industry. MSc Report; Dept. of Civil Engineering, Loughborough University of Technology, UK.

Hoaglin, D.C., Mosteller, F. and Tukey, J.W. (1983) *Understanding Robust and Exploratory Analysis*, John Wiley & Sons, Inc., New York.

Hudson, K.W. (1978) DHSS expenditure forecasting method. *Chartered Surveyor – Building and Quantity Surveying Quarterly* **5**, 42–5.

Hudson, K.W. and Maunick, J. (1974) Capital expenditure forecasting on health building schemes, or a proposed method of expenditure forecast. Research report, Surveying Division, Research Section, Department of Health and Social Security, UK.

Ireland, V. (1983) The role of managerial actions in the cost time and quality of performance of high rise commercial building projects. PhD Thesis, University of Sydney, Australia.

Jepson, W.B. (1969) Financial control of construction and reducing the element of risk. *Contract Journal*, April 24, 862–4.

Kennedy, W.B., Anson, M., Myers, K.A. and Clears, M. (1970) Client time and cost control with network analysis. *The Building Economist* **9,** 82–92.

Kerr, D. (1973) Cash flow forecasting. MSc report; Department of Civil Engineering, Loughborough University of Technology, UK.

McCaffer, R. (1979) Cash flow forecasting. *Quantity Surveying*, August, 22–6.

Nazem, S.M. (1968) Planning contractor's capital. *Building Technology and Management* **6,** 256–60.

Peer, S. (1982) Application of cost-flow forecasting models. *Journal of the Construction Division*, ASCE, Proc. Paper 17128, 108 (CO2), pp. 226–32.

Peterman, G.G. (1973) A way to forecast cash flow. *World Construction*, October, 17–22.

Reinschmidt, K.F. and Frank, W.E. (1976) *Journal of the Construction Division*, ASCE, Proc. Paper 12610, 102 (CO4), pp. 615–27.

Runyan, W.M. (1983) Idiographic goals and methods in the study of lives. *Journal of Personality* **51,** 413–37.

Singh, S. and Woon, P.-W. (1984) Cash flow trends for high rise building projects. In *Organising and Managing Construction*. Proceedings of the 4th International Symposium on Organisation and Management of Construction, University of Waterloo, Canada.

Spirer, H.F. (1975) *Business Statistics: a Problem Solving Approach*. Richard D. Irwin, Illinois.

Tucker, S.N. and Rahilly, M. (1982) A single project cash flow model for a microcomputer. *Building Economist*, December, 109–15.

Velleman, P.F. and Hoaglin, D.C. (1981) *Applications, Basics and Computing of Exploratory Data Analysis*. Duxbury Press, Boston, Massachusetts.

Weisberg, S. (1980) *Applied Linear Regression*. John Wiley & Sons, New York.

Wonnacott, T.H. and Wonnacott, R.J. (1972) *Introductory Statistics for Business end Economics*, John Wiley & Sons, New York.