Multiplicative regression models of the relationship between accounting numbers and market value

Michael Falta & Roger J. Willett[1]
Department of Accountancy & Business Law, University of Otago, New Zealand.
And
School of Accounting and Corporate Governance, University of Tasmania, Australia
e-mail: michael.falta@otago.ac.nz, roger.willett@utas.edu.au

[1] In alphabetical order. Based on papers presented to the Australasian Meeting of the Econometric Society Brisbane, University of Queensland, 3 - 6 July, 2007; European Accounting Association 32nd Annual Congress in Tampere, Finland, 12-15 May, 2009.

Multiplicative regression models of the relationship between accounting numbers and market value

Michael Falta & Roger J. Willett[2]
Department of Accountancy & Business Law, University of Otago, New Zealand.
And
School of Accounting and Corporate Governance, University of Tasmania, Australia
e-mail: michael.falta@otago.ac.nz, roger.willett@utas.edu.au

Abstract

The validity of ordinary least squares (OLS) estimates of relationships between accounting numbers and market value made in capital market research (CMR) using linear, additive models is questioned. Multiplicative models are argued to be more consistent with underlying economic theory for long-lived firms. Annual cross-section and firm-specific dynamic regression models of market on accounting values are estimated in levels and returns, using a selected panel of 30 of some of the largest long-lived USA firms over a 50 year period. Multiplicative models of levels data produce markedly improved statistical specifications compared to additive forms. Lags are also shown to be necessary to produce well-specified models of the relationship between accounting numbers and market value. Deflated returns models based on additive models are shown to suffer from additional problems of statistical inference. The consequences of using misspecified additive models of the relationship between accounting numbers and market value, when data generating processes (DGPs) incorporate multiplicative relationships, is that the size effect of the coefficients is misinterpreted. This is illustrated using analysis and computational experiments. Attention is drawn to the importance of the assumption of homogeneous firm parameters in cross-section estimation.

Key words: *Regression analysis; Capital markets research; Misspecification; Multiplicative functional form*

[2] In alphabetical order.

1. Introduction

This paper questions the validity of using additive, linear regression models of the relationship between accounting numbers and market value in CMR. Two common additive model specifications are compared to a multiplicative specification.[3] It is argued that the multiplicative specification provides an explanation of the relationship between accounting numbers and market value for an important class of large, long-lived USA firms more consistent with fundamental economic theory and empirical evidence.

To test this proposition, a selected sample of panel data for 30 such firms over a 50 year period is used to estimate additive cross-sectional and dynamic models of the relationship between accounting data in levels and market value, similar to those used by Barth et al. (1998) and others. These are shown to be statistically misspecified. Furthermore, specification is markedly improved by a multiplicative formulation of these models with lagged variables, similar to the type reported in Alexander et al (2011) and Cooke et al (2008). If underlying data generating processes have the same functional form and parameters are constant, so that the OLS linear cross-section model is appropriate, and if there is a long-run relationship between market and accounting values, well-specified returns regressions must also take the form of a cross-sectional version of a multiplicative error correction model between these variables. In these circumstances, the ubiquitous market returns models created by deflating the additive model by opening market value as in Easton and Harris (1991) must also be misspecified.

The consequences of misspecification using the additive linear model in our firm sample are illustrated by analysis and computational experiments. The slope coefficient of the accounting variable estimated using an additive linear model, when the DGP is multiplicative, can be quite different from the true value of the real parameter of interest.

[3] By additive (multiplicative) functional forms we mean those forms in which the relationship between the independent variables and between these and the error term is additive (multiplicative).

In the case of returns regressions, the additive model suffers from additional problems. Dividing its variables by opening market value makes the returns model statistically inconsistent with the levels model from which it is derived. It also produces greater than assumed risks of incorrectly inferring the existence of a significant, but actually non-existent, relationship between the dependent and independent variables. Inferential statistics such as the *t*-statistic and R^2 become unreliable, with non-standard sampling distributions. Furthermore, using larger samples to estimate the incorrectly specified models exacerbates these problems. Obviously, this has important implications for inferences made on the basis of the large cross-sections of data often used in CMR.

It is not proposed that all firms' DGPs can be modelled by a multiplicative form similar to that described in this paper. However, no matter what is the form of such DGPs, any useful cross-section model of the relationship between accounting numbers and market value of a sample of firms must be capable of producing valid and reliable estimates of their average parameter values. Hence, there must be some commonality of functional form for the DGPs of firms in the sample if cross-section analysis is to work at all. If the relationship between accounting numbers and market value in larger samples of firms can be approximately represented by a multiplicative statistical model, as in the sample studied here, additive models are clearly wrong. Alternatively, if some DGPs are multiplicative and others are additive, or are more generally of still different forms, then they should not be included without question in large, randomly selected samples with the expectation that cross-section models will yield valid and reliable estimates of parameters of interest.

The next section reviews some prior research in CMR relevant to the issues outlined above. Section 3 describes the economic reasoning behind the multiplicative model of the relationship between accounting numbers and market value. Section 4 explains the data sources and methods used to analyse the sample data. Section 5 compares cross-section and dynamic, sample estimates and specification statistics for additive and multiplicative models of the relationship between accounting numbers and market value. Section 6 describes some of the implications of misspecifying multiplicative relationship between accounting numbers and market value by additive models, assuming underlying firm parameters are constant. Section 7 summarises the conclusions of the paper and discusses additional problems associated with the assumption of firm parameter homogeneity, which is implicit in much cross-section analysis in CMR.

2. Related prior research

Kothari (2001) and Richardson & Tinaikar (2003) review ‘valuation and fundamental analysis research’ relevant to this area of the literature. In the now vast literature using cross-section regression models of the type discussed in this paper, two are particularly relevant as a point of comparison for the alternative multiplicative formulation we propose. Barth et al. (1998) approach the problem of how to obtain meaningful and reliable estimates of the relationship between ‘levels’ of market and accounting values by formulating a type of model, commonly seen in the CMR literature, which is here referred to as an ‘additive linear’ model. Easton and Harris (1991) use ‘returns’ formulations of similar levels models, in which the variables are ‘deflated’ by dividing through by opening period share price or market value. The returns formulation is either used for substantive reasons (i.e., the interest is in market returns rather than market values) or in an attempt to address a perceived problem with ‘scale’ in levels regressions (as recommended by Christie, 1987).[4]

[4] Easton and Sommers (2003) take the unusual step of deflating by closing market value, claiming that this reduces or eliminates bias in coefficient estimates in these models.

Generic descriptions of models of the type used in the Barth and Easton papers cited above are, respectively,

$$M_{i,t} = a_{i,t} + b_{i,t}A_{i,t} + u_t \tag{2.1}$$

and,

$$\frac{M_{i,t}}{M_{i,t-1}} = a'_{i,t} + b'_{i,t}\frac{A_{i,t}}{M_{i,t-1}} + w_t. \tag{2.2}$$

$M_{i,t}$ is the market value of firm *i* at time *t*; $A_{i,t}$ is a vector of accounting and other firm *i*-specific variables at *t*; the estimates of the parameters *a*, *b*, *a'* and *b'* are vectors of constant coefficients for all *i*; and u_t and w_t are white noise error terms. In cross-section applications the 'true' parameters may vary with *t*. In dynamic models the parameters may vary with *i*. Usually, one or both of these assumptions underpin inference, even though it may be admitted that their failure can lead to inaccuracies in estimation (e.g. the 'scale' arguments contained in Easton and Sommers (2003), Barth and Kallapur (1996) and Barth and Clinch (2009)).

As noted above, the model in Expression (2.1) is usually said to be in *levels* of the dependent variable while the *returns* regression of Expression (2.2) is associated with first time differences of market value. Further differencing of market value leads to the concept of abnormal returns, relative to expected returns. If $A_{i,t}$ is the net book value of assets, for example, Expression (2.2) regresses proportional raw returns on the book to market ratio. Many variants of these regressions are used in both cross-section and dynamic analysis, including forms of Expression (2.1) regressing returns rather than market value (e.g. Fama and French, 1992; Kothari and Shanken, 1997; Chen and Zhang, 2007).[5]

The basic justification for using regression models of the relationship between accounting numbers and market value in CMR is that they estimate the size of a change

[5] Dividends, for example, are frequently used to adjust either right-hand-side or the left-hand-side or both sides of Expression (2.2).

in the regressand, conditional on a unit change in the regressor. On purely descriptive grounds, Taylor's theorem can be used to justify using a linear model as an approximation to a more complex non-linear model.[6] Theory is then needed merely to identify likely candidates for the explanatory variables. More formal approaches attempt to derive a precise form for the relationship between accounting numbers and market value from underlying economic assumptions. Among these, Ohlson (1995) suggests net book value and abnormal returns as components in the vector $A_{i,t}$, but the form of the relationship implied by the standard interpretations of the theory is still additive linear. Burgstahler and Dichev (1997) and more generally Ashton et al. (2004) develop theories that imply non-linear forms for the relationship between accounting numbers and market value. However, virtually all empirical models in published CMR to date assume that the error term enters additively, with normality of the error term also assumed to allow inferences to be made about the statistical significance of model coefficients.

Pinning down the form of a generally acceptable, sensible, stable relationship between accounting numbers and market value regression empirically has proved difficult. Tests of Ohlson's 1995 model in its additive form tend to be inconclusive or negative (e.g. Myers, 1999; Dechow et al., 1999; Morel, 2003). Studies such as Easton and Harris (1991) and Easton et al. (1992) evidence variation in estimated coefficient values and inferential statistics over time in large annual cross-sections of data. This has been taken to imply that parameters in the underlying DGPs vary over time, adducing an economic interpretation to the variation in magnitudes and significance of book value and earnings coefficients (e.g. Collins et al., 1997; Barth et al., 1998; Francis and Schipper, 1999; Ryan and Zarowin, 2003).

[6] Econometrically, other interpretations are possible but the approximate linear model is probably the most common in the accounting literature (Hendry et al., 1984).

The absolute magnitudes of estimated parameter coefficients are less discussed than are changes in their relative magnitudes over time, level of significance and expected signs (e.g. Chen and Zhang, 2007). This avoids dealing with the central issue of fundamentals, since the absolute magnitude of the coefficients is exactly what is necessary to be estimated in order to understand the origin of firm value. Studies such as Kothari and Shanken (1997) and Barth et al. (1998) provide descriptive statistics for sample averages of relevant ratios (e.g. the market to book ratio) but the sense of the magnitude of estimated coefficients relative to the descriptive ratios is not extensively analysed, especially in the context of returns regressions.

It is difficult to get a general impression of the magnitude of the coefficients on book value and earnings across the studies cited, because they address different issues and use different variables and models. However, an example of how the failure to consider the absolute magnitudes of coefficient estimates obscures the central issue in fundamentals research can be seen by considering the estimates given in Easton and Sommers (2003). In their model, the accounting vector $A_{i,t}$ in Expressions (2.1) and (2.2) contains book value and annual earnings. The average coefficient estimates over the period from 1963 to 1999 for these two regressors in Expression (2.1) are 0.34 and 9.9, respectively. The sense of these estimates relative to the unconditional sample averages of the market to book and price earnings ratios, respectively, is unclear because there is no theory relating their magnitudes with one another, given that book value and earnings may be expected to be correlated.[7] What is puzzling, however, is the average magnitude of the estimated coefficients in the deflated model, Expression (2.2), over the same period of time. These are, book value 0.39 and earnings 1.49. Apparently, by dividing through

[7] If book value and earnings are not correlated, one would expect to recover the average market to book and price earnings ratios in the case of each coefficient (assuming the linear additive model is correct and all relevant explanatory variables are included in the model). Neither of the two estimated coefficients in Easton and Sommers (2003) appears to reflect likely market to book ratios.

the regressors by opening market value, the relative importance of a change in earnings compared to a change in book value has fallen by a factor of over six. It is not clear why the magnitudes of the coefficients should be altered by the deflation process. If the model is correctly specified they should not be affected.

Misspecification is suspected in such regression models because of uncertainties of the type just described. The most commonly mentioned possible causes of misspecification cited in the CMR literature are correlated omitted variables, heteroskedasticity, influential variables, outliers and a 'scale effect' (e.g. Kothari and Shanken, 1997; Kothari and Shanken, 2003; Barth et al., 1998; Easton and Sommers, 2003; Barth and Clinch, 2009). The acceptance of correlated omitted variables, as being a key reason for misspecification, is one of the main reasons CMR researchers perform sensitivity tests on coefficient estimates by including and excluding different variables in their models. Heteroskedasticity is usually dealt with by tests and adjustments, such as those recommended by White (1980). Influential variables and outliers are a particular feature of deflated models due to the distributional properties of ratios. The former may be identified, for example, by using studentised residuals (Easton and Sommers, 2003). Outliers can be treated by Winsorisation (e.g. Kothari and Shanken, 1997).[8]

Attempting to cure misspecification in the ways just described, can, especially if done piecemeal and without consideration for the possible effect on the distributional properties of the inferential statistics that are used to determine if a model is sound, especially *t*-statistics and R^2 (e.g. Brown et al., 1999), obscure true relationships and lead to a proliferation of different explanations that change with each new sample of data. This is likely to occur, for instance, if variables are added to a model in an *ad hoc* fashion to

[8] Recently, tests comparing different models on the basis of econometric criteria have been used to support model choice (Chen and Zhang, 2007). Although not misspecification tests as such, encompassing tests are useful measures for informing choice between valid models but they are not much use if none of the models compared are valid.

address the possibility of correlated omitted variables. Heteroskedasticity may be apparent, rather than real, caused by substantive misspecification in the functional form of a model.[9] Superficially dealing with heteroskedasticity by using adjusted *t*-values may encourage accepting an estimate as statistically significant when it is, in fact, not an estimate of the parameter of interest at all. Eliminating influential variables and outliers supposes that the extreme observations are unrepresentative of the population of interest, which may result in throwing out informative data. Why, for instance, might researchers dispose of observations as outliers when the data are transformed by deflation, when prior to the transformation they are retained in a model testing a similar hypothesis?

The issue of scaling, as discussed by Barth and Kallapur (1996), Easton and Sommers (2003) and Barth and Clinch (2009), and the focus on qualitative indications and significance tests in Chen and Zhang (2007) illustrate the lack of clarity in the literature about misspecification in regression models. Size or scale has long been supposed to be a problem in regression models of the relationship between accounting numbers and market value (Fama and French, 1992; Chen and Zhang, 2007). Barth and Clinch (2009) believe that the problem can be addressed by searching for deflators that eliminate an underlying factor that scales regressor variables. Easton and Sommers (2003) consider that scale simply is market value and that using the latter as a deflator eliminates scale by definition. Yet the scale problem cannot be a problem simply with the size of the explanatory variables. It is a basic fact of regression analysis that variability in regressor values is a good not a bad thing, from the point of view of acquiring estimation precision. If there is something wrong with scale in this sense, therefore, it is symptom, not a cause of the problem.

[9] Real or 'true' heteroskedasticity is a property of the model error term. It leads to inefficiency but not bias and inconsistency. Misspecification, for instance by correlated omitted variables, leads to biased and inconsistent estimates, not simply inefficient ones.

Specification tests of models should be comprehensive, covering as many assumptions of the maintained regression model as possible. This means testing for autocorrelation and normality of residuals, as well as heteroskedasticity, and performing tests of functional form and parameter constancy. In dynamic modelling, cointegration between the market and accounting variable is also a crucial point to check. In particular, it is important to check that cross-section models are consistent with the DGPs generating the data used in the cross-section analysis.

We adopt this comprehensive approach to testing the specification of the additive linear models in Expressions (2.1) and (2.2) following the approach to dynamic modelling detailed in Alexander et al (2011). The results are reported in Section 5, after a more detailed account of the statistical methods we use. In the next section we provide a theory to explain why we believe additive models do not form the basis for an accurate description of the relationship between accounting numbers and market value.

3. Theory

It is questionable if the assumption of an additive linear model discussed in the previous section is reasonable in the context of the pricing of financial securities. Most empirical and theoretical evidence suggests that, as security prices increase, their variance increases, and that their distribution is approximated by the lognormal distribution (e.g. Black and Scholes, 1973). Consequently, it is natural to ask what part accounting information might play in generating market values that have such characteristics in theory.

The approach we take to answering this question is based upon market value being the aggregate of two causes: accounting effects in which there are possibly short, as well as, long-run elements; and a short-run trading effect on the volatility of market value reflecting other information impacting on the demand and supply of shares in the stock market. As in fundamental economic theory, we assume that in the long-run, financial market returns equate to underlying real accounting returns (Miller and Modigliani, 1961). Consequently, the long-run effects of our model are captured in the accounting component and the non-accounting, short-run effects are centered on zero.

A simple functional form for the long-run relationship between accounting numbers and market value of the i^{th} firm in period t, compatible with lognormal market value and the above assumptions is,

$$M_{i,t} = f_{i,t}(A_{i,j,t})\omega_{i,t} \tag{3.1}$$

where $f_{i,t}(A_{i,j,t})$ is some valuation function of j accounting variables $A_{i,t}$, $\omega_{i,t}$ representing identically and independently distributed residual effects with a median value of 1, and both $f_{i,t}(A_{i,j,t})$ and $\omega_{i,t}$ are lognormal. For the rest of this paper, we make the constraint that the functional form of f is constant across firms and over time, so that the i and t subscripts can be dropped.

As an approximation to the process generating $f(A_{i,j,t})$ that will conform to the requirements specified for Expression (3.1) we assume a weighted geometric average of n accounting levels variables $A_{i,j,t}$, j = 1 to n, and for all i and t,

$$M_{i,t} = k_{i,t}\left(\prod_j A_{i,j,t}^{\beta_{i,j,t}}\right)\omega_{i,t} \tag{3.2}$$

where $k_{i,t}$ is a scaling factor, $\Sigma_j\beta_{,i,j,t} = 1$ and each $A_{i,t}$, for any set of j accounting variables, is generated from a joint lognormal distribution. $M_{i,t}$ in the long-run relationship (3.2) is therefore lognormal, assuming $f(A_{i,j,t}) = k_{i,t}\left(\prod_j A_{i,j,t}^{\beta_{i,j,t}}\right)$ and $\omega_{i,t}$ are independent, as required by the assumptions of OLS.

The long-run market-book relationship in Expressions (3.1) and (3.2) gives a straightforward interpretation of the model variables as combining multiplicatively, with an error term that behaves ‘intensively’ magnified by scale. The form is similar to the Cobb-Douglas production function in economics (Cobb and Douglas, 1928). If it is well-specified and the data are positive, the model can be estimated by OLS by taking logs of the variables in the model, as detailed in the next section. Log transformations are used routinely in specific and *ad hoc* contexts to achieve particular outcomes in CMR (e.g. to make a variable appear more Gaussian) but the simple expedient of transforming all variables to logs is not common.

Due to time delays, the disclosure of accounting information data could act as a filtration on the hypothesised diffusion processes, increasing the likelihood that the presumed long-run effect of accounting on market values, represented by the coefficient estimates $k_{i,t}$ and $\beta_{i,j,t}$ in Expression (3.2), may actually embody short-run effects. If such is the case, lags on the model variables may be required to provide the stationary properties necessary for valid assessment of long- and short-run effects. A great deal of empirical evidence shows that relationships between suitably transformed time series of data in economics are often statistically well-specified by an autoregressive distributed lag (Hendry et al., 1984). This suggests that the relationship between accounting numbers and market value for the i^{th} firm should in the first instance be checked for lags, i.e., be specified as

$$M_{i,t} = k_{i,t} M_{i,t-1}^{\alpha_{i,t}} \prod_j \left(A_{i,j,t}^{\beta_{i,j,t}} A_{i,j,t-1}^{\beta_{i,j,t-1}} \right) \omega_{i,t} \tag{3.3}$$

Here, the terms are as in Expression (3.2) and $\alpha_{i,t}$ is an autoregressive coefficient. It is assumed, for expositional purposes and for reasons of applicability to the present case, that single lags on the market and accounting variables are sufficient to produce a white noise error term. It is well-known that Expression (3.3) in the time domain and assuming it is well-specified, entails an error correction relationship between $M_{i,t}$ and $A_{i,t}$. The error correction model form for Expression (3.3) is the multiplicative returns formulation

$$\frac{M_{i,t}}{M_{i,t-1}} = \left(\frac{\kappa_{i,t}\prod_j A_{i,j,t-1}^{\varphi_{i,j,t-1}}}{M_{i,t-1}}\right)^{\lambda_{i,t}} \prod_j \left(\frac{A_{i,j,t}}{A_{i,j,t-1}}\right)^{\beta_{i,j,t}} \omega_{i,t} \qquad (3.4)$$

where the term in the first parentheses represents the imbalance in last year's long-run relationship between market and an accounting value that 'error corrects' the market value to its long-run equilibrium value $\kappa_{i,t}\prod_j A_{i,j,t-1}^{\varphi_{i,j,t-1}}$ at that date. The long-run parameters are defined by the original 'short-run' parameters in Expression (3.3) as $\kappa_{i,t} = k_{i,t}{}^{(1-\alpha_{i,t})}$ and $\varphi_{i,j,t} = (\beta_{i,j,t} + \beta_{i,j,t-1})/(1-\alpha_{i,t})$. The error correction coefficient $\lambda_{i,t}$ is defined as $(1-\alpha_{i,t})$. In a CMR context the long-run parameters measure the permanent or persistent impact on market value of fluctuations in the accounting variable.

In the cross-section, estimating the slope coefficient in Expression (2.1) does not provide an estimate of the accounting response coefficient, if Expression (3.3) is the correct common form of the firm DGPs. Estimating (2.2) in the same circumstances creates an extraneous constant term and confuses short-run error correction and long-run effects in the estimates of the slope coefficients. Estimates of the overall accounting 'response coefficients', the parameters *b* and *b'* in Expressions (2.1) and (2.2), respectively, are biased and inconsistent. These differences are not trivial. They have obvious policy, as well as, academic implications. We assess the statistical validity of the multiplicative model, relative to the additive model as an accurate description of long-lived firms in Section 5.

The empirical hypotheses to be tested in Section 5 with respect to both levels and returns may be summarised as follows. Hypothesis 1: Cross-section and dynamic additive linear models, as represented in Expressions (2.1) and (2.2), estimated on samples of large, long-lived US firms are statistically misspecified; Hypothesis 2: Cross-section and dynamic multiplicative models, as represented in Expressions (3.3) and (3.4), estimated on samples of large, long-lived US firms are not statistically misspecified. The data and methods we use to test these hypotheses are discussed next.

4. Data and method

This section of the paper describes the data and statistical methods used to estimate and test the specification of the cross-section and dynamic regression models of the relationship between accounting numbers and market value reported in Section 5. The sample data are selected, non-randomly, from Compustat files to include firms that approximate the long-lived firms assumed in the theory described in Section 3. All have 31 December financial year ends. At the time the sample was first obtained (2001) these were all the firms from the Standard and Poor's (S&P) 500 that had historical data over the 50 year period with no missing values. Given that they all existed over the 50 year period we consider them to be 'survivors'. The identity of the 30 firms in the sample is evident from the Tables later in the paper (e.g. Table 4).

The three variables, M_t, B_t and E_t used in the models are as observed at the balance sheet date of 31 December. Market value M_t is calculated as the common share price at the financial year end multiplied by the number of outstanding shares at the same date. These are raw data, unadjusted for splits, i.e., the values that would have been observed in the share markets at 31 December in each year. Reported book value of net assets B_t is the 'Common Equity' data field in Compustat. Reported accounting earnings E_t is Compustat's 'Net Income' variable. The accounting variables are as reported in the annual balance sheet date of 31 December. The simplicity of the data fields makes it easy to replicate all of the estimates reported in this paper. More specific information about the sources of data is given in Table 1.

[INSERT TABLE 1 ABOUT HERE]

Both cross-section and dynamic models use standard OLS estimation, inferential and diagnostic techniques. A summary description of the of the Augmented Dickey Fuller (ADF) tests for unit roots in the dynamic models and of the lesser-known misspecification tests are contained in Table 2. Full details are contained in Alexander et al (2011).

[INSERT TABLE 2 ABOUT HERE]

The multiplicative models of market value and market returns in Expressions (3.3) and (3.4) can be estimated directly using non-linear techniques or with OLS by taking logs of the variables. The latter approach is used because it provides easy access to well-tried estimation methods and standard inferential and specification tests. This requires, however, that the data points are positive. Out of the 1,500 data points from the entire period of analysis, there are five and sixty instances of negative book values and earnings, respectively. In such cases the log values corresponding to these negative values are set to equal -10.

With positive data, Expression (3.3) is transformed in logs to

$$\ln(M_{i,t}) = \ln(k_{i,t}) + \alpha_{i,t}\ln(M_{i,t-1}) + \sum_{\mathrm{j}} \beta_{i,j,t}\ln(A_{i,j,t}) + \sum_{\mathrm{j}} \beta_{i,j,t-1}\ln(A_{i,j,t-1}) + \ln(\omega_{i,t}) \qquad (4.1)$$

and the returns error correction formulation in Expression (3.4) to

$$\ln(\Delta M_{i,t}) = \sum_{\mathrm{j}} \beta_{i,j,t}\ln(\Delta A_{i,j,t}) - \lambda_{i,t}\{\ln(M_{i,t-1}) - [\ln(\kappa_{i,t}) + \varphi_{i,t}\sum_{\mathrm{j}} \ln(A_{i,j,t-1})]\} + \ln(\omega_{i,t}). \qquad (4.2)$$

If Expression (4.1) is statistically well-specified, OLS in a dynamic regression produces consistent estimates of its parameters with a bias in the estimate of the autoregressive term reducing as the length of the sample period increases. Running the ‘static’ regression, i.e., Expression (4.2) over time without any lagged variables, gives consistent OLS estimates of the long-run parameters κ and φ. However, these estimates are biased at finite sample lengths. Consequently, it is safer to include lagged variables in the dynamic specification of the relationship between accounting numbers and market value than to exclude them.

As explained in the previous section, if all firm DGPs in the sample have the form of Expression (3.3) with stable parameters over time and if the other assumptions of OLS are valid (this requires identical parameters for all firms) cross-sectional forms of Expressions (4.1) and (4.2) should return estimates of those parameters with the same expected value as the average dynamic estimates, perturbed only by random variation. More realistically, if the variation in DGP parameters across firms is sufficiently small, this result will continue to hold to an approximation, in that the average cross-section estimate is then close to the average dynamic estimate. In such circumstances, therefore, which we will refer to as the 'homogeneous parameter' case, it is possible to reliably assess the dynamics of the relationship between accounting numbers and market value by cross-sectional means alone. This is the case assumed in most cross-section work in CMR.

Virtually nothing is known about how small 'sufficiently small' needs to be to produce acceptably close estimates of averages of the parameters of interest in cross-section models having the form of Expressions (4.1) and (4.2). More is known about the likelihood of the cross-section, restricted, static form of Expression (3.3) and its log transformation, Expression (4.1), yielding unbiased estimates of the average long-run effect in levels of market value (e.g. Pesaran and Smith, 1995). Recently, 'random parameter' models have been estimated using simulated maximum likelihood and other techniques in panel data contexts (e.g. Greene, 2008: 222-43 and 619-23). Even here, though, it is assumed that there is a common functional form, like Expression (3.3) for all DGPs in a sample.

In this paper, since we use OLS, our estimates assume constant parameters between firms and over time. However, the estimation results can be assessed to see if these assumptions are reasonable in the context of the sample. We do this in the next section.

5. Misspecification in linear additive models of 'levels'

This section examines the sample data to test the hypotheses stated at the end of Section 3. We first consider the visual patterns in the data to demonstrate that these are consistent with the assumptions of the theory described in Section 3. Then we report the results of diagnostic tests of the additive and multiplicative models of the relationship between accounting numbers and market value showing the former are misspecified.

Figure 1 provides views of time series patterns of averages, across the 30 firms, of market value, net book value and earnings. Also shown is the geometric mean of book and earnings displaying clear evidence of an exponential growth pattern over the 50 year sample period (Figure 1, left). In the figure (right), we also exhibit corresponding histograms of the relative frequency distributions. Their patterns well-approximate lognormal distributions.

[INSERT FIGURE 1 ABOUT HERE]

Figure 2 shows the cross-section relationships between market value and three sets of accounting values based upon time averaged data for each individual firm. The cross-section relationship between market value and the accounting variables is approximately linear, whether or not the variables are transformed to logs. The impact of growth in the DGP of these variables is seen in the cross-section by *apparent* heteroskedasticity in the scatter plots of market value with the accounting regressors (Figure 2, top row).

[INSERT FIGURE 2 ABOUT HERE]

The patterns in the average data are representative of the individual series, whether this is over time by individual firms or in the cross-section by year across firms. Time averaging subdues the impact of heteroskedasticity in the cross-section.[10] The appearance of heteroskedasticity in the untransformed numbers and its disappearance when they are logged is more obvious in the annual data. This is true of virtually every one of the 50 years in the sample.

Figure 3 displays histograms of the raw and log data, averaged for each firm across time, for each of the variables. The raw data are more skewed and the log transform gives a more normal shape to them. The patterns satisfactorily approximate the lognormal and normal distributions, respectively.

[INSERT FIGURE 3 ABOUT HERE]

Barth et al. (1998) modelled 'levels' relationships between market and accounting values as follows,

$$M_{i,t} = k_{i,t} + \beta_{i,1,t} B_{i,t} + \beta_{i,2,t} E_{i,t} + \varepsilon_t. \tag{5.1}$$

$M_{i,t}$ is market value and $B_{i,t}$ the book value of net assets of firm i at time t, $E_{i,t}$ is annual earnings of firm i during t and t-1, and ε_t is a white noise error term. The result of using this model and generate statistically reliable estimates of coefficients is shown in Table 3, where it is applied in the cross-section to the 30 firms in our sample in each of the 50 years.

[10] The long windows used in Easton et al. (1992) probably have this effect.

[INSERT TABLE 3 ABOUT HERE]

In many years the coefficients test as being statistically significant, sometimes highly so. The R^2 are invariably high, usually above 70% and often over 80%. However, the coefficient estimates fluctuate from year to year, often changing sign. Such instability in coefficient estimates could be an indication of a misspecified model rather than being due to underlying systematic factors generating the data. The diagnostic statistics in the last three columns of Table 3 confirm that, as a model of the data, Expression (5.1) is statistically misspecified.

Estimating Expression (5.1) as individual firm *i* time series (dynamic approach) produces the same statistical indications of misspecification (see Table 4). Thus, based on the frequency of flagging significance of the specification tests, Hypothesis 1 cannot be rejected.

[INSERT TABLE 4 ABOUT HERE]

Running Model (5.1) in logs,

$$M_{i,t} = e^{k_{i,t}} B_{i,t}^{\beta_{i,1,t}} E_{i,t}^{\beta_{i,2,t}} \omega_t \tag{5.2}$$

i.e., implying a multiplicative process produces the results shown in Table 5 for the levels approach. Now the coefficients show a more stable and consistent pattern, summing to approximate unity in each year. Book values increase from negative to positive values during the sample period and the coefficient on earnings falls from values mostly in excess of 1 until 1980 to mostly small positive values after 1990. Earnings are more often statistically significant in the early years and book in the later years but that feature is affected by the closeness of the coefficients to zero in different years. The diagnostic tests flag misspecification in the model no more than might be expected by chance (at the five per cent level).

[INSERT TABLE 5 ABOUT HERE]

Dynamic regressions using the log-form of Expression (5.1) produce specification statistics that appear to show a well-specified model. However, in this case an additional test for non-stationarity of the residuals is required to be confident that the long-run relationship between the market and accounting variables is not spurious (e.g. see Greene, 2008: 756). ADF tests flag non-stationarity in 21 of the 30 firm models (see Table 6, last column). Adding a single lag for each variable to Model (5.2),

$$M_{i,t} = e^{k_{i,t}} M_{i,t-1}^{\alpha_{i,t}} B_{i,t}^{\beta_{i,1,t}} B_{i,t-1}^{\beta_{i,1,t-1}} E_{i,t}^{\beta_{i,2,t}} E_{i,t-1}^{\beta_{i,2,t-1}} \omega_t, \tag{5.3}$$

including the dependent variable, cures the problem of non-stationarity. The resulting model is well-specified and provides statistical evidence of a non-spurious, long-run relationship between the market and accounting values. Hypothesis 2 also, therefore, cannot be rejected.

[INSERT TABLE 6 ABOUT HERE]

Given that there are sixty occasions when the log of negative earnings is replaced by a near zero positive value, we check the robustness of the misspecification statistics in Tables 3 to 6 by comparing them to those produced by models in which book value is the only independent variable. In this case the strategy of replacing log book value by -10 occurs only five times. However, the incidence of misspecification remains noticeably higher in the additive model, despite the introduction of a lag on book value, see Tables 8 and 9 in the Appendix[11]. Consequently, the presence of negative earnings does not affect our conclusions regarding Hypotheses 1 and 2.

[11] Omission of a lag on book value as in (5.1) further increases the number of flagged misspecification tests.

6. Discussion

The clearest evidence for multiplicative models of the relationship between accounting numbers and market value lies in the apparent heteroskedasticity seen in scatter plots of cross-sections of market and accounting levels data (Figure 2). The figure shows the sample average market and accounting data. Data for *all* firms in the sample show similar patterns. The heteroskedasticity can be explained as the consequence of misspecification of a multiplicative model by an additive model. It causes bias and inconsistency in OLS estimates not just inefficiency, as would be the case of true non-spherical disturbances of the error term in a regression model (see Thursby, 1982). This point is demonstrated in the following.

Assume that data are generated by the simple multiplicative process

$$M_t = \ kA_t\omega_t, \tag{6.1}$$

where the variables are as previously defined, k is a scaling parameter and, in particular, ω_t is lognormal LN(0,1). Then the expected OLS estimate of the slope or response coefficient derived from a model incorrectly using Expression (2.1) is

$$\mathbb{E}\{\hat{b}\} = k\mathbb{E}\{\omega_t\}, \tag{6.2}$$

where $\mathbb{E}$ is the expectation operator. In this case, since $\mathbb{E}\{\omega_t\} = e^{\left(\mu+\frac{1}{2}\sigma^2\right)} = e^{\frac{1}{2}} \cong 1.65$,

$$\mathbb{E}\{\hat{b}\} = 1.65k. \tag{6.3}$$

Thus, if we assume k to equal 3, for example, wrongly using Expression (2.1) to estimate Expression (6.1), we obtain an estimate of the response coefficient of changes in market values to changes in accounting values of ~4.95. This bias is substantial and is not alleviated by increasing the sample size.

The consequences of such misspecification for inference and why large samples might produce misleading confidence in incorrect results are illustrated in the results of a computational experiment mirroring the situation described above. Figure 4 shows histograms from a Monte Carlo simulation experiment (Experiment 1) in which data are generated from the distributions with the parameters described in Table 7, Panel A. The sampling distributions for the coefficients and inferential statistics of the incorrectly specified linear additive model are highly skewed and located near to 5 for the distribution of possible slope values.

[INSERT FIGURE 4 AND TABLE 7 ABOUT HERE]

Despite the quite different conclusions that might be drawn from Expressions (2.1) and (6.1) about the relationship of market to accounting values, the *t*-statistic on the slope coefficient in the former model is significant at least at the 5% level in samples of 30 firms approximately 35% of the time. In samples of 300 firms, the slope coefficient continues to remain significant at about the same level. The between sample variation in the inferential statistics diminishes with increased sample size but is large, even in sample sizes of 300 firms, producing the impression of unstable coefficients. The average sample value of the *t*-statistic for the incorrectly specified model depends upon the particular values of the parameters of the underlying DGP and does not diminish with increases in sample size, as would happen with a correctly specified model. Therefore, if that value happens to be in excess of a chosen critical limit, the risk of being wrongly convinced that an incorrect model of this form is correct will increase with the size of the sample. If a model is misspecified, therefore, large sample sizes are not always beneficial.

If, as in the case of our sample, Expression (3.3) tests as being a well-specified model of the data, dividing Expression (2.1) by M_t to give Expression (2.2) provides an approximation to (3.4) - the error correction ‘returns’ form of Expression (3.3). The approximation is caused by M_t/M_{t-1} being close to unity, due to the high serial correlation between M_t and M_{t-1}. M_t/M_{t-1} is therefore close to its log value. OLS calculates an estimate of M_t/M_{t-1}, i.e., $a + bA_t/M_{t-1}$ that is also close to unity, so that this also is close to its log value. Additionally, on the assumption that M_t and A_t are, in fact, jointly log-normal, then both M_t/M_{t-1} and A_t/M_{t-1} are log-normal. Further, log-normal distributions with these characteristics have shapes similar to a normal distribution. Consequently, Expression (2.2) may appear to be well-specified, because it mimics the log form of Expression (3.3). However, the interpretation of the coefficients in Expression (2.2) should then be the same as for the log form of Expression (3.4). The accounting response coefficient in Expression (2.2) is not b but turns out to be a function of a and b. This observation is generally consistent with the order of the published parameter values estimated using the deflated additive linear models of the relationship between accounting numbers and market value (e.g. Easton and Sommers, 2003).

To these potentially inappropriate interpretations of the coefficients in Expression (2.2) must be added the possibility that it omits information in the lagged value of A_t. If it does so, short- and long-run effects of accounting on market value are ignored, so that, even if an otherwise correct interpretation is given to the coefficients, misleading conclusions about permanent and transitory effects of accounting for market value may be drawn.

An additional problem with using Expression (2.2) to draw inferences about the relationship between accounting numbers and market value is its propensity to create spurious relationships between the dependent and independent variables in the model. The root cause of the problem is simple: if two independent random variables are divided by a third, common random variable, correlations may be established between the resulting ratios that, in a regression analysis of one ratio on the other, may give the incorrect impression of a systematic relationship between the original numerator random variables. That this concern about validity of the model in Expression (2.2) has potentially serious consequences is illustrated by a second simulation experiment.

The main parameters of the second simulation experiment (Experiment 2) are described in Table 7, Panel B. This experiment produces three random walks. One, M_t, interpreted as market value, is generated to be independent of the other two. The second and third, B_t and E_t, interpreted respectively as book values and earnings, although unrelated to M_t, are related to each other. The starting value of E_t in the first period is 5% of the value of the starting value of B_t. Otherwise E_t is a simple random walk, $E_t = E_{t-1} + e_t$, where the error term e_t is a standard, white noise process. B_t is calculated as $B_{t-1} + E_t + v_t$, where v_t is also a standard white noise process. This defines a 'stochastic clean surplus' relationship between B_t and E_t. Lagged values of M_t are used to deflate all variables. Then M_t/M_{t-1} is regressed on B_t/M_{t-1} and E_t/M_{t-1} in accordance with the form of Expression (2.2).

Estimating this model produces an intercept term that is significant 95% of the time, which might be expected. However, it also produces wildly fluctuating parameter estimates for the earnings and book values coefficients which, although centering on their true values, have significant *t*-statistics at the 5% level between 25% and 30% of the time, on average. The average value of R^2 for this model is 44%. This is despite the fact that neither regressor is related to the dependent variable. Figure 5, top row, shows the erratic time sequence behaviour of the coefficient estimates and their associated inferential *t* and R^2 statistics. Figure 5, bottom row, displays some of the sampling distributions of these statistics, illustrating their non-standard forms.

[INSERT FIGURE 5 ABOUT HERE]

Expression (2.2) may in any given situation, correctly model to some degree of approximation, the relationship between returns and accounting numbers. However, the deflation adjustment makes it difficult or impossible to identify if the estimated relationship is real or spurious.

7. Conclusion

The empirical results in Section 5 and the discussion in Section 6 demonstrate that reported connections between market and other values in cross-section regression analysis ought to be treated with caution. Apparently strong, though intermittent relationships appearing in regressions of the type shown in Expressions (2.1) and (2.2) are unreliable unless supported by rigorous diagnostic testing and examination of the possibilities of artificial statistical relationships that are artefacts of technique. It is also important to investigate the underlying dynamics that drive the DGPs of each set of individual observations, to appreciate what cross-section parameter estimates are likely to represent, as some kind of average of the sample of interest.

We have assumed existence of a common form of DGP in the firms of our sample. This is consistent with the empirical evidence reported in Section 5. However, we have also assumed that the parameters of these DGPs are homogeneous, in the sense defined in Section 4. The accuracy of this assumption affects the reliability of OLS estimates, even if these are made with otherwise valid, well-specified cross-section models. We examine the effect of parameter heterogeneity in Falta and Willett (2009).

8. References

Alexander, D., Falta, M. and Willett, R. J. (2011) 'Using forecasting criteria to identify value relevance in the relationship between accounting numbers and market value'. *Abacus,* Forthcoming

Ashton, D., Cooke, T., Tippett, M. and Wang, P. (2004). 'Linear information dynamics, aggregation, dividends and 'dirty surplus' accounting'. *Accounting and Business Research*, 34(4): 277-299.

Barth, M. E. and Clinch, G. (2009). 'Scale effects in capital markets-based accounting research'. Available at SSRN: http://ssrn.com/abstract=201228.

Barth, M. E., Beaver, W. H. and Landsman, W. R. (1998). 'Relative valuation roles of equity book value and net income as a function of financial health'. *Journal of Accounting and Economics*, 25: 1–34.

Barth, M. E. and Kallapur, S. (1996). 'The effects of cross-sectional scale differences on regression results in empirical accounting research'. *Contemporary Accounting Research* 13(2): 527-567.

Black, F. and Scholes, M. (1973). 'The pricing of options and corporate liabilities'. *Journal of Political Economy*, 81(3): 637-654.

Brown, S., Lo, K. and Lys, T. (1999). 'Use of R^2 in accounting research: Measuring changes in value relevance over the last four decades'. *Journal of Accounting and Economics*, 28: 83-115.

Burgstahler, D. C. and Dichev, I. D. (1997). 'Earnings, adaptation and equity value'. *Accounting Review*, 72(2): 187-215.

Chen, P. and Zhang, G. (2007). 'How do accounting variables explain stock price movements? Theory and evidence'. *Journal of Accounting and Economics*, 43: 219-244.

Christie, A. A. (1987). 'On cross-sectional analysis in accounting research'. *Journal of Accounting & Economics*, 9: 231-258.

Cobb, C. W. and Douglas, P. H. (1928). 'A theory of production'. *American Economic Review*, 18(1): 139-65.

Collins, D. W., Maydew, E. L. and Weiss, I. S. (1997). 'Changes in the value-relevance of earnings and book values over the past forty years'. *Journal of Accounting and Economics*, 24: 39-67.

Cooke, T. E., Omura, T., and R. J. Willett, 'Consistency, Value Relevance and Sufficiency of Book for Market Values in Five Japanese Conglomerates Over the Period 1950-2004', *Abacus*, Vol. 45, No. 1, 2009.

Dechow, P. M., Hutton, A. and Sloan, R. G. (1999). 'An empirical assessment of the residual income model'. *Journal of Accounting and Economics*, 26: 1-34.

Dickey, D. A. and Fuller, W. A. (1981). ‘Likelihood ratio statistics for autoregressive time series with a unit root’. *Econometrica*, 49: 1057-1072.

Doornik, J. A. and Hansen, H. (1994) ‘An omnibus test for univariate and multivariate normality’. Working paper, Nuffield College.

Easton, P. D. and Harris, T. S. (1991). ‘Earnings as an explanatory variable for returns’. *Journal of Accounting Research*, 29(1): 19-36.

Easton, P. D., Harris, T. S. and Ohlson, J. A. (1992). ‘Aggregate accounting earnings can explain most of security returns’. *Journal of Accounting & Economics*, 15: 119-142.

Easton, P. D. and Sommers, G. A. (2003). ‘Scale and the scale Effect in market-based accounting research’. *Journal of Business Finance & Accounting*, 30(1&2): 25-55.

Falta, M. and Willett, R. J. (2009). ‘Parameter heterogeneity in cross-section models of the relationship between accounting numbers and market value’. Working Paper, Department of Accountancy and Business Law, University of Otago, New Zealand, available on http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1011585

Fama, E. F. and French, K. R. (1992). ‘The cross-section of expected stock returns’. *Journal of Finance*, 47: 427-465.

Francis, J. and Schipper, K. (1999). ‘Have financial statements lost their relevance?’. *Journal of Accounting Research*, 37(2): 319-352.

Greene, W. H. (2008). *Econometric Analysis*, 6th Edition. Pearson Prentice Hall.

Harvey, A. C. (1990). *The Econometric Analysis of Time Series*, 2nd Edition. Philip Allan.

Hendry, D. and Doornik, J. A. (2001). *Empirical Econometric Modelling Using PcGive*, Vol. I. Timberlake Consultants Ltd.

Hendry, D., Pagan, A. R. and Sargan, J. D. (1984). 'Dynamic specification', in Hendry, D. (ed.) *Econometrics: Alchemy or Science?* (2000). Oxford University Press. Reprinted from Griliches, Z. and Intriligator, M. D. (eds.) *Handbook of Econometrics*, Vol II, Chap. 18. (1984). Amsterdam, Elsevier.

Kothari, S. P. (2001). 'Capital markets research in accounting'. *Journal of Accounting and Economics*, 31: 105-231.

Kothari, S. P. and Shanken, J. (1997). 'Book-to-market, dividend yield, and expected market returns: A time-series analysis'. *Journal of Financial Economics*, 44: 169-203.

Kothari, S. P. and Shanken, J. (2003). 'Time-series coefficient variation in value-relevance regressions: a discussion of Core, Guay, and Van Buskirk and new evidence'. *Journal of Accounting and Economics*, 34: 69-87.

Miller, M. and Modigliani, F. (1961). 'Dividend policy, growth and the valuation of shares'. *Journal of Business*, 34(4): 411-433.

Morel, M. (2003). 'Endogenous parameter time series estimation of the Ohlson model: Linear and nonlinear analysis'. *Journal of Business, Finance & Accounting*, 30(9&10): 1341-1362.

Myers, J. N. (1999). 'Implementing residual income valuation with linear information dynamics'. *Accounting Review*, 74(1): 1-28.

Ohlson, J. (1995). 'Earnings, book values and dividends in equity valuation' *Contemporary Accounting Research*, 11(2): 661-687.

Pesaran, M. H. and Smith, R. (1995). 'Estimating long-run relationships from dynamic heterogeneous panels'. *Journal of Econometrics*, 68: 79-113.

Ramsey, J. B. (1969). 'Tests for specification error in classical linear least squares regression analysis'. *Journal of the Royal Statistical Society*, B, 31: 350-371.

Richardson, G. and Tinaikar, S. (2004). 'Accounting based valuation models: what have we learned?'. *Accounting and Finance*, 44: 223-255.

Ryan, S. G. and Zarowin, P. A. (2003). 'Why has the contemporaneous linear returns-relation declined?', *Accounting Review*, 78(2): 523-552.

Thursby, J. G. (1982). 'Misspecification, heteroskedasticity, and the Chow and Goldfeld-Quandt tests'. *Review of Economics and Statistics*, 64(2): 314-321.

White, H. (1980). 'A heteroskedasticity-consistent covariance matrix estimator and a direct test for heteroskedasticity'. *Econometrica*, 48: 817-38.

Table 1
Definitions and sources of data used in empirical models.

Data	Sources
Earnings, dividends and book value of net assets	As defined by Compustat annual data item numbers A172, A21 and A60 respectively. Sources: Compustat tapes: 1955-1998; Mergent Online 1999 to 2002; Company website 2005.
Market value	Defined as share price at fiscal year-end (A199) multiplied by the number of shares outstanding. Sources: Compustat 1955 to 1998; Datastream 1999 to 2002; Company website 2005.

Table 2
Statistical tests for misspecification and unit roots.

Test	Reference	Comments
RESET	Ramsey (1969)	Test null of correct specification against alternative that squares of estimates of the dependent have been omitted (as per Hendry and Doornik (2001: 263).
Heteroskedasticity	White (1980)	Tests null of unconditional heteroskedasticity against alternative that the variance of the error process depends upon squares of the regressors as well as the regressors themselves. Hendry and Doornik (2001: 262).
Autocorrelation	Harvey (1990)	*F*-test form for unconditional autocorrelation. Null hypothesis is of no autocorrelation of residuals. Hendry and Doornik (2001: 260).
Normality	Doornik and Hansen (1994)	χ^2-test. Null hypothesis is normality of residuals. Hendry and Doornik (2001: 261).
Autoregressive distributed lag	Dickey and Fuller (1981)	ADF tests with constant. Null is non-stationarity. Hendry and Doornik (2001: 230-1).

Table 3
Cross-section models, estimated using Expression (5.1), for 30 firms based on yearly data from 1955 to 2004. Specification tests (cf. Table 2) indicate statistical significance at the 1%-level (**) and 5%-level (*).

Year	k	p(k)	β_1	p(β_1)	β_2	p(β_2)	R^2	Specification tests		
								Normality	Hetero-skedasticity	RESET
1955	-258.08	0.17	8.18	0.0	-15.12	0.01	0.93	**	**	**
1956	145.96	0.35	-1.03	0.68	21.31	0.10	0.93	**	**	**
1957	183.96	0.31	0.04	0.99	12.91	0.30	0.87	**	**	**
1958	32.46	0.80	-4.31	0.0	54.57	0.0	0.97	**	**	**
1959	412.36	0.15	-4.36	0.21	45.37	0.03	0.89	**	**	**
1960	334.18	0.39	4.20	0.14	-10.02	0.54	0.71	**	**	**
1961	846.87	0.08	-7.35	0.10	67.58	0.02	0.77	**	**	**
1962	-115.91	0.74	6.33	0.0	-14.73	0.05	0.87	**	**	**
1963	-3.13	0.99	5.51	0.01	-8.38	0.35	0.88	**		**
1964	80.70	0.80	3.75	0.01	1.46	0.79	0.94	**	**	**
1965	-239.51	0.52	7.48	0.0	-13.41	0.01	0.94	**	**	**
1966	-58.24	0.89	9.02	0.0	-31.14	0.0	0.85	**	**	*
1967	555.59	0.57	4.74	0.36	-7.29	0.80	0.66	**	**	**
1968	1100.60	0.21	-5.38	0.23	47.49	0.06	0.72	**	**	**
1969	1085.76	0.33	-4.44	0.34	44.06	0.11	0.62	**	**	**
1970	-717.00	0.01	0.28	0.24	33.22	0.0	0.97	**	**	**
1971	339.26	0.73	5.13	0.07	-11.32	0.47	0.74	**	**	**
1972	381.68	0.75	8.96	0.03	-31.58	0.16	0.70	**	**	
1973	692.16	0.47	8.41	0.04	-33.72	0.10	0.64	**	**	
1974	-1.78	0.99	-0.60	0.0	16.71	0.0	0.95		*	
1975	138.94	0.49	-0.64	0.0	20.32	0.0	0.98	*	*	
1976	-297.92	0.70	3.61	0.03	-6.43	0.41	0.86	**	**	*
1977	465.66	0.64	-0.12	0.95	9.67	0.31	0.77	**	**	**
1978	2049	0.01	-6.89	0.0	40.53	0.0	0.85	**	**	**
1979	-110.40	0.79	-3.55	0.0	28.76	0.0	0.94	**	**	**
1980	-381.92	0.25	1.00	0.0	6.11	0.0	0.96		**	**
1981	-87.19	0.83	0.49	0.0	6.60	0.0	0.93	*	**	**
1982	-119.91	0.76	0.35	0.01	11.26	0.0	0.98	**	**	*
1983	1598.73	0.10	-1.55	0.01	18.49	0.0	0.93	**	**	**
1984	1906.3	0.03	-1.81	0.0	17.58	0.0	0.94	**	**	**
1985	2040.49	0.05	-2.04	0.0	23.45	0.0	0.94	**	**	**
1986	1112.41	0.14	-1.48	0.0	24.42	0.0	0.95		*	**
1987	-789.01	0.50	-1.47	0.0	22.17	0.0	0.89		**	**
1988	-319.54	0.80	-1.88	0.0	22.45	0.0	0.89	**	**	**
1989	1942.25	0.29	-0.53	0.18	15.22	0.0	0.78	**	**	
1990	4088.39	0.02	0.67	0.0	6.73	0.0	0.80	**		**
1991	2398.78	0.38	2.07	0.0	11.03	0.0	0.75			
1992	9285.60	0.01	1.75	0.0	0.27	0.64	0.38	*		
1993	3088.13	0.20	2.88	0.0	4.36	0.0	0.79	**	*	
1994	1917.47	0.45	1.10	0.0	9.61	0.0	0.80	**	**	
1995	6057.41	0.19	0.08	0.96	12.85	0.01	0.64	**	**	*
1996	5476.47	0.26	-1.44	0.23	23.74	0.0	0.79	**	*	
1997	6201.38	0.46	2.58	0.12	13.54	0.01	0.69	**	**	
1998	1144.51	0.88	1.51	0.29	27.34	0.0	0.86	*		
1999	-18,183.20	0.14	7.50	0.0	9.78	0.11	0.79	**	*	**
2000	1231.95	0.92	0.56	0.74	30.08	0.0	0.78	**		
2001	603.82	0.94	4.02	0.0	12.85	0.0	0.90	*		
2002	13,152.40	0.01	1.46	0.02	12.42	0.0	0.90		**	**
2003	3106.86	0.57	2.98	0.0	7.02	0.01	0.91	**		
2004	5487.93	0.23	1.27	0.01	12.98	0.0	0.94		**	*

Table 4
Dynamic models, estimated using Expression (5.1), for 30 firms based on data from 1955 to 1994. Compare Table 2 for details on specification tests.

Firm	k	p(k)	β_1	p(β_1)	β_2	p(β_2)	R^2	Specification Tests			
								Auto-regression	Normality	Hetero-skedasticity	RESET
Abbott	197.22	0.69	-1.89	0.24	24.19	0.00	0.96	**	**	**	
Bausch	-102.31	0.12	3.02	0.00	-0.10	0.96	0.90		**	**	*
Baxter	240.31	0.14	1.96	0.00	0.74	0.48	0.93	**	**	**	*
Bristol	-835.53	0.21	2.73	0.01	10.20	0.00	0.94	**	**	**	**
Coca Cola	860.79	0.56	-4.03	0.02	33.51	0.00	0.91	**	**	**	
Colgate	-643.18	0.08	2.60	0.00	5.33	0.08	0.74	*	**	**	*
Cooper	-102.25	0.36	1.95	0.00	0.60	0.49	0.92	*	*	**	*
Corning	369.37	0.19	2.85	0.00	-5.27	0.10	0.67	**		*	**
Du Pont	4383.16	0.01	1.49	0.00	-1.60	0.13	0.62	**	*	**	*
Eaton	-216.63	0.16	1.99	0.00	-0.56	0.63	0.73	**	**		**
General Electric	-1022.09	0.50	4.05	0.00	-6.86	0.12	0.93	**		**	**
General Motors	19444.40	0.00	0.02	0.85	0.12	0.62	0.01	**		*	*
Georgia Pacific	584.73	0.01	1.58	0.00	-1.82	0.08	0.75	**			
Gillette	-859.44	0.03	1.91	0.04	19.87	0.00	0.85	**	**	**	
Goodyear	740.50	0.09	0.86	0.00	-0.05	0.96	0.29	**	*		
Hercules	480.03	0.02	1.07	0.00	-0.16	0.88	0.47	**	**		
Ingersoll	213.73	0.19	1.47	0.00	-0.79	0.52	0.62	**	**		**
IBM	16694.10	0.00	1.13	0.00	2.80	0.00	0.78	**		*	
International Paper	215.19	0.33	1.28	0.00	-1.90	0.03	0.89	**		**	**
Johnson & Johnson	-200.97	0.81	2.30	0.05	10.46	0.02	0.90	**	**	**	*
Lilly	27.84	0.95	2.24	0.00	7.75	0.01	0.91	**		**	*
Merck	937.22	0.24	-3.03	0.00	30.00	0.00	0.94		**	**	
Motorola	-510.62	0.17	0.60	0.20	17.78	0.00	0.94	**		**	**
Pfizer	-625.39	0.43	2.22	0.05	8.95	0.12	0.79	**	*	**	*
Raytheon	47.37	0.67	1.30	0.00	3.71	0.04	0.96	*	**		
Rohm	-88.93	0.53	2.48	0.00	-2.49	0.20	0.78	**		*	**
Schering	268.97	0.30	-0.90	0.05	18.99	0.00	0.94	**		*	**
Tektronix	144.75	0.26	1.20	0.00	-0.42	0.61	0.81	**	*	**	
UST	-29.61	0.88	1.50	0.00	-1.09	0.03	0.89	**	**		
United Technologies	-126.01	0.34	0.03	0.98	17.85	0.00	0.93	**	**	**	

Table 5
Cross-section models, estimated using Expression (5.2), for 30 firms based on yearly data from 1955 to 2004. Specification tests (cf. Table 2) indicate statistical significance at the 1%-level (**) and 5%-level (*).

Year	k	p(k)	β_1	p(β_1)	β_2	p(β_2)	R^2	Specification tests		
								Normality	Hetero-skedasticity	RESET
1955	2.43	0.0	-0.03	0.86	1.12	0.0	0.97			
1956	1.71	0.0	0.30	0.11	0.83	0.0	0.95			
1957	1.84	0.01	0.13	0.64	1.01	0.0	0.92			
1958	2.83	0.0	-0.08	0.72	1.19	0.0	0.95			
1959	3.60	0.0	-0.47	0.13	1.56	0.0	0.93			
1960	3.68	0.0	-0.37	0.18	1.41	0.0	0.91			
1961	4.07	0.0	-0.37	0.02	1.38	0.0	0.96			
1962	3.44	0.0	-0.28	0.14	1.31	0.0	0.96		*	
1963	1.48	0.01	0.69	0.0	0.39	0.0	0.90			
1964	3.88	0.0	-0.52	0.03	1.56	0.0	0.96			
1965	4.12	0.0	-0.56	0.04	1.57	0.0	0.94			
1966	3.61	0.0	-0.37	0.30	1.36	0.0	0.87		*	
1967	3.72	0.0	-0.22	0.37	1.12	0.0	0.89		*	
1968	3.68	0.0	-0.18	0.48	1.14	0.0	0.90			
1969	3.79	0.0	-0.31	0.40	1.28	0.0	0.82			
1970	2.38	0.0	0.13	0.50	0.98	0.0	0.89			
1971	1.65	0.01	0.85	0.0	0.13	0.0	0.82			
1972	4.28	0.0	-0.57	0.15	1.56	0.0	0.86			
1973	4.22	0.0	-0.73	0.18	1.69	0.0	0.79			
1974	2.47	0.01	-0.27	0.49	1.35	0.0	0.80			
1975	1.51	0.02	0.49	0.02	0.56	0.01	0.87			
1976	0.60	0.21	0.99	0.0	0.04	0.12	0.90			
1977	2.73	0.0	-0.16	0.48	1.14	0.0	0.93			
1978	2.77	0.0	-0.31	0.28	1.30	0.0	0.91			
1979	2.67	0.0	-0.19	0.51	1.15	0.0	0.87			
1980	0.74	0.35	0.93	0.0	0.05	0.15	0.76			
1981	1.88	0.01	0.53	0.01	0.35	0.04	0.80			
1982	1.08	0.14	0.89	0.0	0.05	0.14	0.81			
1983	1.37	0.02	0.86	0.0	0.06	0.05	0.86			
1984	2.22	0.0	0.43	0.0	0.47	0.0	0.91		*	
1985	2.35	0.0	0.35	0.02	0.59	0.0	0.89			
1986	2.78	0.0	0.47	0.01	0.38	0.02	0.79		**	
1987	2.84	0.0	0.26	0.19	0.61	0.0	0.78		*	
1988	2.74	0.0	-0.01	0.78	0.96	0.0	0.92			
1989	4.23	0.0	0.51	0.0	0.11	0.02	0.59		*	
1990	2.65	0.02	0.72	0.0	0.09	0.01	0.58			
1991	2.25	0.08	0.83	0.0	0.07	0.02	0.56			
1992	3.78	0.0	0.69	0.0	0.04	0.01	0.60			
1993	3.40	0.0	0.73	0.0	0.04	0.03	0.61			
1994	2.53	0.01	0.78	0.0	0.09	0.02	0.72			
1995	2.96	0.01	0.77	0.0	0.05	0.27	0.61			
1996	2.82	0.0	0.29	0.02	0.68	0.0	0.87			
1997	1.92	0.11	0.95	0.0	0.04	0.13	0.66			
1998	1.01	0.40	1.03	0.0	0.09	0.06	0.70			
1999	1.23	0.23	0.49	0.0	0.68	0.0	0.78			
2000	2.09	0.04	0.38	0.01	0.72	0.0	0.82			
2001	1.52	0.20	0.96	0.0	0.09	0.0	0.77			
2002	9.04	0.0	0.10	0.04	0.09	0.0	0.51			**
2003	8.64	0.0	0.14	0.01	0.08	0.09	0.42			**
2004	5.65	0.0	0.51	0.0	0.04	0.18	0.64		*	**

Table 6
Dynamic models, estimated using Expression (5.3), for 30 firms using data from 1955 to 1994. NSWL (Non-stationary without lags), i.e., model is non-stationary unless lags are added. All models test stationary by ADF tests when lags are present.

Firm	k p(k)	α p(α)	β_1 p(β_1)	$\beta_{1,t-1}$ p($\beta_{1,t-1}$)	β_2 p(β_2)	$\beta_{2,t-1}$ p($\beta_{2,t-1}$)	R^2	Specification Tests				
								Auto-regression	Normality	Hetero-skedasticity	RESET	ADF
Abbott	0.86 0.21	0.67 0.00	0.08 0.88	0.00 1.00	0.30 0.25	-0.05 0.85	0.99					NSWL
Bausch	-1.28 0.17	0.49 0.00	4.01 0.01	-3.08 0.03	-0.14 0.39	-0.14 0.38	0.96	*				
Baxter	0.46 0.02	0.81 0.00	1.12 0.00	-0.97 0.00	0.00 0.81	0.00 0.79	0.99					NSWL
Bristol	0.02 0.98	0.88 0.00	1.73 0.00	-1.53 0.01	0.18 0.61	-0.25 0.29	0.99				**	
Coca-Cola	3.53 0.01	0.62 0.00	-0.02 0.97	-0.78 0.14	1.21 0.00	-0.25 0.56	0.98					NSWL
Colgate	0.14 0.84	0.92 0.00	0.58 0.14	-0.51 0.12	0.02 0.90	0.01 0.94	0.96					NSWL
Cooper	0.13 0.42	0.15 0.36	1.08 0.00	-0.17 0.63	0.00 0.90	-0.02 0.40	0.98					
Corning	0.56 0.38	0.83 0.00	1.66 0.25	-1.55 0.27	-0.02 0.56	0.01 0.64	0.84		*			NSWL
Du Pont	0.16 0.81	0.84 0.00	0.10 0.76	0.06 0.87	0.00 0.98	0.00 0.78	0.88	*	*			NSWL
Eaton	0.02 0.96	0.74 0.00	0.77 0.06	-0.47 0.23	-0.02 0.14	-0.03 0.03	0.95					NSWL
GE	-0.23 0.81	0.81 0.00	1.25 0.34	-0.80 0.52	-0.06 0.86	-0.22 0.39	0.95		*		**	NSWL
GM	3.15 0.03	0.71 0.00	-0.09 0.38	0.06 0.53	0.01 0.24	0.01 0.40	0.48					NSWL
Georgia Pacific	0.59 0.08	0.77 0.00	0.95 0.02	-0.77 0.05	-0.02 0.10	0.00 0.87	0.97					
Gillette	0.13 0.75	0.94 0.00	0.01 0.60	-0.02 0.27	0.06 0.50	0.05 0.61	0.93					NSWL
Goodyear	1.64 0.03	0.69 0.00	-0.22 0.62	0.34 0.42	0.02 0.28	-0.06 0.00	0.75					
Hercules	0.84 0.17	0.79 0.00	0.00 1.00	0.11 0.87	-0.01 0.47	0.00 0.94	0.86					
Ingersoll	1.01 0.07	0.68 0.00	0.57 0.52	-0.37 0.67	-0.02 0.49	0.01 0.55	0.85				*	NSWL
IBM	2.51 0.00	0.62 0.00	-0.04 0.92	0.19 0.58	0.02 0.05	0.00 0.95	0.94					NSWL
International Paper	0.37 0.43	0.67 0.00	1.08 0.23	-0.72 0.38	-0.07 0.51	-0.04 0.59	0.90				*	
Johnson & Johnson	1.78 0.14	0.83 0.00	0.46 0.56	-0.83 0.18	0.08 0.79	0.36 0.18	0.98					
Lilly	1.10 0.06	0.84 0.00	0.45 0.45	-0.60 0.21	0.30 0.10	-0.05 0.77	0.98					
Merck	1.77 0.01	0.71 0.00	-0.22 0.42	-0.04 0.87	1.00 0.00	-0.53 0.17	0.98					NSWL
Motorola	0.17 0.76	0.50 0.00	2.09 0.07	-1.46 0.16	-0.01 0.95	-0.15 0.38	0.96					
Pfizer	0.85 0.31	0.67 0.00	-0.73 0.35	0.98 0.20	0.30 0.33	-0.27 0.49	0.97					NSWL
Raytheon	0.78 0.09	0.34 0.03	1.68 0.01	-1.23 0.03	0.12 0.05	0.09 0.16	0.98					
Rohm	0.11 0.81	0.87 0.00	0.58 0.51	-0.45 0.59	0.00 0.97	0.01 0.31	0.89					NSWL
Schering	0.62 0.15	0.85 0.00	-0.23 0.52	0.25 0.46	0.97 0.00	-0.88 0.01	0.98					NSWL
Tektronix	0.34 0.26	0.78 0.00	0.27 0.59	-0.06 0.90	0.01 0.77	-0.02 0.43	0.94		*			NSWL
UST	-0.15 0.62	0.67 0.00	1.28 0.04	-0.92 0.13	0.01 0.60	-0.01 0.29	0.95					NSWL
United Technologies	1.14 0.33	0.60 0.00	-0.07 0.91	0.00 1.00	1.38 0.04	-0.91 0.25	0.99				**	
Average	0.77 0.36	0.71 0.01	0.67 0.40	-0.51 0.41	0.19 0.46	-0.11 0.48	0.92					

Table 7

Experiment 1 (Panel A): Illustration of the effect on OLS coefficient estimates and inferential statistics of modelling 30 and 300 firms each, using a Monte Carlo simulation with 1,000 and 10,000 replications, respectively, a DGP that is multiplicative of the form $M_t = 3A_t^{\beta t}\omega_t$, where $\beta_t = 1$, $\ln(A_t) = a_t = N(5,1)$ and $\ln(\omega_t) = N(0,1)$, by two models: the correctly specified model (model in logs), $m_t = \ln(M_t) = \ln(k) + \beta_t a_t + \ln(\omega_t)$ and incorrectly specified model (linear, raw data model), $M_t = \alpha_t + \beta_t A_t + u_t$. The standard deviation (SD) given is the between sample standard deviation. We give the conclusion in the table.

Experiment 2 (Panel B): Demonstration of the potential effect of deflating variables in a regression model by opening market value, i.e., we investigate if the division of the regressand and regressors by the lagged values of the regressand in a cross-section OLS model may cause a spurious regression relationship to be observed. The DGP for market value is assumed to be independent of both E_t and B_t. The subscript '0' denotes starting time. The first value of E_t is generated as 5% of the first value of B_t. Thereafter E_t is generated as a random walk. B_t is generated initially from a uniform distribution between 0 and 10. Thereafter B_t is the sum of its previous value plus the current value of E_t and a standard normal white noise term. B_t and E_t are thus related by a 'stochastic clean surplus relation' but have no influence on M_t. Method used is Monte Carlo simulation with 10,000 repetitions. The SD given is based on the Monte Carlo repetitions. None of the coefficients, α_t, β_{1t} and β_{2t} will be statistically significant at the 5% level in 95 out of 100 samples and R^2 will be close to zero.

Panel A	Model in logs					Linear, raw data model				
	k	β_t	$t(k)$	$t(\beta_t)$	R^2	α_t	β_t	$t(\alpha_t)$	$t(\beta_t)$	R^2
30 firm average	1.086	1.002	1.167	5.490	0.501	-3.457	4.976	0.324	1.659	0.40
30 firm SD	0.970	0.190	1.050	1.475	0.129	760.690	4.370	1.40	1.457	0.227
300 firm average	1.102	1.0	3.745	17.320	0.50	2.413	4.928	0.450	1.643	0.328
300 firm SD	0.295	0.058	1.023	1.422	0.041	478.423	2.320	2.087	0.773	0.121

Conclusion:
Estimating a DGP that is multiplicative in regressors and error term by a simple linear additive model leads to biased coefficient estimates that fluctuate widely between samples. Although larger sample sizes reduce the degree of fluctuation, they do not eliminate the bias in the estimates leading to the incorrect estimate appearing to be increasingly statistically significant as the sample size grows.

Panel B	DGPs	Model: Variables deflated by opening market value as recommended by Christie (1987).
	$M_t = M_{t-1} + N(0,1)$, $M_0 = U(0,10)$ $E_t = E_{t-1} + N(0,1)$, $E_0 = 0.05B_0$ $B_t = B_{t-1} + E_t + N(0,1)$, $B_0 = U(0,10)$	$M_t/M_{t-1} = \alpha_t + \beta_{1t}E_t/M_{t-1} + \beta_{2t}B_t/M_{t-1} + \varepsilon_t$

Year	α	$t(\alpha)$	β_1	$t(\beta_1)$	β_2	$t(\beta_2)$	Adjusted R^2
	[Average (SD)] applicable for all parameters						
1995	1.01 (0.16)	13.42 (7.15)	0.02 (0.26)	0.20 (8.19)	0.0 (0.01)	-0.30 (7.22)	0.50 (0.35)
1996	0.99 (0.22)	13.27 (7.36)	-0.05 (0.27)	-0.57 (16.02)	0.0 (0.01)	0.14 (7.32)	0.52 (0.34)
1997	1.01 (0.35)	13.64 (7.30)	0.0 (0.39)	0.01 (7.91)	0.0 (0.02)	-0.22 (8.33)	0.53 (0.34)
1998	1.00 (0.14)	13.45 (7.67)	0.0 (0.24)	0.53 (12.78)	0.0 (0.01)	-0.27 (8.54)	0.50 (0.34)
1999	1.00 (0.14)	13.54 (7.18)	0.0 (0.24)	-0.17 (9.91)	0.0 (0.01)	0.20 (9.58)	0.49 (0.34)
2000	0.99 (0.44)	13.54 (7.31)	0.0 (0.25)	-0.27 (8.70)	0.0 (0.01)	-0.24 (12.78)	0.50 (0.34)
2001	0.99 (0.16)	13.56 (7.58)	0.0 (0.29)	-0.27 (13.95)	0.0 (0.01)	-0.27 (10.26)	0.51 (0.34)
2002	1.00 (0.55)	13.91 (8.34)	0.0 (0.44)	0.27 (4.30)	0.0 (0.02)	0.24 (41.02)	0.47 (0.34)
2003	1.02 (0.55)	13.29 (8.34)	0.02 (0.44)	-0.02 (4.30)	0.0 (0.02)	0.57 (41.02)	0.46 (0.34)
2004	1.01 (0.30)	13.44 (8.19)	0.0 (0.25)	-0.03 (4.33)	0.0 (0.01)	-0.58 (125.23)	0.46 (0.34)

Conclusion:
Despite the average value of the coefficients being close to their correct values ($\alpha = 1$, $\beta_1 = \beta_2 = 0$) the variance of the *t*-statistics is high causing the *t*-statistic to behave erratically between samples. The non-standard form of the distributions of the coefficients and the inferential statistics can be seen from the histograms in Figure 2, Panel 2. Due to this, the coefficients on earnings and book value test significantly different from zero in 24% and 35.6% of all samples. R^2 exceeds 90% in over 15% of all samples.

Figure 1

Scatter and relative frequency plots of 30 firms averaged sample data. *Left*: Growth patterns of net book value, earnings, the geometric mean of net book value and earnings, and market value for the period from 1955 to 2004. Given are also results from exponential, $y = a \cdot exp(bx)$, and linear, $y = a+bx$, fits to the data, for the latter only the net book value fit is shown. *Right*: Corresponding relative frequency distributions of dollar amounts using eight intervals each. Given are also results from logarithmic, $y=a+b \cdot ln(x)$, fits to the data.

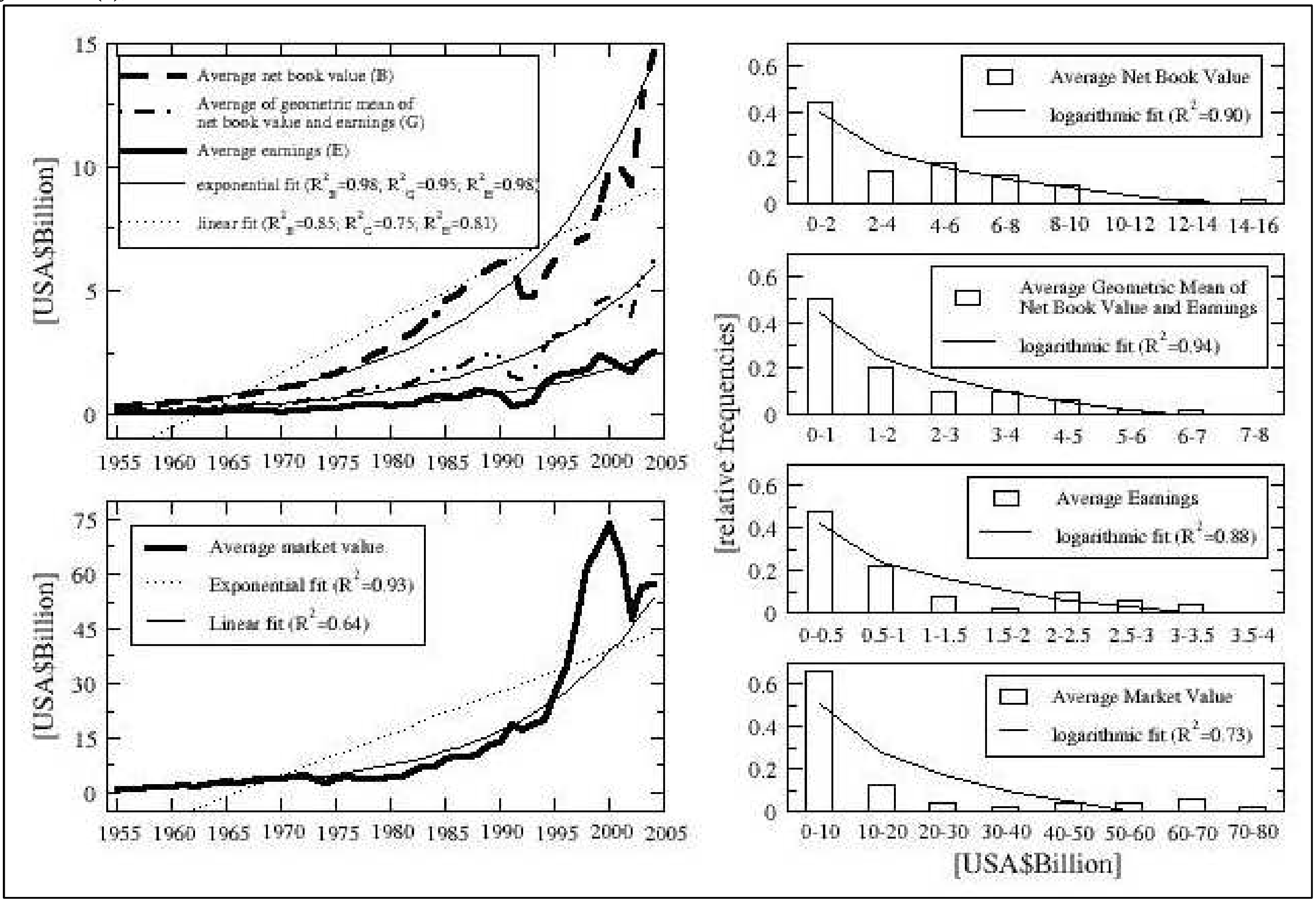


Figure 2

Cross-section scatter plots in raw data and logs of time averaged sample data over the estimation period 1955-1994. The larger R^2 values from linear fitting and visual inspection favour the logarithmic scale for OLS estimation. Also visible is the good correspondence between earnings and market value, an indication as to which accounting variable eventually transpires most frequently in the top-down variable elimination process.

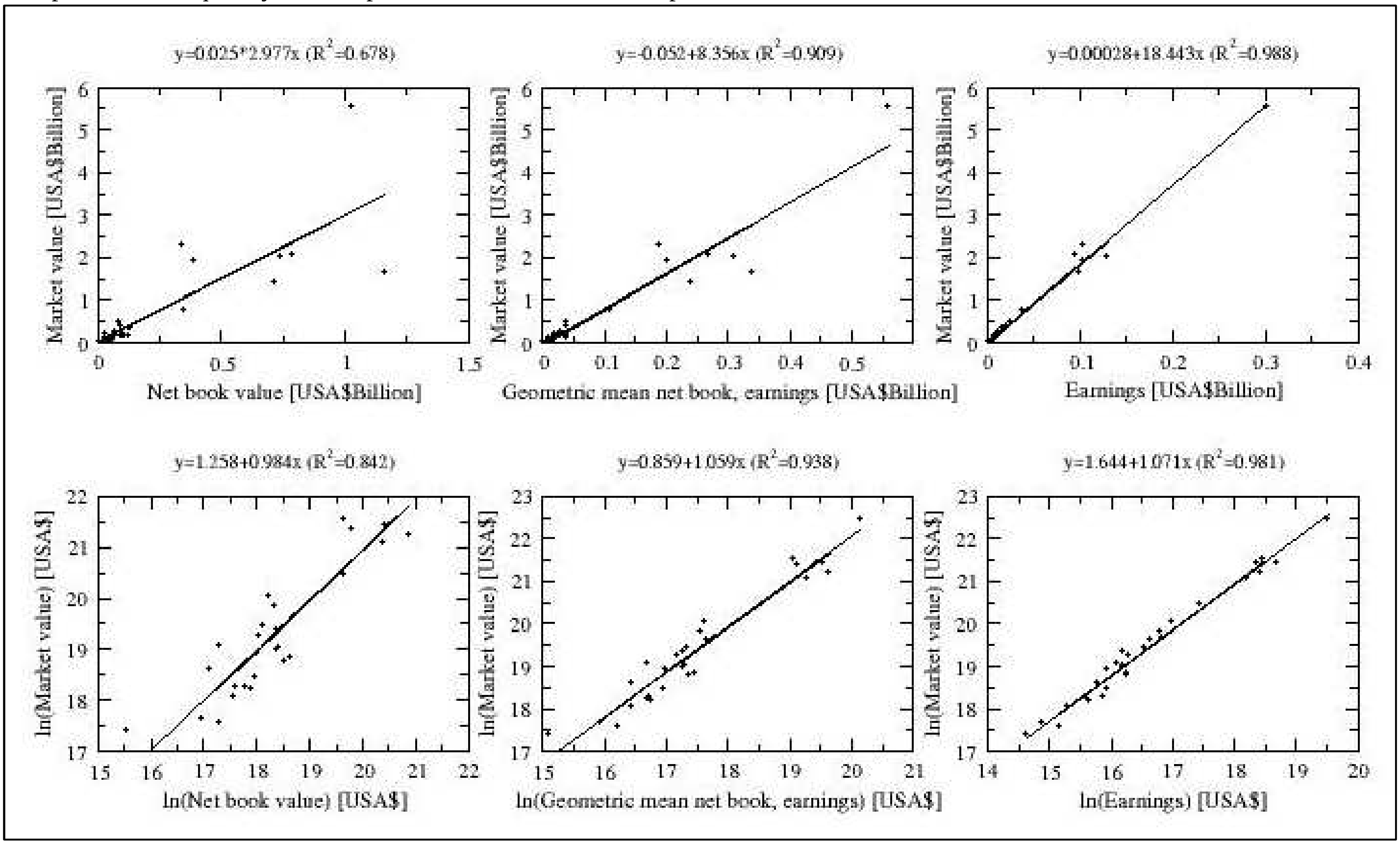


Figure 3

Histograms of raw and logged time averaged sample data over the estimation period 1955-1994. Details of fits are also displayed: For the raw data, we used linear logarithmic forms and for the logged data, the Normal distribution.

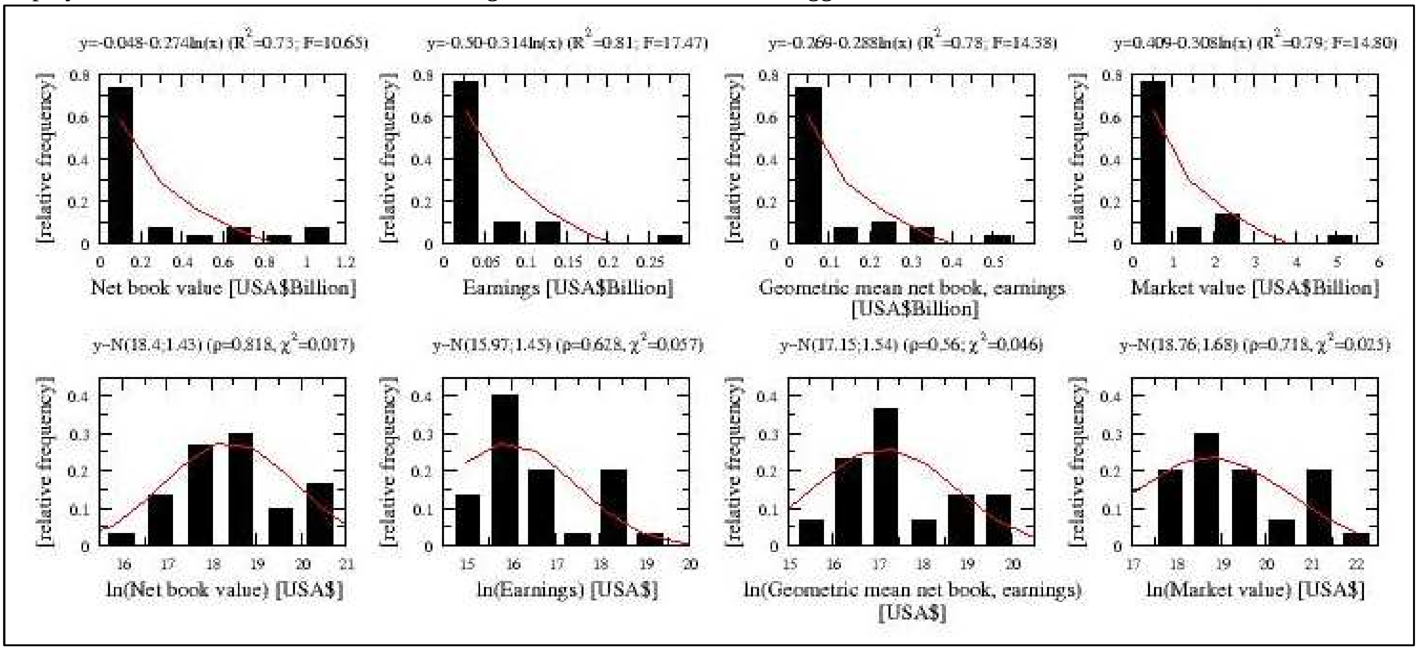


Figure 4
Experiment 1 (cf. Table 7). Histograms of coefficients, t-statistics and R^2 from Monte Carlo simulation. *Left half*: data based on 1,000 replications for 30 firms. *Right half*: data based on 10,000 replications for 300 firms.

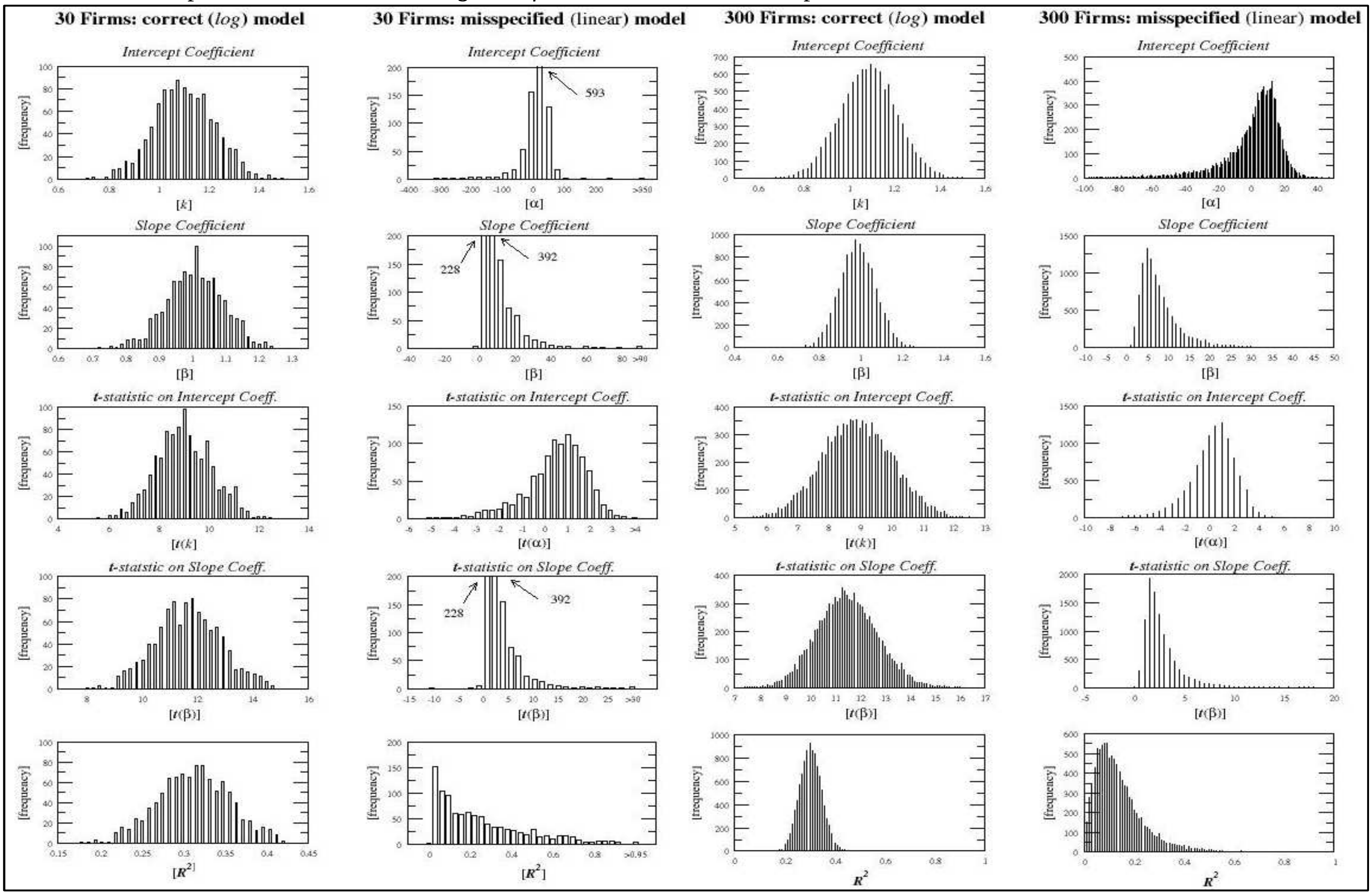
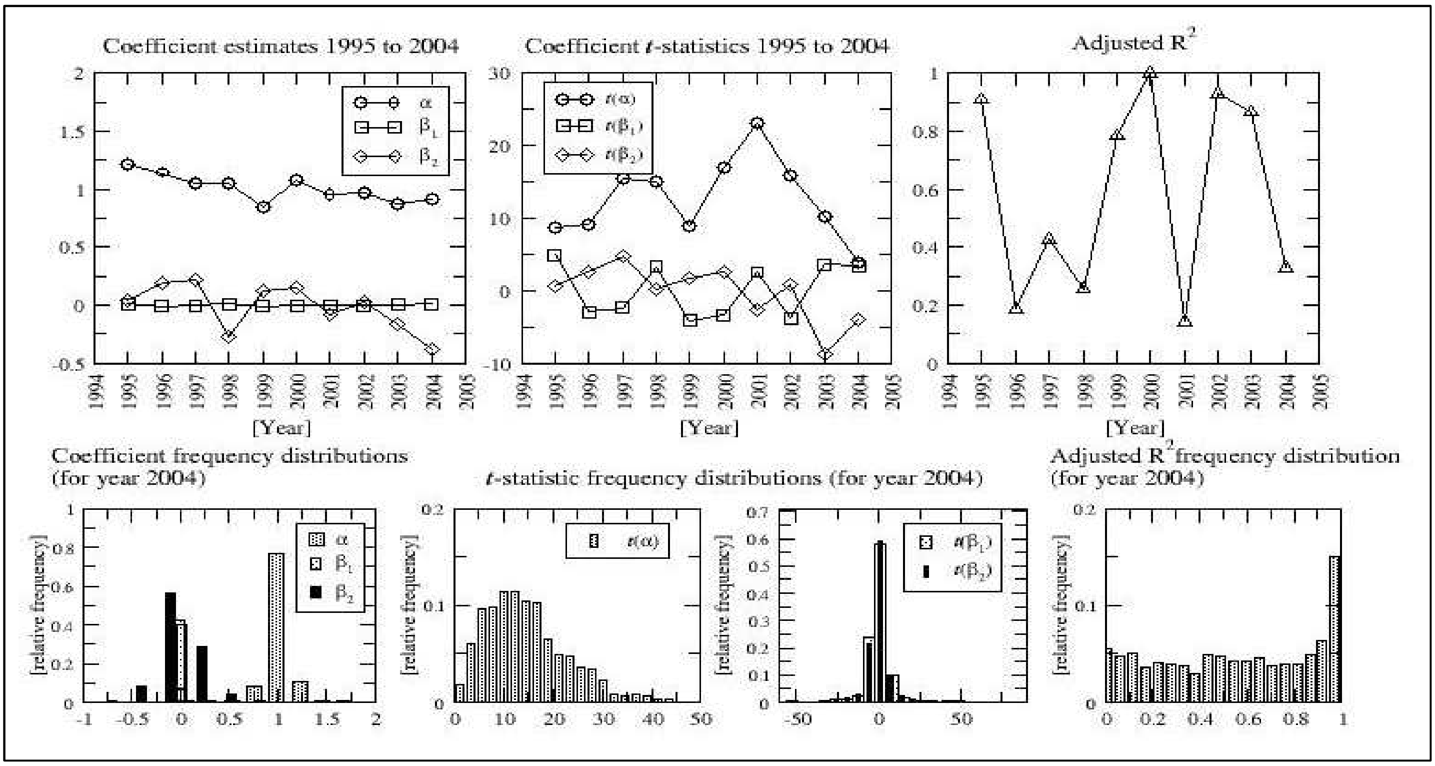


Figure 5
Experiment 2. Behaviour of coefficients, t-statistics and adjusted R^2 produced by deflating Expression (5.1) by $M_{i,t-1}$, i.e., $M_{i,t}/M_{i,t-1} = k_{i,t} + \beta_{i,1,t}E_{i,t}/M_{i,t-1} + \beta_{i,2,t}B_{i,t}/M_{i,t-1} + \varepsilon_t$. *Top row*: Results for a simulated ten year period from, e.g., from 1994 to 2004. *Bottom row*: Results for t=2004.



Appendix 1

Table 8
Results for specification tests (cf. Table 2) of cross-section models, estimated using simplified Expressions (5.1) and (5.2), $M_{i,t}=k_{i,t}+\beta_{i,1,t}B_{i,t}+\varepsilon_t$ and $M_{i,t} = e^{k_{i,t}}B_{i,t}^{\beta_{i,1,t}}\omega_t$, respectively, for 30 firms based on yearly data from 1955 to 2004. Statistical significance levels given are at the 1%-level (**) and 5%-level (*).

Year	Additive model			Multiplicative model		
	Specification tests					
	Normality	Hetero-skedasticity	RESET	Normality	Hetero-skedasticity	RESET
1955	**	**	**			
1956	**	**	**			
1957	**	**	**			
1958	**	**	**			
1959	**	**	**			
1960	**		**			
1961	**		**			
1962	**	**	**			
1963	**	*	**			
1964	**	**	**			
1965	**	**	**			
1966	**	**	**			
1967	**	**	**			
1968	**	**	**			
1969	**	**	**			
1970	**	**	**			
1971	**	**	**			
1972	**	**	**			
1973	**	**	**			
1974	**	**	**			
1975	**	**				
1976	**	**				
1977	**	**				
1978	**	**	*			
1979	**	**	*			
1980	**	**				
1981	**	**				
1982	**	**	*			
1983	**	**	*			
1984	**	**	*			
1985	**	**				
1986	**	**				
1987	**	**				
1988	**	**			*	**
1989		**	*		**	*
1990						
1991	**					
1992			*			
1993	**					
1994	**					
1995	**					
1996	**					
1997	*					
1998		**				
1999	*	**	**			
2000	**	**		*		
2001	**	*				
2002	*	**	**			**
2003	**	**				**
2004	**	**			**	**

Table 9

Results for specification tests (cf. Table 2) of dynamic models, estimated using simplified Expressions (5.1) and (5.2) but including a lag in (5.1): $M_{i,t}=k_{i,t}+\beta_{i,1,t}B_{i,t}+\beta_{i,1,t-1}B_{i,t-1}+\varepsilon_t$ and $\boldsymbol{M_{i,t}=e^{k_{i,t}}B_{i,t}^{\beta_{i,1,t}}B_{i,t-1}^{\beta_{i,1,t-1}}\omega_t}$, respectively, for 30 firms based on data from 1955 to 1994. Statistical significance levels given are at the 1%-level (**) and 5%-level (*).

Firm	Additive model				Multiplicative model			
	Specification Tests							
	Auto-regression	Normality	Hetero-skedasticity	RESET	Auto-Regression	Normality	Hetero-Skedasticity	RESET
Abbott		**	*					
Bausch	**	**	**		*	**		
Baxter		**						
Bristol	*	**	**					*
Coca Cola		**				**		
Colgate		*	*				*	
Cooper	**	**	**					
Corning		**				*		
Du Pont	**		**		*	*		
Eaton		*						
General Electric	**	*	**	*		*		**
General Motors		*						
Georgia Pacific			**					
Gillette		**	**					
Goodyear		*				*		
Hercules		*	**					
Ingersoll	**		**					*
IBM								
International Paper			**					**
Johnson & Johnson		**	**					
Lilly		**	*					
Merck	**	**	*	**				
Motorola	**	**	**					
Pfizer	**	**	**					
Raytheon		**	**					
Rohm								
Schering		**	*					
Tektronix		*	**			*		
UST		**	*					
United Technologies		**	**	**				**