Accounting based valuation: a simultaneous equations model for forecasting earnings to proxy for 'other information'

Iris Bergmann[1] · Wolfgang Schultze[1]

Abstract This paper develops and tests a simultaneous equations model (SEM) for extending accounting based valuation models used in empirical studies. Rather than using analysts' forecasts, we derive forecasts of operating income from the SEM to calculate the 'other information' variable in the Ohlson (Contemp Account Res 11:661–687, 1995) model. The SEM forecasts are based on observable data contained in the firms' reporting, like order backlog, and other publicly available information. The SEM produces more accurate out-of-sample forecasts of operating income compared to simple benchmark models particularly in years around economic changes and instability, like the years 2001 and 2009. Integrating the SEM forecast as 'other information' in market value regressions significantly increases the explanatory power compared to simpler versions without or with single information proxies for 'other information'. Finally, we find that the SEM forecast is able to explain a major portion of the information advantage of analysts relevant for explaining market values.

Keywords Accounting based valuation · Analysts' earnings forecasts · Ohlson model · Simultaneous equations

JEL Classification M40 · M41 · C53

1 Introduction

This paper develops and tests a simultaneous equations model for specifying 'other information' in accounting based valuation (ABV) models. ABV models are used in empirical capital market research for examining the association between accounting

✉ Wolfgang Schultze
wolfgang.schultze@wiwi.uni-augsburg.de

[1] School of Accounting and Control, University of Augsburg, Universitaetsstr. 16, 86159 Augsburg, Germany

information and market values. These studies aim at improving our understanding of the information processing in the capital markets, in particular, of how capital market participants process information and what kind of information they use (e.g. Richardson and Tinaikar 2004; Imam et al. 2008; Tsay et al. 2008; Anandarajan et al. 2011; Agostino et al. 2011; Tswei and Kuo 2012). Such studies have largely relied on the Ohlson (1995) model as a theoretical framework (Callen and Morel 2001; Penman and Yehuda 2009; Ashton et al. 2011; Clubb 2013; Rees and Valentincic 2013; Ashton and Wang 2015; Kuo 2017). Assuming that future residual earnings and other value relevant information follow a stochastic process, the model describes market value as a linear function of current accounting numbers and 'other information'.

A particular challenge for the application of the Ohlson model is specifying the 'other information' component of the model, reflecting value-relevant information not captured by financial accounting (e.g. Barth et al. 2005). Ohlson (2001) suggests using analysts' forecasts as a proxy for 'other information' because they are readily available and comprise information beyond historical financial statements (Brown 1993; White et al. 2003). Many studies have used this approach and find that adding analysts' forecasts significantly improves the explanatory power of accounting information for market values (e.g. Dechow et al. 1999; Choi et al. 2006; Higgins 2011). Beyond such studies interested in the link between accounting information and stock returns, research is especially interested in whether and to what extent market participants consider non-accounting information in order to generate earnings forecasts and estimate firm values (Bradshaw 2011). Based on the notion that the relevance of financial statement information for market values has been declining (Brown et al. 1999), a large body of literature analyzes the relevance of non-accounting information, such as intangibles (Hirschey et al. 2001). For example, Barth et al. (1998) analyze brand value as non-financial information, or Hirschey et al. (2001) use patent data as a proxy for 'other information' in the Ohlson model. However, including such non-financial information in ABV models is difficult. Myers (1999) shows that direct, single items of non-accounting information, like e.g. order backlog, do not adequately represent 'other information'. In this paper, we therefore present a different approach by using a forecasting model to aggregate single items of other information into an earnings forecast that can then be used as a proxy for 'other information' in the same manner as suggested by Ohlson (2001).

It is well established in the literature that analysts' earnings forecasts are superior to time-series models' forecasts due to information and timing advantages of analysts (e.g. Fried and Givoly 1982; Brown et al. 1987; Higgins 2011). That is, analysts have access to and use additional and timelier information than time-series models, which rely on historical accounting numbers only. Hence, the advantage of analysts primarily stems from their use of other, additional information beyond financial statement information. However, little is known about what kind of additional information is considered by analysts and how it is processed in order to generate forecasts and derive stock recommendations (Bradshaw 2011). In this paper, we shed light on this question by analyzing the use of 'other information', that is, publicly available information beyond financial statement information, for deriving earnings forecasts and market valuation.

Because single items of non-accounting information, like order backlog, do not adequately represent 'other information' (Myers 1999), we use a comprehensive set of publicly available information outside the financial statements together with financial accounting information to derive earnings forecasts, which can then be used to determine the 'other information' variable in the Ohlson-model. We replicate the forecasting process presumably followed by financial analysts as described in standard textbooks on financial statement analysis and valuation (e.g. Lundholm and Sloan 2007; Penman 2010). In these 'forecasting frameworks', the forecasting process is not linear but characterized by interdependencies. A simultaneous equations model (SEM) is needed to account for these interdependencies and produce comprehensive earnings forecasts (Tsay et al. 2008).

Based on the residual income model, we identify sales, operating expenditures and net operating assets as central value drivers of future (residual) income. These value drivers (endogenous variables) are modeled with a set of simultaneous equations. External determinants (exogenous variables) of the endogenous variables are derived from theory and empirical evidence and can be classified into macroeconomic (e.g. GDP growth), industry-specific (e.g. market concentration) and firm-specific variables (e.g. order backlog). They are used to simultaneously derive forecasts of future sales, costs, capital investments, and ultimately (residual) income. We estimate the SEM and produce operating income forecasts for our sample comprising all US industrial firms included in the Compustat database that report order backlog data in the period of 1991–2010.

First, we analyze the forecasting performance of the SEM. The errors of the SEM forecasts are compared to the errors produced by simple benchmark models (autoregressive model, random walk model) as well as analysts' forecasts in order to evaluate the forecasting performance. We find evidence that in years of economic disturbances (2001 and 2009) the SEM produces more accurate forecasts of operating income than simple benchmark models. In years of stable and continuous development, the SEM performs similarly well as the autoregressive model and better than a random walk. These results indicate that the application of the more complex model is in order particularly in volatile times to increase forecast accuracy, consistent with prior research (Bryan and Tiras 2007). Comparison with analysts' forecasts of operating income shows no superiority of the SEM, but in volatile times, especially in the year 2009, the SEM's and analysts' errors converge. Hence, in settings with no analysts' forecasts available, the SEM will provide better forecasting results than simple models when the forecasting period is characterized by economic disturbances.

Second, we integrate the SEM forecast as a proxy for 'other information' in market value regressions based on the Ohlson (1995) model. We find that this model version exhibits greater explanatory power for current market values than versions without 'other information' or with single information proxies for 'other information'. This result demonstrates the usefulness of the SEM to systematically integrate 'other information' and the interdependencies between the variables. We show that the SEM forecast is able to explain value beyond book value and operating income, indicating the relevance of the other information contained in our model.

Third, we analyze the relation between the analysts' forecasts and the SEM forecasts. The results show that the SEM forecasts capture a large portion of the information advantage contained in analysts' forecast for explaining market values. This result indicates that analysts use the textbooks' 'forecasting frameworks' to some extent, but their operating income forecasts anchor substantially on the last period's value of operating income, consistent with prior literature (Lambert et al. 2012).

The paper contributes to the literature on ABV models by providing an approach to determine the 'other information' component in the Ohlson (1995) model without a need for analysts' forecasts. Analysts' forecasts are often not available for small and foreign firms. Also, using analysts' forecasts cannot provide a link between single items of non-accounting information, such as intangibles, and market valuation. The SEM used in this study provides such a link by aggregating single items of non-accounting information, such as order backlog or market share, into a forecast. This approach can readily be extended to other similar information items like brand value, R&D expenditures and the like. Richardson and Tinaikar (2004, p. 228) find that: "Since the main contribution of the Ohlson (1995) and Feltham and Ohlson (1995, 1996) models has been to derive a parsimonious but theoretically supported pricing relation between accounting realizations and stock price (or returns), the onus lies with those who seek to modify the LID [linear information dynamics] to get more complex pricing relations to demonstrate the benefit of that additional complexity, given the research question". In this respect, we demonstrate that other information can be proxied by a simultaneous model incorporating single items of other publicly available non-accounting information, such as order backlog. The SEM is able to capture the interdependencies in a more appropriate way than do simple sequential models. It combines single items of non-accounting information in a sensible way. The SEM produces forecasts that are closer to analysts' than other models, particularly when times are volatile.

The second contribution of this paper is to the literature on analysts' forecasts (e.g. Bradshaw et al. 2012; Brown et al. 2013; Wang 2014). Our paper is one of the first to empirically model the forecasting process followed by analysts. Empirical studies trying to open the 'black box' of information processing by describing the forecasting process of capital market participants are rare. For example, Lambert et al. (2012) examine if analysts use the standard textbook framework to predict future earnings. Their findings imply that analysts do not use the 'forecasting framework' in the short-run and that forecast errors are associated with departures from that framework. The authors (p. 31) suggest that "future studies could focus on an attempt to forecast future earnings using the framework to improve forecast accuracy in relation to analysts' forecasts." In this respect, our paper presents a model that replicates some of the forecasting process, based on specific, publicly available information. By comparing the SEM forecasts with the analysts' forecasts, we obtain insights into the analysts' use of available information and the framework. In particular, the model helps to explain the information advantage of analysts found in many studies by linking items of information available beyond the financial statements to the forecast. We show that the SEM forecast is able to capture a major portion of information contained in analysts' forecasts beyond current earnings. Our approach helps to link single components of information to the aggregate earnings forecast of analysts. While analysts seem to be using the 'forecasting framework', some portion of their forecasts remain unexplained. Possible reasons could be analysts' industry- or firm-specific expertise and access to private information.

The remainder of the paper is structured as follows. Section 2 gives an overview of the related literature. In Sect. 3 the simultaneous equations model is derived. Section 4 describes the sample and reports descriptive statistics. Sections 5.1 and 5.2 present the results concerning model estimation and forecasting evaluation. In Sect. 5.3 the forecasts are integrated into the Ohlson (1995) model and tested for empirical validity. The association between the SEM forecasts and the analysts' forecasts are analyzed in Sect. 5.4. Section 6 presents sensitivity analyses. Finally, Sect. 7 concludes the paper.

2 Related literature

The Ohlson (1995) model serves as a theoretical framework for empirical studies analyzing the relationship between accounting numbers and firm value. It describes market value as a linear function of current earnings, book value, and other value-relevant information by assuming that future residual earnings and other value-relevant information follow a stochastic process. The 'other information' component presents a particular challenge for the empirical implementation of the model. Ohlson (1995, p. 668) describes 'other information' "as summarizing value relevant events that have yet to have an impact on the financial statements. Such information bears upon future (abnormal) earnings independently of current and past (abnormal) earnings." Barth et al. (2005) determine 'other information' as the difference of current market value of equity and predicted market value based on a regression excluding other information. Shen and Stark (2013) use the residual of a regression of last period's market value on last period's accounting variables as a proxy for current 'other information'. Myers (1999) uses order backlog because it is readily available, but finds that this does not improve residual income predictions. Many other studies (e.g. Dechow et al. 1999; Choi et al. 2006; Higgins 2011) use analysts' consensus earnings forecasts to measure 'other information' as suggested by Ohlson (2001). While these approaches to proxy for 'other information' help in improving the explanatory power of the regression, they do not contribute to our understanding of the information used by market participants beyond current accounting numbers. They do not reveal what constitutes the 'other information' term and hence, cannot contribute to our understanding of how market participants process information.

Analysts have often been the object of studies interested in the latter question. Research on the forecasting quality of analysts' earnings forecasts in comparison to time-series models' forecasts culminated in the 1980s with the overall conclusion that analysts' forecasts are superior to time-series models' forecasts.[1] This superiority has been found to be due to timing and information advantages of financial analysts, i.e., that they have access to and use additional and timelier information than time-series models (Fried and Givoly 1982; Brown et al. 1987). Analysts make use of other (non-accounting) information beyond financial statement information in order to improve their forecasts. Consequently, including analysts' forecasts in valuation models will quite necessarily lead to higher explanatory power. Accordingly, Richardson and Tinaikar (2004, p. 228) find that it "is a tautology that one can outperform a parsimonious relation when analyst forecasted abnormal earnings are added to the model containing only accounting realizations".

Bryan and Tiras (2007) confirm the usefulness of including analysts' earnings forecasts as 'other information' in the Ohlson model but show that the explanatory power of the regression varies with the quality of the information environment. In poor information environments, analysts base their forecasts more on other information than past earnings and book value. Hence, in poor information environments, the Ohlson model including analysts' forecasts has a greater explanatory power than a model with only earnings and book value or a model with only analysts' forecasts. Volatile times or times of economic disturbances cause a poor information environment and hence, one can conclude from their study that in these times, it is especially useful to include additional information beyond accounting numbers in order to increase explanatory power for market values.

However, relatively little research exists about how analysts actually generate their forecasts and what information is used by them. Apart from some content analyses (Previts et al. 1994; Rogers and Grant 1997) and surveys (Orens and Lybaert 2010), evidence from archival studies is sparse. Lambert et al. (2012) investigate whether analysts use stylized

[1] Recent evidence by Bradshaw et al. (2012) questions the widely accepted superiority of analysts' forecasts over time-series models' forecasts. They find that under certain conditions, i.e. for small firms and long horizons, random walk EPS forecasts are more accurate than analysts' EPS forecasts. For large and stable firms, over short horizons, however, analysts' forecasts are superior to random walk forecasts.

forecasting frameworks to predict future earnings by analyzing the association between the different components of the forecasting framework and analysts forecast errors. They find that analysts do not fully use the forecasting framework in the short-run and forecast errors are associated with departures from that framework. In the long-run, the components of the forecasting framework are informative for the forecasts. However, analysts do not generate forecasts based on the framework. Lambert et al. (2012) conclude that future studies should attempt to forecast future earnings using the forecasting framework to improve forecast accuracy in relation to analysts' forecasts.

In this paper, we develop a model that more accurately describes the forecasting process analysts presumably follow. We expect that our model is able to generate earnings forecasts that are superior to simple time-series-model's forecasts. When these forecasts are integrated into market value regressions, the explanatory power should increase. Furthermore, we expect that the forecasts based on the forecasting framework are able to explain analysts' forecasts. Due to the information advantages of analysts, we do not expect to generate a higher forecast accuracy than analysts, but to be able to explain some of the information advantages of analysts' forecasts.

3 Model framework

Textbooks on financial statement analysis and valuation (e.g. Penman 2010) suggest using forecasting frameworks in order to produce systematic and comprehensive earnings forecasts. Assuming that analysts produce forecasts in a systematic manner, we use these frameworks as a basis for our model building. According to these frameworks, earnings are not directly predicted. Instead, their value drivers are predicted, which are then combined to pro-forma financial statements. Due to the integrated nature of the balance sheet, income and cash flow statement, value drivers need to be forecasted simultaneously. We specify our model as a simultaneous equations model and determine the value drivers within in the model. These endogenous variables are defined based on the relationships described in the financial analysis literature. In addition, we derive exogenous variables from the forecasting literature, which are not determined within the model but required to forecast the endogenous variables.

3.1 Forecasting process

The typical starting point of the forecasting framework is the sales forecast (income statement), which is then followed by a forecast of the cost structure to predict profit margins and determine earnings. The investment forecast helps to predict the asset turnover, which determines future net assets (balance sheet). The amount of future net assets in turn determines the amount of future sales that can be generated. Furthermore, the depreciation charges to these assets enter the income statement as costs. An isolated prediction of the earnings number would neglect the interdependencies between the income statement and the balance sheet, and generate wrong estimates (White et al. 2003; Lundholm and Sloan 2007). Finally, the payout policy and financing activities need to be predicted for the calculation of future debt and interests (Penman 2010).

Forecasts are no ends in themselves but are used as inputs to valuation models in order to estimate firm value. The type of valuation model determines the input. We focus on the residual income valuation model (RIM) because of its connection to the Ohlson model. It expresses firm value as the present value of expected future residual income plus book value of capital. We build our model to forecast residual operating income, focusing on operating activities to generate value.

$$V_0^{NOA} = NOA_0 + \sum_{t=1}^{\infty} \frac{RI_t}{(1+r)^t} \tag{3.1}$$

where NOA = net operating assets, r = cost of capital, t = time index, RI = residual operating income, V^{NOA} = market value of operations (value of the firm)

Residual operating income (RI) can be split into its value drivers: sales (SALES), operating expenditures (OPEX) and net operating assets (NOA).

$$RI_t = OI_t \times (1 - tax) - r \times NOA_{t-1} = (SALES_t - OPEX_t) \times (1 - tax) - r \times NOA_{t-1} \tag{3.2}$$

where OI = operating income, $OPEX$ = operating expenditures, $SALES$ = sales, tax = tax rate.

We assume that tax rates are constant in the future. Forecasting NOA is not directly necessary for calculating future RI, but forecasting the other value drivers is because of their interdependencies. NOA consists of noncurrent assets (NCA) and net operating working capital (NWC). NCA is defined as NCA in the last period plus increases and decreases of last period's assets. NCA thus links the balance sheet to the income statement and statement of cash flows.

$$NOA_t = NCA_{t-1} + CAPEX_t - D_t + NWC_t$$

where $CAPEX$ = capital expenditures, D = depreciation, NCA = noncurrent assets, NWC = net operating working capital

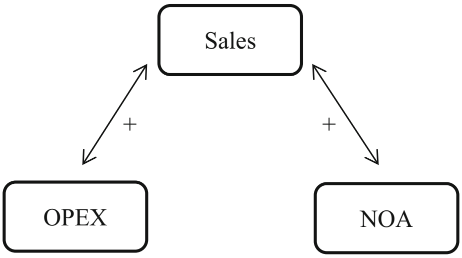
To predict next period's NOA, capital expenditures (CAPEX), depreciation of assets (D) and NWC are forecasted. The cost of capital is defined as the weighted average cost of capital (WACC), i.e. the required return on equity and debt as the total value of operations. Assuming a constant leverage ratio facilitates forecasting because constant WACC can be applied. To complete the pro forma balance sheet, income statement and cash flow statement, future payout ratios and the cost of debt capital need to be forecasted. However, when using RIM for forecasting and valuation, intrinsic value estimates are solely based on accounting measures and the forecast of SALES, OPEX and NOA is sufficient for valuation. Hence, we will focus on these when creating the forecasting model.[2]

3.2 Simultaneous equations model

The value drivers are linked via the pro forma financial statements and characterized by interactions. We model these interactions with a set of equations constituting the simultaneous equations model, where the dependent variables are jointly determined by several independent variables. The endogenous variables are explained within the model and are correlated with the error term (endogeneity). The values of the exogenous variables are determined outside the system of equations and are thus not correlated with the error term. They influence the endogenous variables but not vice versa. For this reason, the lagged endogenous variables are also considered to be exogenous to the model (Pindyck and Rubinfeld 1998).

[2] For the remaining equations to set up complete financial statements see the "Appendix".

Fig. 1 Interdependency between the endogenous variables



An important condition for the simultaneous model is that each equation has its own causal interpretation, that is, that one equation is not just a transformation of the other. In such a case, the same variables would be contained in both equations. For our model, this condition implies that each dependent variable is influenced by different factors. Based on the findings in the preceding section, a simple simultaneous model consists of three components that need to be forecasted: SALES, OPEX and NOA. In particular, we assume interdependencies between SALES and OPEX and between SALES and NOA. On the one hand, the amount of SALES that should be produced determines the NOA that need to be put in place to generate SALES. On the other hand, the NOA that are put in place reflect the ability to generate SALES. In case of capacity constraints, fewer SALES can be produced. In line with Lin (1992) we assume that OPEX and SALES are interrelated, because operating expenses will increase if sales increase and sales will increase if operating expenses, e.g. marketing expenditures, increase. OPEX and NOA are considered to be indirectly related through SALES. This relationship is thus not explicitly included in the model (Fig. 1).[3]

Based on these interdependencies, we can set up the following structural model:

$$SALES_{it} = f(OPEX_{it}, NOA_{it}, SALES_{it-1}, EV) + \varepsilon_{it} \tag{3.3}$$

$$OPEX_{it} = f(SALES_{it}, OPEX_{it-1}, EV) + \eta_{it} \tag{3.4}$$

$$NOA_{it} = f(SALES_{it}, NOA_{it-1}, EV) + \upsilon_{it} \tag{3.5}$$

where EV = exogenous variables, i = firm index, ε, η, υ = error terms structural model

The structural model consists of an equation for each value driver, which depends on the other value drivers (endogenous variables), its lagged value (lagged endogenous variables) and further exogenous variables. Besides these three stochastic Eqs. (3.3) to (3.5), the model is composed of the following deterministic Eqs. (3.6) and (3.7) that are necessary to calculate OI and RI.

$$OI_{it} = (SALES_{it} - OPEX_{it}) \tag{3.6}$$

$$RI_{it} = OI_{it} \times (1 - tax) - r \times NOA_{it-1} \tag{3.7}$$

Only stochastic equations need to be estimated. However, the OLS-method would produce biased and inconsistent estimates when the endogenous variables serve as

[3] However, including this relationship in the model would not alter the final reduced form of Eqs. (3.8)–(3.10).

independent variables. One way to estimate the structural model is to solve it for each of the endogenous variables to obtain its reduced form where feedbacks no longer exist. In the reduced form of the model, the endogenous variables are a function of the predetermined variables and the error terms only (Pindyck and Rubinfeld 1998). Standard techniques like OLS can be applied to estimate the reduced form. The estimated model is then used to generate the forecasts. For forecasting purposes, the coefficients of the variables are of minor interest, but for interpretation purposes, the coefficients of the structural form are necessary. Two-stage least squares (2SLS) and three-stage least squares (3SLS) regressions are useful estimation procedures for obtaining the values of the structural parameters.

To eliminate endogeneity, linear transformations of the Eqs. (3.3) to (3.5) yields the reduced form of the equations model:

$$SALES_{it} = f(SALES_{it-1}, NOA_{it-1}, OPEX_{it-1}, EV) + \kappa_{it} \tag{3.8}$$

$$OPEX_{it} = f(SALES_{it-1}, NOA_{it-1}, OPEX_{it-1}, EV) + \lambda_{it} \tag{3.9}$$

$$NOA_{it} = f(SALES_{it-1}, NOA_{it-1}, OPEX_{it-1}, EV) + \mu_{it} \tag{3.10}$$

where κ, λ, μ = error terms reduced model

3.3 Exogenous variables

For each of the value drivers SALES, OPEX and NOA, we derive the exogenous variables from the literature separately. We further classify them into macroeconomic, industry-specific and firm-specific variables.

3.3.1 Sales

The sales forecast is the first and most important step because other forecasts directly rely on it (Koller et al. 2010). Sales are dependent on product prices and quantities and driven by, for example, competition, product substitutes, brand association, and patent protection (Penman 2010). There is a large number of macroeconomic factors that influence future sales of all firms in the economy. So far, empirical analysis makes only little use of these factors in forecasting (Richardson et al. 2010). As suggested by Richardson et al. (2010), we select macroeconomic variables based on economic reasoning concerning their influence on value drivers of residual income.

Qualitative variables (or indicators) are used in empirical research to provide early signals of changes in the economy. They are intended to measure the sentiment and expectations of producers and/or consumers concerning the future economic development. Examples are the *Purchasing Managers Index (PMI),* the *University of Michigan Consumer Sentiment Index (MCSI)* and the *OECD composite leading indicator (CLI)*, where the latter comprise the other two indicators. We expect a positive influence of the change in the sentiment on future sales.

Important quantitative macroeconomic variables used in empirical studies (e.g. Lin 1992; Lev and Thiagarajan 1993) are the inflation rate and the change in *gross domestic product (GDP).* The change in *GDP* indicates the overall rate at which the economy is growing and is expected to positively influence sales. The inflation rate, defined as the change in the *consumer price index (CPI)*, indicates a general rise in price levels. In nominal terms, revenue will increase. But in real terms, sales may decline because uncertainty is higher when inflation is high. We use industry-specific variables that are relevant for all industries, but calculated at industry level. The *growth of a market* is measured as the growth rate of sales within an industry and is expected to be positively related to future sales. Further variables are derived from Porter's (1998) concept of the five forces. Of these, rivalry among existing firms and threat of new entrants relate to sources of direct competition (Lundholm and Sloan 2007) and can be proxied by market concentration and barriers to entry. Market concentration can be measured by the *Herfindahl index*, calculated as the sum of squared market shares of all firms in an industry. Barriers to entry can be measured by *capital intensity*, i.e. depreciation expenditures divided by sales. The accompanying industry measure is the sales weighted average of the particular intensity (Cheng 2005a). For high levels of competition, i.e. low barriers to entry and a low concentration, prices decline and profitability converges towards the cost of capital. Thus, a positive relation between barriers to entry (concentration) and sales (and especially profit margin) is assumed.

Firm-specific determinants are relevant if firms are not homogenous within industries (Cheng 2005a). Due to different strategies that firms may adopt within an industry, firm-specific variables are also useful for predicting future sales. The most important exogenous variable considered in this paper is *order backlog*. Order backlog is a voluntarily disclosed firm-specific indicator for future sales. As order backlog follows a random walk process (Rajgopal et al. 2003), we assume that current period's order backlog is positively related to next period's sales. The *market share* can be considered a direct indicator for future sales. A high current market share also implies high sales in the future. It is measured as the ratio of a firm's sales to industry sales. *Capital intensity* can also be calculated on firm-level basis and measures the minimum required capital (Cheng 2005a).

3.3.2 Operating expenditures (profit margin)

For a given sales level, OPEX determine the profit margin. Hence, in order to identify adequate exogenous variables, we analyze determinants influencing the profit margin. In general, profit margin is driven by production technology, economies of scale, learning, competitiveness in labor and supplier markets (Penman 2010), market power, and competition.

Macroeconomic variables that can be used to forecast operating expenses are changes of cost indices like the *Producer Price Index (PPI)*. In contrast to CPI, PPI is based on prices producers pay and mainly has an impact on expenses. We assume a positive relationship between change in PPI and OPEX. On the industry-level, *market growth* is also considered to influence margins. If market growth is small, the competition for market shares will be stronger, leading to smaller profit margins. On the firm-level, the *market share* can be used for forecasting profit margins, because a high market share leads to economies of scale in purchasing, production, marketing, etc. (Cheng 2005a) and hence, lower costs and higher margins.

3.3.3 Net operating assets (asset turnover)

Asset turnover indicates to what extent net assets are able to generate sales and is thus driven by the production technology of the firm. Besides sales, changes in working capital and long-term assets needs to be predicted to calculate asset turnover. Long-term assets, like property, plant, and equipment (PP&E), are determined by capital expenditures, which are driven, for example, by the introduction of new products and the direction of the future business. Working capital can further be split into accounts payable, inventories and accounts receivable. The amount of accounts payable is driven by supplier power and material costs, accounts receivable by customer power and their solvency, product quality and the distribution policy and inventories by the quality management and logistics of the firm (Penman, 2010).

We use the *federal funds rate (fedrate)*[4] to predict net assets. An increasing fedrate is an indicator of an increase of other interest rates, like interest rates on mortgages, loans or savings deposits. Higher interest rates on debt capital will increase the cost of capital which in turn will decrease the number of viable investment opportunities. Hence, we expect a negative relationship between the current fedrate and future assets. We also expect a negative effect on sales because consumer investments will decrease if interest rates increase (Lundholm and Sloan 2007).

4 Data and descriptive statistics

The sample comprises all US industrial firms (active and inactive) included in the Compustat database that report order backlog data within the period of 1991 to 2010. All variables are deflated by lagged total assets to control for heterogeneity of the sample firms. Further, the extreme upper and lower 1% of the distribution of each variable is eliminated (trimming) to reduce the influence of outliers on the regressions. In line with other empirical studies, regulated firms, including financial institutions (SICs between 6000 and 6999) and utilities (SICs between 4900 and 4999), are excluded because their operations are markedly different from other firms (Cheng 2005b). SIC codes 9000 and above are excluded (Begley and Feltham 2002) for the same reasons. Firm-years with negative net operating assets in the current or prior period are eliminated. The sample selection process is reported in Table 1.

Net operating assets (NOA) are defined as operating assets (OA) minus operating liabilities (OL), i.e. all assets except cash and marketable securities minus non-interest bearing debt. NOA is calculated in accordance with other studies (Nissim and Penman 2001; Callen and Segal 2005; or Soliman 2008) (item numbers in Compustat in parentheses) as:

$$NOA = OA - OL$$

with

$$\begin{aligned} OA &= \mathit{total\ assets}(TA, \#6) - \mathit{financial\ assets}\,(FA) \\ &\quad FA = \mathit{cash}\ \&\ \mathit{short\ term\ investments}\,(CSTI, \#1) \\ &\qquad + \mathit{other\ investments\ and\ advances}\,(OIA, \#32) \\ OL &= \mathit{total\ liabilities}\,(TL, \#181) - \mathit{financial\ liabilities}\,(FL) \\ &\quad FL = \mathit{long\ term\ debt}\,(LTD, \#9) + \mathit{debt\ in\ current\ liabilities}\,(DCL, \#34) \\ &\qquad + \mathit{minority\ interests}\,(MI, \#38) \end{aligned}$$

NOA is thus defined as:

[4] As the rate is negotiated between the banks and not determined by the Federal Reserve System (FED), we use the effective rate, i.e. the weighted average across all rates.

Table 1 Sample selection procedure

Sample	Firm-year observations
1. US industrial firms from 1991 to 2010	43,645
2. With positive order backlog data	27,941
3. Observations with non-missing data	20,477
4. After trimming	18,920
5. After elimination of industries	18,533
6. With positive net operating assets	18,057

$$NOA = TA - CSTI - OIA - TL + LTD + DCL + MI$$

In Compustat, other investments and advances (OIA) are defined as investments to unconsolidated subsidiaries and affiliates in which the parent company has no control. Like cash & short term investments (CSTI), it is assumed that these investments are not part of the operating assets. Minority interests (MI) are included in NOA because these also have to earn the required return.

Operating expenditures (OPEX) are the sum of costs of goods sold (COGS, #41), selling, general and administrative expenses (SG&A, #189), depreciation and amortization (D, #14). SALES is measured as net sales (SALES, #12), i.e.gross sales reduced by cash discounts, trade discounts, returned sales and allowances for which credit is given to customers. The measurement of the exogenous variables is described in Table 2. If a variable is considered an indicator or its calculation is based on amounts that we aim to forecast (e.g. sales), its lagged value is used to estimate the equations. Otherwise forecasts of these exogenous variables are required. In our model, the latter applies only for the change of GDP, CPI and PPP. We include the actual values of these variables in the model assuming that perfect forecasts are available at the time the forecast is made. The advantage is that forecast errors of the input variables do not influence the forecast of the variable of interest. However, this simplification is only possible in an ex post analysis.

Descriptive statistics for the main variables used in the estimations are shown in Table 3. The correlations between all of the variables are shown in Table 4. Apart from lagged NOA and OPEX, all correlations between the endogenous variables are positive. In addition to high correlations (> 0.5) between current and lagged values, SALES and OPEX are also highly correlated. Further, the qualitative indicators lagged ΔPMI, lagged ΔCLI and lagged ΔMCSI are highly positively correlated. This is not surprising because their common intention is to measure changes in business climate and MCSI and PMI are components of CLI. It follows that it is only useful to include one of the indicators in the empirical study. We also observe high correlations between the qualitative indicators and the quantitative macroeconomic variables ΔCPI, ΔGDP and ΔPPI, because the indicators are supposed to predict changes in the economy. The negative correlations between lagged CapIntF and SALES as well as OPEX are counterintuitive, since higher market barriers to entry are assumed to lead to less competition, higher sales and profit margins respectively, as illustrated in the previous section. A reason for this negative relation may be the measurement of capital intensity (as proxy for barriers to entry) as depreciation and amortization (D&A) divided by sales. Increasing CapIntF may either be due to increasing D&A costs or decreasing SALES (or both). If, for example, amortization of intangibles increases, the future prospects of the firms might be negative and SALES might decrease. And if SALES decrease, OPEX (in terms of COGS) will also decrease.

Table 2 Measurement of exogenous variables

Variable	Definition	Measurement	Expected sign
Macroeconomic			
ΔCPI	Percentage change in consumer price index (Source: U.S. Department of Labor: Bureau of Labor Statistics)	$\Delta CPI_t = \frac{(CPI_t - CPI_{t-1})}{CPI_{t-1}}$	± (SALES)
ΔGDP	Percentage change in gross domestic product (Source: U.S. Department of Commerce: Bureau of Economic Analysis)	$\Delta GDP_t = \frac{(GDP_t - GDP_{t-1})}{GDP_{t-1}}$	+ (SALES)
ΔPPI	Percentage change in producer price index (Source: U.S. Department of Labor: Bureau of Labor Statistics)	$\Delta PPI_t = \frac{(PPI_t - PPI_{t-1})}{PPI_{t-1}}$	+ (OPEX)
Fedrate (lagged)	Federal funds rate (Source: Federal Reserve)	Average of daily effective federal funds rates	− (NOA) − (SALES)
ΔPMI (lagged)	Percentage change in Purchasing Managers Index (Source: ISM)	$\Delta PMI_t = \frac{(PMI_t - PMI_{t-1})}{PMI_{t-1}}$	+ (SALES)
ΔMSCI (lagged)	Percentage change in Consumer Sentiment Index (Source: Survey Research Center: University of Michigan)	$\Delta MSCI_t = \frac{(MSCI_t - MSCI_{t-1})}{MSCI_{t-1}}$	+ (SALES)
ΔCLI (lagged)	Percentage change in OECD composite leading indicator (Source: OECD)	$\Delta CLI_t = \frac{(CLI_t - CLI_{t-1})}{CLI_{t-1}}$	+ (SALES)
Industry-specific			
HI (lagged)	Herfindahl index (market concentration)	$HI_{jt} = \sum MS_{it}^2$	+ (SALES)
CapIntI (lagged)	Capital intensity of the industry (barriers to entry)	$CapIntI_{jt} = \sum\left(\frac{D\&A_{it}}{SALES_{it}} \times MS_{it}\right)$	+ (SALES)
MG (lagged)	Market growth	$MG_{jt} = \frac{\sum SALES_{it} - \sum SALES_{it-1}}{\sum SALES_{it-1}}$	+ (SALES) − (OPEX)
Firm-specific			
OB (lagged)	Order backlog	Compustat item # 98	+ (SALES)
MS (lagged)	Market share	$MS_{it} = \frac{SALES_{it}}{\sum SALES_{it}}$	+ (SALES) − (OPEX)
CapIntF (lagged)	Capital intensity of the firm (barriers to entry)	$CapintF_{it} = \frac{D\&A_{it}}{SALES_{it}}$	+ (SALES)

Industry subscript *j*, firm subscript *i*. The percentage change of the macroeconomic variables is based on annual changes of the averages of monthly values. Markets are classified based on two-digit-SIC codes; MS is calculated based on three-digit-SIC codes. Advertising expenditures are excluded as only some firms disclose this information. *SALES* is sales, *OPEX* is operating expenditures and *NOA* is net operating assets

Table 3 Distribution statistics

Variable	Number of observations	Median	Mean	Minimum	Maximum
TA_{it}	18,057	142	1280	0.202	371,000
$SALES_{it}$	18,057	168	1190	0.274	184,000
NOA_{it}	18,057	77	628	0.001	172,000
$OPEX_{it}$	18,057	157	1080	0.368	177,000
OB_{it}	18,057	39	753	0.002	160,000
$SALES_{it}/TA_{it-1}$	18,057	1.2545	1.3609	0.1665	4.3311
NOA_{it}/TA_{it-1}	18,057	0.6181	0.6172	0.0000	1.8965
$OPEX_{it}/TA_{it-1}$	18,057	1.1858	1.2976	0.2076	4.2969
OB_{it-1}/TA_{it-2}	18,057	0.3086	0.5542	0.0082	5.8681
$SALES_{it-1}/TA_{it-2}$	18,057	1.2815	1.3942	0.1929	4.3422
NOA_{it-1}/TA_{it-2}	18,057	0.6297	0.6320	0.0002	1.8995
$OPEX_{it-1}/TA_{it-2}$	18,057	1.2072	1.3236	0.2201	4.2358
ΔCPI_t	18,057	2.6674	2.5104	−0.3206	3.8154
ΔGDP_t	18,057	3.0710	2.8590	−3.4858	4.8254
ΔPPI_t	18,057	2.3237	2.3428	−8.8080	9.7858
$Fedrate_{t-1}$	18,057	4.21	3.9623	0.16	6.24
ΔPMI_{t-1}	18,057	1.35	0.6911	−15.96	17.1
$\Delta MCSI_{t-1}$	18,057	1.15	0.4916	−25.47	11.47
ΔCLI_{t-1}	18,057	3.1691	2.8549	−2.5770	5.2229
HI_{t-1}	18,057	0.0453	0.0617	0.0204	0.8155
$CapIntI_{t-1}$	18,057	0.0443	0.0486	0.0023	0.3175
MG_{t-1}	18,057	8.3204	8.7292	−86.9933	486.1529
MS_{it-1}	18,057	0.0569	0.6957	0.0001	69.8663
$CapIntF_{it-1}$	18,057	0.0323	0.0405	0	1.0233

TA is total assets (#6). *SALES* is net sales (#12). *NOA* is net operating assets (= #6 − #1 − #32 − #181 + #9 + #34 + #38). *OPEX* is operating expenditures (#41 + #189 + #14). *OB* is order backlog (#98). Undeflated values of *TA*, *SALES*, *NOA*, *OPEX* and *OB* are measured in Million USD. *ΔCPI* is the change in consumer price index [%]. *ΔGDP* is the change in gross domestic product [%]. *ΔPPI* is the change in producer price index [%]. *Fedrate* is the federal funds rate [%]. *ΔPMI* is the change in Purchasing Managers Index [%]. *ΔMCSI* is the change in University Michigan Consumer Sentiment Index [%]. *ΔCLI* is the change in OECD composite leading indicator [%]. *HI* is the Herfindahl index. *CapIntI* is capital intensity of the industry. *MG* is market growth [%]. *MS* is market share [%]. *CapIntF* is capital intensity of the firm

5 Results

5.1 Model estimation

We estimate the reduced form of the SEM including the exogenous variables described in the preceding section. All values are lagged apart from the three quantitative macroeconomic variables ΔGDP, ΔPPI and ΔCPI where actual values in t serve as explanatory variables for the value drivers in t. By doing so, it is assumed that perfect forecasts are available in $t - 1$. We use the generalized least squares (GLS) method allowing estimation in the presence of both heteroskedasticity across panels and autocorrelation within panels. Here, we assume a heteroskedastic error structure with no cross-sectional correlation, i.e.

Table 4 Correlations

	$SALES_{it}$	$SALES_{it-1}$	NOA_{it}	NOA_{it-1}	$OPEX_{it}$	$OPEX_{it-1}$	OB_{it-1}	ΔCPI_t	ΔGDP_t	ΔPPI_t	$Fedrate_{t-1}$
$SALES_{it}$		**0.8059***	*0.2538**	0.0734*	**0.9575***	**0.8057***	*0.3009**	0.0514*	0.1437*	−0.0198*	0.0674*
$SALES_{it-1}$	**0.7938***		0.1632*	0.2496*	**0.7847***	**0.9618***	*0.3540**	0.0383*	0.1191*	−0.0919*	0.1514*
NOA_{it}	0.2479*	0.1405*		**0.6908***	0.1923*	0.1149*	0.0344*	0.0169*	0.1468*	−0.0436*	0.0936*
NOA_{it-1}	0.0367*	*0.2513**	**0.6068***		0.0390*	0.2005*	0.0682*	0.0322*	0.0952*	−0.0765*	0.1564*
$OPEX_{it}$	**0.9642***	**0.7737***	0.1865*	−0.0018		**0.8423***	*0.2767**	0.0450*	0.1377*	−0.0335*	0.0744*
$OPEX_{it-1}$	**0.7947***	**0.9678***	0.0911*	0.2016*	**0.8241***		*0.3373**	0.0343*	0.1168*	−0.0906*	0.1407*
OB_{it-1}	*0.2904**	*0.3365**	−0.002	0.0322*	*0.2778**	*0.3302**		0.0365*	0.0253*	−0.0051	0.0551*
ΔCPI_t	0.0879*	0.0594*	0.0503*	0.0586*	0.0810*	0.0572*	0.0103		−0.1659*	**0.6816***	0.0306*
ΔGDP_t	0.1569*	0.1093*	0.1563*	0.1052*	0.1491*	0.1082*	−0.0162*	*0.3352**		−0.2100*	0.2465*
ΔPPI_t	0.0128	−0.0499*	−0.0145	−0.0370*	0.0021	−0.0487*	0.0071	**0.8053***	0.2077*		*−0.4263**
$Fedrate_{t-1}$	0.0714*	0.1466*	0.0956*	0.1595*	0.0756*	0.1359*	0.0141	0.2044*	0.2141*	−0.2260*	
ΔPMI_{t-1}	0.0624*	0.0182*	0.0369*	0.0049	0.0583*	0.0214*	−0.0036	0.2312*	0.2484*	*0.3258**	*−0.3008**
$\Delta MCSI_{t-1}$	0.1410*	0.1192*	0.1208*	0.0960*	0.1303*	0.1123*	0.0122	*0.4619**	**0.6446***	*0.4033**	0.1666*
ΔCLI_{t-1}	0.1400*	0.1439*	0.1291*	0.1365*	0.1399*	0.1406*	−0.0043	*0.4380**	**0.5911***	0.1249*	*0.4152**
HI_{t-1}	0.1795*	0.1728*	0.0109	0.0118	0.1894*	0.1838*	0.0693*	0.0259*	−0.0101	0.0066	−0.0546*
$CapIntI_{t-1}$	−0.1956*	−0.2005*	−0.0851*	−0.0829*	−0.1841*	−0.1887*	−0.1075*	−0.0450*	−0.0382*	−0.0051	−0.0746*
MG_{t-1}	0.0835*	0.1255*	0.0629*	0.0925*	0.0774*	0.1143*	0.0703*	0.1614*	0.1025*	0.0093	0.2334*
MS_{it-1}	0.0766*	0.0800*	−0.0539*	−0.0498*	0.0670*	0.0717*	0.1311*	−0.0154*	−0.0472*	0.0156*	−0.0377*
$CapIntF_{it-1}$	*−0.3480**	*−0.4008**	−0.0637*	−0.0546*	*−0.2994**	*−0.3364**	−0.1479*	−0.0131	−0.0032	0.0126	−0.0434*

Table 4 continued

	ΔPMI_{t-1}	$\Delta MCSI_{t-1}$	ΔCLI_{t-1}	HI_{t-1}	$CapIntI_{t-1}$	MG_{t-1}	MS_{it-1}	$CapIntF_{it-1}$
$SALES_{it}$	0.0718*	0.1225*	0.1231*	0.2153*	−0.2010*	0.0947*	0.0666*	*−0.4655**
$SALES_{it-1}$	0.0165*	0.1211*	0.1267*	0.1951*	−0.2079*	0.1411*	0.0881*	**−0.5428***
NOA_{it}	0.0357*	0.1089*	0.1207*	0.0568*	−0.0820*	0.0659*	0.0639*	−0.0359*
NOA_{it-1}	−0.0047	0.0923*	0.1119*	0.0475*	−0.0801*	0.0963*	0.0780*	−0.0372*
$OPEX_{it}$	0.0672*	0.1182*	0.1233*	0.2106*	−0.1887*	0.0885*	−0.0108	*−0.4377**
$OPEX_{it-1}$	0.0212*	0.1159*	0.1236*	0.1957*	−0.1951*	0.1275*	0.012	**−0.5002***
OB_{it-1}	0.0036	0.0444*	0.0299*	0.1257*	−0.1376*	0.1102*	0.0581*	−0.2463*
ΔCPI_t	0.1618*	−0.1322*	0.0569*	0.0813*	−0.0528*	0.1912*	−0.0058	−0.0297*
ΔGDP_t	0.1104*	*0.3952**	**0.5495***	−0.0051	−0.0206*	0.1421*	−0.0406*	−0.0104
ΔPPI_t	*0.3584**	−0.0699*	−0.1018*	0.0453*	−0.0101	−0.0078	0.0251*	0.0044
$Fedrate_{t-1}$	*−0.3728**	0.1924*	0.2406*	−0.1103*	−0.0423*	*0.2621**	−0.0367*	−0.0362*
ΔPMI_{t-1}		*0.4757**	**0.6474***	0.1193*	0.0029	0.0018	−0.0104	0.0038
$\Delta MCSI_{t-1}$	**0.6005***		**0.6550***	0.0526*	−0.0224*	0.1647*	−0.0281*	−0.0172*
ΔCLI_{t-1}	**0.5758***	**0.6619***		−0.0094	−0.0224*	0.1935*	−0.0421*	−0.0151*
HI_{t-1}	0.0562*	0.0331*	−0.0094		*−0.3646**	0.0276*	*0.3233**	−0.1998*
$CapIntI_{t-1}$	−0.0256*	−0.0980*	−0.0665*	−0.2297*		−0.0388*	*−0.3255**	*0.4031**
MG_{t-1}	−0.0027	0.1694*	0.2265*	0.0104	−0.0577*		−0.0521*	−0.0542*
MS_{it-1}	−0.0111	−0.0321*	−0.0502*	*0.2813**	−0.1543*	−0.0193*		−0.1021*
$CapIntF_{it-1}$	0.0072	−0.0277*	−0.0187*	−0.0742*	*0.3521**	−0.0303*	−0.0718*	

Bravais-Pearson correlations below the diagonal and Spearman rank correlations above the diagonal. Probabilities reported are: significant at $p < 0.05$*. Correlations > 0.5 and < −0.5 are given in bold; Correlations > 0.25 and < −0.25 are given in italics

SALES is deflated sales. *NOA* is deflated net operating assets. *OPEX* is deflated operating expenditures. *OB* is deflated order backlog. *ΔCPI* is the change in consumer price index. *ΔGDP* is the change in gross domestic product. *ΔPPI* is the change in producer price index. *Fedrate* is the federal funds rate. *ΔPMI* is the change in Purchasing Managers Index. *ΔMCSI* is the change in University Michigan Consumer Sentiment Index. *ΔCLI* is the change in OECD composite leading indicator. *HI* is the Herfindahl index. *CapIntI* is capital intensity of the industry. *MG* is market growth. *MS* is market share. *CapIntF* is capital intensity of the firm

Table 5 GLS SEM (reduced form)

$$SALES_{it} = \alpha_{10} + \alpha_{11}SALES_{it-1} + \alpha_{12}OB_{it-1} + \alpha_{13}OPEX_{it-1} + \alpha_{14}NOA_{it-1} + \alpha_{15}fedrate_{t-1} + \alpha_{16}\Delta GDP_t + \alpha_{17}\Delta PPI_t + \alpha_{18}HI_{t-1} + \alpha_{19}CapIntI_{t-1} + \alpha_{110}MS_{it-1} + \alpha_{111}CaptIntF_{it-1} + \kappa_{it}$$

$$OPEX_{it} = \alpha_{20} + \alpha_{21}SALES_{it-1} + \alpha_{22}OB_{it-1} + \alpha_{23}OPEX_{it-1} + \alpha_{24}NOA_{it-1} + \alpha_{25}fedrate_{t-1} + \alpha_{26}\Delta GDP_t + \alpha_{27}\Delta PPI_t + \alpha_{28}HI_{t-1} + \alpha_{29}CapIntI_{t-1} + \alpha_{210}MS_{it-1} + \alpha_{211}CaptIntF_{it-1} + \lambda_{it}$$

$$NOA_{it} = \alpha_{30} + \alpha_{31}SALES_{it-1} + \alpha_{32}OB_{it-1} + \alpha_{33}OPEX_{it-1} + \alpha_{34}NOA_{it-1} + \alpha_{35}fedrate_{t-1} + \alpha_{36}\Delta GDP_t + \alpha_{37}\Delta PPI_t + \alpha_{38}HI_{t-1} + \alpha_{39}CapIntI_{t-1} + \alpha_{310}MS_{it-1} + \alpha_{311}CaptIntF_{it-1} + \mu_{it}$$

	Equation (1) $SALES_{it}$ Coef.[std.errors]	Equation (2) $OPEX_{it}$ Coef.[std.errors]	Equation (3) NOA_{it} Coef.[std.errors]
$SALES_{it-1}$	0.4366*** [0.00]	−0.2603*** [0.00]	0.1223*** [0.00]
OB_{it-1}	0.0140*** [0.00]	0.0040*** [0.00]	−0.0048*** [0.00]
$OPEX_{it-1}$	0.3678*** [0.00]	1.0762*** [0.00]	−0.1414*** [0.00]
NOA_{it-1}	−0.3952*** [0.00]	−0.3809*** [0.00]	0.5779*** [0.00]
$Fedrate_{t-1}$	−0.0129*** [0.00]	−0.0077*** [0.00]	−0.0036*** [0.00]
ΔGDP_t	0.0312*** [0.00]	0.0266*** [0.00]	0.0156*** [0.00]
ΔPPI_t	0.0037*** [0.00]	0.0025*** [0.00]	−0.0011*** [0.00]
HI_{t-1}	0.3711*** [0.01]	0.3672*** [0.01]	0.0495*** [0.00]
$CapIntI_{t-1}$	−0.6616*** [0.02]	−0.5113*** [0.02]	−0.3271*** [0.00]
MS_{it-1}	−0.0023*** [0.00]	−0.0028*** [0.00]	−0.0031*** [0.00]
$CapIntF_{it-1}$	−0.7046*** [0.02]	−0.6760*** [0.01]	−0.0940*** [0.00]
constant	0.5021*** [0.00]	0.4558*** [0.00]	0.2671*** [0.00]
No. of Obs.	17,705	17,705	17,705
Wald Chi2	1,650,897	2,625,267	13,218,132

SALES is deflated sales. *NOA* is deflated net operating assets. *OPEX* is deflated operating expenditures. *OB* is deflated order backlog. *ΔGDP* is the change in gross domestic product. *ΔPPI* is the change in producer price index. *Fedrate* is the federal funds rate. *HI* is the Herfindahl index. *CapIntI* is capital intensity of the industry. *MS* is market share. *CapIntF* is capital intensity of the firm. *MG* was excluded in the estimation because this variable is insignificant in all first stage regressions of the structural model. Due to the high correlation between *ΔCPI* and *ΔPPI*, *ΔCPI* was also excluded in Eq. (1)

The smaller sample size (less 352 observations) is due the GLS regressions which require more than one observation per group, i.e. per firm

* $p < 0.10$; ** $p < 0.05$; *** $p < 0.010$

the variance is different for each individual, and AR(1) autocorrelation with panel (or individual) coefficients of the AR(1) process. Table 5 shows the results.

All variables are highly significant and the validity of the model is high. Interpretation of the coefficients in the reduced form may be misleading, as it is derived from the structural model. The interpretation of the effect of a particular independent variable on the dependent variable can be difficult as the coefficients may have the wrong sign (Hanke and Wichern 2009). However, for forecasting purposes, we are not primarily interested in the interpretation of the coefficients but in the prediction of the left hand-side variables.

5.2 Forecasting performance

5.2.1 Benchmark models

Next, we use the estimated model to generate forecasts of the value drivers. In addition to the model including the actual values of the macroeconomic variables, we also estimate and use two further variants, i.e. one model including one of the lagged indicators instead ('indicator' model) and a version including none of the exogenous variables ('simple' model). In order to assess the forecasting performance, we compare operating income (OI) forecasts from the different SEM versions, which are all calculated according to Eq. (3.6), with two models that are extensively used as benchmarks in the analyst forecast literature (e.g. Capstaff et al. 2001; Higgins 2011; Bradshaw et al. 2012), i.e. OI forecasts from a first-order autoregressive model (AR(1)) and a random walk (RW) model. According to the AR(1) model, next period's OI is a function of last period's OI:

$$OI_{it} = f(OI_{it-1}) + \varepsilon_{it} \tag{5.1}$$

Further, a dummy variable (LOSS) is included in the different SEMs and the AR(1) model that equals 1 if OI is negative and 0 otherwise. The dummy variable allows for different intercepts for profit and loss firms. The RW model is described by the equation:

$$OI_{it} = OI_{it-1} + \varepsilon_{it} \tag{5.2}$$

In-sample forecasts do not separate the estimation period from the forecast period whereas out-of-sample forecasts are forecasts for a time period after the estimation period. In this paper we present the errors of both types. The different types of forecast errors are described in Table 6. The in-sample forecasting results for all models are displayed in Table 7. The estimation sample consists of 17,705 observations as 352 observations are dropped in order to perform the GLS regression.

Bias, measured by the signed error (ME), of all forecasting models is close to zero, indicating that the models do not produce systematic over- or underestimations of OI. An ME of zero, however, does not indicate a perfect forecast as negative and positive values may neutralize. The ME of the RW model is distinguishably higher than the MEs of the other models implying that the RW model tends to overestimate future income. Accuracy is measured as the absolute error by median average percentage error (MdAPE), mean absolute error (MAE) and root mean square error (RMSE). MdAPE is used instead of the mean average percentage error (MAPE) as ratios produce outliers when the denominator, deflated OI, is close to zero.

The highest accuracy is obtained by the SEM with additional exogenous variables, irrespective of whether an indicator variable (e.g. lagged ΔMCSI) or current values of macroeconomic variables (ΔGDP, ΔPPI) are included in the regression. Lagged ΔMCSI is

Table 6 Error measures

ME (mean error)	$ME = \frac{1}{n}\sum_{t=1}^{n}(\hat{x}_t - x_t)$
MAE (mean absolute error)	$MAE = \frac{1}{n}\sum_{t=1}^{n}\lvert\hat{x}_t - x_t\rvert$
MAPE (mean absolute percentage error)	$MAPE = \frac{1}{n}\sum_{t=1}^{n} APE = \frac{1}{n}\sum_{t=1}^{n}\left\lvert\frac{\hat{x}_t - x_t}{x_t}\right\rvert \cdot 100\%$
MdAPE (median absolute percentage error)	$MdAPE = Median(APE)$
RMSE (root mean squared error)	$RMSE = \sqrt{\frac{1}{n}\sum_{t=1}^{n}(\hat{x}_t - x_t)^2}$

Table 7 In-sample forecast errors OI

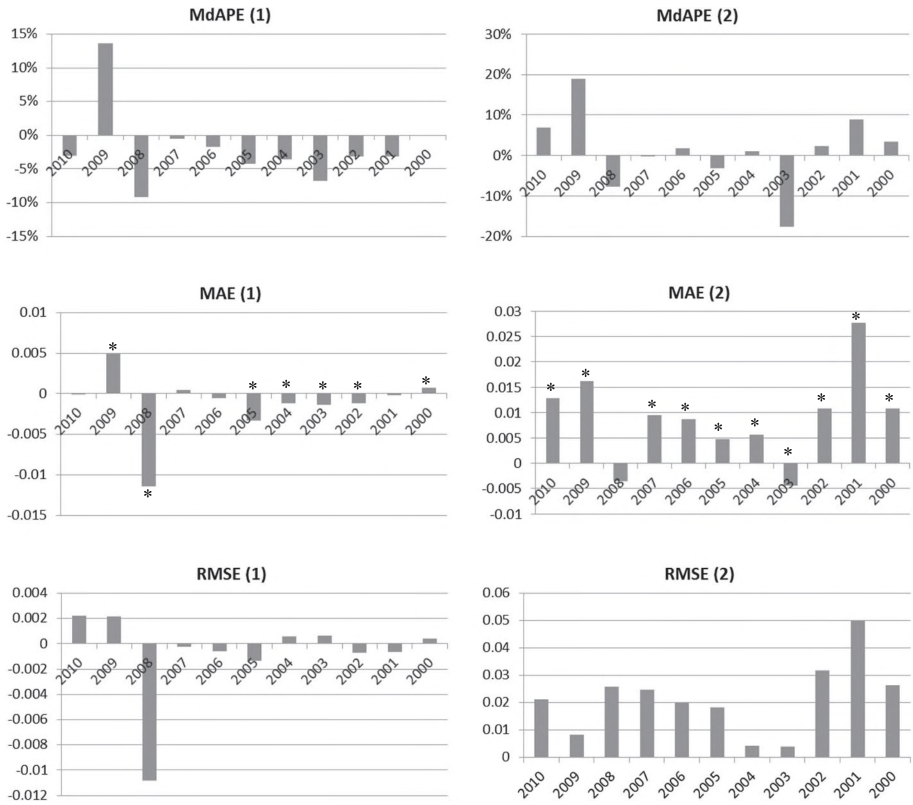
	MdAPE (%)	ME	MAE	RMSE
SEM	38.05	0.0011	*0.0619*	*0.1032*
SEM—'indicator' model (MCSI)	*37.95*	0.0001	0.0620	0.1034
SEM—'simple' model	38.66	*0.0000*	0.0622	0.1036
AR(1) model	39.18	0.0013	0.0625	0.1039
RW model	42.92	0.0074	0.0756	0.1306

The SEM considers the exogenous variables inclusive actual values of ΔGDP and ΔPPI. The SEM 'indicator' model includes MCSI as indicator instead of ΔGDP and ΔPPI. The SEM 'simple model' includes none of the exogenous variable using Eqs. (3.8)–(3.10). The AR(1) model is estimated according to (5.1), the RW model simply calculates the forecast according to (5.2). The lowest model error in each category is highlighted in italics

The t test rejects the null hypothesis that the MAE of the SEM is equal to (larger than) the MAE of the autoregressive model at $p > 0.000$. For the ME, the null hypothesis that the error of the SEM is larger can be rejected at $p > 0.1$. The null hypothesis cannot be rejected for the MAPE. Comparing the errors from the SEM with the errors of the RW model, the null hypothesis that the errors of the SEM are equal (larger) can be rejected for the MAE and ME at $p > 0.000$ but not for the MAPE. However, the MAPE is very prone to OI outliers and hence not always reliable

chosen as indicator variable because it leads to better forecasting results than the other indicators (lagged ΔCLI, lagged ΔPMI). The second best model is the SEM without additional exogenous variables followed by the AR(1) model. We also modify the AR(1) model by (i) including order backlog as additional variable and (ii) disaggregating OI into SALES and OPEX (not reported). However, neither (i) nor (ii) yields improvement over the simple AR(1) model. The RW model produces distinguishably worse forecasting results than the other models.

Next, we compare the out-of-sample forecasting errors of OI produced by the models. We only refer to the SEM version including the macroeconomic variables in the following because it produces the lowest in-sample forecast errors. As is common in the out-of-sample forecasting literature, we use rolling horizons to determine the estimation periods. First, we keep the length of the estimation period constant (in the following 'constant estimation period') (e.g. Kesavan et al. 2010) and use a seven-year estimation period to produce forecasts for the year following the estimation period. For example, the first estimation period includes the years 2003–2009 and the forecast is derived for the year 2010. The next estimation period is 2002–2008 and 2009 is the forecasting year. The seven-year period is chosen in order to obtain forecasts for the year 2000 when the dot-com



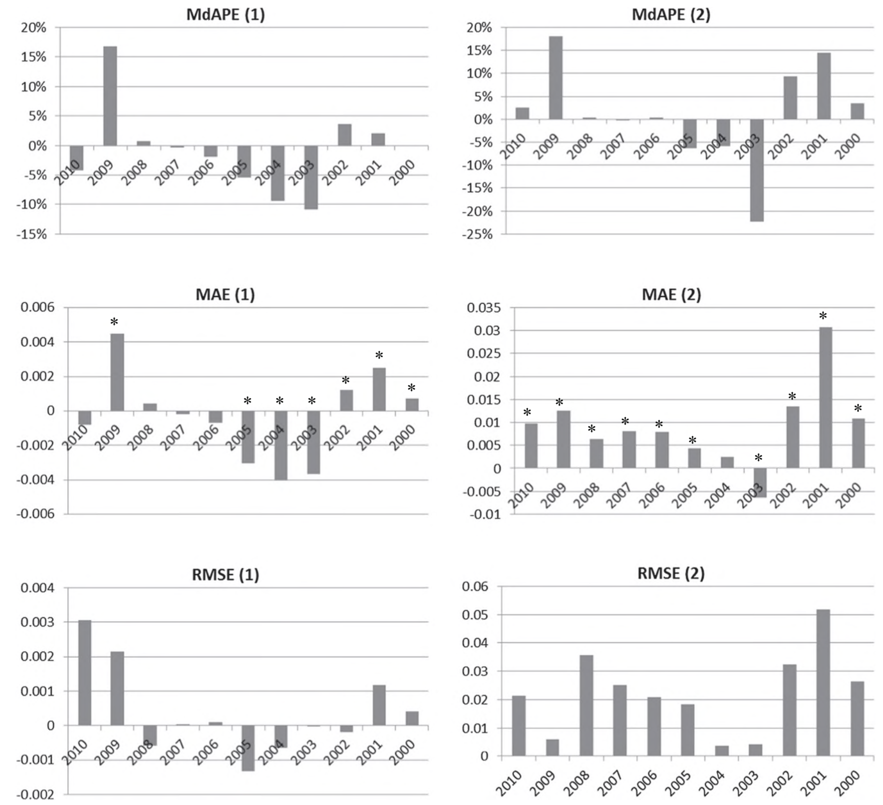
# of observations	2010	2009	2008	2007	2006	2005	2004	2003	2002	2001	2000
used for prediction	629	725	721	764	801	813	862	857	920	971	1064
used for estimation	5372	5540	5800	6021	6380	6786	7284	7717	8093	8425	8620

Fig. 2 Differences of forecasting errors of OI between AR(1) model and SEM (1) and RW model and SEM (2) based on constant estimation periods

bubble burst. The use of lagged values of the endogenous variables in the regressions requires lagged data for two periods because all values are deflated by lagged total assets. Hence, the data period reduces to the years 1993–2010 and a maximum of 7 years (1993–1999) can be used in the out-of-sample forecasting for the year 2000.

Second, we use all available data and increase the length of the period with every estimation, that is, no observations are dropped when the estimation period moves one period forward (in the following ‘varying estimation period’) (see e.g. Wu and Hu 2009). For example, the first estimation period includes the years 1993–2009 and the forecast is derived for the year 2010. The next estimation period includes 1993–2008 to derive the forecast for the year 2009. Both procedures generate forecasts for the same number of observations. The difference is the number of observations used for model estimation. For each procedure, eleven forecasts are generated for the years 2000–2010 that are compared to actual values in order to assess forecasting accuracy and bias.

Figure 2 presents the results for the constant estimation period. The differences between the forecasting errors (MdAPE, MAE and RMSE) of the benchmark models, the AR(1) and



# of observations	2010	2009	2008	2007	2006	2005	2004	2003	2002	2001	2000
used for prediction	629	725	721	764	801	813	862	857	920	971	1064
used for estimation	17071	16342	15637	14865	14083	13259	12426	11559	10644	9684	8620

Fig. 3 Differences of forecasting errors of OI between AR(1) model and SEM (1) and RW model and SEM (2) based on varying estimation periods

RW model respectively, and the SEM are displayed. A positive difference indicates a superior forecasting performance of the SEM and a negative difference a superior forecasting performance of the respective benchmark model. The results point out that the superiority of a model depends on the period for which the forecast is made.

Comparing the AR(1) model with the SEM shows that the SEM produces lower forecast errors for the financial crisis year 2009. This holds true for all error measures. For the year 2000, the beginning of the dotcom crisis, the SEM also outperforms the AR(1) model but the difference is smaller. In the other years of the forecasting period, the AR(1) model outperforms the SEM in about two-thirds of the cases.

Comparing the RW model with the SEM shows that the SEM is superior in most of the years. The RMSE, for example, is smaller for the SEM in every forecasting year. In comparison to the RW model, the error difference is greatest in the year 2009 and also in the year 2001. Compared to the SEM and the AR(1) model, the RW model only produces lower errors in terms of MAE and MdAPE in the year 2003.

Figure 3 presents the results for varying estimation periods. The results are similar to those based on constant estimation periods. The SEM produces lower forecast errors than a RW forecast in most of the years and lower errors than the AR(1) model in the years 2001 and 2009 for all error measures. For MdAPE and MAE, this also holds true for the years 2002 and 2008 and for RMSE for the year 2010. MAE and RMSE are also lower for the SEM in the year 2000. Irrespective of the error measure, for the period between these two crises, the AR(1) model's forecasting ability is better for operating income. Again, the RW model only produces lower errors (MAE and MdAPE) in the year 2003. The results indicate that in volatile times, i.e. years of or after economic disturbances, it may be useful to include additional information in the forecasting model in order to increase forecast accuracy.

5.2.2 Analysts' forecasts

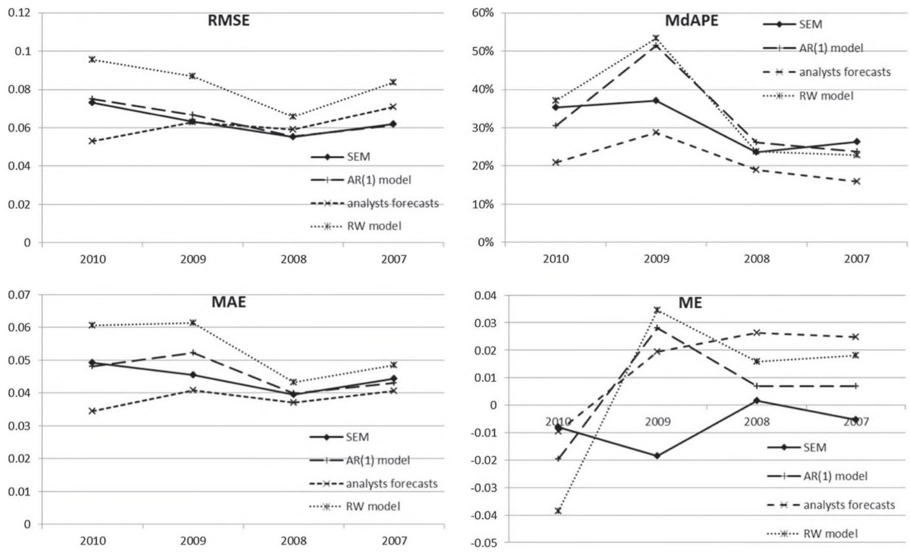
We compare the forecast errors produced by the SEM with those of analysts. We obtain analysts' forecasts of EBIT[5] and SALES from the Institutional Brokers Estimates System (I/B/E/S) in Thompson Reuters EIKON for 2007–2010 and 2005–2010 respectively. We use the mean consensus forecasts from the summary history database. The original forecasts are deflated by lagged total assets for comparability with the SEM forecasts. Consensus analysts' forecasts are available monthly until the end of the forecast period. As a firm announces its financial statements several days after the fiscal year end (FYE), we use the first forecast that is produced after the earnings announcement date (EAD) to ensure that the data set available is the same for the SEM and the analysts.

Analysts' forecasts are not available for all firms in the sample. This reduces the prediction sample in order to enable a fair comparison. Analogously, the estimation sample is reduced to firms that have analysts' forecasts in the forecast period. To be included in the sample, a firm needs to be followed by analysts in at least one of the years 2007–2010. For example, the prediction sample reduces to 353 observations for the forecast year 2010 and the estimation sample (2000–2009) to 3532 observations. As EBIT forecasts are only available from 2007, the estimation sample does not need to be restricted to seven years (in order to generate forecasts for the years starting from 2000). Hence, the estimation period is augmented to ten years for the following analyses.

In contrast to the out-of-sample forecasting results, the forecast errors are smaller when only firms that are covered by analysts are included. The reason might be that on average these firms are larger, measured by the mean of total assets and sales compared to the full sample. Hence, these firms exhibit more stable earnings time-series that are easier to predict.

All exogenous variables are included in the estimations, but not all are significant in every estimation period, like lagged fedrate. The results (see Fig. 4) show that for MdAPE and MAE, forecast errors of analysts are the lowest in all years. For RMSE, this superiority is only given in 2010. The RW model performs worst for almost all error measures and years, except for MdAPE in 2008 and 2007. In the main financial crisis year 2009, the SEM beats both the AR(1) and RW model. Further, the forecast errors of analysts and the SEM are very similar in this year. In 2007 and 2008, the overall error level is lower and in line with the results of the benchmark model comparisons. In years following a stable economic development, additional variables do not improve forecasting performance and the simpler beat the richer models. Interestingly, for 2010, all error measures are smaller for the SEM

[5] EBIT is earnings before interests and taxes and equals operating income (OI) as it is defined in this paper.



# of observations	2010	2009	2008	2007
used for prediction	353	333	302	293
used for estimation	3,532	3,378	3,248	3,104

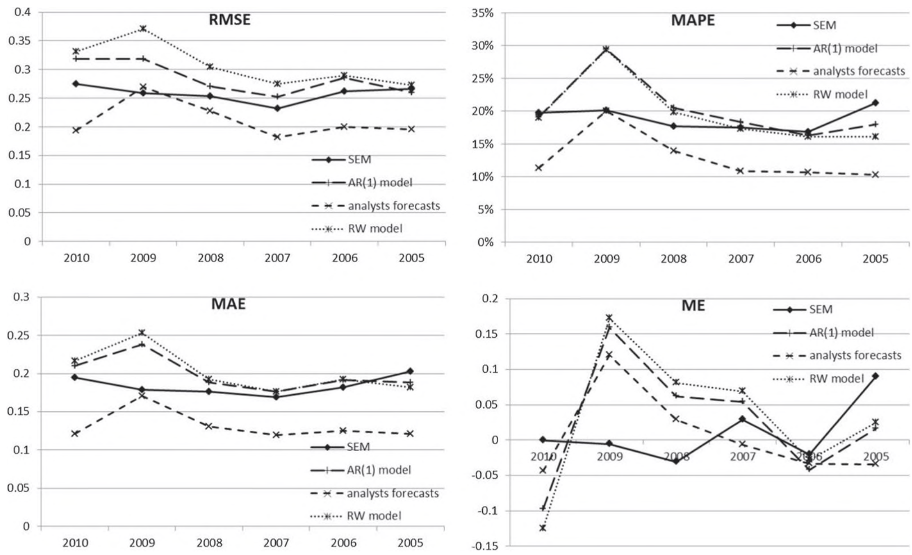
Fig. 4 Forecast errors of OI forecasts produced by SEM, AR(1) model, RW model and analysts

without additional variables than for the SEM including these variables, indicating that more information does not improve the forecasting results. For 2007 through 2009, the opposite is true. The MEs of the SEM are different to the MEs of the other forecasts. In the years 2007–2009, the SEM always underestimates and the other models always overestimate actual OI. However, the ME of the SEM is always closest to zero. In 2007 and 2008, analysts produced the most optimistic forecasts. In 2010, all MEs are negative, but the SEM still produces the smallest ME.

Next, we compare SALES forecasts (see Fig. 5). Analysts' SALES forecasts are available during 2005–2010. A firm needs to be followed by analysts in at least one of the years to be included in the sample. Like for OI, analysts' forecasts are the most accurate among the different models based on MAPE[6] and MAE in all years. In 2009 the SEM produces the lowest RMSE and also MAPEs and MAEs that are quite close or almost identical to that of the analysts. The AR(1) and RW models perform significantly worse in 2009. For the other years, the errors for the benchmark models and the SEM converge. In 2005, the SEM produces the highest errors (apart from RMSE). Similar to the OI results, the ME for the SEM is closest to zero for 2009 and 2010.

Interestingly, the relative performance measures indicate that including additional variables in the SEM improves the results only in 2009, while in all other years the contrary holds true. We conclude that in volatile times, external determinants provide additional information useful for forecasting. In times of stable development, forecasts of simple models are not significantly worse (or even better) than more sophisticated models and, hence, simple models are sufficient. In years of economic disturbances, however, it is

[6] As outliers are a not a major concern for the SALES forecast, MAPE is reported. MdAPE yields similar results.



# of observations	2010	2009	2008	2007	2006	2005
used for prediction	492	536	506	500	502	481
used for estimation	5,252	5,097	4,955	4,778	4,578	4,337

Fig. 5 Forecast errors of SALES forecasts produced by SEM, AR(1) model, RW model and analysts

worth the effort to build more complex models und use additional information for forecasting next period's sales. Under such conditions, the errors of the SEM are very close or even better than those of analysts. In line with Bryan and Tiras (2007), our findings imply that 'other information' is especially useful to improve forecasting accuracy in poor information environments.

5.3 Market value regression and SEM-based 'other information'

In this section, we integrate the SEM forecast as a proxy for 'other information' in a market value regression based on the Ohlson (1995) model. Ohlson (1995) finds that market value can be expressed as a linear function of current book value, current residual income and 'other information':

$$V_{it}^{E} = B_{it} + \alpha_1 RI_{it} + \alpha_2 v_{it} \tag{5.3}$$

with

$$\alpha_1 = \frac{\omega}{(1+r) - \omega}$$

$$\alpha_2 = \frac{(1+r)}{(1+r-\omega)(1+r-\gamma)}$$

Ohlson (2001) shows that the 'other information' term v_{it} can be expressed as "next period's expected residual income adjusted for ωx_t^a" Ohlson (2001, p. 113), that is, the forecast for next period's residual income (RI^F_{t+1}) less the information that is already included in current earnings, i.e. the prediction of next-period earnings based on an AR(1) process (RI^{AR}_{t+1}) (Ohlson 2001, Eq. 6):

$$v_{it} = \widehat{RI^F_{it+1}} - \widehat{RI^{AR}_{it+1}} \tag{5.4}$$

Ohlson (2001) finds that this expression for 'other information' can be used directly in the Ohlson (1995) model (5.3), yielding an expression of market value as a linear function of book value, current residual income and expected residual income. By using forecasted earnings, the other information term becomes independent from the information contained in current earnings, satisfying the orthogonal characteristic of the other unspecified information in the Ohlson (1995) model (Ohlson 2001). In our analysis, we use two different proxies for expected residual income: analysts' forecasts and SEM forecasts. Ohlson (2001) suggests using analysts' forecast as a proxy for next period's expected residual income because they are readily observable. As an alternative, we use the forecast generated by the SEM and calculate 'other information' as the difference between the SEM forecast and the forecast produced by the autoregressive model. Depending on whether the SEM forecast or the analysts' forecast is included, $\widehat{RI^F_{it+1}}$ is defined as $\widehat{RI^{SEM}_{it+1}}$ or $\widehat{RI^{AF}_{it+1}}$:

$$v_{1it} = \widehat{RI^{SEM}_{it+1}} - \widehat{RI^{AR}_{it+1}}$$

$$v_{2it} = \widehat{RI^{AF}_{it+1}} - \widehat{RI^{AR}_{it+1}}$$

where $\widehat{RI^{AF}_{it+1}}$ = forecast of residual income by analysts, $\widehat{RI^{SEM}_{it+1}}$ = forecast of residual income by the SEM, $\widehat{RI^{AR}_{it+1}}$ = forecast of residual income based on the autoregressive model, with $\widehat{RI^{AR}_{it+1}} = \alpha_0 + \alpha_1 OI_{it}$.

Residual income is calculated according to Eq. (3.2). For the tax rate, we use the maximum federal corporate tax rate (Biddle, Bowen, and Wallace 1997; Begley and Feltham 2002). The tax rate is set to zero if the firm reports a loss in the respective year. For the cost of capital the following formula is applied:

$$r = r_E \times \frac{V^E}{V^{NOA}} + r_{PS} \times \frac{V^{PS}}{V^{NOA}} + (1 - tax) \times r_D \times \frac{V^D}{V^{NOA}}$$

where r_{PS} = cost of preferred stock, V^{PS} = market value of preferred stock.

The cost of equity r_E is estimated according to the CAPM. We use the 10-year Treasury bond yield of the respective year as the risk free rate (Koller, Goedhart, and Wessels 2010) and add a constant equity risk premium of 6% (Penman and Sougiannis 1998; Francis, Olsson, and Oswald 2000; Nissim and Penman 2001). The cost of preferred stock r_{PS} is calculated by dividing preferred dividends (#19) by preferred stock (#130) (Francis, Olsson, and Oswald 2000). The cost of debt r_D can be calculated as the ratio of interests (#15) and interest bearing debt (#9 + #34). If the firm has a noncontrolling (minority) interest in a subsidiary, the income attributable to the noncontrolling interest (#49) has to be added to the enumerator and the noncontrolling interest (#38) to the denominator.[7] Like Francis, Olsson, and Oswald (2000), we set the upper bound on the cost of debt and the cost of

[7] ASC 810-10-45 requires that non-controlling interests are reported within equity. For calculation of the WACC it does not matter where these interests are considered as long as the weights correspond to the respective cost rates.

preferred stock equal to the cost of equity and the lower bound equal to the risk-free rate. For estimating the target capital structure we use current market capitalization, i.e. common stock price at FYE (#199) multiplied by the number of shares outstanding (#25), as proxy for market value of common equity V^E and assume that this capital structure is going to be constant in the future. Market values of preferred stock V^{PS} and interest bearing debt (including non-controlling interests) V^D are proxied by their book values.[8] Hence, the difference between NOA and V^{NOA} is only by the inclusion of market value of common equity instead of book value. If no market values are available, book values are used instead (provided that they are positive).

Next, we compare the explanatory power of different specifications of the 'other information' term with a model that includes only lagged book value and earnings (M1). As a simple approximation of 'other information' we follow Myers (1999) and use order backlog as a direct proxy (M2) because it is contained in the firms' reports but is not part of the financial statements. Order backlog is considered to be the most important source of additional information among other firm-specific external determinants (see Sect. 3.3). Myers (1999) uses order backlog as a proxy for the 'other information' in the LID of the Feltham and Ohlson (1995) model and finds that it can have a positive or negative effect on next-period residual income, but the overall effect is small. Possible reasons may be that order backlog is a better indicator of next-period sales instead of residual income because order backlog directly translates into sales in the following periods whereas residual income is also affected by other variables (like operating costs and costs for the use of capital). Furthermore, order backlog alone may not be sufficient to capture the entire concept of 'other information'.

In M3 and M4, we specify the 'other information' term based on forecasted earnings as described in (5.2). In M3, we use the SEM forecast to determine 'other information'. For comparison, we also include analysts' forecasts as proxies in M4 and expect that this specification exhibits the greatest power for explaining market values. The reason is that analysts' forecasts are superior to other forecasts in capturing forward-looking information. Also, stock recommendations are based on analyst forecasts, which in turn influence the formation of market values.

The resulting regression models are:

$$\text{M1}: \; (V^E_{it+3m}) = \delta_0 + \delta_1 B_{it} + \delta_2 RI_{it} + \varepsilon_{it} \tag{5.5}$$

$$\text{M2}: \; (V^E_{it+3m}) = \delta_0 + \delta_1 B_{it} + \delta_2 RI_{it} + \delta_3 OB_{it} + \varepsilon_{it} \tag{5.6}$$

$$\text{M3}: \; (V^E_{it+3m}) = \delta_0 + \delta_1 B_{it} + \delta_2 RI_{it} + \delta_3 v_{1it} + \varepsilon_{it} \tag{5.7}$$

$$\text{M4}: \; (V^E_{it+3m}) = \delta_0 + \delta_1 B_{it} + \delta_2 RI_{it} + \delta_3 v_{2it} + \varepsilon_{it} \tag{5.8}$$

where $v_{1it} = \widehat{RI^{SEM}_{it+1}} - \widehat{RI^{AR}_{it+1}}$, $v_{2it} = \widehat{RI^{AF}_{it+1}} - \widehat{RI^{AR}_{it+1}}$.

We expect that the 'other information' variable is significant and that the explanatory power increases from M1 to M4. In particular, we expect an increase from M2 to M3 because the SEM forecast contains additional information beyond order backlog and it is derived from a model that accounts for the interdependencies between the value drivers of

[8] In contrast to debt, book values of preferred stock and non-controlling interests typically do not equal their market values. Book values are used as proxies due to data constraints. Hence, $V^D = \#9 + \#34 + \#38$ and $V^{PS} = \#130$.

Table 8 Market value regressions (OLS)

M1: $V^E_{it+3m} = \delta_0 + \delta_1 B_{it} + \delta_2 RI_{it} + \varepsilon_{it}$
M3: $V^E_{it+3m} = \delta_0 + \delta_1 B_{it} + \delta_2 RI_{it} + \delta_3 v_{1it} + \varepsilon_{it}$
M2: $V^E_{it+3m} = \delta_0 + \delta_1 B_{it} + \delta_2 RI_{it} + \delta_3 OB_{it} + \varepsilon_{it}$
M4: $V^E_{it+3m} = \delta_0 + \delta_1 B_{it} + \delta_2 RI_{it} + \delta_3 v_{2it} + \varepsilon_{it}$

2006–2009

V^E_{it+3m}	M1 Coef./[std.errors]	M2 Coef./[std.errors]	M3 Coef./[std.errors]	M4 Coef./[std.errors]
B_{it}	1.7785*** [0.16]	1.7767*** [0.16]	1.8730*** [0.16]	1.8970*** [0.15]
RI_{it}	5.2320*** [0.75]	5.1563*** [0.77]	5.1887*** [0.76]	5.5520*** [0.81]
OB_{it}		0.0388 [0.03]		
v_{1it}			7.2703*** [1.41]	
v_{2it}				4.8769*** [0.90]
Constant	0.3603*** [0.09]	0.3357*** [0.10]	0.2820*** [0.09]	0.2047** [0. 90]
No. of obs.	1193	1193	1193	1193
R-sq	0.3423	0.3433	0.3736	0.3945
Vuong's (1989) Z-statistic	M1 versus M2 0.70	M2 versus M3 2.47**	M3 versus M4 1.35	

OLS regressions with heteroskedasticity-robust standard errors. B_t is book value of common equity (annual data item #60) in t deflated by lagged total assets. V^E_{t+3m} is deflated market value of common equity in t where the market value of common equity is calculated as common stock price three months after FYE (quarterly data item #14) multiplied by the number of shares outstanding (quarterly data item #61). RI_t is operating income after depreciation (annual data item #178) deflated by lagged total assets multiplied with $(1 - tax)$, and minus cost of capital (WACC) applied on lagged deflated NOA; OB_t is order backlog (annual data item #98) deflated by lagged total assets; v_{1t} is measured as $\frac{\widehat{RI^{SEM}_{t+1}}}{TA_t} - \frac{\widehat{RI^{AR}_{t+1}}}{TA_t}$, v_{2t} is measured as $\frac{\widehat{RI^{AF}_{t+1}}}{TA_t} - \frac{\widehat{RI^{AR}_{t+1}}}{TA_t}$. The lower number of observations (1193) compared to the analysts' EBIT forecast sample used in the second part of Sect. 5.2 (353 + 333 + 302 + 293 = 1281) is due to the generation of the leading other information variables and the lower number of RI observations

* $p < 0.10$; ** $p < 0.05$; *** $p < 0.01$

residual income. As outlined above, the explanatory power should also increase from M3 to M4. The estimations are based on the analysts' EBIT forecast sample described in the second part of Sect. 5.2. The forecasts of one year are included in a market value regression of the previous year, e.g. the forecast for 2010 is used as input in the regression of 2009 market value. The availability of analysts' EBIT forecasts from 2007 to 2010 allows us to conduct market value regressions for the period 2006–2009.

The results displayed in Table 8 show that the coefficients of v_{1it} and v_{2it} are highly significant, emphasizing the incremental explanatory power of the 'other information' in the regression. Order backlog is positive but insignificant, indicating that the variable alone is not sufficient as a proxy for other value relevant information, consistent with Myers (1999). As expected, the R^2 of the regressions exhibit the following order: M1 < M2 < M3 < M4. We use the test of Vuong (1989) to compare the explanatory power (R^2) of two differently specified regressions from the same sample.[9] We find that the

[9] For a detailed discussion of the Vuong (1989) test, see Dechow (1994), Appendix 2. The Vuong test cannot be applied for comparing M1 and M2 as this test is only valid for non-nested models.

difference between the R^2 of M2 and M3 is highly significant at $p < 0.05$. These results show that the SEM forecasts are able to explain value beyond the information contained in current book value, current operating income and ‘other information’ simply proxied by order backlog. The highest R^2 is produced by the regression containing analysts’ forecasts. This result is due to the information advantage of analysts, who use a more comprehensive information set than ours. However, the difference between the R^2 of M3 and M4 is not significant at conventional levels, indicating that the ‘other information’ derived from analyst forecasts is not significantly more useful for explaining market values in the regression compared to the ‘other information’ derived from the SEM. This implies that a systematic forecasting procedure like the SEM is useful to produce forecasts that are able to serve as proxies for ‘other information’ in cases when analysts’ forecasts are not available or when single items of information are of particular interest that can be integrated in the SEM forecast. While order backlog is not significant when used as a direct proxy, it is an important input in the SEM forecast, which in turn is highly significant.

The question remains whether and to what extent the SEM forecast is able to capture part of the information advantage of analysts. The next section addresses this question.

5.4 Explanatory power of SEM forecasts for analysts’ forecasts and analysts’ information advantage

If the forecasting framework is relevant to analysts, there should be an association between the forecast based on the framework and the analysts’ forecast. In this case, one could conclude that analysts use the framework. Regressing the analysts’ forecasts on the SEM and the benchmark model forecasts (AR(1) and RW) yields the results displayed in Table 9. They show that the SEM forecast is able to explain analysts’ forecasts. The R^2 of the regression on the SEM forecast is significantly greater than of the AR(1) forecast. However, the association between the RW forecast and the analysts’ forecast is the strongest. These results indicate that analysts’ OI forecasts anchor substantially on the value of OI in the previous period, which is in line with the results of Lambert et al. (2012).

The results in Sect. 5.2 imply that analysts’ forecasts are superior to time-series model’s forecasts and the SEM forecasts. This superiority may be due to information advantages of analysts. Hence, we analyze whether the SEM forecast is able to explain the information advantage of analysts. For addressing this question, we investigate whether the SEM forecasts possess explanatory power for the analysts’ information advantage (AA) as proposed by Kross et al. (1990):

$$AA_{it+1} = \delta_0 + \delta_1 \widehat{OI}_{it+1}^{SEM} + \varepsilon_{it} \tag{5.9}$$

We measure AA as the difference of the absolute value of the forecast error of the benchmark AR(1) model, and the absolute value of the analysts’ forecast error based on Kross et al. (1990):

$$AA_{it} = AE_{it}^{AR} - AE_{it}^{AF} \tag{5.10}$$

where AA_{it} = analyst earnings forecast advantage, AE_{it}^{AR} = absolute value of the forecast error in AR(1) model, AE_{it}^{AF} = absolute value of the forecast error of analysts’ forecast.

The AA is positive on average, indicating that the error of the analyst forecast is smaller than that of the AR(1) model.

Table 9 Analysts' forecasts regressions on models' forecasts

M5: $\widehat{OI_{it+1}^{AF}} = \delta_0 + \delta_1 \widehat{OI_{it+1}^{AR}} + \varepsilon_{it}$
M6: $\widehat{OI_{it+1}^{AF}} = \delta_0 + \delta_1 \widehat{OI_{it+1}^{SEM}} + \varepsilon_{it}$
M7: $\widehat{OI_{it+1}^{AF}} = \delta_0 + \delta_1 \widehat{OI_{it+1}^{RW}} + \varepsilon_{it}$

2007–2010

$\widehat{OI_{it+1}^{AF}}$	M5 Coef./[std.errors]	M6 Coef./[std.errors]	M7 Coef./[std.errors]
$\widehat{OI_{it+1}^{AR}}$	0.8168*** [0.02]		
$\widehat{OI_{it+1}^{SEM}}$		0.7738*** [0.02]	
$\widehat{OI_{it+1}^{RW}}$			0.6823*** [0.01]
Constant	0.0266*** [0.00]	0.0395*** [0.00]	0.0380*** [0.00]
No. of obs.	1281	1281	1281
R-sq	0.5792	0.5896	0.6224
Vuong's (1989) Z-statistic	M5 versus M6 1.67*	M6 versus M7 1.62	

Reported R-squares are unadjusted. $\widehat{OI_{t+1}^{AF}}$ is the analysts' operating income forecast; $\widehat{OI_{t+1}^{AR}}$ is the autoregressive operating income forecast; $\widehat{OI_{t+1}^{SEM}}$ is the SEM operating income forecast; $\widehat{OI_{t+1}^{RW}}$ is the RW operating income forecast (= last period's value of operating income). All values are deflated by lagged total assets

* $p < 0.10$; ** $p < 0.05$; *** $p < 0.01$

The regression results in Table 10 show a highly significant ($p < 0.01$) coefficient of the OI forecast produced by the SEM implying that the SEM forecast is able to explain parts of the analyst information advantage. The overall model is highly significant, indicating a high relevance of the SEM. The relatively low R^2 is in line with the results of Kross et al. (1990) and indicates that there are other factors beyond those captured in the SEM forecast relevant for explaining the analyst information advantage.

We further analyze whether the SEM forecast is able to capture information contained in analysts' forecasts that is relevant for explaining market values beyond information contained in financial statements. We analyze how much of the information advantage relevant for explaining market values is captured by the 'other information' derived from the SEM forecast. Given that analysts' forecasts are richer than the SEM forecasts because of analyst information advantage, we use the 'other information' based on analysts' forecasts (v_{2it}) as defined in Sect. 5.3 as the starting point. This 'other information' term is split into two parts, i.e. one part that can be explained by the SEM (v_{1it}) and one part that cannot be explained by the SEM (v_{3it}): $v_{2it} = v_{1it} + v_{3it}$.

$$\begin{aligned} v_{2it} &= \widehat{RI_{it+1}^{AF}} - \widehat{RI_{it+1}^{AR}} \\ &= \left(\widehat{RI_{it+1}^{AF}} - \widehat{RI_{it+1}^{SEM}}\right) + \left(\widehat{RI_{it+1}^{SEM}} - \widehat{RI_{it+1}^{AR}}\right) \\ &= v_{3it} + v_{1it} \end{aligned} \tag{5.11}$$

The first part of Eq. (5.11) is the difference between the analysts' forecast and the SEM forecast, i.e. the part that cannot be explained by the SEM (v_{3it}). The second part is the

Table 10 Analyst advantage regression (OLS)

$AA_{it+1} = \delta_0 + \delta_1 \widehat{OI_{it+1}^{SEM}} + \varepsilon_{it}$	
2007–2010	
AA_{it+1}	Coef. [std.errors]
$\widehat{OI_{it+1}^{SEM}}$	0.1092*** [0.02]
Constant	−0.0011 [0.00]
No. of obs.	1281
R-sq	0.04
F(1,1279)	21.37***

OLS regressions with heteroskedasticity-robust standard errors; reported R-squares are unadjusted. AA_{t+1} is the analyst information advantage measured as the difference of the absolute value of the AR(1) forecast error and the absolute value of the analysts' forecast error. $\widehat{OI_{t+1}^{SEM}}$ is the SEM operating income forecast

* $p < 0.10$; ** $p < 0.05$; *** $p < 0.01$

difference between the SEM forecast and the autoregressive forecast, i.e. the part that can be explained by the SEM, i.e. v_{1it}. We regress the 'other information' based on analysts' forecasts v_{2it} on the 'other information' based on SEM forecasts v_{1it} to obtain the part that cannot be explained by the SEM v_{3it} as the residual from the regression. These residuals are then integrated as v_{3it} in the market value regression jointly with v_{1it}. This approach is a simple extension of the basic Ohlson (2001) framework and enables us to analyze the different parts of the 'other information' in order to draw inferences about the ability of the SEM forecast to capture part of the explanatory power of analysts forecasts. We expect that the coefficient and t-value of v_{1it} in relation to v_{3it} increase with the ability of the SEM to explain market value.

The results displayed in Table 11 confirm these conjectures. The coefficient and significance level of v_{1it} (part of 'other information' that can be explained by the SEM) (coefficient 8.76, t-value 9.07) is higher in relation to v_{3it} (coefficient 4.90, t-value 7.28) (part of 'other information' that cannot be explained by the SEM) and a Wald Chi Squared Test indicates that their difference is significant (F-value 13.76). This result shows that the SEM forecast is able to capture a large fraction of the difference between the analysts' and the AR(1) forecast, that is, 'other information'. However, due to the information advantage of analysts, a portion relevant for explaining market value remains unexplained by the SEM forecast.

6 Sensitivity analyses

Finally, we conduct additional analyses to analyze the sensitivity of our results to model specifications. We set up and test further variants of the SEM assuming either (i) a constant asset turnover or (ii) a constant profit margin. The system of equations is then reduced to two stochastic equations and the missing variable is calculated (i) by applying the expected SALES growth rate on current NOA in order to determine expected NOA or (ii) by applying the current profit margin on expected SALES in order to determine OI directly. However, these simpler variants of the SEM do not exhibit a superior forecasting performance than the original SEM. Especially the model variant assuming a constant profit margin generates forecasts that are only as accurate as RW forecasts.

Table 11 Market value regression (OLS) with v_{1t} and v_{3t}

M4*a*: $V^E_{it+3m} = \delta_0 + \delta_1 B_{it} + \delta_2 RI_{it} + \delta_3 v_{1it} + \delta_4 v_{3it} + \varepsilon_{it}$	
2006–2009	
V^E_{it+3m}	M4a Coef. [std.errors]
B_{it}	1.7585*** [0.15]
RI_{it}	5.3915*** [0.78]
v_{1it}	8.7560*** [0.97]
v_{3it}	4.8995*** [0.67]
Constant	0.4808 [0.09]
No. of obs.	1193
R-sq	0.45

OLS regressions with heteroskedasticity-robust standard errors; reported R-squares are unadjusted. B_t is book value of common equity (annual data item #60) in t deflated by lagged total assets; V^E_{t+3m} is deflated market value of common equity in t whereas the market value of common equity is calculated as common stock price at FYE (annual data item #199) multiplied by the number of shares outstanding (annual data item #25); d_t is dividends on common equity (annual data item #21) paid in t deflated by lagged total assets; E_t is net income (annual data item #172) deflated by lagged total assets; v_{1t} is measured as $\widehat{\frac{RI^{SEM}_{t+1}}{TA_t}} - \widehat{\frac{RI^{AR}_{t+1}}{TA_t}}$, v_{3t} is defined as $\widehat{\frac{RI^{AF}_{t+1}}{TA_t}} - \widehat{\frac{RI^{SEM}_{t+1}}{TA_t}}$ and is measured as the residual from a regression of $v_{2t} = \widehat{\frac{RI^{AF}_{t+1}}{TA_t}} - \widehat{\frac{RI^{AR}_{t+1}}{TA_t}}$ on v_{1t}

The fewer number of observations (1193) compared to the analysts' EBIT forecast sample used in the Sect. 5.2 (353 + 333 + 302 + 293 = 1281) is due to the generation of the leading other information variables and the lower number of RI observations

* $p < 0.10$; ** $p < 0.05$; *** $p < 0.01$

Concerns that the results might be driven by the use of actuals of the macroeconomic variables (GDP and PPI growth) can be eliminated by using forecasts. Real GDP levels and growth rate forecasts are obtained from the Federal Reserve Bank of Philadelphia's Survey of Professional Forecasters.[10] All results remain stable when forecasts of the annual growth rate of real GDP are included in the SEM.

Alternative deflators such as lagged common equity do not change the main inferences drawn from the out-of-sample analysis. For the in-sample analysis, the error difference between the SEM and the AR(1) model is not as clear as for lagged total assets. MdAPE is lower for the SEM, ME, MAE, and RMSE is lower for the AR(1) model.

We repeat the analysis from Sect. 5.3 using the returns specification of Easton (1999), regressing returns on earnings, change in earnings and change in 'other information'. M1 excludes 'other information', M2 includes change in OB, M3 includes change in 'other information' based on the SEM forecast and M4 includes change in 'other information' based on analysts' forecasts. The returns regression yields significant coefficient estimates of the 'other information' variables. The R^2 of M3 is significantly larger at $p < 0.01$ than the R^2 of M2 and also significantly larger at $p < 0.1$ than the R^2 of M4. According to these results, the SEM forecast is able to explain returns even better than the analysts' forecast.

[10] We use real GDP annual growth rate forecasts released at the beginning of the respective year. PPI forecasts are not available.

7 Conclusions

This paper presents an alternative approach for incorporating 'other information' in the Ohlson (1995) model. Instead of using analysts' forecasts (Ohlson 2001), we derive forecasts of operating income from publicly available information based on the forecasting framework described in standard textbooks. Due to the interdependencies between the value drivers of operating income, i.e. sales, operating expenditures and net operating assets, a simultaneous equations model (SEM) is developed which allows us to produce forecasts as inputs to accounting based valuation models.

We find that the SEM forecast is able to explain market value beyond current book value, current earnings and single information proxies for 'other information' like order backlog. The model provides an approach to link different aspects of other information contained in a firm's non-financial communication with an accounting based valuation model. The results emphasize the usefulness of a comprehensive forecast derived from a model that accounts for the interrelations between the different value drivers for explaining market values.

Second, we find that the SEM forecasts are able to explain the analysts' forecasts and that they capture a major portion of the information advantage of analysts. The result implies that while analysts seem to use the 'forecasting framework' described in textbooks, there remains an unexplained portion of their forecasts. A possible reason is that analysts may have industry/firm-specific expertise or access to private information.

When comparing out-of-sample forecasting results of the SEM with simple benchmark models (AR(1) model, RW model), we find evidence that in years around economic changes and instability, the SEM produces more accurate forecasts of operating income. In years of stable and continuous development, the SEM performs as well as the autoregressive model and better than a random walk. The SEM forecasts are not able to beat analysts' forecasts, but in volatile times, the errors of the SEM and analysts converge. These results add to the findings of Bryan and Tiras (2007) by showing that including other value relevant information in the forecasting model in poor information environments, i.e. volatile times, is useful in order to increase forecast accuracy.

Our model is very general and applied equally to all industries. Industry characteristics may cause different model parameters and hence, industry specific models may be able to produce more precise forecasting results. Future research could attempt to build models that incorporate industry-specific determinants like store growth for retailers (Kesavan et al. 2010), load factors for airlines or capacity utilization for manufacturers (Richardson et al. 2010). In addition, our results could be complemented by experimental studies on how analysts generate their forecasts and whether they use forecasting frameworks or not.

Acknowledgement We gratefully acknowledge helpful comments by Anne Wyatt, Stephen Salter, Thorsten Sellhorn, Tami Dinh, Andreas Weiler, Marco Wilkens, Maria Assel and workshop participants at the University of Augsburg, participants at the Eden doctoral and young scholar seminar on visualizing, measuring and managing intangibles 2010 in Catania, the EAA Annual Congress 2013 in Paris and 2014 in Tallinn, the AAA Annual Meeting in Anaheim and the AAA IAS Midyear Meeting in San Antonio. We gratefully acknowledge financial support by the Wissenschaftliche Gesellschaft für Prüfung und Controlling, Augsburg e.V. and the Research Center for Global Business Management, University of Augsburg.

Appendix

The cost of debt capital or borrowing cost determines next-period's net financial expenses (NFE), i.e. interest income less interest expense.

$$NFE_t = r_D \times (1 - tax) \times NFO_{t-1}$$

where NFO = net financial obligations, NFE = net financial expenses

Earnings are defined as forecasted OI (after tax) less NFE.

$$E_t = OI_t \times (1 - tax) - NFE_t$$

The payout ratio is applied to the earnings forecast in order to determine net dividends.[11] Net financial obligations (NFO), i.e. financial assets less financial obligations or liabilities, can be calculated as[12]

$$NFO_t = NFO_{t-1} + NFE_t - FCF_t + d_t$$

$$\text{with}\, FCF_t = OCF - ICF = OI_t \times (1 - tax) - NOA_t + NOA_{t-1}$$

where FCF = free cash flow, OCF = operating cash flow, ICF = investing cash flow, d = net dividends.

Free cash flow (FCF) (after tax) can be determined by cash flow from operations less cash investment, but also by operating income (after tax) less the change in NOA. Thus, FCF can be directly derived from forecasts of accounting measures and one does not need to forecast cash flows.

Finally, book value of equity is the residual amount calculated as

$$B_t = NOA_t - NFO_t = B_{t-1} + E_t - d_t$$

Taken these forecasts together, complete pro forma balance sheet, income statement and cash flow statement can be set up.

References

Agostino M, Drago D, Silipo DB (2011) The value relevance of IFRS in the European banking industry. Rev Quant Financ Acc 36:437–457

Anandarajan A, Francis B, Hasan I, John K (2011) Value relevance of banks: global evidence. Rev Quant Financ Acc 36:33–55

Ashton D, Wang P (2015) Conservatism in residual income models: theory and supporting evidence. Account Bus Res 45:387–410

Ashton D, Peasnell K, Wang P (2011) Residual income valuation models and inflation. Eur Account Rev 20:459–483

Barth ME, Clement MB, Foster G, Kasznik R (1998) Brand values and capital market valuation. Rev Acc Stud 3:41–68

Barth ME, Beaver WH, Hand JRM, Landsman WR (2005) Accruals, accounting-based valuation models, and the prediction of equity values. J Account Audit Finance 20:311–345

[11] Strictly speaking, net dividends are dividend payments (including share repurchases) less capital contributions from shareholders. These items have to be considered when forecasting the payout ratio.

[12] This identity is also known as the cash conservation equation (Penman 2010, p. 236). If it is assumed that the change of cash and cash equivalents is zero, FCF can be expressed as the sum of net transactions with equity investors (d) and net transactions with debt investors (NFE − ΔNFO).

Begley J, Feltham GA (2002) The Relation between market values, earnings forecasts, and reported earnings. Contemp Account Res 19:1–48
Biddle GC, Bowen RM, Wallace JS (1997) Does EVA® beat earnings? Evidence on associations with stock returns and firm values. J Account Econ 24:301–336
Bradshaw MT (2011) Analysts' forecasts: what do we know after decades of work? Working paper, Boston College. http://ssrn.com/abstract=1880339
Bradshaw MT, Drake MS, Myers JN, Myers LA (2012) A re-examination of analysts' superiority over time-series forecasts. Rev Acc Stud 17:944–968
Brown LD (1993) Earnings forecasting research: its implications for capital markets research. Int J Forecast 9:295–320
Brown LD, Griffin PA, Hagerman RL, Zmijewski ME (1987) Security analyst superiority relative to univariate time-series models in forecasting quarterly earnings. J Account Econ 9:61–87
Brown S, Lo K, Lys T (1999) Use of R 2 in accounting research: measuring changes in value relevance over the last four decades. J Account Econ 28:83–115
Brown LD, Huang K, Pinello AS (2013) To beat or not to beat? The importance of analysts' cash flow forecasts. Rev Quant Financ Acc 41:723–752
Bryan DM, Tiras SL (2007) The influence of forecast dispersion on the incremental explanatory power of earnings, book value, and analyst forecasts on market prices. Account Rev 82:651–677
Callen JL, Morel M (2001) Linear accounting valuation when abnormal earnings are AR(2). Rev Quant Financ Acc 16:191–204
Callen JL, Segal D (2005) Empirical tests of the Feltham-Ohlson (1995) model. Rev Acc Stud 10:409–429
Capstaff J, Paudyal K, Rees W (2001) A comparative analysis of earnings forecasts in Europe. J Bus Finance Account 28:531–562
Cheng Q (2005a) What determines residual income? Account Rev 80:85–112
Cheng Q (2005b) The role of analyst' forecasts in accounting-based valuation: a critical evaluation. Rev Acc Stud 10:5–31
Choi YS, O'Hanlon JF, Pope PF (2006) Conservative accounting and linear information valuation models. Contemp Account Res 23:73–101
Clubb C (2013) Information dynamics, dividend displacement, conservatism, and earnings measurement: a development of the Ohlson (1995) valuation framework. Rev Acc Stud 18:360–385
Dechow PM (1994) Accounting earnings and cash flows as measures of firm performance: the role of accounting accruals. J Account Econ 18:3–42
Dechow PM, Hutton AP, Sloan RG (1999) An empirical assessment of the residual income valuation model. J Account Econ 26:1–34
Easton PD (1999) Security returns and the value relevance of accounting data. Account Horiz 13:399–412
Feltham GA, Ohlson JA (1995) Valuation and clean surplus accounting for operating and financial activities. Contemp Account Res 11:689–731
Feltham GA, Ohlson JA (1996) Uncertainty resolution and the theory of depreciation measurement. J Account Res 34:209–234
Francis J, Olsson P, Oswald DR (2000) Comparing the accuracy and explainability of dividend, free cash flow, and abnormal earnings equity value estimates. J Account Res 38:45–69
Fried D, Givoly D (1982) Financial analysts' forecasts of earnings: a better surrogate for market expectations. J Account Econ 4:85–107
Hanke JE, Wichern DW (2009) Business forecasting, 9th edn. Prentice Hall, Upper Saddle River
Higgins HN (2011) Forecasting stock price with the residual income model. Rev Quant Financ Acc 36:583–604
Hirschey M, Richardson VJ, Scholz S (2001) Value relevance of nonfinancial information: the case of patent data. Rev Quant Financ Acc 17:223–235
Imam S, Barker R, Clubb C (2008) The use of valuation models by UK investment analysts. Eur Account Rev 17:503–535
Kesavan S, Gaur V, Raman A (2010) Do inventory and gross margin data improve sales forecasts for U.S. public retailers? Manag Sci 56:1519–1533
Koller T, Goedhart M, Wessels D (2010) Valuation: measuring and managing the value of companies, 5th edn. Wiley, Hoboken
Kross W, Ro B, Schroeder D (1990) Earnings expectations: the analysts' information advantage. Account Rev 65:461–476
Kuo CY (2017) Is the accuracy of stock value forecasting relevant to industry factors or firm-specific factors? An empirical study of the Ohlson model. Rev Quant Financ Acc 49:195–225
Lambert D, Matolcsy Z, Wyatt A (2012) How do analysts forecast earnings? Working paper University of Queensland

Lev B, Thiagarajan SR (1993) Fundamental information analysis. J Account Res 31:190–215

Lin WT (1992) Analysis and forecasting of income statement account balances: the dynamic interdependency and ARIMA approaches. J Forecast 11:283–307

Lundholm RJ, Sloan RG (2007) Equity valuation and analysis, 2nd edn. McGraw-Hill/Irwin, New York

Myers JN (1999) Implementing residual income valuation with linear information dynamics. Account Rev 74:1–28

Nissim D, Penman SH (2001) Ratio analysis and equity valuation: from research to practice. Rev Acc Stud 6:109–154

Ohlson JA (1995) Earnings, book values, and dividends in equity valuation. Contemp Account Res 11:661–687

Ohlson JA (2001) Earnings, book values, and dividends in equity valuation: an empirical perspective. Contemp Account Res 18:107–120

Orens R, Lybaert N (2010) Determinants of sell-side financial analysts' use of non-financial information. Account Bus Res 40:39–53

Penman SH (2010) Financial statement analysis and security valuation, 4th edn. McGraw-Hill/Irwin, New York

Penman SH, Sougiannis T (1998) A comparison of dividend, cash flow, and earnings approaches to equity valuation. Contemp Account Res 15:343–383

Penman S, Yehuda N (2009) The pricing of earnings and cash flows and an affirmation of accrual accounting. Rev Acc Stud 14:453–479

Pindyck RS, Rubinfeld DL (1998) Econometric models and economic forecasts, 4th edn. McGraw-Hill/Irwin, New York

Porter ME (1998) The competitive advantage: creating and sustaining superior performance. Free Press, New York

Previts GJ, Bricker RJ, Robinson TR, Young SJ (1994) Content analysis of sell-side financial analyst company reports. Account Horiz 8:55–70

Rajgopal S, Shevlin T, Venkatachalam M (2003) Does the stock market fully appreciate the implications of leading indicators for future earnings? Evidence from order backlog. Rev Acc Stud 8:461–492

Rees W, Valentincic A (2013) Dividend irrelevance and accounting models of value. J Bus Finance Account 40:646–672

Richardson G, Tinaikar S (2004) Accounting based valuation models: what have we learned? Account Finance 44:223–255

Richardson S, Tuna I, Wysocki P (2010) Accounting anomalies and fundamental analysis: a review of recent research advances. J Account Econ 50:410–454

Rogers RK, Grant J (1997) Content analysis of information cited in reports of sell-side financial analysts. J Financ Statement Anal 3:17–30

Shen Y, Stark AW (2013) Evaluating the effectiveness of model specifications and estimation approaches for empirical accounting-based valuation models. Account Bus Res 43:660–682

Soliman MT (2008) The use of DuPont analysis by market participants. Account Rev 83:823–853

Tsay RS, Lin YM, Wang HW (2008) Residual income, value-relevant information and equity valuation: a simultaneous equations approach. Rev Quant Financ Acc 31:331–358

Tswei K, Kuo CY (2012) A study of stock price behavior in Taiwan via residual income valuation theory and structural identification. Rev Pac Basin Financ Mark Polic 15:1250016

Vuong QH (1989) Likelihood ratio tests for model selection and non-nested hypotheses. Econometrica 57:307–333

Wang P (2014) On the relevance of earnings components in valuation and forecasting. Rev Quant Financ Acc 42:399–413

White GI, Sondhi AC, Fried D (2003) The analysis and use of financial statements, 3rd edn. Wiley, New York

Wu JL, Hu YH (2009) New evidence on nominal exchange rate predictability. J Int Money Finance 28:1045–1063