ORIGINAL RESEARCH

A nonlinear dynamic approach to cash flow forecasting

Yang Pang[1] · Shimeng Shi[2] · Yukun Shi[3] · Yang Zhao[4]

Abstract
We propose a novel grey-box model to capture the nonlinearity and the dynamics of cash flow model parameters. The grey-box model retains a simple white-box model structure, while their parameters are modelled as a black-box with a Padé approximant as a functional form. The growth rate of sales and firm age are used as exogenous variables because they are considered to have explanatory power for the parameter process. Panel data estimation methods are applied to investigate whether they outperform the pooled regression, which is widely used in the extant literature. We use the U.S. dataset to evaluate the performance of various models in predicting cash flow. Two performance measures are selected to compare the out-of-sample predictive power of the models. The results suggest that the proposed grey-box model can offer superior performance, especially in multi-period-ahead predictions.

Keywords Cash flow growth · Cash flow prediction · Grey-box model · Panel data model

JEL Classification C14 · C33 · C53 · M41

✉ Yukun Shi
  Yukun.shi@glasgow.ac.uk

  Yang Pang
  pangyang@china-inv.cn

  Shimeng Shi
  shimeng.shi@xjtlu.edu.cn

  Yang Zhao
  yangzhao@cufe.edu.cn

1 Department of Public Equity, China Investment Corporation, 12/F New Poly Plaza, No.1 Chaoyangmen Beidajie, Dongcheng District, Beijing 100010, People's Republic of China

2 International Business School Suzhou, Xi'an Jiaotong-Liverpool University, Business Building (BS), South Campus, No.8 Chongwen Road, Suzhou 215123, People's Republic of China

3 Adam Smith Business School, University of Glasgow, Glasgow G12 8QQ, UK

4 Chinese Academy of Finance and Development, Central University of Finance and Economics, Haidian District, Beijing 100081, People's Republic of China

1 Introduction

Future cash flows are critical for the survival of corporations. Reliable and accurate cash flow forecasting is important for academics and practitioners. For example, the value of a firm could be estimated by the sum of discounted future cash flows generated during its lifetime. One of the primary inputs of this valuation method is future cash flows. Also, when firms have larger accruals, more heterogeneous accounting choices than their peers in the same industry, higher earnings volatility, higher capital intensity, or poorer financial health, financial analysts prefer to provide cash flow forecasts to help their clients to make better investment decisions (Defond and Hung, 2003). Compared with accruals, cash flows are more difficult to be manipulated in earnings management; therefore, they could be used to monitor earnings transparency (McInnis and Collins, 2011). However, it is challenging to model the dynamics of future cash flows, which may be partially attributed to limited data and theory.

Before the 1980s, cash flows were indirectly estimated by deducting accruals and non-cash items from earnings, where measurement errors are inevitable (Drtina and Largay, 1985). Greenberg et al. (1986) use earnings and cash flow (both lagged) as two independent variables to predict cash flows. They suggest that lagged earnings perform better than lagged cash flows, both for one-year ahead and multi-year (2 and 3 years) ahead cash flow predictions. Wilson (1987) designs a multiple regression model which includes some lagged variables as predictors, e.g., earnings, cash flows, capital expenditures, and accruals. After the 1980s, cash flows data became publicly available.[1] Hence, an increasing number of papers about cash flows emerged. Finger (1994) examines the incremental predictive power of earnings to cash flows. Distinct from the previous research, she uses the unit-root test to examine the stationarity of these two time series. For 75% of the firms, their earnings and cash flows series are nonstationary, i.e., they follow a random walk process. Therefore, it may be better to use the difference instead of the level of earnings and cash flows series in the predictive study.

Based on the assumption of a random walk sales process, Dechow et al. (1998) (hereafter DKW model) develop a model of earnings, cash flows, and accruals. They also investigate the correlations between changes in earnings, cash flows, and accruals and their autocorrelations. The important message of their study is that, unlike earnings, a univariate time-series model is not sufficient to model cash flows, because it omits the other possible predictors, e.g., accruals. They also find that current earnings do outperform current cash flows in predicting future cash flows. Besides the random walk assumption, the DKW model also assumes that earnings and working capital accrual items are constant proportions of sales. Based on these assumptions, earnings could be the optimal predictor of future cash flows. However, because of managerial behaviours and other factors, earnings and working capital accruals may not always have a linear relationship with sales. Under a more empirical framework, Lorek and Willinger (2009) compare the performance of two single variable models in cash flow prediction, using earnings and cash flow of the last period respectively as the predictive variable. The two models are estimated in both cross-sectional and time-series ways, the former of which has a restrictive assumption that the parameters on the predictive variable are constant among firms. Moreover, the

---

[1] Canada was the first country that required cash flow disclosure in 1985. In the U.S, a statement of cash flow had become a compulsory part of financial reports since 1987, when Statement of Financial Accounting Standards (SFAS) No. 95 was published.

performances of the models are compared both in-sample and out-of-sample. They find that using past cash flow as the predictive variable and estimating the model on a firm-specific basis shows a better result in out-of-sample prediction. Ball and Nikolaev (2020) also study the predictive power of earnings for future cash flow with various methods. First, they use different definitions of earnings to construct the predictive variable, finding that. In model estimation, they apply fixed effects in pooled regression, which is expected to perform better than simple cross-sectional regression. Their empirical study support that earnings are better predictors for future cash flow, as long as the earnings are measured as operating cash flows adding working capital accruals. Gordon et al. (2017) also suggest that the results of cash flow prediction models are sensitive to accounting classification choices.

Based on the DKW model, Barth et al. (2001) (hereafter BCN model) decompose earnings into cash flows and accruals. However, unlike DKW who apply the model to the individual firms, BCN assume that the profit-generating processes of the sample firms are homogeneous and use pooled regression estimation method. BCN show that their revised model better fits the empirical data. Lorek and Willinger (2010) compare the BCN model with the DKW model. They suggest that DKW's firm-by-firm estimation strategy generates more accurate predictions. Cheng and Hollie (1996, 2008) disaggregate cash flows into core and non-core cash flows. The core components are generated from sales, cost of goods sold, and operating and administrative expenses. The non-core components are interest, taxes, and others. Cheng and Hollie find that the core components have higher persistence than the non-core components. Compared with the BCN model, the model with disaggregated cash flow components can improve cash flow prediction. Orpurt and Zang (2009) confirm that disaggregating cash flows may provide more useful information, which is further supported by Farshadfar and Monem (2013a and 2013b). Farshadfar and Monem (2011) focus on accruals and separate them into two components, i.e., discretionary and non-discretionary accruals. The two components of accruals are expected to make different contributions to cash flow prediction. Because discretionary accruals are more persistent, discretionary accounting could be employed to enhance earnings' predictive power for future cash flows. However, their conclusions need further verification. Bostwick et al. (2016) show that in addition to the BCN model, adding goodwill impairments to the predictor sets also improves the accuracy of cash flow prediction. Nallareddy et al. (2020) also find that disaggregating accruals into sub-components perform better than using aggregated accruals. More importantly, their study shows that the predictive ability of cash flows and accruals can differ over time.

The previous literature above indicates that the cash flow process is complicated. Hence, extra accounting information might be used as exogenous variables to better predict future cash flows. Also, heterogeneity in firms' business models and operating activities is not considered by the prior studies. Instead of using firm-by-firm estimation or pooled regression estimation, a panel data model which combines time-series and cross-sectional analysis could help resolve this issue. Linear models are commonly used by existing research. Although they are easy to understand, parsimonious, and have predictive power, they are inadequate to capture the nonlinearity of cash flow dynamics. Nonlinear models have more complicated structures and sometimes add computational burdens. However, if the nonlinear models could provide better forecasting performance in empirical applications, it would

not be difficult for researchers to make the trade-off between simplicity and accuracy. In addition, the cash flow generating process may vary with a firm's life cycle. The close link between cash flow patterns and firms' life cycle stages are well documented in financial accounting textbooks and previous literature. For example, Dickinson (2011) argues that the operating, investing, and financing cash flows behave differently across various life cycle stages. For cash flows from operations, introduction firms may generate negative cash flows due to the lack of customers and experience. However, in the growth and maturity stages, their operating cash flows would become positive and grow at different rates. In the decline stage, the growth rates of firms' operating cash flows decrease, and the firms may even suffer from negative cash flows again. Although the investing and financing cash flows are not the focus of our study, their patterns may also vary with different life-cycle stages, which can be explained by economic theories such as managerial optimism and pecking order theory. In addition, Hovakimian (2009) suggests that cash flow is associated with a firm's life cycle, according to the corporate life cycle hypothesis. Firms are normally characterized by low cash flows when they are at their early stage, and they are often able to generate greater cash flows when they become more mature. They argue that firms may experience growing cash flows and decreasing growth opportunities over their life cycle. Therefore, the existing static models could be extended to dynamic models to improve cash flow prediction performance. Also, in terms of selecting the optimal prediction model, there is no consensus regarding the criteria. Both in-sample fitting and out-of-sample prediction should be considered to evaluate one model's forecasting performance.

This paper attempts to address the above-mentioned research gaps. The main contributions of this paper are as follows. First, based on the DKW model and the BCN model, we introduce panel data models to allow for the heterogeneity in firms' business activities. Also, motivated by the grey-box model developed by Tan and Li (2002), we suggest potential improvements to incorporate the dynamic and nonlinear components in the panel data models.[2] The parameters in the static panel data models are treated as time-varying (TV) state variables and there are no linearity restrictions on the TV parameters. The nonlinearity of the parameters is captured by a function of the Padé approximant. Two exogenous variables are employed as input variables that are considered to have explanatory power for the parameter process. Moreover, we apply our models in the U.S. market and show that their prediction performances are better than that of the existing models. Finally, we discuss the long-term cash flow forecasting and the criteria adopted to evaluate the models' performance. The idea of the vector autoregressive (VAR) model is analogously extended to a nonlinear form, so that our advanced models are applicable in the multiple-period setting. Regarding model selection, although it is difficult to draw a consistent conclusion based on multiple criteria, we could at least approach a balance point.

The structure of this paper is as follows. Section 2 focuses on the model developments that fill in the gaps of the prior literature. In Sect. 3, various cash flow prediction models are empirically applied to the U.S. dataset, and the empirical results are discussed. Section 4 is the concluding remarks.

---

[2] According to Tan and Li (2002), white-box models refer to a set of linear ordinary differential equations (ODEs) which are based on unrealistic assumptions. Examples of black-box models include nonlinear autoregressive moving average models and artificial neural network models. The complexity of black-box models limits their applications. Combining the characteristics of white-box models with that of black-box models, grey-box models take advantage of prior physical knowledge of complex systems and information in the dataset. Therefore, grey-box models could achieve a good balance between complexity and accuracy.

2 Model specification

2.1 Dynamic cash flow prediction model

In this section, we start with DKW's model and then describe the reasons why it is critical to propose grey-box models. The DKW model makes two major assumptions. First, sales follow a random walk. Second, costs and working capital accruals (i.e., accounts receivables, accounts payables, and inventory) are constant proportions of sales. Under these assumptions, the best forecast of cash flow is:

$$E_t[CF_{t+k}] = EARN_t, k = 1, 2, ..., n \tag{1}$$

where *CF* denotes net operating cash flow and *EARN* denotes earnings. In their empirical study, DKW use the firm-specific regression below:

$$CF_{i,t+k} = \gamma_{i,0} + \gamma_{i,1} CF_{i,t} + \gamma_{i,2} EARN_{i,t} + \varepsilon_{i,t} \tag{2}$$

where $\gamma_{i,0}$, $\gamma_{i,1}$ and $\gamma_{i,2}$ are time-invariant parameters. Individual firms are denoted by the subscript *i*. DKW's model suggests that earnings, as a naive predictor, work better than cash flows itself, because earnings take account of the information of accrual terms. However, DKW's random walk assumption on sales may be too strict. As shown in Sect. 3.3, sales do not necessarily follow a random walk, but have a predictable growth pattern. To generalise the model, we introduce an additional variable r_t, which is the growth rate of sales. Re-derive the DKW model as follows. First, net operating cash flows is the difference between cash received and cash paid out[3]:

$$\begin{aligned} CF_t &= (SALES_t - \Delta AR_t) - (PURCHASE_t - \Delta AP_t) \\ &= (SALES_t - \Delta AR_t) - (COST_t + \Delta INV_t - \Delta AP_t) \\ &= EARN_t - \Delta WC_t \end{aligned} \tag{3}$$

Assume that costs, accounts receivables, accounts payables, and inventory are constant proportions of sales, so are the earnings and working capital accruals:

$$EARN_t = \alpha SALES_t \tag{4}$$

$$WC_t = \beta SALES_t \tag{5}$$

where α and β are constants. Define r_t to be the growth rate of sales and the relation holds:

$$SALES_t = (1 + r_t)SALES_{t-1} \tag{6}$$

The recursive relationships for earnings and working capital accruals can be derived as:

$$\begin{aligned} EARN_t &= \alpha SALES_t = \alpha(1 + r_t)SALES_{t-1} \\ &= (1 + r_t)EARN_{t-1} \end{aligned} \tag{7}$$

---

[3] Please refer to Table 1 for variable definitions.

Table 1 Descriptive statistics for the variables in cash flow prediction model

	Mean	Standard Error	Median	Mode	Standard Deviation	Sample Variance	Kurtosis	Skewness	Range	Minimum	Maximum
CF	0.0683	0.0004	0.0631	0.0000	0.1318	0.0174	26.6543	−0.1282	6.3604	−1.8893	4.4711
DA	0.0459	0.0001	0.0359	0.0000	0.0461	0.0021	122.0470	6.8188	1.6376	−0.0034	1.6342
ΔAR	0.0142	0.0002	0.0068	0.0000	0.0722	0.0052	34.8779	0.7104	2.8081	−1.3301	1.4779
ΔAP	0.0077	0.0002	0.0033	0.0000	0.0483	0.0023	160.6502	2.0117	4.1359	−1.9632	2.1727
ΔINV	0.0094	0.0002	0.0008	0.0000	0.0592	0.0035	148.3367	2.2320	5.3354	−2.2495	3.0859
$OTHER$	0.0214	0.0003	0.0196	0.0000	0.0892	0.0080	165.1803	−2.1291	7.4330	−4.9773	2.4557

The variables are defined as follows: CF is the net operating cash flow; DA is the depreciation and amortisation; ΔAR is the changes in account receivable; ΔAP is the changes in account payable; ΔINV is the Changes in inventory; $OTHER$ is the other accruals, calculated as earnings before interest, tax, depreciation and amortisation (EBITDA) – (CF+ΔAR+ΔINV – ΔAP – DA). All variables are deflated by average total assets of each firm.Table 2: The cash flow prediction model parameters estimated with panel data methods

$$\Delta WC_t = \beta \Delta SALES_t = \beta r_t SALES_{t-1}$$
$$= \beta r_t(1 + r_{t-1})SALES_{t-2} = \left(\frac{r_t}{r_{t-1}} + r_t\right)\Delta WC_{t-1} \quad (8)$$

Therefore, Eq. (3) can be rewritten as:

$$CF_t = EARN_t - \Delta WC_t = (1 + r_t)EARN_{t-1} - (\frac{r_t}{r_{t-1}} + r_t)\Delta WC_{t-1}$$
$$= (1 + r_t)EARN_{t-1} - (1 + r_t)\Delta WC_{t-1} + (\frac{r_{t-1} - r_t}{r_{t-1}})\Delta WC_{t-1} \quad (9)$$
$$= (1 + r_t)CF_{t-1} + (\frac{r_{t-1} - r_t}{r_{t-1}})\Delta WC_{t-1}$$

Combining (8) and (9), we have the following relationship based on the information set at time t:

$$E_t\left(\begin{bmatrix} CF_{t+1} \\ \Delta WC_{t+1} \end{bmatrix}\right) = E_t\left(\begin{bmatrix} 1 + r_{t+1} & \frac{r_t - r_{t+1}}{r_t} \\ 0 & \frac{r_{t+1}}{r_t}(1 + r_t) \end{bmatrix}\right) \times \begin{bmatrix} CF_t \\ \Delta WC_t \end{bmatrix} \quad (10)$$

When r_t is zero, i.e., sales follow a random walk, Eq. (10) reduces to the DKW model described by Eq. (1). More importantly, because r_t is not a constant, Eq. (10) shows a dynamic relationship between future cash flows and accrual terms. Equation (10) also implies the heterogeneity of individual firms in predicting cash flows. The parameters estimated for individual firms might be different, which mainly comes from the growth stages they are in. As predicted by Eq. (10), when firms are in the early stages, their sales growth rates are usually high. The difference between the parameters of lagged cash flows and lagged accruals is large, which implies that a small portion of future cash flows is predicted by lagged accruals. As firms become mature, their growth slows down, and the explanatory power of lagged cash flows and lagged accruals would gradually converge. This conjecture is supported by empirical findings in Sect. 3.3. Moreover, at time t, r_t is observable, while r_{t+1} is unknown. Based on our relaxed assumption that sales may have a predictable growth pattern, r_{t+1} could be expressed as a linear or non-linear function of r_t. Therefore, Eq. (10) could be reduced to the version which only depends on r_t. It implies that the coefficients of CF_t and ΔWC_t are essentially governed by r_t.

The BCN model suggests that the components of accrual terms have different effects, and therefore it extends Eq. (2) to:

$$CF_{t+1} = \beta_0 + \beta_1 CF_t + \beta_2 \Delta INV_t + \beta_3 \Delta AP_t + \beta_4 \Delta AR_t$$
$$+ \beta_5 DEP_t + \beta_6 AMORT_t + \beta_7 OTHER_t + \varepsilon_{t+1} \quad (11)$$

where *DEP* denotes depreciation, *AMORT* denotes amortisation, and *OTHER* denotes other accruals. Cheng and Hollie (2008) further extend the BCN model by disaggregating cash flow into components, which involves more regressors. Due to data availability, the number of firms that are eligible for individual estimation may be small. This could be a reason why BCN and Cheng and Hollie use cross-sectional regression rather than comply with the DKW framework that estimates model parameters individually. Empirically, when

considering individual effects, panel data models are more appropriate than cross-sectional regression. As seen, Eq. (11) is still a simple, parsimonious, and static model, which can be used as a benchmark.

2.2 Panel data method

In panel data analysis, we face the problem of dealing with data of two dimensions. This study mainly focuses on accounting variables at firm level. Therefore, the time-series dimension of the panel is relatively short due to the low frequency of a firms' financial information disclosure, but the cross-sectional dimension is large as there are many firms in the market. A question naturally arises that there may exist heterogeneity across different groups and the methods of model estimation should adapt to this situation accordingly. For instance, the fixed effect and random effect models are linear estimators used in the situation when there is heterogeneity only in the intercept term among the groups while the factor loadings on the predictive variables remain homogeneous. Equation (11) then becomes:

$$
\begin{aligned}
CF_{i,t+1} = &\beta_{i,0} + \beta_1 CF_{i,t} + \beta_2 \Delta INV_{i,t} \\
&+ \beta_3 \Delta AP_{i,t} + \beta_4 \Delta AR_{i,t} + \beta_5 DEP_{i,t} \\
&+ \beta_6 AMORT_{i,t} + \beta_7 OTHER_{i,t} + \varepsilon_{i,t+1}
\end{aligned}
\tag{12}
$$

Note that subscript i for individual effects only appears in the intercept $\beta_{i,0}$. If the firms are completely heterogeneous, there is no gain in analysing panel data, and it would be optimal to undertake individual estimation. It is unrealistic to assume homogeneity across firms. First, the sizes and/or business scales of firms vary, which often implies that their financial variables are unlikely to follow the same statistical distribution. Firms with different sizes also have different exposures to business risks. However, companies may at least share some similarities. For example, managers learn from leading firms no matter what industries they specialise in.

Heterogeneity in firm size causes the potential problem that for different firms the unexplained parts of the dependent variable may not be identically distributed. In Eq. (12), this problem is reflected in two aspects: the intercept may differ in level and the error term may have different variances. Difference in intercept would bias the parameter estimation by cross-sectional regression if the individual effects are correlated with the independent variables. In predictive applications, we need to consider and distinguish the individual effects of each firm, otherwise the prediction would be significantly biased. The individual intercepts can be solved by adding dummy variables, such as in Barth, et al. (2001), which is applicable but brings in too many parameters. Alternatively, the intercept term of individual effect can be eliminated by using one of the following two approaches –- demean and first difference.

Before introducing the two methods, it should be noted that the different variances of the error terms also have an impact on the estimation procedure. A regression-based estimation procedure aims to minimise the sum of squared errors or another error norm. Therefore, firms with large variance in the error term tend to dominate the results, especially when the degree of heterogeneity is high. A solution provided in the BCN paper is to deflate all the variables by a size-related variable, such as the total assets or shares outstanding. In this paper, average total asset is used as the deflator.

Demean. If (12) holds, for each individual firm i the following holds too:

$$\overline{CF}_i = \bar{\beta}_{i,0} + \beta_1 \overline{CF}_i + \beta_2 \overline{\Delta INV}_i + \beta_3 \overline{\Delta AP}_i$$
$$+ \beta_4 \overline{\Delta AR}_i + \beta_5 \overline{DEP}_i + \beta_6 \overline{AMORT}_i + \beta_7 \overline{OTHER}_i + \bar{\varepsilon}_i \tag{13}$$

Deduct (13) from (12), and we have:

$$
\begin{aligned}
CF_{i,t+1} - \overline{CF}_i =& \beta_{i,0} - \bar{\beta}_{i,0} + \beta_1(CF_{i,t} - \overline{CF}_i) \\
&+ \beta_2(\Delta INV_{i,t} - \overline{\Delta INV}_i) + \beta_3(\Delta AP_{i,t} - \overline{\Delta AP}_i) \\
&+ \beta_4(\Delta AR_{i,t} - \overline{\Delta AR}_i) + \beta_5(DEP_{i,t} - \overline{DEP}_i) \\
&+ \beta_6(AMORT_{i,t} - \overline{AMORT}_i) \\
&+ \beta_7(OTHER_{i,t} - \overline{OTHER}_i) + \varepsilon_{i,t+1} - \bar{\varepsilon}_i
\end{aligned}
\tag{14}
$$

Removing the group mean from each variable eliminates the individual effect of the intercept term but keeps the other parameters unaffected (actually the demeaning has brought endogeneity into the model, which will be further discussed later). The individual $\beta_{i,0}$ could be calculated by manipulating Eq. (14) as:

$$\widehat{\beta}_{i,0} = \overline{CF}_{i,t} - (\beta_1 \overline{CF}_{i,t-1} + \beta_2 \overline{\Delta INV}_{i,t-1} + \beta_3 \overline{\Delta AP}_{i,t-1}$$
$$+ \beta_4 \overline{\Delta AR}_{i,t-1} + \beta_5 \overline{DEP}_{i,t-1} + \beta_6 \overline{AMORT}_{i,t-1} + \beta_7 \overline{OTHER}_{i,t-1}) \tag{15}$$

First Difference. An alternative approach is to take the first difference for all the variables, and the model becomes:

$$
\begin{aligned}
\Delta CF_{i,t+1} =& \beta_1 \Delta CF_{i,t} + \beta_2 \Delta^2 INV_{i,t} + \beta_3 \Delta^2 AP_{i,t} \\
&+ \beta_4 \Delta^2 AR_{i,t} + \beta_5 \Delta DEP_{i,t} + \beta_6 \Delta AMORT_{i,t} \\
&+ \beta_7 \Delta OTHER_{i,t} + \Delta \varepsilon_{i,t+1}
\end{aligned}
\tag{16}
$$

where Δ^2 denotes the second order difference i.e. Δ_t-Δ_{t-1}. By differencing, the intercept, which is assumed to differ across individuals but to be constant over time, has been eliminated. Using the least squared (LS) method to estimate the parameters in (14) and (16) is unable to solve the endogeneity problem brought by the inclusion of AR term (for details see e.g., Cameron and Trivedi, 2005). Endogeneity causes inconsistency in the estimation of parameters. We apply the general method of moments (GMM) to obtain theoretically consistent estimated parameters. The GMM estimator is applied to Eq. (16), and it takes the form of:

$$\widehat{\beta}_{AB} = \left[\left(\sum_{i=1}^{N} \tilde{\mathbf{X}}_i' \mathbf{Z}_i \right) \mathbf{W}_N \left(\sum_{i=1}^{N} \mathbf{Z}_i' \tilde{\mathbf{X}}_i \right) \right]^{-1} \left(\sum_{i=1}^{N} \tilde{\mathbf{X}}_i' \mathbf{Z}_i \right) \mathbf{W}_N \left(\sum_{i=1}^{N} \mathbf{Z}_i' \tilde{\mathbf{y}}_i \right) \tag{17}$$

where $\tilde{\mathbf{X}}_i$ is a $(T-2) \times (K+1)$ matrix with all regressors including the lagged dependent variable for the ith company and $\tilde{\mathbf{y}}_i$ is a $(T-2) \times 1$ vector of the ith company's dependent variable. $\mathbf{Z}_i$ is a $(T-2) \times r$ matrix of instrumental variables (IV). $\mathbf{W}_N$ is weighting matrix. The GMM estimator is used for the first time in cash flow prediction. It is compared empirically with the demean method to see whether this theoretically more consistent model can provide better prediction as well.

2.3 Nonlinear dynamic (grey-box) cash flow prediction model

Equation (10) describes a dynamic model that may better predict future cash flows. This section introduces a nonlinear dynamic model. This model suggests that the parameters' dynamics are nonlinear and unknown. Re-write Eq. (11) by allowing for time-varying parameters:

$$
\begin{aligned}
CF_{i,t+1} = {}& \beta_{i,t,0} + \beta_{i,t,1}CF_{i,t} + \beta_{i,t,2}\Delta INV_{i,t} \\
& + \beta_{i,t,3}\Delta AP_{i,t} + \beta_{i,t,4}\Delta AR_{i,t} + \beta_{i,t,5}DEP_{i,t} \\
& + \beta_{i,t,6}AMORT_{i,t} + \beta_{i,t,7}OTHER_{i,t} + \varepsilon_{i,t+1}
\end{aligned}
\tag{18}
$$

Each parameter in Eq. (18) is controlled by a process:

$$
\beta_{i,t,j} = F\big(z_{t+1}, z_t\big), j = 0, \dots, 7
\tag{19}
$$

where z_{t+1} and z_t are dynamic processes and determine the dynamics of $\beta_{i,t,j}$. As shown in Eq. (10), sales growth rate r_t could serve as one of the proxies for z_t. The form of F is unknown, but we could use a nonlinear approximation function to fit it. There are several options for such functions. For instance, a neural network is considered as a universal approximator since it can approximate any function (Cybenko, 1989). Taylor series and Fourier series could approximate functions with any degree of accuracy. Considering that higher degrees of complexity usually cost efficiency, this paper adopts the Padé approximant (Tan and Li, 2002) for F because it requires fewer coefficients while maintains sufficient accuracy. Besides sales growth rate, we also use firm age as the input variable of the black-box model. Sales growth and firm age are commonly employed as the proxies for life cycle (e.g., Anthony and Ramesh, 1992). We use the sales growth rate r_{t+1} and r_t as a candidate for z_{t+1} and z_t as a starting point. r_{t+1}, however, is not observable at time t. If we assume that r_{t+1} is predicted by r_t, either by linear or nonlinear predictive functions, Eq. (19) can be reduced to a function of a single variable r_t and expressed as:

$$
\beta_{i,t,j} = F(r_{i,t}) = \frac{a_0 + a_1 r_{i,t} + a_2 r_{i,t}^2}{1 + a_3 r_{i,t} + a_4 r_{i,t}^2} j = 0, \dots, 7
\tag{20}
$$

where a_0 to a_4 are the coefficients to capture the dependence of the parameter $\beta_{i,t,j}$ on r_t. The functional form in Eq. (20) is a Padé approximant of order 2/2 (take the polynomials of the variable r_t up to order 2, both in the numerator and the denominator). In Eq. (18), we have 8 parameters, and hence there are 40 Padé approximant coefficients in total to be determined in this model. Increasing the order of Padé approximant would improve the precision of data fitting but also inevitably require more coefficients to be estimated. We use order 2/2 to reach an optimal balance between gain and cost. A simple example can illustrate how Eq. (20) capture the dynamics of $\beta_{i,t,j}$. Assume that $\beta_{i,t,1}$ follows exactly the form of Eq. (10) and a firm's growth rate decays at a constant rate, e.g., $r_{t+1} = 0.9r_t$. Therefore, $\beta_{i,t,1}$ in Eq. (20) could be calibrated as $(1 + 0.9r_t)$, which has a dynamic pattern as plotted in Fig. 1.

Equation (19) and (20) show an application of a black-box model, the antonym of a white-box model. The latter one is mostly common in physics and engineering disciplines where physical laws are well known and applied without any uncertainty. In social science, however, we need to consider uncertainty, structural changes, etc. Hence, the exact forms of functions describing the complex interactions among variables may not be available. Nevertheless, we
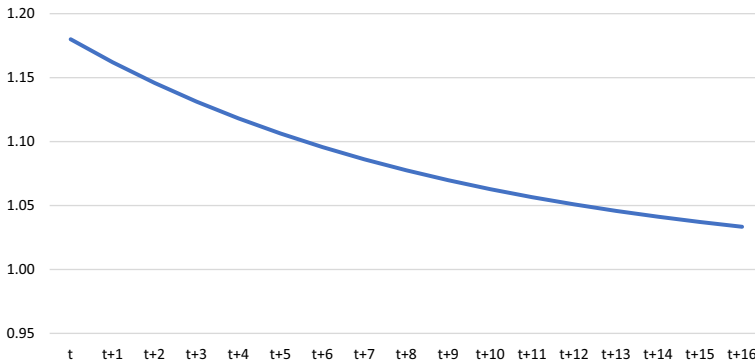
Fig. 1 An example of the dynamics of $\beta_{i,t,1}$ in Eq. (20) Note: Assume that the initial growth rate of a firm is 20%

may use data and black-box models to approximate the relationships between variables with certain accuracy. In the spirit of Tan and Li (2002), the joint model of Eq. (18) and (19) forms a grey-box system. The black-box model, i.e., Eq. (19), governs the dynamics and the nonlinearity of model parameters of the white-box model, i.e., Eq. (18). Each of the 8 parameters in Eq. (18) takes the function described by Eq. (20), and thus there are 40 Padé approximant coefficients to be estimated. We estimate them by minimising the sum squared prediction errors of all observations. With these coefficients, each parameter β could be calculated accordingly with respect to the level of r_t or firm age, and then the prediction of cash flows could be obtained.

2.4 Long-Term Prediction

A natural extension of the one-period BCN model is to increase the lag length for multiperiod ahead predictions. The main drawback of this option is that the maximum period ahead that can be predicted is critically limited by data availability. Stock valuation is considered as the aggregation of all cash flows that will be received in the future discounted back to the current time. In principle, the cash flows need to be predicted to infinite future. In univariate time series models, it is simple to derive multiple-period models by the recursive results. In multi-variable models, however, the effects of the other variables need to be considered as well. For simplicity, those variables can be treated as exogenous. Alternatively, long-term forecast can be achieved by estimating predictive models of all those variables and recursively forecast them into any period in the future. In linear models, the Vector Autoregressive (VAR) model (Sims, 1980) is the corner stone and has the form:

$$\mathbf{y}_{i,t} = \boldsymbol{\beta}\mathbf{y}_{i,t-1} + \boldsymbol{\nu}_{i,t} \tag{21}$$

where $\mathbf{y}_{i,t}$ is the vector of all relevant variables in the model; $\boldsymbol{\beta}$ is the parameter matrix identifying the predictive system; $\boldsymbol{\nu}_{i,t}$ is the disturbance vector for all variables. Once $\boldsymbol{\beta}$ is estimated, all variables could be predicted by the following relationship:

$$\widehat{\mathbf{y}}_{i,t+k} = \boldsymbol{\beta}^k\mathbf{y}_{i,t} \tag{22}$$

where k denotes the number of periods ahead required to be predicted. VAR is not limited by the length of data samples and hence is a more flexible tool. The dynamic models, either linear or nonlinear, take the form below instead:

$$\mathbf{y}_{i,t} = \boldsymbol{\beta}_t \mathbf{y}_{i,t-1} + \boldsymbol{\nu}_{i,t} \tag{23}$$

Hence, the prediction takes the form of:

$$\widehat{\mathbf{y}}_{i,t+k} = \prod_{n=1}^{k} \boldsymbol{\beta}_{t+n} \mathbf{y}_{i,t} \tag{24}$$

2.5 Model performance evaluation

For individual time-series prediction, it is straightforward to compare the performance of different models. In most cases, error-based criteria are commonly used, such as mean squared errors (MSE), mean absolute error (MAE), etc. There are also non-parametric measures. The model that generates smaller prediction errors is preferred. It is not trivial to select one criterion or multiple criteria to evaluate these models, especially in the panel data setting. In panel data, there are many individual firms, which makes the comparison results contradictive. In practice, the results of specific individuals are of more concern than that of the aggregate group, and firm-based specific prediction is of more value.

Usually, it is difficult to judge two models according to the aggregated measure, e.g., SSE of all firms. It is highly likely that one model that produces a smaller SSE performs worse for half of the sample firms. Therefore, the two-dimensional feature of panel data poses another requirement on predictive models in practice, i.e., generality. A good model should fulfil the aggregated accuracy and show superior power for as many individuals as possible. Studies in the early periods, e.g., Ball and Watts (1972), calculate the average rank of each model in fitting each observation or each group in general as a measure to evaluate the models. This paper also adopts the rank measure as a criterion for judging models' performance.

3 Empirical study of the U.S. market

We study the annual data of the U.S. firms because the U.S. is the world's largest economic entity and the empirical evidence drawn from the U.S. data could be treated as a benchmark. The variables used in this study are all public financial information, which are directly available in firms' annual reports. All accounting data are obtained from the WRDS Compustat database. The data cover all the listed firms in the U.S., spanning the period from 1957 to 2013.

Before proceeding to the application of cash flow prediction models, it is helpful to have a general impression about the cash flow process first. There is no doubt that each sample firm has its particular cash flow paths. Larger firms tend to generate higher cash flows and vice versa, which makes the cash flows of different firms incomparable. Therefore, the first step in comparing them is to normalise the cash flow series of each firm. We accomplish it by deflating each firm's cash flows by their initial positive cash flow observation. Negative
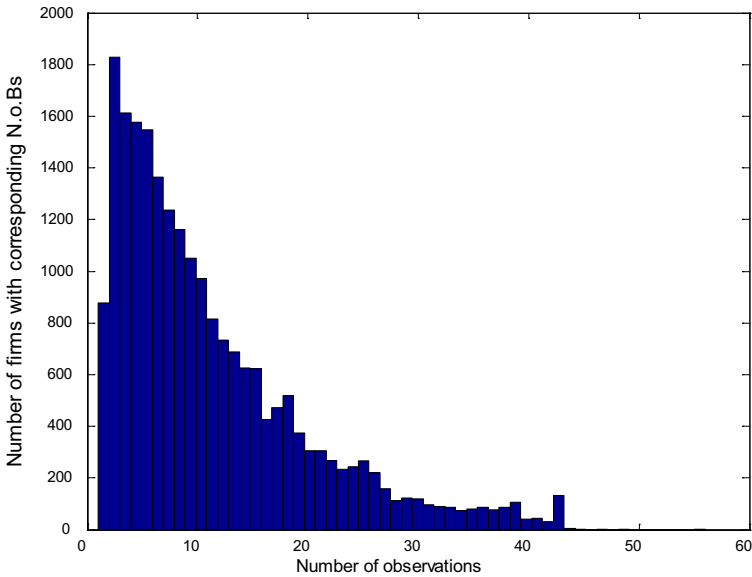
Fig. 2 The distribution of the number of firms with different sample lengths. Note: There are 21,905 firms with at least one available cash flow observation that is either disclosed or indirectly estimated and 239,835 firm-year cash flow observations in total from 1957 to 2013 for the entire sample

observations, if they appear at the beginning of any firm's cash flow series, are excluded. In such way, every firm has the same starting point, i.e., one unit of cash flow, no matter when it starts to operate. This process is less influenced by specific year effects because the time when each firm starts to enter the sample is diversified. The indicator for time is not the absolute year, e.g., 1987, 1990, or 2015, but is denoted as the number of years ahead of each firms' beginning time. Thus, firms with various sizes and/or different ages could be compared. The results should therefore be more general than alternatively bringing together observations of the same particular period.

Cash flow disclosure was not compulsory until 1987. The DKW paper attempts to estimate cash flow indirectly from balance sheet and income statement in order to increase the available sample with tolerable measurement and calculation errors. In this study, cash flows before 1987 are estimated indirectly following the DKW approach. Operating cash flow is estimated using the general formula: net income + depreciation and amortisation – changes in non-cash current assets + changes in current liability. Using the data after 1987, the correlation between actual cash flow and estimated cash flow is 0.82. Therefore, it seems that the level of estimation error is tolerable.

3.1 Trend of cash flow levels

There are 21,905 firms with at least one available cash flow observation that is either disclosed or indirectly estimated and 239,835 firm-year cash flow observations in the entire sample. The distribution of available observations of each firm is shown in Fig. 2. The firm with the most observations is IBM that has 56 observations. Most sample firms have less than 20 observations. Both the upper and lower 1 percent of the normalised cash flows
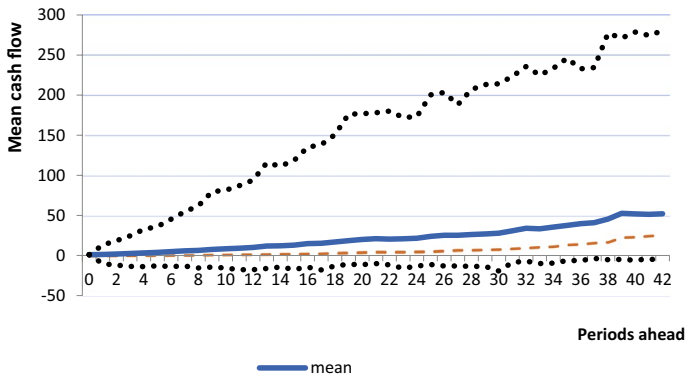
Fig. 3 Mean, median and the 2.5–97.5 percentiles of the normalised cash flow series of all firms. Note: For each leading period, i.e., period ahead of the starting time, there will be a distribution of cash flow levels. This figure shows the mean, median and 95% range (from 2.5 to 97.5 percentiles) up to 42 periods ahead of starting time. The results beyond 43 periods ahead are not shown due to very few observations (no more than 16)

are excluded. For each leading period, i.e., the period ahead of the starting time, there is a distribution of cash flow levels. The mean, median and 95% range (from 2.5 to 97.5 percentiles) are shown in Fig. 3. The figure only shows 42 periods ahead of starting time, as the number of sample firms available reduces along with the x-axis. The results beyond 43 periods ahead are based on very few observations (no more than 16), so that they are not meaningful. In the chart, there is an obvious and almost monotonic upward trend for the mean and median, the latter being lower, cash flow series. The mean level of cash flow in period 42 is nearly 52, which implies an annual growth rate of 10%. Similarly, the median level of cash flow in period 42 is 26, implying an annual growth rate of 8%.

However, because of survivorship bias, it is early to conclude that the U.S. firms have high cash flow growth rates. From Fig. 3, it is clearly seen that the cash flow distributions of each period are asymmetric. The asymmetry increases with periods. In 42 years, there are firms whose cash flow increased to 250 times but there is no firm whose cash flow decreased to 250 times. Managers tend to eliminate the possibility of symmetric cash flow distributions. Firms that incur losses one year after another may not be allowed to lose permanently. Hard decisions may be made, e.g., the management might be replaced, go into bankruptcy, or be taken over. If the firms could not recover, they are likely to exit the business. Therefore, it is possible that firms' cash flows grow permanently to a very high level, but it is impossible that firms' cash flows decrease permanently. This study conducts a simulated illustration to show the effect of survivorship bias. Assume that cash flow follows a random walk process, and its noise term follows a normal distribution with zero mean and variance of 16:

$$CF_t = CF_{t-1} + \varepsilon_t$$
$$CF_0 = 1 \qquad\qquad (25)$$
$$\varepsilon_t \sim N(0, 16)$$

Theoretically, this process has an expectation of 1, i.e., the initial value, in any period. However, a quit rule to the process is set: stop if there are 5 negative numbers in a row. The maximum length for each simulated series is 42. The simulation is run for 10,000 times.
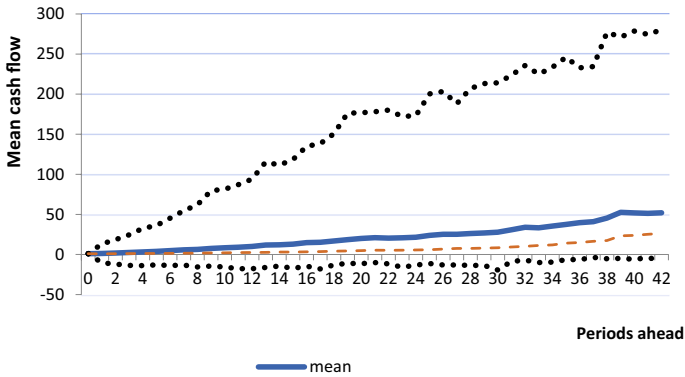
Fig. 4 Mean, median and the 2.5–97.5 percentiles of the simulated cash flow series. The distribution of the simulated sample is illustrated in Fig. 4. The maximum length for the simulation is 42. The simulation is run for 10,000 times the number of sample firms

The distribution of the simulated sample is illustrated in Fig. 4. There are some similar features between Fig. 3 and 4. The simulated data also shows an increasing pattern for the mean level of cash flow series even though they are generated from a random walk process. These results suggest that survivorship bias has an upward effect on the general conclusion drawn from the sample. The true expectation of cash flow trend may not be as high as shown in Fig. 3.

To take the survivorship bias into account, the survival rates of the firms are calculated. For each period from 1 to 42, the survival rate is calculated by dividing the number of firms whose number of observations of the specified length is available by the number of firms that appear early enough in the sample to provide the required number of observations. For example, there are 17,043 firms providing observations of one year ahead of their initial time. To examine the 1-year survival rate, the denominator is the number of firms of which the first observations appeared before (including) year 2011, which is 18257. Thus, the 1-year survival rate is 93.35%. To calculate the 42-year survival rate, the numerator is the number of firms that provides observations of 42 years ahead of their initial time, i.e., 294, and the denominator is the number of firms that started before (including) 1970, i.e., 2015. The rates indicate the proportion of firms that survives in the list for a certain length of period, and this is depicted in Fig. 5. The survival rates suggest that in 10 years less than half of the U.S. firms remain in the market; in 42 years, less than 15%. Therefore, it should be realised that when calculating the mean of cash flow series with the surviving sample firms, a large proportion of bad cases are not contained in the sample. To make an adjustment, the mean of cash flow series is multiplied by the survival rates, which should be a better way to describe the true unconditional expectation of the cash flow pattern. The adjusted mean series are depicted in Fig. 6, along with the original cash flow mean series. The adjusted mean suggests that it is more appropriate to expect, in general, that a firm's cash flow could grow to 7.58 times, rather than 52 times, of its original value in 42 years. This implies an annual growth rate of roughly 5%.
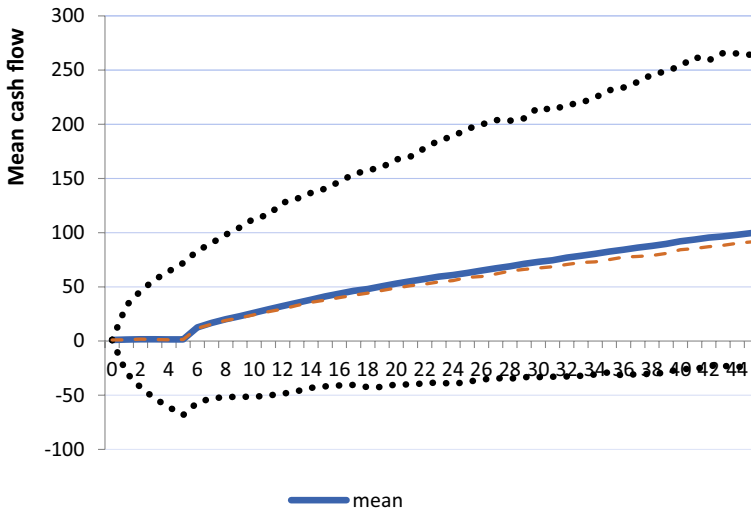
Fig. 5 Survival rates of U.S. sample firms. Note: This figure depicts the survival rates of the US firms. The rates indicate the proportion of firms that are listed at any time in the market for a certain length of period
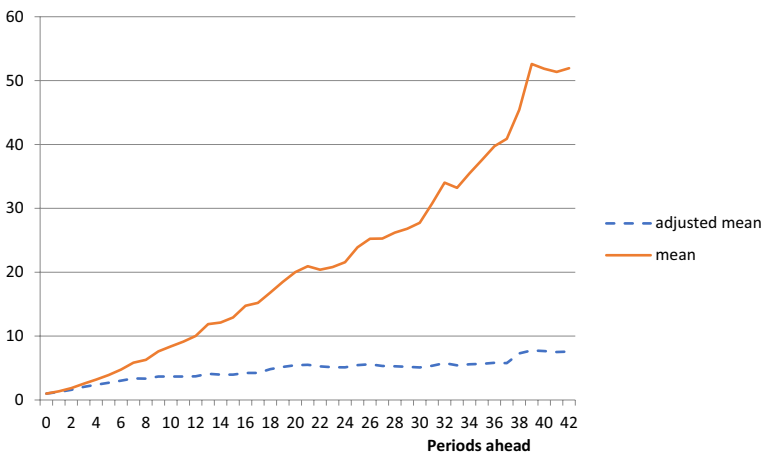

Fig. 6 The mean cash flow series adjusted for survivorship bias. To describe the true unconditional expectation of the cash flow pattern make an adjustment, the mean of cash flow series is multiplied by the survival rates. Figure 6 plots the adjusted mean cash flow series along with the originally calculated mean cash flow series

3.2 Parameter estimation in the static model using different methods

From this section on, the public available cash flow data are used in the modelling of cash flow. We follow the criteria listed in the BCN paper to exclude observations if they belong to any of the following categories:

- Financial services firms (SIC codes 6000-6999);

Table 2 he cash flow prediction model parameters estimated with panel data methods.

	Pooled Regression		Demean	Difference	A−B
	(1987–1996)	(1987–2005)	(1987–2005)	(1987–2005)	(1987–2005)
β_1	0.613***	0.692***	0.392***	− 0.303***	0.197***
	(39.94)	(82.46)	(33.02)	(− 20.75)	(6.55)
β_2	0.247***	0.323***	0.273***	0.0785***	0.0742*
	(10.81)	(17.61)	(15.68)	(4.70)	(2.51)
β_3	− 0.534***	− 0.594***	− 0.491***	− 0.240***	− 0.292***
	(− 13.05)	(− 23.24)	(− 18.90)	(− 12.05)	(− 8.10)
β_4	0.399***	0.447***	0.416***	0.199***	0.195***
	(19.49)	(27.53)	(22.17)	(11.49)	(6.71)
β_5	0.253***	0.140***	0.293***	0.0581	0.0374
	(8.40)	(7.06)	(7.99)	(1.72)	(0.73)
β_6	0.290***	0.351***	0.338***	0.142***	0.132***
	(16.35)	(24.10)	(19.53)	(7.99)	(4.47)
β_7	0.00336**	0.00439***			
	(3.17)	(5.27)			
N	27,630	63,120	63,120	52,884	52,884

Statistics in parentheses. ***, **, and * denote the significance level at 1, 5, and 10%, respectively. Model estimated: $CF_{i,t+1} = \beta_{i,0} + \beta_1 CF_{i,t} + \beta_2 \Delta INV_{i,t} + \beta_3 \Delta AP_{i,t} + \beta_4 \Delta AR_{i,t} + \beta_5 DA_{i,t} + \beta_6 OTHER_{i,t} + \varepsilon_{i,t+1}$.

- Sales less than \$10 million;
- Share price less than \$1;
- Earnings or cash flow in the extreme upper and lower 1 percent of their respective distributions.

This provides a sample of 99,845 firm-year observations. Table 1 provides the descriptive statistics of cash flow, depreciation and amortisation, changes in account receivable, changes in account payable, changes in inventory and other accruals, all variables deflated by average total assets of each firm. On average, cash flow deflated by average total assets is about 0.07 for the sample firms, but the dispersion is very large with a standard deviation of 0.13. There are also special cases in the sample where the minimum and maximum cash flow observations are greater than the firms' average total assets in magnitude.

The equation to be estimated is:

$$CF_{i,t+1} = \beta_{i,0} + \beta_1 CF_{i,t} + \beta_2 \Delta INV_{i,t} + \beta_3 \Delta AP_{i,t} + \beta_4 \Delta AR_{i,t} \\ + \beta_5 DA_{i,t} + \beta_6 OTHER_{i,t} + \varepsilon_{i,t+1} \tag{26}$$

where *DA* denotes depreciation and amortisation. Note the intercept term is assumed to be identical for all firms in the pooled regression. First, to compare with the results in the BCN paper, pooled regression is applied for data between 1987 and 1996. For the rest of the paper, the whole sample is partitioned into two subsamples: data from 1987 to 2005 is used for in-sample estimation, and data from 2006 to 2013 is used for out-of-sample prediction performance comparison. Parameters in Eq. (26) are then estimated using four different methods: pooled regression, demean, first difference, and Arellano-Bond estimator.

The estimated results are summarised in Table 2. Numbers in parentheses are *t* statistics based on heteroskedastic robust standard error. The second column shows the results for the period between 1987 and 1996. The number of sample observations is 27,630. The results are very close to that of the BCN paper. All the selected variables are both statistically and economically significant and the signs of the parameters are consistent with those reported in the BCN paper. For the rest of the table, estimation period is from 1987 to 2005. Column 3 lists the results of pooled regression, which do not deviate much from that using the shorter period of data. The fourth, fifth and sixth columns are the estimators considering individual effects, where the intercept terms vary across firms but do not stay constant over time. Therefore, the intercept terms are not shown in the table. Column 4 gives the estimation results applying demean method. There is a major difference in the AR parameter, i.e., β_1, between the results by this method and pooled regression. Pooled regression, which ignores individual effect, tends to bias parameters upwards, therefore the demean method shows that AR parameter is 0.392, much lower than 0.61 in the second column and 0.69 in the third column, both of which are estimated by pooled regression.

Column 5 provides the results estimated using the first difference of the variables. This method results in a negative autocorrelation in the AR term. The AR parameter shown in column 5 is negative and statistically significant at 0.01 level. The negative parameter is due to endogeneity caused by the first differences and thus it is not consistent. Conclusions drawn from the other parameters are generally consistent with that in the previous 3 columns, except for the depreciation and amortisation term (β_5). This term is no long statistically significant when using the first difference estimator. Results of using the Arellano-Bond estimator are reported in column 6, supporting the insignificance of depreciation and amortisation. The Arellano-Bond estimator applies the GMM method, which is implemented by assigning all independent variables as instrumental variables (IV). It is considered that the Arellano-Bond estimator would take account of the endogeneity brought on by taking the first difference for the variables. Therefore, the AR parameter in column 6 is positive in contrast with column 5. It is hard to make a sound conclusion based on the results. However, the results in Table 2 suggest that pooled regression, which is widely used in the extant literature, biases the AR parameter upwards and gives readers the false impression that the cash flows are very persistent. In addition, there is no clear conclusion about whether depreciation and amortisation are significant for the cash flow model.

3.3 Parameters estimation in the grey-box models

The grey-box model does not assume the parameters to follow a linear random process, but it attempts to capture the parameters' dynamics and heterogeneity by a deterministic function of some exogenous variable. The grey-box model can be written in the form of:

$$CF_{i,t+1} = \beta_{i,t,0} + \beta_{i,t,1}CF_{i,t} + \beta_{i,t,2}\Delta INV_{i,t} + \beta_{i,t,3}\Delta AP_{i,t} + \beta_{i,t,4}\Delta AR_{i,t}$$
$$+\beta_{i,t,5}DA_{i,t} + \beta_{i,t,6}OTHER_{i,t} + \varepsilon_{i,t+1} \quad (27)$$
$$\boldsymbol{\beta_{i,t}} = \mathbf{F}(z_t)$$

Each parameter is assumed to be a function of variable z and the function is captured by a Padé approximant. In the previous sections, we have shown that sales growth rate might have explanation power for the parameters' dynamics. Therefore, the lagged sales growth rate is a candidate for variable z:
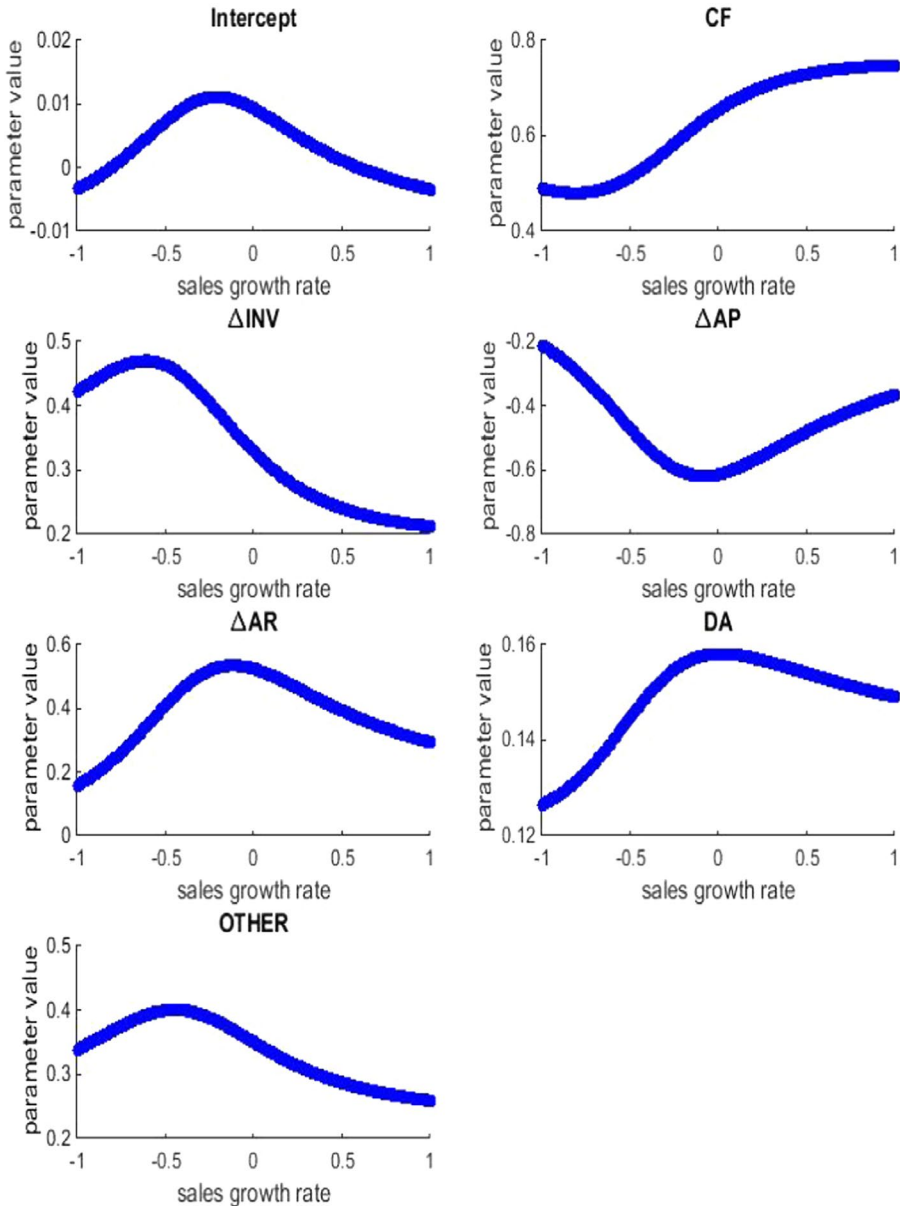
Fig. 7 The association of parameter values and sales growth rates by grey-box model. Figure 7 plots how the 7 parameters in Eq. (27) vary with the lagged sales growth rates. The growth rates take the values from −1 that means an extreme scenario of sales dropping to zero to 1 that means that sales double

$$\beta_{i,t} = \mathbf{F}(r_{t-1}) \tag{28}$$

Using the in-sample data, the coefficients in the Padé approximant are estimated by minimising the sum squared error of prediction errors in the model, i.e., Eq. (27). For
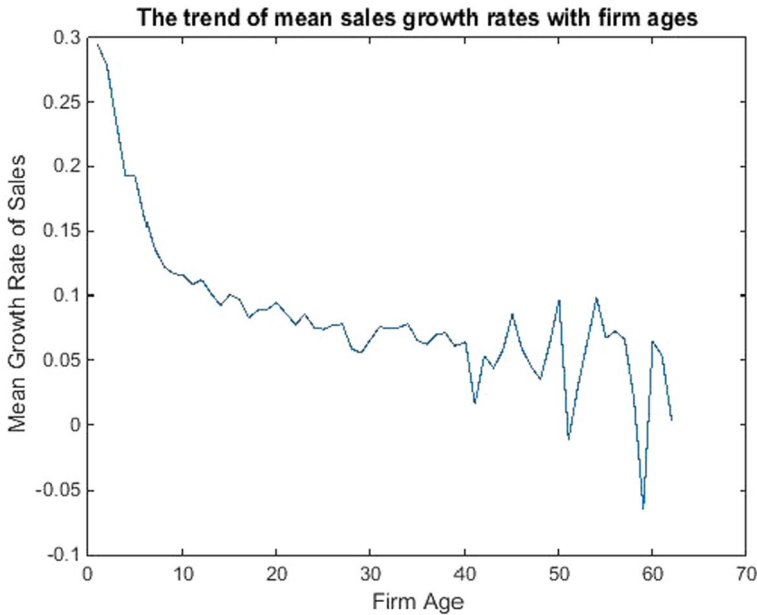
Fig. 8 The relationship between mean sales growth rate and firm age. Fig. 8 depicts the mean growth rate of sales for each firm age. Firm age is calculated as the number of years ahead of that firm's first observation in the sample

each parameter, there are 5 coefficients to be determined. These coefficients are used to numerically replicate the unknown functional form of **F**. The functional form of **F** can be shown in a graphic form. Figure 7 plots how the 7 parameters in Eq. (27) vary with the lagged sales growth rates. The growth rates take the values from -1 (i.e., sales drop to none) to 1 (i.e., sales double). Sales could grow more than double and have no upper limit in theory. However, it does not occur frequently. All seven parameters show nonlinear patterns. An interesting phenomenon in the charts is that as growth rates of sales get higher, the effects of the predictors tend to decline except the AR term. As a result, the distances between the AR parameter and the others get greater. When the growth rate approaches 0, the parameters on cash flow, changes in accounts payable and changes in accounts receivable converge in absolute values. This finding implies that for mature firms which have relatively lower sales growth rates, the gain of disaggregating earnings into components to predict cash flow becomes smaller than growing firms.

There is one drawback in selecting lagged sales growth rate in the cash flow prediction model: multiple periods ahead prediction requires to predict sales growth rates, which brings in more complexity. Firm age could be an alternative proxy for growth rates. Use it as an input variable of Eq. (28):

$$\beta_{i,t} = F(AGE_{i,t}) = \frac{a_0 + a_1 AGE_{i,t} + a_2 AGE_{i,t}^2}{1 + a_3 AGE_{i,t} + a_4 AGE_{i,t}^2} \tag{29}$$

Firms' growth rates tend to decline as time goes by, which implies a negative relation between firms' ages and their growth rates. This conjecture is supported by empirical
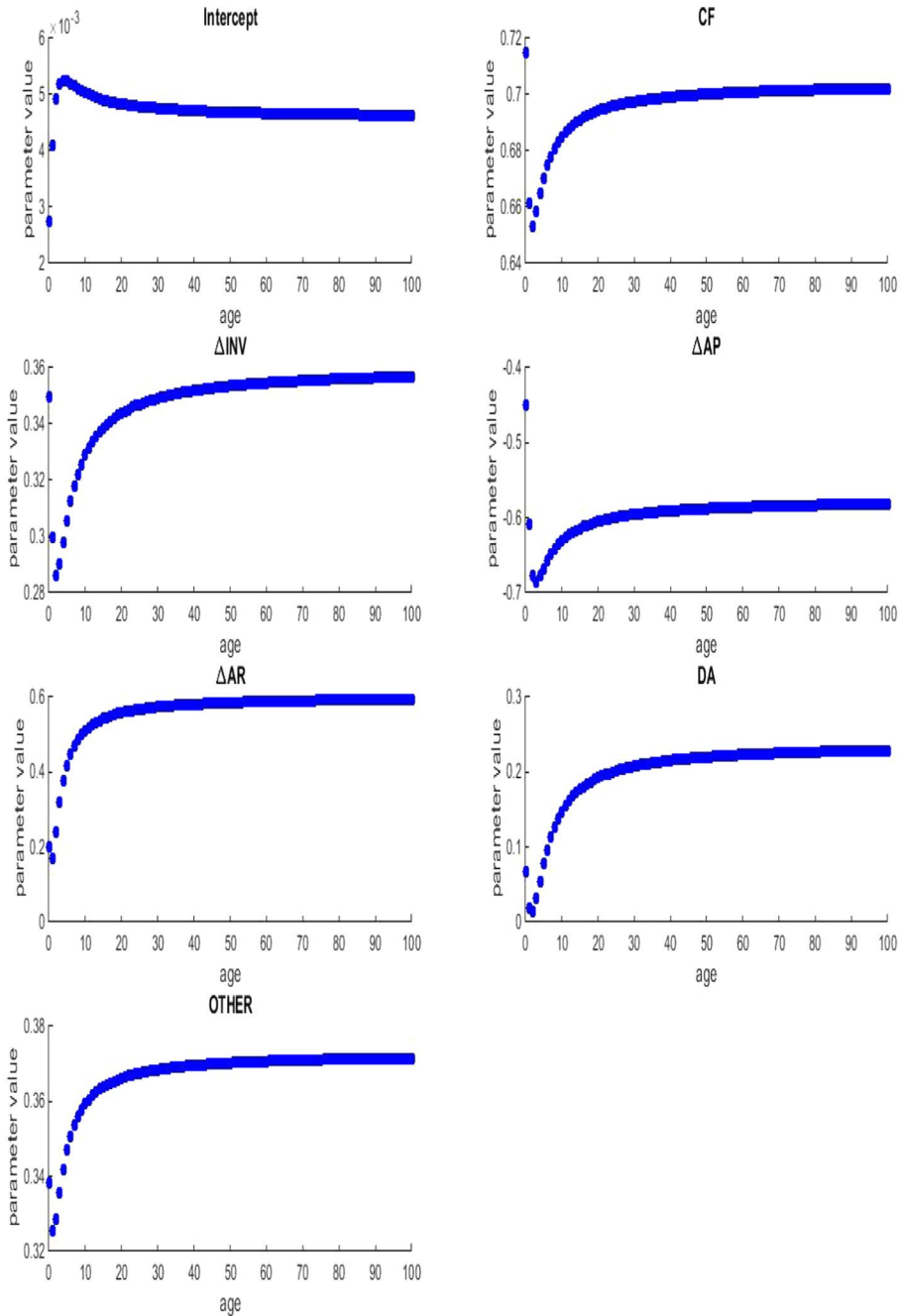
Fig. 9 The evolution of parameters values with increasing firm age by grey-box model (U.S. listed firms). Fig. 9 reports the evolution of parameters values with firm age (up to age of 100) by grey-box model

results. Using all sample data, mean growth rate of sales is calculated for each age, which is plotted in Fig. 8. Firm age is calculated as the number of years ahead of that firm's first observation in the sample because firm age is not available in the database. There is a clear declining trend of mean sales growth rates along with the firm ages. The growth rates gradually drop in the first 10 years. After 20 years, the growth rates remain above 5 percent. After 40 years, the growth rates become spiky, probably due to small sample size. After 60 years, the trend of growth is not known. Age seems to be an appropriate proxy for growth. Moreover, firm age is simpler and requires no prediction.

The functions of Eq. (29) are fitted in the same way as lagged growth rates, and the results are plotted in Fig. 9. The parameters do not change monotonically with age. They tend to reach their extreme values in the early ages and then approach some fixed levels after 20 to 30 years. It implies that for firms that are older than 20 years, the grey-box model may be no different from the simple pooled regression model in prediction cash flow.

3.4 In-sample fitness to data of different models

The above sections have presented the estimation results of various models. The performance of models could be examined by comparing their data fitting ability and also out-of-sample performance. This section briefly shows the in-sample results of each model. The in-sample fitness is examined based on the two measures, i.e. mean squared error and average rank.

In general, more complicated models and/or models with more parameters are considered to better fit the in-sample data but there is risk of over-fitting. The most complicated model of this study is the grey-box model. However, the linear panel models have more parameters than the grey-box model for taking account of individual effects. Nonetheless, grey-box model has the advantage in making predictions for firms whose individual effect is not easy to calculate–- consider a firm with only one observation for instance. The models for comparison are the random walk model (Model 1), the theoretical DKW model that says the prediction of future cash flow is current cash flow plus the changes in working capital terms (Model 2), the BCN model estimated by pooled regression (Model 3), the panel model that assumes homogeneous and constant parameters except the intercept term estimated using demean (Model 4), difference (Model 5) and Arellano-Bond estimators (Model 6), the grey-box model using sales growth rates (Model 7) and firm age (Model 8) as additional input variables.

The MSE and average ranks of each model are calculated and shown in Table 3 (Panel A). The results are based on 62,927 firm-year observations. The second row lists the resulting mean squared errors (MSE) and the numbers in the third row are the average ranks of each model. For both measures, a smaller number indicates better performance. The panel Models 4 and 6 have produced lowest MSE as they calculate individual effects for each firm. Model 5, however, has obtained a higher MSE than Model 4 and 6. Recall that the AR parameter estimated by Model 5 is negative due to the bias introduced by taking the first difference of variables, and the in-sample results suggest that the biased estimation might not make proper predictions. Despite that the panel Models 4 and 6 have lower MSE, their average ranks are among the highest tier, only lower than Model 5. Therefore,

Table 3 The in-sample fitness and out-of-sample cash flow prediction performance

Models	1	2	3	4	5	6	7	8
Panel A: The in-sample fitness of various cash flow models (U.S. firms)								
MSE	0.0101	0.0125	0.0076	0.0053	0.0078	0.0057	0.0075	0.0075
Av. Rank	4.9777	5.5063	5.6012	5.8425	5.9641	5.8913	5.2823	4.7504
Panel B: The out-of-sample one-period-ahead cash flow prediction performance of various models (U.S. firms)								
MSE	0.015	0.015	0.0113	0.014	0.0376	0.0198	**0.0111**	**0.0109**
Av. Rank	**4.9476**	5.4358	5.8546	6.1324	6.2193	5.8686	5.3747	**4.5557**
Panel C: The out-of-sample multi-period-ahead cash flow prediction performance of various models (U.S. firms)								

Models		1	2	3	4	5	6	7	8
2006~ 2012	MSE	0.0205	0.0208	**0.0190**	0.0217		0.0262	0.0239	**0.0177**
	Av. Rank	**4.3647**	4.3991	4.5984	4.6508		4.3927	4.4868	**4.2927**
2007~ 2012	MSE	0.0227	0.0227	**0.0214**	0.0244		0.0278	0.0268	**0.0199**
	Av. Rank	4.4292	**4.3701**	4.5719	4.6276		4.377	4.4751	**4.3338**

Bold numbers indicate two models that perform best in that criterion. Model 1: random walk model; Model 2: theoretical DKW model; Model 3: BCN model (pooled regression), Model 4: BCN model (demean); Model 5: BCN model (difference); Model 6: BCN model (Arellano-Bond estimators); Model 7: grey-box model using sales growth rates; Model 8: grey-box model using firm age. Bold numbers indicate two models that perform best in that criterion

the panel models may be inferior in data fitting. Another point is that Model 6 using the Arellano-Bond estimator that is considered consistent does not outperform Model 4. Model 3 assumes total homogeneity even for the intercept term and is estimated simply by pooled regression. Although it has higher MSE than the panel models, the lower average rank of Model 3 indicates more general description of the cash flow process, which is inconsistent with the expectation from an econometric perspective because the results estimated by pooled regression without considering individual effects are likely to be biased. Model 1 and 2 predict cash flow in a naive way and are thus selected to be benchmark models. Their MSEs are relatively large compared with the others but their average ranks are lower than Models 3, 4, 5, and 6. It is noteworthy that Model 2 has poorer in-sample data fitting than Model 1. It implies that including accrual terms may not make better prediction than the simple random walk model. The random walk model performs very well based on average rank criterion, second only to Model 8. Model 7 and Model 8 are grey-box models. They fit the data very well. The MSEs of these two models are comparable with that of pooled regression, lower than the models with individual effect. However, the grey-box model with firm age as the input variable has the lowest average rank of all 8 models. Model 7 has the third lowest average rank, only higher than Model 1 and 8.

In summary, the grey-box models generally are the best form of model to fit the data in-sample comparing with other options. Firm age as the black-box input variable seems to work better than sales growth rates. The random walk model, though producing high prediction error, may better describes the cash flow process than some parameterised models. To gain deeper knowledge of the models' performance, out-of-sample test is conducted.

3.5 Out-of-sample prediction performance of different models

As shown in Table 3 (Panel A), the grey-box model provides promising performance. Not only the MSE of the two grey-box models are as low as that of pooled regression, their average ranks are also among the best models. For practical prediction, out-of-sample examination is more important. The models perform well in-sample may not necessarily extend their superiority to the out-of-sample period. In this particular application, i.e. cash flow prediction, one-period-ahead and multi-period-ahead predictions are both important and useful, therefore, this section will test the two types of predictions separately.

3.5.1 One-period-ahead prediction

The data after 2005 is used for out-of-sample test purpose. For panel models, i.e. Model 4, 5, and 6, the application of them in the out-of-sample period required that the individual intercept values for the target firm are available, which exclude observations of firms that do not appear before 2005.

The MSE and average rank for the whole out-of-sample data are calculated and listed in Table 3 (Panel B). The calculations are based on 17,965 firm-year observations. For each criterion, the best two models are labelled by bold numbers. The panel Model 4, 5, and 6 have higher in-sample MSE. They also have the poorest performance of all the 8 models in the one-period-ahead forecast based on average rank. Model 5 has the poorest prediction. Model 4 has the least MSE among the three panel models but shows the highest average rank of all eight models. Model 6 that applies Arellano-Bond estimator does not appear to be a good predictive model according to both measures. Grey-box models prove their power in this comparison, especially for Model 8 which has both the lowest MSE and the lowest average rank. Model 7 has the second lowest MSE and its average rank is in the middle position of the eight models. The model that has the second-best average rank is the benchmark Model 1, which has the same level of MSE as Model 2. Model 3 has provided medium performance, better than the panel models but worse than the grey-box models.

3.5.2 Multi-period-ahead prediction

Predictions beyond one period are made recursively by extending the predictions of cash flow to all the predictive variables and thus use the predicted variables to make further periods' predictions. Therefore, the workload of making multi-period-ahead prediction increases as there are 6 predictors in the model which are to be predicted into the future for longer-term use.

In this study, the predictive variables are predicted using the same way, i.e., the same independent variables, model structure, and estimation methods. Model 1 and 2 are exceptions as their multi-period-ahead predictions are simply the last in-sample observation of cash flow or cash flow plus changes in working capital accruals. Grey-box model 7 is not suitable for the multi-period task because the input variable, i.e., sales growth rates, requires more prediction, which adds extra complexity to the model. Therefore, only grey-box model 8 can be utilised and it replaces model 7. There are 7 models in total to compete in multi-period-ahead prediction in out-of-sample setting.

The parameters used for prediction are the in-sample estimated results and the predictors take their initial values in year 2005. Table 3 (Panel C) reports the models' performance in the multi-period-ahead setting. The predictions are compared with the sample data from 2006 to 2012. Therefore, the models' performance in the predictions of one to up to seven years ahead could be examined. The results favour the grey-box model as it outperforms all other models based on both criterion. The panel models generate higher MSE than simpler models 1, 2, and 3 and the more complicated grey-box model. It seems that considering individual effects does not help make prediction in practice, which is a counterintuitive conclusion. Model 3 has the second lowest MSE, and Model 1 has the second lowest average rank. They are simple enough but provide good prediction results in practice.

Another comparison is made by excluding the observations in 2006 (to focus on multi-period-ahead results), and the results are shown in the 3rd and 4th rows in Table 3 (Panel C). Grey-box model still performs best in both criteria. The conclusions in general do not change much for the sub-period only except that Model 2 outperform Model 1 when data in 2006 is excluded, which suggests that the DKW assertion that earnings make better prediction for cash flows than cash flows per se is more descriptive in the long run.

3.6 Robustness Check

To check the robustness of the proposed model, two more datasets from different economy, i.e. U.K. and China, are examined. In both markets, we find that grey-box models could provide impressive and encouraging predictions for future cash flow, especially for the shorter future. Simple pooled regression model also offers competitive performance and accurate predictions. Detailed discussions are presented in Appendix. The empirical results shown in the Appendix are supportive to the model's robustness in different economic environment.

4 Conclusion

This paper proposes a cash flow forecast model which captures the nonlinearity and dynamics of the cash flow process. Our model incorporates heterogeneity across firms in cash flow prediction as we allow for a panel data setting which has both time-series and cross-sectional dimensions. The nonlinearity is captured numerically by a black-box model, and the linear form is captured by a white-box model. Therefore, our model is considered as a grey-box model, and it achieves a good balance between prediction accuracy and model complexity. Moreover, to incorporate the dynamics of cash flows, the parameters of the panel data models are treated as time-varying. No linearity restrictions are imposed on these time-varying parameters.

To evaluate the performance of our new model, we conduct an empirical analysis of the U.S. data and compare various models' forecasting performance by using multiple criteria. Both in-sample and out-of-sample performances are examined, the latter of which includes not only the one-period-ahead prediction but also multi-period-ahead predictions. The models used in the study are classified into two categories, i.e., panel data models that consider individual effects and grey-box models. In general, our empirical results show that the proposed grey-box model consistently outperforms other models both in-sample and out-of-sample, especially in multi-period-ahead predictions. In particular, we show that our model performs better not only before the global financial crisis, but also during the crisis period.

This paper helps improve our understanding of the cash flow generating process. It suggests the importance of considering the nonlinearity and dynamics of model parameters when predicting cash flows. The empirical results show that sales growth rate and firm age may be responsible for the variability of model parameters and the change in the importance of regressors. In firms' early stages, their growth rates are high, the predictive power of lagged cash flow is higher than that of accruals. As firms become mature, firms' growth rates decline. The predictive power of lagged cash flow and accrual terms tend to converge. Our grey-box mode benefits practitioners (e.g., investors, analysts, and auditors) and researchers in the following contexts: (1) to better forecast future cash flows, which are key inputs of valuation models; (2) to provide an example of retaining a simple model structure but allowing for dynamics and nonlinearity of parameters.

Appendix

U.K. market

Data for U.K. listed firms are obtained from the Datastream. The accounting information is collected from annual reports which span the period from 1989 to 2015. All sampled firms are traded on the London Stock Exchange. Similar to the U.S. sample, all financial firms in sectors of banking, investment or insurance are excluded. There are initially 1272 firms that have available observations for cash flow from operating activities. The age of U.K. firms is calculated from the date of incorporation of that particular firm up to the year of data availability from the Financial Analysis Made Easy (FAME) database. After exclusion of firms that have unavailable or inconsistent ticker symbols recorded in the two databases, firm ages could be calculated for 1009 firms. A little different from the U.S. data, cash flow prediction for U.K. firms uses depreciation and amortisation as two separate predictive variables, mainly due to variable availability for the different databases used in both studies.

Similar to the U.S. analysis, models placing different assumptions or specifications on the parameters are selected, which include the random walk model (Model 1), the theoretical DKW model that suggests the prediction of future cash flow is current cash flow plus the changes in working capital terms (Model 2), the BCN model estimated by pooled regression (Model 3), the panel model that assumes homogeneous and constant parameters

Table 4 The in-sample fitness of various cash flow prediction models (U.K. listed firms)

Models	1	2	3	4	5	6	7
MSE	0.0069	0.0076	0.0054	0.0037	0.0038	0.0052	0.0052
Av. Rank	4.1445	4.5199	4.7268	4.8838	4.9644	4.3383	3.6943

Note: Model 1: random walk model;

Model 2: theoretical DKW model;

Model 3: BCN model (pooled regression),

Model 4: BCN model (demean);

Model 5: BCN model (Arellano-Bond estimators);

Model 6: grey-box model using sales growth rates;

Model 7: grey-box model using firm age

Table 5 The out-of-sample (one period) performance of various cash flow prediction models (U.K. listed firms)

Models	1	2	3	4	5	6	7
MSE	0.0136	0.0117	0.0101	0.0125	0.0128	0.0101	0.0102
Av. Rank	3.9406	4.4810	4.8730	4.8873	4.8256	4.3574	3.7708
Models' ranks in the two measures of performance							
Rank(MSE)	7	4	1	5	6	2	3
Rank(Av. Rank)	2	4	6	7	5	3	1

Note: Model 1: random walk model;

Model 2: theoretical DKW model;

Model 3: BCN model (pooled regression),

Model 4: BCN model (demean);

Model 5: BCN model (Arellano-Bond estimators);

Model 6: grey-box model using sales growth rates;

Model 7: grey-box model using firm age

Table 6 The out-of-sample (multi-period) performance of various models (U.K. listed firms)

	Models	1	2	3	4	5	6
2008	MSE	0.0154	0.0162	**0.0132**	0.0163	0.0164	**0.0134**
	Av. Rank	**3.5361**	3.8578	4.2013	4.2407	4.1028	**3.8643**
2009	MSE	0.0202	0.0218	**0.0165**	0.0208	0.0206	**0.0164**
	Av. Rank	**3.8206**	4.0175	4.0175	4.1772	4.0153	**3.9628**
2010	MSE	0.0211	0.0242	**0.0180**	0.0227	0.0224	**0.0179**
	Av. Rank	**3.8568**	3.9802	3.9846	**3.9515**	3.9537	4.0925
2011	MSE	0.0220	0.0257	**0.0190**	0.0256	0.0253	**0.0187**
	Av. Rank	**3.8838**	3.9795	**3.9499**	3.9704	3.9522	4.0752
2012	MSE	0.0250	0.0273	**0.0214**	0.0269	0.0267	**0.0212**
	Av. Rank	**3.8333**	4.1765	4.0515	3.8995	**3.7941**	4.0270
2013	MSE	0.0205	0.0218	**0.0195**	0.0203	0.0204	**0.0199**
	Av. Rank	**3.8385**	4.0708	**3.8669**	3.8725	3.9972	4.1416
2014	MSE	0.0221	0.0252	**0.0205**	0.0245	0.0243	**0.0209**
	Av. Rank	**3.7726**	4.1204	4.0368	3.9097	**3.8562**	4.0903

Bold numbers indicate two models that perform best in that criterion

Model 1: random walk model;

Model 2: theoretical DKW model;

Model 3: BCN model (pooled regression),

Model 4: BCN model (demean);

Model 5: BCN model (Arellano-Bond estimators);

Model 6: grey-box model using firm age

except the intercept term estimated using demean (Model 4) and Arellano-Bond estimators (Model 5), the grey-box model using sales growth rates (Model 6) and firm age (Model 7) as additional input variables. Observations before 2006 are used for in-sample estimation and observations after 2006 are used for out-of-sample performance comparison.

The in-sample fitness of various models is compared by two criteria, i.e. mean squared error (MSE) and average rank of each model. A smaller number indicates a model's better performance in data fitting. The results of the 7 models are shown in Table 4. Model 1 fits the in-sample data better than Model 2 in both criteria, which is against the assertion made in Dechow et al. (1998) that cash flow plus changes in working capital terms better predict future cash flow than cash flow alone. The random walk model empirically generates lower prediction errors and its average rank also suggests that the random walk process describes the data well even for individual observations. Although Model 3 has lower MSE than Model 1 and 2 as estimated by OLS method, the average rank of Model 3 is higher. Model 4 and 5 have similar performance, the former performing marginally better than the latter. Both of them, taking account of individual effects of firms, have much lower MSE than other models, but their average ranks are among the highest of all. The two grey-box models have quite balanced performance. Their MSE is comparable with the simple linear regression, i.e. Model 3, and they both have lower average ranks. Model 6 that applies sales growth rates as the black-box input has the third best average rank, only worse than Model 1 and Model 7. Model 7 that uses firm age has obtained the lowest average rank among all models.

Observations from 2007 to 2015 are used for out-of-sample prediction comparison. The one-period-ahead predictions made by the 7 models are compared in Table 5. Although Model 1 provides a lower in-sample MSE than Model 2, it has a higher MSE than Model 2 in the out-of-sample period. It however still has lower average rank than most of the other models. The lowest MSE is obtained by Model 3, i.e. the pooled regression model, but it has the worst performance in individual predictions as implied by its highest average rank. Model 4 and 5 have similar performance, none of which is impressive in either criterion. Model 6 and 7, i.e. the grey-box models, have good, balanced performance in the in-sample period. Their MSE is as low as that generated by Model 3 and the average ranks of the grey-box models are among the best of all 7 models. The lower half of Table 5 shows the ranks of each model according to their performance by each criterion. Model 1 and 3, although having leading performances in one criterion (Model 1 in average rank and Model 3 in MSE), have the poorest performance in the other criterion. Only the grey-box models could perform well for both criteria. Therefore, it can be concluded that their predictions are accurate and generally applicable for one-period-ahead use.

Table 6 summarises the performance of the models in predicting cash flows of multiple periods ahead. Model 6 that applies sales growth rates as additional input variable is not applicable in this test because future sales growth rates are unknown *ex ante*. Therefore, the multi-period competition is among the other 6 models. The in-sample period ends in 2006 and therefore the predictive variables take their values in 2006 if available. The predictions made in 2007 are actually one-period-ahead predictions and thus not included in Table 6 Year 2015 is not listed either for there are only 3 observations. Bold numbers indicate the two models performing best for that criterion, which appear 9 times for Model 3 and Model 6. These two models outperform the other models in accuracy as they always provide the least MSE. Model 1 has bold numbers shown 7 times, all in the average rank criterion. Model 2 has shown none. Model 1 is also better than Model 2 by the criterion of

Table 7 The in-sample fitness of various cash flow prediction models (China data)

Models	1	2	3	4	5	6	7	8	9	10
MSE	0.0086	0.0080	0.0053	0.0041	0.0042	0.0052	0.0041	0.0042	0.0052	0.0055
Av. Rank	5.2605	6.0713	6.2420	6.3027	6.0512	6.1217	6.3869	6.4778	5.9825	4.7520

Note: Model 1: random walk model;

Model 2: theoretical DKW model;

Model 3: BCN model (pooled regression),

Model 4: BCN model (demean);

Model 5: BCN model (Arellano-Bond estimators);

Model 6: CH model (pooled regression),

Model 7: CH model (demean);

Model 8: CH model (Arellano-Bond estimators);

Model 9: grey-box model using sales growth rates;

Model 10: grey-box model using firm age

MSE in every year of the out-of-sample period. The evidence is consistent with the in-sample result and is against the DKW assertion. Model 4 and 5 do not provide impressive predictions, which is consistent with the results of one-period-ahead predictions. Comparing Model 3 and 6, the former performs better in predicting cash flows of further future. In the prediction for 2011 and 2013 (corresponding to 5 and 7 years ahead respectively), Model 3 has been one of the two best models by both criteria while Model 6 has such achievement for 2008 and 2009 (2 and 3 years ahead).

Chinese market

The financial data of the sample firms is obtained from the RESSET database. The sample firms are limited to Chinese A stock firms and non-bank firms. The number of sample firms is 3460. Annual cash flow data is collected from 1997 to 2015. Thus, there are 36,047 net operating cash flow observations in total, from which the trend of cash flows' evolution could be examined. These cash flow observations are divided by firms' contemporary total assets, forming cash to asset ratio and the values of the ratio of the observations are sorted into sub-samples according to the firms' ages in the particular observation year.

The RESSET database provides the direct method statement of cash flow, hence with the additional data it could be examined whether the disclosed cash flow components could improve the predictability of future cash flows. The disaggregated cash flow model will be in the form of:

$$CF_{t+1} = \beta_0 + \beta_1 CF\_Crec_t + \beta_2 CF\_Cpaid_t + \beta_3 CF\_NCrec_t + \beta_4 CF\_NCpaid_t$$
$$+ \beta_5 \Delta INV_t + \beta_6 \Delta AP_t + \beta_7 \Delta AR_t + \beta_8 DEP_t + \beta_9 AMORT_t + \beta_{10} OTHER_t + \varepsilon_{t+1}$$
$$\text{(A.1)}$$

where *CF_Crec* denotes cash received from the sales of goods and rendering of services, *CF_Cpaid* denotes cash paid for goods and services, *CF_NCrec* denotes other cash receipt calculated as sub-total cash inflows from operating activities minus *CF_Crec* and

CF_NCpaid denotes other cash paid out calculated as sub-total cash inflows from operating activities minus *CF_Cpaid*. Thus, net operating cash flow is disaggregated into four components. The models examined in this section are the random walk model (Model 1), the theoretical DKW model (Model 2), the BCN model estimated by pooled regression (Model 3), demean method (Model 4) and GMM method (Model 5), equation (A.1) estimated by pooled regression pooled regression (Model 6), demean method (Model 7) and the GMM method (Model 8), the grey-box model with sales growth rate (Model 9) and firm ages (Model 10) separately as additional input variable to the black-box. Models 6, 7 and 8 have more predictive variables than the other models. Their performance can be directly compared with Model 3, 4 and 5 respectively to see whether the disaggregation of net cash flow will benefit the forecast.

The in-sample fitness of all the models are compared in Table 7. The two measures, i.e. MSE and average rank, are used to compare the models' performance. The models that generate the lowest MSE are Models 4 and 7, followed by Models 5 and 8. It is consistent with expectation because they estimated intercept terms for individual firms. Apart from them, the lowest MSE is generated by Model 6, the cash disaggregation model, and Model 9, the grey-box model using sales growth rates. Comparing Models 6 and 3, it can be seen that the disaggregation of cash flow results in better fitness of the in-sample data, as both measures of Model 6 are superior to Model 3. Model 9 has not only low MSE but also low average rank. Its average rank is only higher than Model 1 and Model 10. The comparison suggests that grey-box models in general better fit the in-sample data.

The data from 2007 to 2015 is used of out-of-sample comparison. The one-period-ahead results are summarised in Table 8. The MSE and average ranks are calculated on 10,104 observations. Model 6 has the lowest MSE. Besides, it outperforms Model 3 in both criteria. Similarly, Models 7 and 8 respectively outperform Models 4 and 5. The comparison shows the incremental power of using disaggregated components of net cash flow as predictors. Model 10 has the lowest average rank, followed by the random walk model and then Model 9. Model 6 has a moderate average rank although it performs best according to the MSE criterion. Model 1's low average rank suggests that it could well describe the general pattern of the sample firms' cash flow dynamics. However, it has the largest MSE of all

Table 8 The out-of-sample (one period) performance of various cash flow prediction models (China data)

Models	1	2	3	4	5	6	7	8	9	10
MSE	0.0283	0.0260	0.0231	0.0270	0.0274	0.0224	0.0261	0.0261	0.0224	0.0233
Av. Rank	4.9053	5.9974	6.3955	6.5928	6.4578	6.3325	6.3822	6.4176	5.8698	4.3095
Rank(MSE)	10	5	3	8	9	1	6	7	2	4
Rank(Av. Rank)	2	4	7	10	9	5	6	8	3	1

Note: Model 1: random walk model;

Model 2: theoretical DKW model;

Model 3: BCN model (pooled regression),

Model 4: BCN model (demean);

Model 5: BCN model (Arellano-Bond estimators);

Model 6: CH model (pooled regression),

Model 7: CH model (demean);

Model 8: CH model (Arellano-Bond estimators);

Model 9: grey-box model using sales growth rates;

Model 10: grey-box model using firm age

Table 9 The out-of-sample (multi-period) performance of various cash flow prediction models (China data) with adjustment made to the grey-box model

Models	1	2	3	4	5	6	7	8	9
MSE	0.0306	0.0300	0.0294	0.0297	0.0301	0.0291	0.0296	0.0300	0.0293
Av. Rank	4.7684	5.4712	5.7305	6.0113	5.5127	5.5742	5.9149	5.6597	4.7772
Rank(MSE)	9	7	3	5	8	1	4	6	2
Rank(Av. Rank)	1	3	7	9	4	5	8	6	2

Note: 7146 observations; year 2007 excluded;

Model 1: random walk model;

Model 2: theoretical DKW model;

Model 3: BCN model (pooled regression),

Model 4: BCN model (demean);

Model 5: BCN model (Arellano-Bond estimators);

Model 6: CH model (pooled regression),

Model 7: CH model (demean);

Model 8: CH model (Arellano-Bond estimators);

Model 9: Combinative model of grey-box model using firm age and Model 3

the models. Therefore, no single model could outperform others in both criteria. The Grey-box models have encouraging performance as they have low MSE and low average rank as well. Although Model 9 does not outperform Model 6 in the MSE measure, it is worth noting that Model 6 has more variables than Model 9 and hence more detailed information is exploited by Model 6. Model 9 still has an MSE which is very close to that of Model 6 despite leaving fewer predictors. Model 1 is the simplest random walk model, which is usually used as benchmark model. In this case, Model 10 is the only model that beats Model 1 in both criteria. If Model 2, the theoretical DKW model, is used as benchmark, there will be only two models that could beat it in the two measures, i.e. Model 9 and Model 10.

There is an adjustment made to the grey-box model in multi-period prediction, which is to predict the variables apart from cash flow with a simpler model, e.g. Model 3, while keep using the grey-box model to predict cash flow variable (a combinative model). The result in applying the combinative model is listed in Table 9. The grey-box model using sales growth rate is excluded. The data used is from 2008 to 2015 and year 2007 is excluded. The MSE and average ranks of the models are calculated over 7146 observations. The combinative model performs well in terms of both measures but does not rank the first in either. Model 6 has the lowest MSE and Model 1 has the lowest average rank. As the same as in the one-period prediction test, none of the models could outperform the others in both criteria, but Model 9 provides the most balanced performance. Although it has higher MSE than Model 6, fewer variables are adopted by Model 9 and thus Model 9 has an advantage over Model 6 that it can apply to the situation where direct method disclosure of cash flow is unavailable. On the other hand, the differences in the MSE generated by the models are small and negligible. For instance, the MSE of Model 6 is 0.0291, merely 1% smaller than that of Model 9 (0.0293) and 5% smaller than that of Model 1 (0.0306). MSE could be used to measure the predictability of cash flows. Higher MSE usually indicates more difficult predictions. Comparing the MSE in multi-period performance of various dataset in this paper, it can be seen that the MSE of China data is generally

higher than that of U.S. and U.K. data, which implies that firms in the former samples have less predictability in future cash flows.

References

Albrecht WS, Lookabill LL, McKeown JC (1977) The time-series properties of annual earnings. J Account Res 15(2):226–244

Anthony JH, Ramesh K (1992) Association between accounting performance measures and stock prices: a test of the life cycle hypothesis. J Account Econ 15(2–3):203–227

Arellano M, Bond S (1991) Some test of specification for panel data: monte Carlo evidence and an application to employment equations. Rev Econ Stud 1:277–297

Ball R, Nikolaev VV (2020) FASB was right: Earnings beat cash flows when predicting future cash flows. Chicago Booth Research Paper, 20–23

Ball R, Watts R (1972) Some time series properties of accounting income. J Financ 27(3):663–681

Barth ME, Cram DP, Nelson KK (2001) Accruals and the prediction of future cash flows. Account Rev 76(1):27–58

Beaver WH (1970) The time series behavior of earnings. J Account Res 8:62–99

Bernard VL, Stober TL (1989) The nature and amount of information in cash flows and accruals. Account Rev 64(4):624–652

Bostwick ED, Krieger K, Lambert SL (2016) Relevance of goodwill impairments to cash flow prediction and forecasting. J Acc Audit Financ 31(3):339–364

Box GE, Jenkins GM, Reinsel GC (1970) Time series analysis: forecasting and control. Holden-Day, San Francisco

Cameron AC, Trivedi PK (2005) Microeconometrics: methods and applications. Cambridge University Press, New York, NY

Cheng CA, Hollie D (1996) The persistence of cash flow components into future cash flows. Working paper, available at SSRN 541293.

Cheng CA, Hollie D (2008) Do core and non-core cash flows from operations persist differentially in predicting future cash flows? Rev Quant Financ Acc 31(1):29–53

Cybenko G (1989) Approximation by superpositions of a sigmoidal function. Math Control Signals Systems 2(4):303–314

Dechow PM, Kothari SP, Watts L (1998) The relation between earnings and cash flows. J Account Econ 25(2):133–168

Dechow PM, Richardson SA, Tuna I (2003) Why are earnings kinky? an examination of the earnings management explanation. Rev Acc Stud 8(2):355–384

DeFond ML, Hung M (2003) An empirical analysis of analysts' cash flow forecasts. J Account Econ 35(1):73–100

Dickinson V (2011) Cash flow patterns as a proxy for firm life cycle. Account Rev 86(6):1969–1994

Dopuch N, Watts R (1972) Using time-series models to assess the significance of accounting changes. J Account Res 10(1):180–194

Drtina RE, Largay JA III (1985) Pitfalls in calculating cash flow from operations. Account Rev 60(2):314–326

Farshadfar S, Monem R (2011) Discretionary accruals and the predictive ability of earnings in the forecast of future cash flows: evidence from Australia. Corp Ownersh Control 9(1):597–608

Farshadfar S, Monem R (2013a) The usefulness of operating cash flow and accrual components in improving the predictive ability of earnings: a re-examination and extension. Accounting & Finance 53(4):1061–1082

Farshadfar S, Monem R (2013b) Further evidence on the usefulness of direct method cash flow components for forecasting future cash flows. Int J Account 48(1):111–133

Finger CA (1994) The ability of earnings to predict future earnings and cash flow. J Account Res 32(2):210–223

Fuller RJ, Huberts LC, Levinson MJ (1992) It's not higgledy-piggledy growth! J Portfolio Manag 18(2):38–45

Gordon EA, Henry E, Jorgensen BN, Linthicum CL (2017) Flexibility in cash-flow classification under IFRS: determinants and consequences. Rev Acc Stud 22(2):839–872

Greenberg RR, Johnson GL, Ramesh K (1986) Earnings versus cash flow as a predictor of future cash flow measures. J Acc Audit Financ 1(4):266–277

Hales J, Orpurt SF (2013) A review of academic research on the reporting of cash flows from operations. Account Horiz 27(3):539–578

Hovakimian G (2009) Determinants of investment cash flow sensitivity. Financ Manage 38(1):161–183

Jones JJ (1991) Earnings management during import relief investigations. J Account Res 29(2):193–228

Kaplan SN, Ruback RS (1995) The valuation of cash flow forecasts: An empirical analysis. J Financ 50(4):1059–1093

Kim M, Kross W (2005) The ability of earnings to predict future operating cash flows has been increasing—not decreasing. J Account Res 43(5):753–780

Kryzanowski L, Mohsni S (2013) Growth of aggregate corporate earnings and cash-flows: Persistence and determinants. Int Rev Econ Financ 25(1):13–23

Lee CWJ, Chen C (1990) Structural changes and the forecasting of quarterly accounting earnings in the utility industry. J Account Econ 13(2):93–122

Little IM (1962) Higgledy piggledy growth. Bull Oxf Univ Inst Econ Stat 24(4):387–412

Lookabill LL (1976) Some additional evidence on the time series properties of accounting earnings. Account Rev 51(4):724–738

Lorek KS, Willinger GL (2007) The contextual natural of the predictive power of statistically-vased quarterly earnings models. Rev Quant Financ Acc 28:1–22

Lorek KS, Willinger GL (2008) Time-series properties and predictive ability of quarterly cash flows. Adv Account 24(1):65–71

Lorek KS, Willinger GL (2009) New evidence pertaining to the prediction of operating cash flows. Rev Quant Financ Acc 32(1):1–15

Lorek KS, Willinger GL (2010) Time series versus cross-sectionally derived predictions of future cash flows. Adv Account 26(1):29–36

Manegold JG (1981) Time-series properties of earnings: a comparison of extrapolative and component models. J Account Res 19(2):360–373

McInnis J, Collins DW (2011) The effect of cash flow forecasts on accrual quality

Nallareddy S, Sethuraman M, Venkatachalam M (2020) Changes in accrual properties and operating environment: implications for cash flow predictability. J Account Econ 69(2–3):101313

Orpurt SF, Zang Y (2009) Do direct cash flow disclosures help predict future operating cash flows and earnings? Account Rev 84(3):893–935

Salamon GL, Smith ED (1977) Additional evidence on the time series properties of reported earnings per share: comment. J Financ 32(5):1795–1801

Sims CA (1980) Macroeconomics and reality. Econometrica 48(1):1–48

Tan KC, Li Y (2002) Grey-box model identification via evolutionary computing. Control Eng Pract 10(7):673–684

Watts RL, Leftwich RW (1977) The time series of annual accounting earnings. J Account Res 15(2):253–271

Wilson GP (1987) The incremental information content of the accrual and funds components of earnings after controlling for earnings. Accounting Review 62(2):293–322