Journal of Accounting and Economics 53 (2012) 504–526

Contents lists available at SciVerse ScienceDirect

Journal of Accounting and Economics

journal homepage: www.elsevier.com/locate/jae

The implied cost of capital: A new approach ☆

Kewei Hou [a,*], Mathijs A. van Dijk [b], Yinglei Zhang [c]

[a] Fisher College of Business, The Ohio State University, 2100 Neil Avenue, Columbus, OH 43210, United States
[b] Rotterdam School of Management, Erasmus University, 3000 DR Rotterdam, The Netherlands
[c] School of Accountancy, Chinese University of Hong Kong, Shatin, N.T., Hong Kong

ARTICLE INFO

Article history:
Received 24 February 2010
Received in revised form
30 November 2011
Accepted 9 December 2011
Available online 21 December 2011

JEL classification:
G12
G14
G29
G31
G32
M40
M41

Keywords:
Cross-sectional earnings model
Earnings forecasts
Expected returns
Implied cost of capital
Asset pricing tests

ABSTRACT

We use earnings forecasts from a cross-sectional model to proxy for cash flow expectations and estimate the implied cost of capital (ICC) for a large sample of firms over 1968–2008. The earnings forecasts generated by the cross-sectional model are superior to analysts' forecasts in terms of coverage, forecast bias, and earnings response coefficient. Moreover, the model-based ICC is a more reliable proxy for expected returns than the ICC based on analysts' forecasts. We present evidence on the cross-sectional relation between firm-level characteristics and ex ante expected returns using the model-based ICC.

1. Introduction

Estimating a firm's expected stock return (or cost of equity capital) is essential for studying the relation between firm-level (risk) characteristics and expected returns—a central theme in finance and capital markets research in accounting. Expected returns also play a key role in firm valuation, capital budgeting, and other corporate finance settings, and are

☆ We thank John Core (editor), Mozaffar Khan (referee), Gary Biddle, Zhihong Chen, Patricia Dechow, Peter Easton, John Griffin, Zhaoyang Gu, Hao Jiang, Bin Ke, Charles Lee, Clive Lennox, Christian Leuz, Roger Loh, Jim Ohlson, Chul Park, Lubos Pástor, Gordon Phillips, Scott Richardson, K.R. Subramanyam, Siew Hong Teoh, Huai Zhang and seminar participants at Erasmus University, Georgia Tech, Hong Kong University, Hong Kong University of Science and Technology, Limperg Institute, Nanyang Technological University, Singapore Management University, Tsinghua University, University of California at Irvine, University of California at Los Angeles, University of Southern California, the 2010 Western Finance Association Meetings, the 2010 Financial Management Association Meetings, the 2010 Asian Financial Association Conference, the 13th Conference of the Swiss Society for Financial Market Research, the Autumn 2010 Inquire U.K. Seminar, and the 2010 State Street Global Markets European Quantitative Forum for helpful comments and suggestions. We thank Karen Lin and Chunquan Zhou for research assistance. We are grateful to Inquire UK and the Research Grants Council (RGC) of Hong Kong for funding support of this project.
* Corresponding author. Tel.: +1 614 292 0552; fax: +1 614 292 2418.
E-mail address: hou.28@osu.edu (K. Hou).

0165-4101/$ - see front matter © 2011 Elsevier B.V. All rights reserved.
doi:10.1016/j.jacceco.2011.12.001

important to investment management practices such as portfolio allocation, performance evaluation, active risk control, and style/attribution analysis.

Prior academic studies almost exclusively rely on ex post realized returns to measure ex ante expected returns. However, as many researchers (e.g., Blume and Friend, 1973; Sharpe, 1978; Froot and Frankel, 1989; Elton, 1999) have pointed out, realized returns are a noisy proxy for expected returns. For example, Elton (1999) demonstrates that average realized returns can deviate significantly from expected returns over prolonged periods of time. Expected returns can also be estimated using asset pricing models such as the CAPM and the Fama and French (1993) three-factor model, but those estimates too are based on realized returns. Moreover, they are notoriously imprecise (see, e.g., Fama and French, 1997).

To address the deficiencies of the expected return estimates based on realized returns, recent accounting and finance studies (e.g., Gordon and Gordon, 1997; Claus and Thomas, 2001; Gebhardt et al., 2001; Easton, 2004; Ohlson and Juettner-Nauroth, 2005) propose an alternative approach to estimate expected returns: the implied cost of capital (ICC).[1] The ICC of a firm is the internal rate of return that equates the firm's stock price to the present value of expected future cash flows. In other words, the ICC is the discount rate that the market uses to discount the expected cash flows of the firm. The main advantage of the ICC is that it does not rely on noisy realized returns or on any specific asset pricing model. Instead, it derives expected return estimates directly from stock prices and cash flow forecasts.

The idea behind the ICC is simple and intuitively appealing. As a result, the ICC has been widely used in both finance and accounting research.[2] The common approach in this literature is to use analysts' earnings forecasts to proxy for cash flow expectations.[3] However, recent empirical evidence suggests that the performance of the analyst-based ICC as a proxy for expected returns is less than fully satisfactory. A priori, a reliable expected return proxy should positively predict future realized returns.[4] Several studies (e.g., Gebhardt et al., 2001; Easton and Monahan, 2005; Guay et al., 2011) examine the relation between the analyst-based ICC and future realized returns and find only mixed results. For example, Easton and Monahan (2005) show that the analyst-based ICC has little predictive power for future realized returns after controlling for cash flow news and discount rate news. They conclude that the analyst-based ICC is not a reliable proxy for expected returns and attribute the lack of reliability to the quality of analysts' earnings forecasts.

There are other concerns about the analyst-based ICC. One such concern is that even though analysts' forecasts are widely used by researchers and practitioners, they also exhibit important biases. A large body of research (e.g., Francis and Philbrick, 1993; Dugar and Nathan, 1995; McNichols and O'Brien, 1997; Lin and McNichols, 1998; Easton and Sommers, 2007) documents that analysts tend to be overly optimistic in their forecasts, likely the result of the conflicts of interest they are subject to. Furthermore, Abarbanell and Bushee (1997) and Francis et al. (2000) find large valuation errors when analysts' forecasts are used in valuation models.

A second major concern is coverage. The IBES analyst data are only available after the late 1970s, and small firms and financially distressed firms are underrepresented (La Porta, 1996; Hong et al., 2000; Diether et al., 2002). In addition, for many firms with analyst data, earnings forecasts beyond the second year or long-term growth forecasts (which are required by some of the commonly used ICC models) are not available. This is especially true in the earlier years. As a result, the analyst-based ICC has limited cross-sectional and time-series coverage, which can impede the investigation of questions that require a long time-series of expected return estimates or expected return estimates for small and distressed firms.

In this paper, we propose a new approach to estimate the ICC. We use earnings forecasts generated by a cross-sectional model instead of analysts' forecasts to proxy for cash flow expectations. Previous studies (e.g., Fama and French, 2000, 2006; Hou and Robinson, 2006; Hou and van Dijk, 2011) show that cross-sectional models are able to explain a large fraction of the variation in expected profitability across firms. We estimate model-based earnings forecasts for up to five years into the future and then use those earnings forecasts to compute the ICC for more than 170,000 firm-year observations over the period 1968–2008.

A major advantage of our model-based approach is that it uses the large cross-section of individual firms to compute earnings forecasts and therefore generates statistical power while imposing minimal survivorship requirements. Our approach allows us to compute the earnings forecasts and ICC for any firm with publicly traded equity and information on a limited number of accounting variables. Hence, the cross-sectional coverage of our model-based earnings forecasts and ICC is much larger than the coverage of analysts' forecasts and the analyst-based ICC. In addition, we are able to estimate the model-based earnings forecasts and the model-based ICC for earlier periods during which the IBES analyst data are not available.

We show that our cross-sectional earnings model captures significant variation in earnings performance across firms. The average R^2's of the regressions forecasting one-, two-, and three-year ahead earnings are 86%, 81%, and 78%,

[1] See Easton (2007) and Richardson et al. (2010) for reviews of this literature.

[2] For example, the ICC has been used to test the tradeoff between risk and return (Gebhardt et al., 2001; Pástor et al., 2008; Chava and Purnanandam, 2010; Lee et al., 2009) and to study the impact of corporate governance and disclosure (Botosan, 1997; Botosan and Plumlee, 2002; Francis et al., 2005b; Ashbaugh-Skaife et al., 2009), legal institutions and market regulations (Hail and Leuz, 2006), cross-listings (Hail and Leuz, 2009), taxes (Dhaliwal et al., 2005), earnings smoothness (Francis et al., 2004), accruals quality (Francis et al., 2005a; Core et al., 2008), and accounting restatements (Hribar and Jenkins, 2004) on a firm's cost of capital.

[3] One exception is Allee (2010), who uses time-series earnings forecasts based on an exponential smoothing method to estimate the ICC.

[4] Lee et al. (2010) formally derive this property using a simple return decomposition framework based on Campbell (1991).

respectively. The forecasts generated by the model are on average less accurate than analysts' forecasts, but exhibit much lower levels of forecast bias, and, more importantly, much higher levels of earnings response coefficient (ERC) than analysts' forecasts. The latter finding suggests that the earnings forecasts from the cross-sectional model represent a better proxy for market expectations of future earnings. This is in contrast to the previous earnings forecasting literature which generally concludes that analysts' forecasts are superior to forecasts from time-series models (see, e.g., Brown et al., 1987).[5]

We compute five individual ICC estimates (based on five commonly used ICC models) and a composite ICC estimate (the average of the five individual ICC estimates) using the model-based earnings forecasts. For comparison purposes, we also compute the equivalent ICC estimates using the IBES consensus analyst forecasts. Following Gebhardt et al. (2001), Easton and Monahan (2005), and Guay et al. (2011), we evaluate the quality of the model-based ICC and the analyst-based ICC by examining their relation with future realized returns. We find that the model-based ICC is a strong positive predictor of future realized returns. A decile spread portfolio that goes long in stocks with the highest composite model-based ICC and short in stocks with the lowest composite model-based ICC produces significantly positive average buy-and-hold returns of 10% to 12% per annum for holding periods of one, two, and three years after portfolio formation. In contrast, the average return of the spread portfolio based on the composite analyst-based ICC is less than 5% per annum and statistically insignificant for each of the three holding periods. Furthermore, the differences in average returns between the spread portfolios based on the composite model-based ICC and the composite analyst-based ICC are economically large and statistically significant. Hence, the model-based ICC is a more reliable predictor of future stock returns than the analyst-based ICC.

Our results are robust to alternative specifications of the cross-sectional earnings model (e.g., including additional accounting variables as earnings predictors or estimating the earnings model using scaled earnings instead of dollar earnings), to adjusting for the predictable component of analysts' forecast bias, and to specific methods used to compute the ICC. Furthermore, we show that the performance of both the model-based earnings forecasts and the model-based ICC relative to their analyst-based counterparts is stronger for firms with a poorer information environment (smaller, younger firms, firms with higher idiosyncratic volatility, lower analyst coverage, more volatile earnings, poorer accruals quality, or lower past returns).

Our approach to estimate the ICC has important implications for many key issues in accounting and finance. We use our model-based ICC to re-examine the cross-sectional relation between expected returns and a variety of firm-level characteristics (risk proxies) that have been shown to predict average realized returns. Our analysis indicates that inferences about the cross-section of expected returns are sensitive to the choice of expected return proxy (average ex post realized returns vs. ex ante model-based ICC).

The rest of the paper is organized as follows. Section 2 introduces the data and the cross-sectional earnings model, and discusses the estimation of the ICC. Section 3 compares the performance of the earnings forecasts generated by the cross-sectional model to that of analysts' forecasts. Section 4 evaluates the performance of the model-based ICC and compares it to that of the analyst-based ICC. Section 5 examines the cross-sectional relation between a number of firm-level characteristics and expected returns using the model-based ICC. Section 6 discusses a number of additional robustness checks. Section 7 concludes.

2. Data and empirical methodology

2.1. Data and sample selection

Our sample includes all NYSE, Amex, and Nasdaq listed securities with sharecodes 10 or 11 (i.e., excluding ADRs, closed-end funds, and REITs) that are at the intersection of the CRSP monthly returns file from July 1963 to June 2009 and the Compustat fundamentals annual file from 1963 to 2009. Our results are robust to excluding utilities and financials from the analysis. We use the following variable definitions. Earnings is income before extraordinary items from Compustat. Book equity is Compustat stockholder's equity. Total assets and dividends are also from Compustat. Prior to 1988, accruals are calculated using the balance sheet method as the change in non-cash current assets less the change in current liabilities excluding the change in short-term debt and the change in taxes payable minus depreciation and amortization expense. Starting in 1988, we use the cash flow statement method to calculate accruals as the difference between earnings and cash flows from operations.[6] We also obtain consensus analyst forecasts and the corresponding actual earnings from the IBES summary files.

2.2. Cross-sectional earnings model

To forecast earnings of individual firms, we use a model that is based on an extension and variation of the cross-sectional profitability models in Fama and French (2000, 2006), Hou and Robinson (2006), and Hou and van Dijk (2011). Previous studies on model-based earnings forecasts (e.g., Brown and Rozeff, 1978; Fried and Givoly, 1982;

[5] A recent paper by Bradshaw et al. (2011) shows that analysts' superiority over time-series forecasts is negligible for smaller and younger firms, and over longer horizons.

[6] See Hribar and Collins (2002) for details. Our results are robust to using the balance sheet method for the entire sample period.

Brown et al., 1987; O'Brien, 1988) tend to focus on time-series models fit separately to individual firms. To enhance power, empirical tests are usually restricted to firms with long earnings histories. This data requirement introduces survivorship bias to the tests. In addition, estimates based on these individual time-series models are not very precise. An important advantage of our cross-sectional approach is that it provides statistical power without imposing strict survivorship requirements.

Specifically, for each year between 1968 and 2008, we estimate the following pooled cross-sectional regressions using the previous ten years of data:

$$E_{i,t+\tau} = \alpha_0 + \alpha_1 A_{i,t} + \alpha_2 D_{i,t} + \alpha_3 DD_{i,t} + \alpha_4 E_{i,t} + \alpha_5 Neg\,E_{i,t} + \alpha_6 AC_{i,t} + \varepsilon_{i,t+\tau}, \tag{1}$$

where $E_{i,t+\tau}$ denotes the earnings of firm i in year $t+\tau$ ($\tau=1$ to 5), $A_{i,t}$ is the total assets, $D_{i,t}$ is the dividend payment, $DD_{i,t}$ is a dummy variable that equals 1 for dividend payers and 0 otherwise, $Neg\ E_{i,t}$ is a dummy variable that equals 1 for firms with negative earnings and 0 otherwise, and $AC_{i,t}$ is accruals. All explanatory variables are measured as of year t.

The main difference between Eq. (1) and the cross-sectional models used in prior studies (e.g., Fama and French, 2000) is that we use Eq. (1) to forecast dollar earnings while previous papers use cross-sectional models to forecast profitability (earnings scaled by total assets). We focus on dollar earnings to make our forecasts comparable with analysts' forecasts. In addition, the ICC literature (e.g., Gebhardt et al., 2001) exclusively uses dollar earnings forecasts to estimate the ICC. That said, we are concerned about potentially overweighting firms with extreme dollar earnings in estimating Eq. (1).[7] To address this concern, we winsorize earnings and other level variables each year at the 1st and 99th percentiles (observations beyond the extreme percentiles are set to equal to the values at those percentiles). We also carry out robustness checks in Section 6 by scaling earnings (and the other level variables in Eq. (1)) using lagged total assets and find similar results.

For each firm i and each year t in our sample, we compute earnings forecasts for up to five years into the future by multiplying the independent variables as of year t with the coefficients from the pooled regression estimated using the previous ten years of data. This is to ensure that our earnings forecasts are strictly out of sample (that is, all information that is required to forecast earnings for years $t+1$ through $t+5$ is available in year t). In addition, we only require a firm to have non-missing values for the independent variables in year t to estimate its earnings forecasts. As a result, the survivorship bias is kept to a minimum.

2.3. Estimating the ICC

The ICC for a given firm is the internal rate of return that equates the current stock price to the present value of expected future cash flows. Previous studies have developed a variety of methods to estimate the ICC. To ensure that our results are not driven by any specific method, our main analyses are based on a "composite" ICC measure that is the average of the following five individual ICC estimates: Claus and Thomas (CT, 2001), Easton (modified price-earnings growth or MPEG, 2004), Gebhardt et al. (GLS, 2001), Gordon and Gordon (Gordon, 1997), and Ohlson and Juettner-Nauroth (OJ, 2005).[8] These individual ICC estimates differ in the use of forecasted earnings, the explicit forecast horizon, and the assumptions regarding short-term and long-term growth rates.[9] They can be broadly grouped into three categories: CT and GLS are based on the residual income valuation model; OJ and MPEG are abnormal earnings growth-based models; Gordon is based on the Gordon growth model. We provide a detailed description of the five individual ICC estimates in Appendix A.

We compute each of the five individual ICC estimates for each firm at the end of June of each year t by using the end-of-June market equity and the model-based earnings forecasts for up to five years into the future. To ensure that the model-based earnings forecasts are based on information that is publicly available at the time of ICC estimation, we impose a minimum reporting lag of three months. That is, we compute the model-based earnings forecasts for firms with fiscal year ends from April of year $t-1$ to March of year t by multiplying their accounting variables with the coefficients from the pooled regression estimated using the previous ten years of data. We then match these earnings forecasts to the market equity at the end of June of year t to estimate the ICC.[10]

In addition to the five individual ICC estimates, we also construct a "composite" ICC measure as the equal-weighted average of the five individual estimates. To maximize coverage, we only require a firm to have at least one non-missing individual ICC estimate to compute its composite ICC. However, our results are robust to requiring firms to have non-missing values for all five individual ICC estimates. In Section 4, we match the ICC estimates for individual firms computed at the end of June of year t with their annual stock returns from July of year t to June of year $t+1$, from July of year $t+1$ to June of year $t+2$, and from July of year $t+2$ to June of year $t+3$ to evaluate the performance of these ICC estimates. Fig. 1 illustrates a timeline for the estimation procedure described above.

For comparison, we also compute analyst-based ICC estimates using the latest consensus analyst forecasts as of June of year t. Relative to our model-based forecasts, analysts clearly have a timing advantage as they have access to information

[7] Profitability regressions can also be dominated by extreme observations created by scaling earnings using assets that are close to zero, unless care is taken to mitigate the influence of these observations.

[8] We also present results based on the individual ICC estimates as robustness checks.

[9] We refer to Easton and Monahan (2005) and Lee et al. (2010) for comprehensive examinations of the various ICC models to date.

[10] We follow previous studies and set individual ICC estimates that are below zero to missing. We also winsorize the ICC estimates annually at the 1st and 99th percentiles to minimize the impact of outliers.

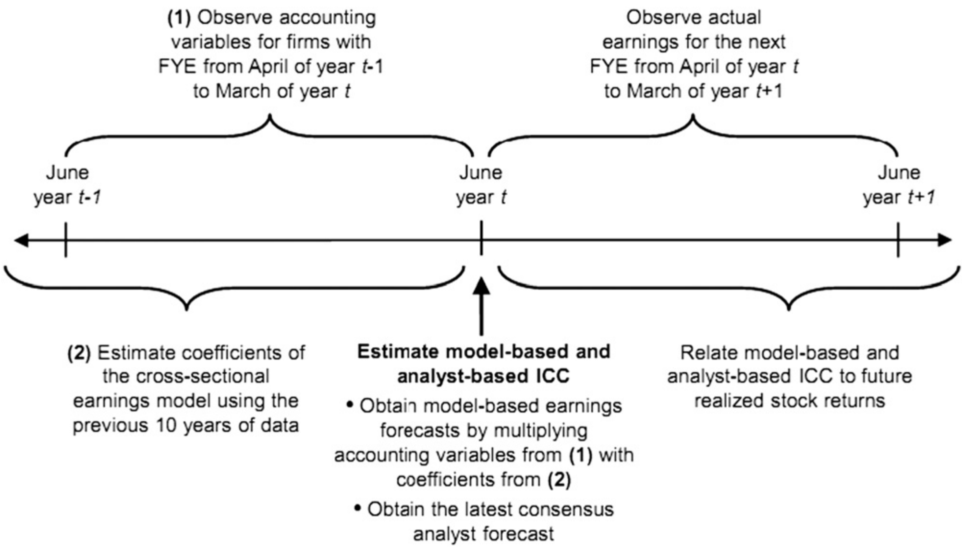

Fig. 1. Timeline of earnings forecasts and ICC estimation. This figure illustrates the timeline of the earnings forecasts and the ICC estimation. At the end of June of each year *t*, we obtain the model-based earnings forecasts for firms with fiscal year ends (FYE) from April of year $t-1$ to March of year *t* as the product of (1) the accounting variables from the most recent FYE (from April of year $t-1$ to March of year *t*, assumed to be known by June of year *t*) and (2) the coefficients of the cross-sectional earnings model estimated using the previous ten years of data (also assumed to be known by June of year *t*). We also obtain the latest consensus analyst forecast as of June of year *t*. We then match the model-based and analysts' forecasts with the corresponding actual earnings for the next FYE (from April of year *t* to March of year $t+1$) to compare their performance in terms of forecast bias, forecast accuracy, and earnings response coefficient (ERC). Similar comparisons are also made using longer term forecasts and actual earnings. We compute the five individual (model-based or analyst-based) ICCs (GLS, CT, OJ, MPEG, and Gordon) and a composite ICC (the average of the five individual ICCs) for each firm using its end-of-June market equity or stock price and the model-based earnings forecasts or the latest consensus analyst forecasts for up to five years into the future. We then match the individual and composite ICCs with annual stock returns from July of year *t* to June of year $t+1$, from July of year $t+1$ to June of year $t+2$, and from July of year $t+2$ to June of year $t+3$ to evaluate the performance of these ICC estimates.

available through June of year *t*, while our model-based earnings forecasts are based on accounting variables dated at least three months ago (firms with March of year *t* fiscal year ends) and as far back as 14 months ago (firms with April of year $t-1$ fiscal year ends). This difference in timing could potentially bias the results against our cross-sectional earnings model. However, we will show in the next section that, despite the timing disadvantage, the earnings forecasts from the cross-sectional model are associated with substantially lower levels of forecast bias and higher levels of earnings response coefficient than analysts' forecasts.

3. Performance of the earnings forecasts based on the cross-sectional model

3.1. Summary statistics and estimates of the cross-sectional earnings model

Panel A of Table 1 presents summary statistics (the time-series averages of the cross-sectional mean, median, standard deviation, and select percentiles) of the variables used in the cross-sectional earnings model (Eq. (1)). Panel B of Table 1 reports the average coefficients from the pooled regressions estimated each year from 1968 to 2008 and their time-series Newey-West *t*-statistics.[11] To conserve space, we only report the results for the one-, two-, and three-year ahead earnings regressions (those for four- and five-year ahead regressions are available upon request). The average coefficient for each of the independent variables maintains the same sign across different forecast horizons. Consistent with the results of Fama and French (2006), Hou and Robinson (2006), and Hou and van Dijk (2011), firm-level earnings are highly persistent. The coefficients on lagged earnings are 0.8304 (*t*-stat of 35.09), 0.7924 (*t*-stat of 25.21), and 0.7871 (*t*-stat of 22.89) for the one-, two-, and three-year ahead earnings regressions, respectively. Future earnings are also significantly positively related to total assets of the firm. In addition, firms that pay out more dividends and firms with lower accruals tend to have higher future earnings. The coefficient on the negative earnings dummy is positive and significant and the coefficient on the dividend dummy is positive but insignificant for all three horizons.

[11] We estimate the regression each year to allow the coefficients to vary over time. We correct for serial dependence by applying the Newey-West procedure to the annual coefficient estimates, which is a common practice in the literature (e.g., Gebhardt et al., 2001; Richardson et al., 2006). See Gow et al. (2010) for a useful discussion of the robustness of this procedure in accounting applications. While we recognize that the Newey-West *t*-statistics may not sufficiently address the time-series dependence in the underlying data, we note that we will only use the coefficient estimates, and not the potentially biased *t*-statistics or standard errors, to compute the model-based earnings forecasts.

Table 1
Cross-sectional earnings regressions, 1968–2008.

Panel A: Summary statistics of the variables in the cross-sectional earnings model

Variable	Mean	1%	25%	Median	75%	99%	STD
E_t	49.07	−174.82	−1.33	3.41	22.78	1269.20	184.81
A_t	1529.78	2.49	34.18	138.48	607.52	36,767.94	5505.94
D_t	19.11	0.00	0.01	0.27	4.32	498.24	69.33
DD_t	0.49	0.00	0.07	0.41	1.00	1.00	0.47
$Neg\ E_t$	0.25	0.00	0.00	0.00	0.63	1.00	0.41
AC_t	−43.14	−1138.99	−17.82	−2.48	0.67	104.81	157.27

Panel B: Coefficient estimates of the cross-sectional earnings model

LHS		Intercept	A_t	D_t	DD_t	E_t	$Neg\ E_t$	AC_t	$Adj.R^2$
E_{t+1}	Coefficient	−1.0311	0.0024	0.2659	0.9019	0.8304	1.3768	−0.0725	0.86
	t-stat	*−2.27*	*7.42*	*8.00*	*1.51*	*35.09*	*2.14*	*−6.23*	
E_{t+2}	Coefficient	−1.0738	0.0051	0.3799	0.6430	0.7924	2.6787	−0.0848	0.81
	t-stat	*−1.81*	*9.15*	*8.85*	*0.66*	*25.21*	*3.43*	*−4.09*	
E_{t+3}	Coefficient	−0.7876	0.0082	0.4653	0.0972	0.7871	3.9879	−0.0932	0.78
	t-stat	*−1.21*	*9.95*	*10.80*	*0.08*	*22.89*	*5.18*	*−3.67*	

Panel A of this table presents summary statistics (the time-series averages of the cross-sectional mean, median, standard deviation, and select percentiles) of the variables used in the cross-sectional earnings model. All variables except DD_t and $Neg\ E_t$ are expressed in $ millions. Panel B of this table reports the average coefficients and their time-series Newey-West *t*-statistics (in *italics*) from pooled regressions estimated each year from 1968 to 2008 using the previous ten years of data. E_{t+1}, E_{t+2}, and E_{t+3} are the one-, two-, and three-year ahead earnings (income before extraordinary items), respectively. A_t is total assets. D_t is the dividend payment. DD_t is a dummy variable that equals 1 for dividend payers and 0 otherwise. $Neg\ E_t$ is a dummy variable that equals 1 for firms with negative earnings and 0 otherwise. AC_t is accruals. Prior to 1988, accruals are calculated using the balance sheet method as the change in non-cash current assets less the change in current liabilities excluding the change in short-term debt and the change in taxes payable minus depreciation and amortization expense. Starting in 1988, accruals are calculated using the cash flow statement method as the difference between earnings and cash flows from operations.

Panel B of Table 1 also shows that the cross-sectional model captures a substantial part of the variation in future earnings performance across firms. The average regression R^2's are 86%, 81%, and 78% for the one-, two-, and three-year ahead earnings regressions, respectively.

3.2. Summary statistics of the model-based and analysts' earnings forecasts

Table 2 reports, for each five-year subperiod and the entire sample period 1968–2008, the time-series averages of the annual mean and median one-, two-, and three-year ahead earnings forecasts based on our cross-sectional model (Panel A) and those of the IBES consensus forecasts (Panel B), as well as the correlations between the model-based and analysts' forecasts (Panel C). We scale the model-based earnings forecasts using a firm's end-of-June market equity and analysts' per share earnings forecasts using the end-of-June stock price to report them in the same units. Panel A reveals a general trend of declining model-based earnings forecasts since the late 1970s, which is consistent with the finding of Fama and French (2004) that US publicly traded firms have become less profitable over time. Analysts' forecasts (Panel B) exhibit a similar time-series trend. Panels A and B also show that the mean (median) model-based earnings forecast tends to be higher (lower) than the mean (median) analyst forecast, especially for the longer-term (two- and three-year ahead) forecasts. However, we want to offer a note of caution about direct comparisons between the model-based and analysts' forecasts here because of the differences in the sample of firms for which each type of forecasts is available, and also because the model-based forecasts are based on Compustat (GAAP) earnings but analysts' forecasts are based on pro forma (Street) earnings (which are purged of transitory or special items and therefore do not necessarily equal GAAP earnings).

Panels A and B of Table 2 also report the average number of firms for which the model-based and analysts' forecasts are available. The difference in coverage between the two types of forecasts is very large. The number of firms for which we can compute the model-based forecasts increases steadily from around 1,750 (the same across all forecast horizons) in the late 1960s/early 1970s to well over 6,000 in the mid to late 1990s, after which the number drops to around 5,000 in recent years (which coincides with a decrease in the number of firms on Compustat during the same time). On the other hand, analysts' forecasts start in 1982, the earliest year for which we can obtain the three-year ahead forecasts from IBES.[12]

[12] The three-year ahead analyst forecast can only be obtained for a substantially smaller number of firms when compared to the one- and two-year ahead forecasts. Even as recently as 2008, the three-year ahead earnings forecast is available for just 2,213 firms, less than two-thirds of the number of firms with one- and two-year ahead earnings forecasts. We follow prior studies in the ICC literature and estimate an imputed forecast for firms with missing three-year ahead analyst forecast using the consensus long-term growth forecast and the two-year ahead forecast. This treatment increases the total number of firm-year observations with available three-year ahead forecast from 26,448 to 74,922 (and from 2,213 to 2,997 in 2008).

Table 2
Summary statistics of earnings forecasts, 1968–2008.

Period	N	E_{t+1} Mean	E_{t+1} Median	E_{t+2} Mean	E_{t+2} Median	E_{t+3} Mean	E_{t+3} Median	Period	N_{t+1}	E_{t+1} Mean	E_{t+1} Median	N_{t+2}	E_{t+2} Mean	E_{t+2} Median	N_{t+3}	E_{t+3} Mean	E_{t+3} Median
Panel A: Model-based forecasts								*Panel B: Analysts' forecasts*									
1968–1973	1748	0.1214	0.1018	0.1741	0.1257	0.2329	0.1516										
1974–1978	3171	0.1988	0.1659	0.2957	0.2075	0.3864	0.2451										
1979–1983	3268	0.1662	0.1380	0.2518	0.1720	0.3293	0.2011										
1984–1988	3793	0.0695	0.0680	0.1455	0.0933	0.2085	0.1164	1982–1988	2662	0.0796	0.0856	2190	0.1115	0.1078	1827	0.1288	0.1229
1989–1993	4803	0.0703	0.0510	0.1917	0.0677	0.2987	0.0902	1989–1993	3218	0.0616	0.0724	2861	0.0917	0.0891	2260	0.1082	0.1012
1994–1998	6311	0.0642	0.0325	0.1242	0.0436	0.1747	0.0589	1994–1998	4697	0.0463	0.0590	4302	0.0695	0.0742	3591	0.0884	0.0863
1999–2003	5794	−0.1078	0.0202	−0.0128	0.0275	0.0871	0.0439	1999–2003	4306	0.0036	0.0507	3905	0.0462	0.0651	3341	0.0636	0.0764
2004–2008	5248	0.0466	0.0244	0.0828	0.0292	0.1408	0.0379	2004–2008	3873	0.0228	0.0505	3756	0.0471	0.0628	3235	0.0637	0.0723
1968–2008	172,432	0.0797	0.0759	0.1571	0.0965	0.2323	0.1190	1982–2008	99,100	0.0455	0.0653	89,454	0.0761	0.0819	74,922	0.0934	0.0941

Panel C: Correlations between model-based forecasts and analysts' forecasts, 1982–2008

		Model-based forecasts			Analysts' forecasts		
		E_{t+1}	E_{t+2}	E_{t+3}	E_{t+1}	E_{t+2}	E_{t+3}
Model-based forecasts	E_{t+1}	1.00					
	E_{t+2}	0.92	1.00				
	E_{t+3}	0.82	0.93	1.00			
Analysts' forecasts	E_{t+1}	0.64	0.62	0.61	1.00		
	E_{t+2}	0.55	0.55	0.57	0.88	1.00	
	E_{t+3}	0.45	0.46	0.48	0.79	0.94	1.00

This table presents summary statistics of the one-, two-, and three-year ahead earnings forecasts based on the cross-sectional earnings model (Panel A) and the IBES consensus analyst forecasts (Panel B), as well as the correlations between the model-based and analysts' forecasts (Panel C). E_{t+1}, E_{t+2}, and E_{t+3} refer to the one-, two-, and three-year ahead earnings that are forecasted. They are scaled by market equity for the model-based forecasts and by stock price for analysts' forecasts. N represents the number of firms for which the model-based forecasts are available (the same across all horizons). N_{t+1}, N_{t+2}, and N_{t+3} represent the number of firms for which analysts' earnings forecasts are available at each horizon. If the three-year ahead earnings forecast is missing from IBES, we use the consensus long-term growth forecast and the two-year ahead forecast to impute the three-year ahead forecast. To conserve space, Panels A and B report the time-series averages of N, N_{t+1}, N_{t+2}, N_{t+3}, and the mean and median forecasts of E_{t+1}, E_{t+2}, and E_{t+3} for each five-year subperiod. The last row of Panels A and B reports the total number of firm-year observations (the time-series sums of N, N_{t+1}, N_{t+2}, and N_{t+3}) and the time-series averages of the mean and median forecasts over the entire sample period.

The initial coverage is limited, especially for the two- and three-year ahead forecasts. The coverage improves over time, but it never reaches the level of the model-based forecasts. Even toward the end of the sample period we are still able to obtain the one-year ahead model-based forecast for around 1,500 more firms per year than the one-year ahead analyst forecast; the gap is even bigger for the two- and three-year ahead forecasts. For the entire 1968–2008 period, we are able to compute the model-based forecasts for 172,432 firm-year observations irrespective of the forecast horizon, while the coverage for analysts' forecasts is only 99,100, 89,454, and 74,922 for the one-, two-, and three-year horizons, respectively.

Panel C of Table 2 reports that the correlations between the model-based and analysts' forecasts are 0.64, 0.55, and 0.48 for the one-, two-, and three-year horizons, respectively (based on the common sample of firm year observations for which both types of forecasts are available). These correlations suggest that although there is significant overlap between the two types of forecasts, they also exhibit substantial differences. We also see in Panel C that the correlations between the same type of earnings forecasts but of different horizons are considerably higher, ranging from 0.82 to 0.93 for the model-based forecasts and from 0.79 to 0.94 for analysts' forecasts.

3.3. Evaluation of the model-based and analysts' earnings forecasts

Table 3 compares the performance of the model-based earnings forecasts to that of analysts' forecasts. It reports the time-series averages (and their associated time-series Newey-West t-statistics) of the annual mean and median forecast bias (Panel A), mean and median forecast accuracy (Panel B), and two different measures of the earnings response coefficient (Panel C) for the model-based and analysts' forecasts, as well as the differences between the two types of forecasts. The table is based on the common sample of firm-year observations for which both types of forecasts are available (the sample period starts in 1982 because of the availability of analysts' forecasts and ends in 2006 because we require actual earnings for up to three years into the future). Each panel compares the performance of the model-based forecasts to that of analysts' forecasts for the one-, two-, and three-year horizons as well as for a "weighted" earnings forecast, which is computed as the sum of the discounted forecasts for the three horizons with an annual discount rate of 10%.

Following prior research, we define the forecast bias as the difference between the actual (realized) earnings and the earnings forecast (model-based or analysts'), scaled by the end-of-June market equity for the model-based forecasts and the end-of-June stock price for analysts' forecasts.[13] A negative bias indicates an optimistic forecast. Panel A of Table 3 confirms the well-established result that analysts' forecasts tend to be overly optimistic, especially at longer horizons.[14] The mean forecast bias of the consensus analyst forecasts is significantly negative at all forecast horizons and the magnitude increases monotonically with the horizon (-0.0268, -0.0379, and -0.0414 for the one-, two-, and three-year ahead forecasts, respectively). The mean bias of the weighted analyst forecast is -0.0708 (t-stat of -6.46). In contrast, the mean bias of the model-based forecasts is substantially smaller in magnitude (-0.0209, -0.0167, and -0.0109 for the one-, two-, and three-year ahead forecasts, respectively, and -0.0228 for the weighted forecast) and is only significantly different from zero for the one-year ahead forecast. The difference in the mean bias between the model-based and analysts' forecasts is 0.0059 (t-stat of 1.10) for the one-year ahead forecast and it increases to 0.0213 (t-stat of 2.77) for the two-year ahead forecast, 0.0305 (t-stat of 2.97) for the three-year ahead forecast, and 0.0480 (t-stat of 3.19) for the weighted forecast. The median forecast bias is smaller in magnitude than the mean bias for both the model-based and analysts' forecasts (especially the former). For the model-based forecasts, the median bias is very close to zero and is not statistically significant at any horizon (the median bias of the weighted forecast is 0.0017 with a t-stat of 0.21). The median bias of analysts' forecasts, though smaller than the mean bias, is still sizable and is significant at all horizons (the median bias of the weighted forecast is -0.0330 with a t-stat of -5.63). The difference in the median bias between the model-based and analysts' forecasts is always positive (and comparable in magnitude to the difference in the mean bias) and highly significant at all horizons (the difference for the weighted forecast is 0.0347 with a t-stat of 11.19).

We define the forecast accuracy as the absolute value of the forecast bias (a smaller number is indicative of a more accurate earnings forecast). Panel B of Table 3 shows that the model-based forecasts are less accurate than analysts' forecasts. For example, the mean accuracy of the model-based forecasts is 0.0837 (t-stat of 9.00), 0.0938 (t-stat of 12.44), 0.1031 (t-stat of 12.44), and 0.1745 (t-stat of 11.10) for the one-, two-, three-year ahead forecasts and the weighted forecast, respectively, compared to 0.0362 (t-stat of 9.08), 0.0532 (t-stat of 11.94), 0.0616 (t-stat of 13.88), and 0.0955 (t-stat of 10.61) for the corresponding analysts' forecasts. The difference in the mean accuracy between the model-based and analysts' forecasts is statistically significant at all horizons. The median forecast accuracy exhibits a similar pattern, although the difference between the two types of forecasts is substantially smaller than that for the mean accuracy at each horizon.

[13] We use GAAP actual earnings (income before extraordinary items from Compustat) as the benchmark for evaluating the model-based earnings forecasts (also based on GAAP earnings) and Street actual earnings provided by IBES for analysts' forecasts (based on Street earnings). This is to ensure that our results are not driven by mixing the actual earnings with earnings forecasts based on a different earnings definition. Nevertheless, we have also performed robustness checks by using either GAAP actual earnings or Street actual earnings as the benchmark for both the model-based forecasts and analysts' forecasts. These results are discussed later in this section.

[14] Several recent studies (e.g., Guay et al., 2011; Mohanram and Gode, 2011) seek to model and remove the bias in analysts' forecasts. In Section 6, we investigate the implications of their approach for our results.

Table 3
Bias, accuracy, and earnings response coefficient (ERC) of earnings forecasts, 1982–2006.

	E_{t+1}			E_{t+2}			E_{t+3}			Weighted earnings forecast		
	Model	Analysts	Difference	Model	Analysts	Difference	Model	Analysts	Difference	Model	Analysts	Difference
Panel A: Bias of model-based forecasts vs. analysts' forecasts												
Mean	−0.0209	−0.0268	0.0059	−0.0167	−0.0379	0.0213	−0.0109	−0.0414	0.0305	−0.0228	−0.0708	0.0480
	−2.57	*−6.80*	*1.10*	*−1.41*	*−7.17*	*2.77*	*−0.72*	*−6.83*	*2.97*	*−0.98*	*−6.46*	*3.19*
Median	0.0031	−0.0041	0.0072	0.0010	−0.0143	0.0153	−0.0001	−0.0216	0.0214	0.0017	−0.0330	0.0347
	1.29	*−3.87*	*4.94*	*0.25*	*−6.01*	*9.70*	*−0.03*	*−6.26*	*11.83*	*0.21*	*−5.63*	*11.19*
Panel B: Accuracy of model-based forecasts vs. analysts' forecasts												
Mean	0.0837	0.0362	0.0475	0.0938	0.0532	0.0406	0.1031	0.0616	0.0415	0.1745	0.0955	0.0789
	9.00	*9.08*	*4.63*	*12.44*	*11.94*	*4.23*	*12.44*	*13.88*	*3.83*	*11.10*	*10.61*	*3.91*
Median	0.0302	0.0105	0.0197	0.0400	0.0242	0.0158	0.0458	0.0340	0.0118	0.0799	0.0496	0.0303
	20.33	*9.02*	*9.85*	*29.61*	*12.94*	*7.27*	*36.86*	*13.11*	*4.58*	*34.14*	*10.75*	*6.07*
Panel C: Earnings response coefficient (ERC) of model-based forecasts vs. analysts' forecasts												
Announcement ERC	0.0662	0.0473	0.0188	0.1101	0.0733	0.0368	0.1487	0.0872	0.0616	0.1279	0.0795	0.0484
	16.36	*12.71*	*6.06*	*22.25*	*14.12*	*4.99*	*19.30*	*13.17*	*6.70*	*16.92*	*9.96*	*4.64*
Annual ERC	0.2471	0.1961	0.0510	0.5543	0.3775	0.1768	0.8703	0.5551	0.3153	0.7773	0.5130	0.2643
	11.41	*16.33*	*2.91*	*13.07*	*23.05*	*4.69*	*12.66*	*18.88*	*5.65*	*11.17*	*18.34*	*4.41*

This table reports the time-series averages of the mean and median forecast bias (Panel A), mean and median forecast accuracy (Panel B), and the announcement and annual earnings response coefficients (ERCs) (Panel C) for the model-based and analysts' forecasts, as well as the differences between the model-based and analysts' forecasts. Newey-West t-statistics are reported in *italics*. The results are based on the common sample of firm-year observations for which both types of forecasts are available (the sample period starts in 1982 because of the availability of analysts' forecasts and ends in 2006 because we require actual earnings up to three years into the future). E_{t+1}, E_{t+2}, and E_{t+3} refer to the one-, two-, and three-year ahead earnings that are forecasted. The weighted earnings forecast is the sum of the discounted forecasts of E_{t+1}, E_{t+2}, and E_{t+3} with a discount rate of 10%. Forecast bias is the difference between actual (realized) earnings and the earnings forecast (model-based or analysts'), scaled by the end-of-June market equity for the model-based forecasts and by the end-of-June stock price for analysts' forecasts. Forecast accuracy is the absolute value of the forecast bias. ERC is estimated in two ways. First, we estimate annual cross-sectional regressions of the sum of the quarterly earnings announcement returns (market adjusted, from day −1 to day +1) over the next one, two, and three years on firm-specific unexpected earnings measured over the same horizon. For the weighted earnings forecast measure, we regress the sum of the earnings announcement returns over the next three years on the sum of the (discounted) unexpected earnings for the next three years. In the second method, we regress the buy-and-hold returns over the next one, two, and three years on the unexpected earnings over the same horizon. We standardize the unexpected earnings to have unit variance each year to make the ERCs comparable between the model-based forecasts and analysts' forecasts.

Although comparing the forecast bias and forecast accuracy provides insights into the properties of the model-based and analysts' forecasts, the key challenge from the perspective of estimating the ICC is to determine which of the two types of forecasts represents a better approximation of market expectations. This is because, by design, a firm's ICC is the discount rate that the market uses to discount its expectations of future earnings of the firm. Since earnings forecasts that are more accurate and/or less biased do not necessarily do a better job capturing the market's expectations of future earnings (see, e.g., Brown, 1993; O'Brien, 1988; Wiedman, 1996), we turn to the earnings response coefficient (ERC) as a more direct way of evaluating how closely the model-based and analysts' forecasts line up with market expectations. The literature on ERC dates back to Ball and Brown (1968), Fried and Givoly (1982), Lev and Ohlson (1982), and Easton and Zmijewski (1989). The ERC captures the reaction of stock prices to unexpected earnings (i.e., the difference between actual and forecasted earnings, or the forecast bias), which should be greater for better proxies for the market's earnings expectations (see, e.g., Brown et al., 1987).

We estimate the ERC of the model-based and analysts' forecasts in two different ways. First, we estimate annual cross-sectional regressions of the sum of the quarterly earnings announcement returns (market adjusted, from day -1 to day $+1$) over the next one, two, and three years on firm-specific unexpected earnings (based on GAAP actual earnings for the model-based forecasts and Street actual earnings for analysts' forecasts) measured over the same horizon.[15] The "announcement ERC", which is the coefficient from the cross-sectional regression, is similar to the one used by Brown et al. (1987) and Easton and Zmijewski (1989) to study the stock market response to earnings announcements.

In the second method, each year we regress the buy-and-hold returns over the next one, two, and three years on the unexpected earnings over the same horizon. This "annual ERC" is motivated by various papers in the accounting literature that use long-term buy-and-hold returns to study the value relevance of earnings (Collins et al., 1994; Hayn, 1995;

[15] For the weighted forecast, we regress the sum of the earnings announcement returns over the next three year on the sum of the (discounted) unexpected earnings for the next three years.

Francis et al., 2005c; Hou et al., 2011) or to compare the performance of analysts' forecasts to that of forecasts based on time-series models (Fried and Givoly, 1982; Bradshaw et al., 2011).

We report the time-series averages of the announcement ERCs and annual ERCs in Panel C of Table 3. To make the ERCs comparable between the model-based and analysts' forecasts, we standardize their corresponding unexpected earnings to have unit variance before running the cross-sectional regression each year.

The announcement ERCs associated with the model-based forecasts are 0.0662, 0.1101, 0.1487, and 0.1279 for the one-, two-, three-year ahead forecasts and the weighted forecast, respectively, all of which are highly significant with a minimum *t*-stat of 16.36. The corresponding announcement ERCs for analysts' forecasts are considerably smaller at 0.0473, 0.0733, 0.0872, and 0.0795 (also all significant).[16] The difference in the announcement ERCs between the model-based and analysts' forecasts is substantial and statistically significant at all horizons as well as for the weighted forecast (all *t*-stats are greater than 4). Thus, the model-based earnings forecasts are associated with significantly greater ERCs around earnings announcements than analysts' forecasts. The annual ERC shows a similar pattern. The annual ERCs associated with the model-based forecasts are 0.2471, 0.5543, 0.8703, and 0.7773 (with a minimum *t*-stat of 11.17) for the one-, two-, three-year ahead forecasts and the weighted forecast, respectively, compared with 0.1961, 0.3775, 0.5551, and 0.5130 (also all significant) for the corresponding analysts' forecasts. The differences in the annual ERCs between the model-based and analysts' forecasts are all highly significant and greater in magnitude than the differences in the announcement ERCs.

To explore the impact of the differences in earnings definitions (GAAP vs. Street) on our results, we re-evaluate the relative performance of the model-based and analysts' forecasts by using the same actual earnings (either GAAP or Street) as the benchmark for both types of forecasts. We find that matching analysts' forecasts (which are based on Street earnings) to GAAP actual earnings causes analysts' forecasts to appear considerably more optimistic, less accurate, and to have lower ERCs compared to matching them to Street actual earnings.[17] As a result, analysts' forecasts are now not only inferior to the model-based forecasts (matched to GAAP actual earnings) in terms of bias and ERC, but are now only slightly more accurate than the model-based forecasts, with the differences in bias and ERCs being economically large and statistically highly significant. On the other hand, matching the model-based forecasts (which are based on GAAP earnings) to Street actual earnings causes the model-based forecasts to appear pessimistic (i.e., a positive forecast bias), slightly less accurate, and to have somewhat lower ERCs compared to matching them to GAAP actual earnings.[18] Consequently, compared to analysts' forecasts (matched to Street actual earnings), the model-based forecasts are still less biased (though their mean bias is now positive) and less accurate, and still have greater announcement and annual ERCs, with the differences in bias, accuracy, and annual ERC all being significant. In sum, this analysis suggests that using the same actual earnings (either GAAP or Street) to evaluate the performance of both the model-based and analysts' forecasts is likely to bias the inferences against the forecasts that are based on a different earnings definition, with the model-based forecasts appearing to be less affected by this concern.

4. Performance of the ICC estimated using the model-based earnings forecasts

The previous section shows that the cross-sectional earnings model is able to explain a large fraction of the variation in future earnings across firms. The earnings forecasts produced by the model are on average less accurate than analysts' forecasts, but are superior in terms of coverage and forecast bias. More importantly, they are associated with greater earnings response coefficients (ERCs), which suggests that the model-based earnings forecasts represent a better proxy for the market's earnings expectations than analysts' forecasts. In this section, we examine the performance of the ICC estimated using the model-based earnings forecasts and compare it to the performance of the ICC based on analysts' forecasts.

4.1. Summary statistics of the model-based and analysts-based ICC estimates

We compute each of the five individual ICC estimates as well as the composite ICC measure for each firm at the end of June of each year using the earnings forecasts generated by the cross-sectional model or the latest consensus analyst forecasts.[19] Table 4 reports, for each five-year subperiod and for the entire sample period 1968–2008, the average number of firms for which we are able to compute the composite model-based ICC (Panel A) and the composite analyst-based ICC

[16] There is a potential measurement issue when we estimate the announcement ERCs for analysts' forecasts, as the analysts may have already observed one or more quarterly earnings announcements (depending on the fiscal year end) by June of year *t* when we obtain the consensus analyst forecast. The announcement ERCs for the model-based forecasts are not affected by this issue because the model-based forecasts are always based on accounting information at the previous fiscal year end. As a robustness check, we exclude quarterly earnings announcements that took place before June of year *t* and re-estimate the announcement ERCs for analysts' forecasts. We find that those ERCs become slightly smaller as a result, and the differences in the announcement ERCs between the model-based and analysts' forecasts are thus bigger than those reported in Table 3. These results are available upon request.

[17] For example, the mean bias, mean accuracy, announcement ERC, and annual ERC of the weighted analyst forecast are −0.1278, 0.1458, 0.0668, and 0.3708, respectively, when we use GAAP actual earnings as the benchmark, compared to −0.0708, 0.0955, 0.0795, and 0.5130, respectively, when we use Street actual earnings as the benchmark (in Table 3).

[18] The mean bias, mean accuracy, and announcement and annual ERCs of the weighted model-based forecast are 0.0478, 0.1870, 0.0874, and 0.6339, respectively, when we use Street actual earnings as the benchmark, compared to −0.0228, 0.1745, 0.1279, and 0.7773, respectively, when we use GAAP actual earnings as the benchmark (in Table 3).

[19] These ICC estimates are available from the authors upon request.

Table 4
Summary statistics of composite ICC estimates, 1968–2008.

Period	*N*	Mean	25%	Median	75%	Period	*N*	Mean	25%	Median	75%
Panel A: Composite model-based ICC						*Panel B: Composite analyst-based ICC*					
1968–1973	1748	0.1266	0.0710	0.1016	0.1506						
1974–1978	3171	0.1871	0.1029	0.1488	0.2242						
1979–1983	3268	0.1647	0.0909	0.1291	0.1865						
1984–1988	3793	0.1362	0.0640	0.0923	0.1481	1982–1988	2602	0.0980	0.0745	0.0926	0.1139
1989–1993	4803	0.1701	0.0543	0.0852	0.1627	1989–1993	3165	0.0913	0.0653	0.0809	0.1017
1994–1998	6311	0.1216	0.0399	0.0624	0.1082	1994–1998	4605	0.0746	0.0551	0.0697	0.0867
1999–2003	5794	0.1689	0.0363	0.0644	0.1185	1999–2003	4173	0.0856	0.0501	0.0688	0.0925
2004–2008	5246	0.1232	0.0344	0.0532	0.0823	2004–2008	3810	0.0712	0.0494	0.0619	0.0776
1968–2008	172,417	0.1492	0.0619	0.0923	0.1477	1982–2008	96,974	0.0852	0.0600	0.0761	0.0959

This table presents summary statistics of the composite model-based ICC (Panel A) and the composite analyst-based ICC (Panel B). The composite ICC (model-based or analyst-based) is the average of five individual ICCs (GLS, CT, OJ, MPEG, and Gordon). The individual ICCs are computed for each firm at the end of June of each year using end-of-June market prices and earnings forecasts (model-based or analysts') for up to five years into the future. To conserve space, the table reports, for each five-year subperiod, the average *N* (number of firms for which we can compute the composite ICC) and the time-series averages of the annual mean, and the 25th, 50th (median), and 75th percentiles of the composite ICC. The last row reports the total number of firm-year observations (the time-series sum of *N*) and the time-series averages of other summary statistics of the composite ICC over the entire sample period.

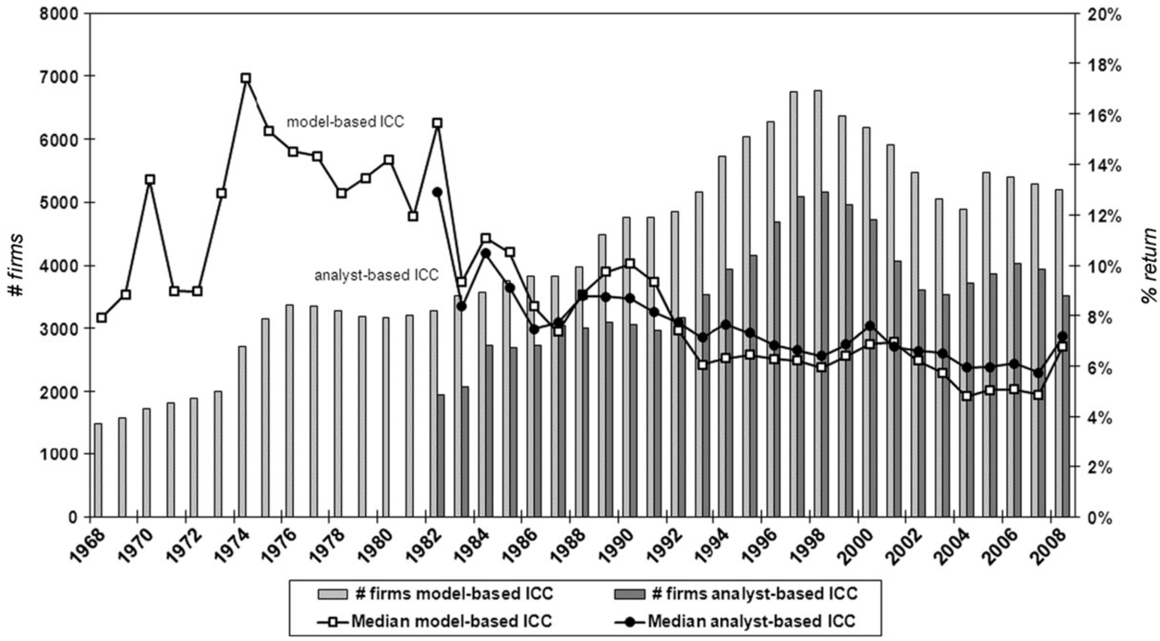

Fig. 2. Time-series plot of coverage and composite model-based and analyst-based ICCs. This figure plots the year-by-year coverage (# firms; bars, left axis) and the median composite model-based and analyst-based ICCs (% return; lines, right axis).

(Panel B), as well as the time-series averages of the annual mean, and the 25th, 50th (median), and 75th percentiles of the two composite ICC measures. Fig. 2 plots the year-by-year coverage and the median composite model-based and analyst-based ICCs.

The coverage of the model-based ICC starts around 1,750 firms (1968–1973 subperiod), quickly increases to around 3,000 firms by the mid 1970s, and peaks at 6,300 firms in the mid to late 1990s. Toward the end of the sample period, the coverage drops to around 5,000 firms. Over the entire sample period 1968–2008, we are able to estimate the composite model-based ICC for a total of 172,417 firm-year observations.

The coverage of the analyst-based ICC, on the other hand, does not start until 1982 and is much more limited than that of the model-based ICC during the entire time period for which both ICCs are available (1982–2008). The total number of firm-year observations for the composite analyst-based ICC is 96,974, less than 60% of the number for the composite model-based ICC. In untabulated results, we find that the difference in coverage between the model-based ICC and the analyst-based ICC is even more pronounced for some of the individual ICC estimates, especially those that require long-term earnings forecasts. For example, the coverage of the model-based GLS ICC is 161,734 firm-year observations, while

Table 5
Relation between composite ICC estimates and future realized returns, 1968–2006.

Decile	$r_{t,t+1}$	*t-stat*	$r_{t,t+2}$	*t-stat*	$r_{t,t+3}$	*t-stat*	$r_{t,t+1}$	*t-stat*	$r_{t,t+2}$	*t-stat*	$r_{t,t+3}$	*t-stat*
	Panel A: Composite model-based ICC (1968–2006)						*Panel B: Composite analyst-based ICC (1982–2006)*					
1	0.0702	*2.02*	−0.0096	*−0.35*	−0.0182	*−0.76*	0.0991	*2.08*	0.0149	*0.43*	0.0066	*0.22*
2	0.0977	*3.18*	0.0386	*1.54*	0.0341	*1.55*	0.1053	*2.72*	0.0398	*1.43*	0.0343	*1.48*
3	0.1117	*3.66*	0.0578	*2.40*	0.0540	*2.55*	0.1248	*3.45*	0.0691	*2.96*	0.0648	*3.54*
4	0.1242	*4.03*	0.0752	*2.92*	0.0723	*3.31*	0.1284	*3.81*	0.0756	*3.55*	0.0718	*3.97*
5	0.1211	*3.89*	0.0774	*3.06*	0.0733	*3.23*	0.1341	*4.32*	0.0861	*4.05*	0.0799	*4.45*
6	0.1305	*3.97*	0.0822	*2.95*	0.0795	*3.19*	0.1541	*4.50*	0.0948	*4.01*	0.0861	*4.08*
7	0.1415	*3.99*	0.0829	*2.84*	0.0784	*3.01*	0.1465	*3.84*	0.0929	*3.26*	0.0853	*3.42*
8	0.1390	*3.69*	0.0785	*2.62*	0.0737	*2.68*	0.1547	*3.68*	0.0963	*3.12*	0.0868	*3.22*
9	0.1566	*3.77*	0.0834	*2.52*	0.0815	*2.72*	0.1451	*3.38*	0.0848	*2.47*	0.0762	*2.52*
10	0.1764	*4.42*	0.1085	*3.47*	0.1022	*3.71*	0.1381	*2.34*	0.0635	*1.46*	0.0519	*1.40*
10-1	0.1062	*3.77*	0.1181	*4.89*	0.1203	*5.39*	0.0391	*0.82*	0.0486	*1.20*	0.0453	*1.33*
	Panel C: 10-1 spread, composite model-based ICC vs. composite analyst-based ICC (common sample period, 1982–2006)											
Model	0.1171	*3.55*	0.1344	*5.11*	0.1330	*5.61*						
Analysts	0.0391	*0.82*	0.0486	*1.20*	0.0453	*1.33*						
Difference	0.0780	*2.79*	0.0858	*3.65*	0.0877	*4.34*						

This table reports the time-series averages of the annualized buy-and-hold returns over the next one, two, and three years (and their Newey-West *t*-statistics in *italics*) of decile portfolios sorted on the composite model-based ICC (Panel A, sample period 1968–2006) and the composite analyst-based ICC (Panel B, sample period 1982–2006), as well as the return spreads between Deciles 10 and 1 (10-1). Panel C reports the average 10-1 return spreads associated with the composite model-based and analyst-based ICC over the common sample period 1982–2006 as well as the differences between the two return spreads (and their associated Newey-West *t*-statistics).

the coverage for the corresponding analyst-based ICC is only 68,414 (about 40% of the coverage of the model-based ICC).[20] The greater coverage of the model-based ICC not only enhances the power of any empirical test that uses the ICC, it also allows us to address research questions that require a long time-series of expected return estimates or expected return estimates for firms that are not covered by analysts.

Fig. 2 shows that there is considerable time-series variation in both the model-based and analyst-based ICCs. The median composite model-based ICC increases from around 8% in the late 1960s to a high of around 17% in 1974, then gradually declines to around 6% by the mid to late 1990s, after which it hovers around 6% for the remainder of the sample period.[21] The median composite analyst-based ICC starts from a high of 13% in 1982, decreases to around 6% by the late 1990s, and remains close to 6% until the end of the sample period. The decline in both the model-based and analyst-based ICCs after the early 1980s coincides with a dramatic increase in the number of newly listed firms on major US exchanges. Fama and French (2004) hypothesize that the increase in the new lists is due to a decline in the cost of equity capital. Our results support this explanation.

4.2. Evaluation of the composite model-based and analysts-based ICC estimates

Following Gebhardt et al. (2001), Easton and Monahan (2005), and Guay et al. (2011), we evaluate the quality of both the model-based and analyst-based ICCs as proxies for expected returns by examining their relation with future realized returns. The idea behind this test is that a reliable expected return estimate should positively predict future realized returns. Furthermore, higher quality expected return estimates should have stronger predictive power for realized returns (see, e.g., Lee et al., 2010).

Table 5 presents the results of this test. At the end of June of each year, we sort firms into decile portfolios based on their composite model-based or analyst-based ICC. We then compute the annualized equal-weighted buy-and-hold returns of each portfolio over the next one, two, and three years.[22] Table 5 reports the time-series averages of the annualized buy-and-hold returns (and their associated time-series Newey-West *t*-statistics) of the decile portfolios sorted on the composite model-based ICC (Panel A, sample period 1968–2006) and the composite analyst-based ICC (Panel B, sample period 1982–2006), as well as the return spreads between the highest-ranked ICC decile (Decile 10) and the lowest-ranked ICC decile (Decile 1).[23]

[20] The coverage of the model-based CT, OJ, MPEG, and Gordon ICCs is 159,758, 136,719, 146,392, and 131,138, respectively, compared with 64,553, 61,502, 79,508, and 85,792 for the corresponding analyst-based ICCs.

[21] We focus on the median ICC here because the mean ICC is likely to be dominated by the large number of small firms, especially for the model-based ICC.

[22] We include realized returns for the second and third year here because the ICC represents a weighted average of the discount rates for all future periods, and should thus predict realized returns beyond the first year.

[23] The sample period for this test ends in 2006 because we require realized returns for three years after portfolio formation.

Table 6
Relation between individual ICC estimates and future realized returns, 1968–2006.

ICC estimates		Model-based ICC GLS	CT	OJ	MPEG	Gordon	Composite	Analyst-based ICC GLS	CT	OJ	MPEG	Gordon	Composite
Panel A: Correlations between individual ICCs (1968–2006)													
Model	GLS	1.00											
	CT	0.85	1.00										
	OJ	0.84	0.83	1.00									
	MPEG	0.65	0.74	0.60	1.00								
	Gordon	0.84	0.82	0.99	0.57	1.00							
	Composite	0.91	0.93	0.95	0.72	0.93	1.00						
Analysts	GLS	0.61	0.38	0.42	0.33	0.41	0.48	1.00					
	CT	0.36	0.45	0.35	0.27	0.34	0.39	0.65	1.00				
	OJ	0.56	0.52	0.58	0.21	0.58	0.59	0.71	0.72	1.00			
	MPEG	0.18	0.13	0.10	0.65	0.05	0.07	0.45	0.37	0.18	1.00		
	Gordon	0.54	0.50	0.56	0.22	0.56	0.57	0.68	0.65	0.99	0.13	1.00	
	Composite	0.47	0.42	0.45	0.37	0.45	0.44	0.83	0.80	0.90	0.51	0.85	1.00

ICC estimates	$r_{t,t+1}$	*t-stat*	$r_{t,t+2}$	*t-stat*	$r_{t,t+3}$	*t-stat*	$r_{t,t+1}$	*t-stat*	$r_{t,t+2}$	*t-stat*	$r_{t,t+3}$	*t-stat*
	Panel B: 10-1 spread, individual model-based ICCs (1968–2006)						*Panel C: 10-1 spread, individual analyst-based ICCs (1982–2006)*					
GLS	0.0736	*2.82*	0.0920	*4.14*	0.0997	*4.72*	0.0301	*0.72*	0.0384	*1.09*	0.0421	*1.28*
CT	0.0889	*3.46*	0.0856	*3.60*	0.0893	*3.86*	0.0014	*0.04*	0.0173	*0.55*	0.0213	*0.76*
OJ	0.0808	*3.25*	0.1052	*5.24*	0.1077	*5.99*	0.0354	*0.72*	0.0724	*1.95*	0.0715	*2.39*
MPEG	0.0732	*2.04*	0.0215	*0.72*	0.0167	*0.63*	−0.0226	*−0.42*	−0.0847	*−1.99*	−0.0825	*−2.26*
Gordon	0.0777	*3.03*	0.1088	*5.30*	0.1123	*6.32*	0.0452	*0.95*	0.0835	*2.25*	0.0799	*2.43*

This table reports the correlations between the individual and composite ICCs (Panel A) and the time-series averages of the 10-1 return spreads (and their Newey-West *t*-statistics) associated with the individual model-based ICCs (Panel B, sample period 1968–2006) and the individual analyst-based ICCs (Panel C, sample period 1982–2006). The correlation between any two ICCs is based on the common sample of observations for which both ICCs are available.

Panel A of Table 5 shows that the composite model-based ICC has significant predictive power for future realized returns. The average annualized buy-and-hold returns increase almost monotonically from Decile 1 to Decile 10, and the average return spreads between the extreme deciles (10–1) are 10.62% (*t*-stat of 3.77), 11.81% (*t*-stat of 4.89), and 12.03% (*t*-stat of 5.39) for the one-, two-, and three-year horizons, respectively, all of which are economically large and statistically significant.

In contrast, Panel B of Table 5 shows that the composite analyst-based ICC has no significant predictive power for future realized returns. The average buy-and-hold returns display an inverted U-shaped pattern across the ICC deciles, and the average return spreads between the extreme deciles, though positive (3.91%, 4.86%, and 4.53% for the one-, two-, and three-year horizons, respectively), are all statistically insignificant (*t*-stats below 1.40). In addition, these return spreads are considerably smaller than the return spreads associated with the composite model-based ICC.

To make sure that the difference in the return predictability between the model-based and analyst-based ICCs is not driven by the different sample periods over which these ICCs are available, we report the average 10-1 return spreads associated with the composite model-based ICC over the common sample period 1982–2006 in Panel C of Table 5 (the spreads for the composite analyst-based ICC are reproduced verbatim from Panel B). We also formally test whether the return spreads associated with the two ICC estimates are statistically different. We find that the average return spreads associated with the composite model-based ICC are slightly bigger for the common sample period (1982–2006) than for the full sample period (1968–2006). More importantly, the differences in the associated return spreads between the composite model-based and analyst-based ICCs are both economically large (7.80%, 8.58%, and 8.77% for the one-, two-, and three-year horizons, respectively) and statistically significant (*t*-stats of 2.79, 3.65, and 4.34, respectively). We conclude that the model-based ICC is a more reliable predictor of future stock returns than the analyst-based ICC.

4.3. Evaluation of the individual model-based and analysts-based ICC estimates

In Table 6, we try to dissect the findings in Table 5 by analyzing the performance of the individual ICC estimates. Panel A of Table 6 reports the correlations between the individual ICCs and between the individual and composite ICCs. The five individual model-based ICCs are generally highly correlated with each other and with the composite ICC. The correlation ranges from a low of 0.57 (between Gordon and MPEG) to a high of 0.99 (between Gordon and OJ).[24] The correlations

[24] MPEG has relatively low correlations with the other model-based ICCs we consider, ranging from 0.57 (with Gordon) to 0.74 (with CT). The correlations between the other model-based ICCs are considerably higher, ranging from 0.82 (between Gordon and CT) to 0.99 (between Gordon and OJ).

between the analyst-based ICCs are also positive but show much greater variation, from a low of 0.13 (between Gordon and MPEG) to a high of 0.99 (between Gordon and OJ).[25]

The correlations between the model-based and analyst-based ICCs (presented in the lower left quadrant of Panel A) are all positive but are considerably lower than those between the model-based ICCs or between the analyst-based ICCs. The correlation ranges from a low of 0.05 (between analyst-based MPEG and model-based Gordon) to a high of 0.65 (between analyst-based MPEG and model-based MPEG). Among these correlations, the diagonal ones (between the model-based and analyst-based ICCs using the same ICC estimation method) tend to be the highest (0.61 for GLS, 0.45 for CT, 0.58 for OJ, 0.65 for MPEG, 0.56 for Gordon, and 0.44 for the composite ICC). However, they are still generally lower than the correlations between the model-based ICCs estimated using different methods. We conclude that the type of earnings forecasts (model-based vs. analysts' forecasts) used to compute the ICC is more important than the specific method (e.g., GLS vs. Gordon) used to compute the ICC.

Panels B and C of Table 6 repeat the return predictability test of Table 5 for the individual ICC estimates. To conserve space, we only report the average 10-1 realized return spreads for the one-, two-, and three-year horizons and their time-series Newey-West *t*-statistics. All of the individual model-based ICCs show an economically large and statistically significant average return spread for the one-year horizon, ranging from 7.32% (*t*-stat of 2.04) for MPEG to 8.89% (*t*-stat of 3.46) for CT. With the exception of MPEG, all of the individual model-based ICCs also produce significant average return spreads of around 10% for the two- and three-year horizons. The average return spreads associated with MPEG are close to zero and not statistically significant over these horizons.

The evidence of return predictability is much weaker for the individual analyst-based ICCs. None of them produces a significant average return spread for the one-year horizon. The average spread ranges from −2.26% (*t*-stat of −0.42) for MPEG to 4.52% (*t*-stat of 0.95) for Gordon. For the two- and three-year horizons, OJ and Gordon do produce significant average return spreads, but both the magnitudes (ranging from 7.15% to 8.35%) and the *t*-stats (from 1.95 to 2.43) of the spreads are smaller than those associated with the corresponding model-based ICCs (spreads from 10.52% to 11.23% and *t*-stats from 5.24 to 6.32). GLS and CT are associated with insignificant average return spreads for the two- and three-year horizons (spreads from 1.73% to 4.21% and *t*-stats from 0.55 to 1.28), whereas MPEG produces significant but negative average return spreads for the two- and three-year horizons (−8.47% with a *t*-stat of −1.99 and −8.25% with a *t*-stat of −2.26, respectively). Finally, unreported results show that the difference in the associated return spreads between the model-based and analyst-based ICCs is statistically significant for most of the individual ICCs and at almost all horizons.

In short, our finding that the model-based ICC is a more reliable predictor of future stock returns than the analyst-based ICC is not driven by any specific method to compute the ICC. The individual model-based ICCs are highly correlated with each other and each of them predicts realized returns more strongly than its analyst-based counterpart. These results, combined with the evidence from the previous sections, suggest that using the model-based earnings forecasts instead of analysts' forecasts can substantially improve the quality of ICC as a proxy for expected stock returns. In the next section, we use these new and improved ICC estimates to study the cross-sectional relations between a broad set of firm-level characteristics and expected stock returns. Many of these relations are difficult to investigate using the analyst-based ICC because they require expected return estimates on firms that are not followed by analysts.

5. Implications for the cross-sectional relation between firm characteristics and expected returns

A large body of literature in asset pricing and capital markets research in accounting aims to uncover what firm-level (risk) characteristics determine the cross-sectional variation in expected stock returns.[26] The vast majority of these studies rely on average realized returns as a proxy for expected returns in the asset pricing tests.[27] In this section, we investigate whether a set of firm characteristics that have been previously shown to predict the cross-section of average realized returns can also explain the cross-sectional variation of ex ante expected returns proxied by the model-based ICC. If the realized return differences associated with these characteristics represent systematic variation in ex ante expected returns, we should expect the differences to also show up in the model-based ICC.

We compute all characteristics based on information available prior to the end of June of each year and match them to the composite model-based ICC estimated at the end of June as well as to realized stock returns from July to June of next year. We examine a total of 14 firm characteristics. Market beta is estimated for each stock at the end of June of each year using the stock's previous 60 monthly returns (24-month minimum). Size is the natural logarithm of the end-of-June market equity. BE/ME is the natural logarithm of the ratio of book equity to market equity at the previous fiscal year end. Leverage is book value of debt divided by book equity. Idiosyncratic volatility is the standard deviation of the market model residuals estimated annually

[25] Again, the analyst-based MPEG shows notably lower correlations with the other analyst-based ICCs (between 0.13 and 0.51).

[26] The characteristics that have been investigated include market beta (Fama and MacBeth, 1973; Fama and French, 1992), size (Banz, 1981; Fama and French, 1992), book-to-market equity (Fama and French, 1992; Lakonishok et al., 1994), leverage (Bhandari, 1988; Fama and French 1992), idiosyncratic volatility (Ang et al., 2006, 2009; Hou and Loh, 2011), distress (Bharath and Shumway, 2008; Campbell et al., 2008), investments (Titman et al., 2004; Fama and French, 2006), asset growth (Cooper et al., 2008), accruals (Sloan, 1996; Hirshleifer et al., in press), net operating assets (Hirshleifer et al., 2004), analyst coverage (Diether et al., 2002; Hou and Moskowitz, 2005), analyst dispersion (Diether et al., 2002), earnings smoothness (Francis et al., 2004; McInnis, 2010), and accruals quality (Francis et al., 2005a; Core et al., 2008).

[27] One notable exception is the stream of research that relates firm characteristics to the cross-section of the analyst-based ICC. See, for example, Gebhardt et al. (2001), Francis et al. (2004), Botosan and Plumlee (2005), Gow et al. (2010), McInnis (2010), and Botosan et al. (2011).

at the end of June using the previous 60 monthly returns (24-month minimum). Distress is the default probability estimated using a dynamic logit model as in Campbell et al. (2008). CAPEX is capital expenditure divided by lagged total assets. Asset growth is the growth rate in total assets. Accruals is accruals divided by lagged total assets. NOA is net operating assets divided by lagged total assets. Analyst dispersion is the standard deviation of analysts' forecasts divided by the end-of-June stock price. Analyst coverage is the number of analysts following the firm as of June. Earnings smoothness is earnings volatility divided by operating cash flow volatility estimated using the previous ten years (five-year minimum) of data. Accruals quality is computed following Francis et al. (2005a) using a cross-sectional regression of accruals on lagged, current, and future cash flows plus the change in sales revenue and gross property, plant, and equipment.[28] Accruals quality is the standard deviation of the firm-level residuals over the previous five years. We multiply earnings smoothness and accruals quality by -1 so that higher values of those variables indicate smoother earnings and higher accruals quality.

To examine the cross-sectional relation between these firm-level characteristics and ex post realized returns as well as ex ante expected returns, we conduct annual firm-level Fama and MacBeth (1973) cross-sectional regressions of realized stock returns and the composite model-based ICC on the characteristics (sample period 1968–2008). Table 7 reports the time-series averages of the annual regression coefficients and their associated time-series Newey-West t-statistics.[29]

Panels A and B of Table 7 report the results of univariate regressions of realized returns (Panel A) and the model-based ICC (Panel B) on each characteristic (there are 14 separate univariate regressions reported in each panel). The results show that market beta is negatively related to both realized returns and the model-based ICC. The average coefficient on market beta is -0.0173 in the realized return regressions but is not significant (t-stat of -1.28). The coefficient in the ICC regressions is greater in magnitude (-0.0369) and is statistically significant (t-stat of -3.54).[30]

The relation between size and the model-based ICC is much stronger than the size effect in realized returns. The coefficient on size in the realized return regressions is -0.0066 with an insignificant t-statistic (-0.98). On the other hand, the coefficient in the ICC regressions is more than five times as large (-0.0376) and is statistically highly significant (t-stat of -13.52). Thus, investors demand high ex ante returns for holding small firms, but only a small fraction of these high ex ante expected returns is realized ex post in our sample period.[31]

The BE/ME effect is significant in both realized returns and the model-based ICC. The coefficient on BE/ME is 0.0555 (t-stat of 5.12) in the realized return regressions and 0.0475 (t-stat of 12.82) in the ICC regressions. Leverage is also positively related to both realized returns and the model-based ICC. The coefficient on leverage is 0.0060 (t-stat of 1.89) in the realized return regressions, compared with a coefficient of 0.0134 (t-stat of 12.73) in the ICC regressions.

Idiosyncratic volatility is negatively related to realized returns but positively related to the model-based ICC (coefficients of -0.2514 and 0.1466 in the realized return regressions and the ICC regressions, respectively), although neither relation is significant (t-stats of -1.19 and 1.43, respectively). Similarly, distress is also negatively related to realized returns and positively related to the model-based ICC (coefficients of -0.0087 and 0.0756, respectively).[32] The relation with realized returns is insignificant (t-stat of -0.34) but the relation with the model-based ICC is highly significant (t-stat of 5.07), which suggests that firms with higher default probabilities are expected to earn significantly higher returns ex ante.

The negative and significant relations between CAPEX, asset growth, accruals, and NOA with realized returns all carry over to the model-based ICC, although the economic magnitudes are attenuated. The coefficients in the realized return regressions are -0.1888 (t-stat of -2.04), -0.1102 (t-stat of -5.69), -0.1329 (t-stat of -2.71), and -0.1120 (t-stat of -4.50), respectively, compared with coefficients of -0.1352 (t-stat of -2.06), -0.0755 (t-stat of -7.58), -0.0669 (t-stat of -3.49), and -0.0571 (t-stat of -4.06), respectively, in the ICC regressions.

Analyst dispersion is negatively and significantly related to realized returns (coefficient of -0.1928 and t-stat of -7.36) but positively and significantly related to the model-based ICC (coefficient of 0.1010 and t-stat of 3.00). On the other hand, analyst coverage is negatively related to both realized returns and the model-based ICC (coefficients of -0.0018 and -0.0046 and t-stats of -1.09 and -4.77, respectively), although only the latter relation is significant. Thus, firms that are followed by fewer analysts or have greater analyst dispersion are expected by investors to earn higher returns.

We find little evidence that earnings smoothness and accruals quality are significantly related to either realized returns or the model-based ICC. The coefficients in the realized return regressions are 0.0088 (t-stat of 0.75) and 0.2751 (t-stat of 0.91), respectively. The corresponding coefficients in the ICC regressions are 0.0001 (t-stat of 0.02) and -0.2754 (t-stat of -1.67), respectively.

[28] All variables in this regression are scaled by lagged total assets. We estimate the regression each year for each Fama and French (1997) industry with at least 20 observations.

[29] Our results are similar if we estimate annual Fama and MacBeth (1973) cross-sectional rank regressions or pooled regressions with two-way clustered standard errors (firm and year).

[30] In general, the coefficients in the ICC regressions are associated with much higher t-statistics because the model-based ICC is substantially less volatile than realized returns.

[31] This result is consistent with the finding of Hou and van Dijk (2011) that negative cash flow shocks to small firms and positive cash flow shocks to big firms after the early 1980s cause the size effect in realized returns to be negligible even though there is a significant size effect in ex ante expected returns.

[32] Dichev (1998) finds a negative relation between distress and realized returns using Altman's Z-score or Ohlson's O-score. Our results are robust to using these alternative distress measures.

Table 7
Relation between firm characteristics, realized returns, and composite model-based ICC estimates, Fama-MacBeth (1973) regressions, 1968–2008.

Beta	Size	BE/ME	Leverage	Idiosyncratic volatility	Distress	CAPEX	Asset growth	Accruals	NOA	Analyst dispersion	Analyst coverage	Earnings smoothness	Accruals quality
Panel A: Univariate regressions of realized returns on firm characteristics													
−0.0173	−0.0066	0.0555	0.0060	−0.2514	−0.0087	−0.1888	−0.1102	−0.1329	−0.1120	−0.1928	−0.0018	0.0088	0.2751
−1.28	*−0.98*	*5.12*	*1.89*	*−1.19*	*−0.34*	*−2.04*	*−5.69*	*−2.71*	*−4.50*	*−7.36*	*−1.09*	*0.75*	*0.91*
Panel B: Univariate regressions of composite model-based ICC on firm characteristics													
−0.0369	−0.0376	0.0475	0.0134	0.1466	0.0756	−0.1352	−0.0755	−0.0669	−0.0571	0.1010	−0.0046	0.0001	−0.2754
−3.54	*−13.52*	*12.82*	*12.73*	*1.43*	*5.07*	*−2.06*	*−7.58*	*−3.49*	*−4.06*	*3.00*	*−4.77*	*0.02*	*−1.67*
Panel C: Multivariate regressions of composite model-based ICC on firm characteristics													
−0.0343	−0.0350	0.0128											
−7.50	*−12.56*	*2.96*											
−0.0344	−0.0351	0.0099	0.0021										
−7.25	*−12.34*	*1.88*	*2.19*										
−0.0005	−0.0466	0.0000		−0.6662									
−0.04	*−10.45*	*0.00*		*−4.28*									
−0.0353	−0.0353	0.0136			−0.0081								
−8.66	*−13.27*	*3.39*			*−1.79*								
−0.0345	−0.0349	0.0129				0.0570							
−7.67	*−12.63*	*3.17*				*1.01*							
−0.0341	−0.0347	0.0115					−0.0088						
−7.51	*−12.32*	*2.58*					*−1.56*						
−0.0344	−0.0350	0.0119						−0.0151					
−7.69	*−12.30*	*2.66*						*−1.40*					
−0.0346	−0.0350	0.0120							−0.0205				
−7.39	*−12.18*	*2.63*							*−2.17*				
−0.0376	−0.0239	0.0069								0.0867			
−6.28	*−5.36*	*1.52*								*2.80*			
−0.0420	−0.0378	0.0041									0.0030		
−6.73	*−4.93*	*0.84*									*2.59*		
−0.0353	−0.0335	0.0070										0.0215	
−8.22	*−12.48*	*1.58*										*8.37*	
−0.0325	−0.0350	0.0012											0.4450
−6.16	*−11.55*	*0.21*											*3.70*
−0.0017	−0.0477	−0.0047	0.0163	−0.5552	−0.0962	0.0412	0.0308	0.0070	−0.0595	0.0621	0.0029	0.0058	0.1968
−0.13	*−5.50*	*−0.49*	*6.44*	*−2.46*	*−3.93*	*1.96*	*5.46*	*0.75*	*−6.40*	*2.56*	*3.45*	*1.71*	*2.88*

This table reports the time-series averages of the coefficients (and their associated time-series Newey-West *t*-statistics in *italics*) from annual Fama-MacBeth (1973) cross-sectional regressions of individual stocks' ex post realized returns (Panel A) or composite model-based ICC (Panels B and C) on various firm-level characteristics. *Beta* is the market beta estimated for each stock at the end of June of each year, using the stock's previous 60 monthly returns (24-month minimum). *Size* is the natural logarithm of the market equity at the end of June. *BE/ME* is natural logarithm of the ratio of book equity to market equity at the previous fiscal year end. *Leverage* is book value of debt divided by book equity. *Idiosyncratic volatility* is the standard deviation of the market model residuals estimated annually at the end of June using the previous 60 monthly returns (24-month minimum). *Distress* is the default probability estimated using a dynamic logit model as in Campbell et al. (2008). *CAPEX* is capital expenditure divided by lagged total assets. *Asset growth* is the growth rate in total assets. *Accruals* is accruals divided by lagged total assets. *NOA* is net operating assets divided by lagged total assets. *Analyst dispersion* is the standard deviation of analysts' forecasts divided by the end-of-June stock price. *Analyst coverage* is the number of analysts following the firm as of June. *Earnings smoothness* is earnings volatility divided by operating cash flow volatility estimated using the previous ten years (five-year minimum) of data. *Accruals quality* is computed following Francis et al. (2005a) using a cross-sectional regression of accruals on lagged, current, and future cash flows plus the change in sales revenue and gross property, plant, and equipment (all variables in the regression are scaled by lagged total assets). We estimate the regression each year for each Fama and French (1997) industry with at least 20 observations. Accruals quality is the standard deviation of the firm-level residuals over the previous five years. We multiply earnings smoothness and accruals quality by −1 so that higher values of those variables indicate smoother earnings and higher accruals quality.

In sum, the univariate Fama-MacBeth regression results suggest that investors demand high ex ante returns for holding firms with small market capitalization, high BE/ME, high leverage, high default probabilities, low investments (CAPEX or asset growth), low accruals, low NOA, high analyst dispersion, or low analyst coverage—possibly because these firms are riskier. For high BE/ME, low investments, low accruals, or low NOA firms, their high ex ante expected returns are also realized ex post in our sample period. The same cannot be said about small firms, or firms with high leverage, high default probabilities, high analyst dispersion, or low analyst coverage, as the coefficients on these characteristics are either insignificant or significant but with the opposite signs in the realized return regressions. For idiosyncratic volatility, earnings smoothness, and accruals quality, we fail to find significant effects in either ex ante expected returns or ex post realized returns, which suggests that these three characteristics are unlikely to be risk proxies.

To further examine the relation between firm-level characteristics and ex ante expected returns, Panel C of Table 7 reports the results of multivariate regressions of the model-based ICC on multiple characteristics at the same time. The first model includes market beta, size, and BE/ME in one regression. The results confirm the univariate findings that the model-based ICC is negatively and significantly related to beta and size, and positively and significantly related to BE/ME. The coefficients on beta and size retain their magnitudes and statistical significance of the univariate regressions, but the coefficient on BE/ME, though still significant, is much smaller in magnitude than the univariate coefficient.

The next 11 models use market beta, size, and BE/ME as controls and add other characteristics one at a time to study their incremental contribution in explaining the variation in the model-based ICC. The regression results show that even after controlling for the beta, size, and BE/ME effects, the model-based ICC continues to be positively and significantly related to leverage and analyst dispersion, and negatively and significantly related to NOA. These results are consistent with the univariate findings. On the other hand, the model-based ICC is now positively and significantly related to analyst coverage, earnings smoothness, and accruals quality, negatively and significantly related to idiosyncratic volatility, and insignificantly related to distress, CAPEX, asset growth, and accruals. These results are different from the univariate findings. Of particular interest are earnings smoothness and accruals quality, for which the univariate regressions fail to uncover significant effects in either realized returns or the model-based ICC. The positive and significant coefficients in the multivariate ICC regressions suggest that after controlling for beta, size, and BE/ME, firms with smoother earnings or higher accruals quality are expected to earn higher returns, a result that seems inconsistent with the risk-based interpretations of these two characteristics. The last model of Panel C is a ‘‘kitchen sink’’ regression in which we include all 14 characteristics.[33] The regression results show that, when considered simultaneously, leverage, CAPEX, asset growth, analyst dispersion, analyst coverage, and accruals quality are positively and significantly related to the model-based ICC; size, idiosyncratic volatility, distress, and NOA are negatively and significantly related to the model-based ICC; the relations between beta, BE/ME, accruals, and earnings smoothness and the model-based ICC are insignificant.

Taken together, the results in Table 7 show that inferences about the cross-section of expected returns are sensitive to the choice of expected return proxy (average ex post realized returns vs. ex ante model-based ICC).

6. Additional robustness tests

In this section, we analyze the performance of the model-based earnings forecasts and the model-based ICC relative to their analyst-based counterparts for different subsamples of firms and different subperiods. We also perform additional tests to ensure that our main findings are robust to alternative specifications of the cross-sectional earnings model and to adjusting analysts' forecasts for the predictable component of their bias.

6.1. Performance of the model-based forecasts and ICC vs. analysts' forecasts and ICC for subsamples and subperiods

In this subsection, we investigate whether there is systematic cross-sectional variation in the relative performance of the model-based earnings forecasts and the model-based ICC vs. analysts' forecasts and the analyst-based ICC. In particular, we are interested in whether the relative performance is related to the information environment of a firm.

At the end of June of each year we sort firms into three groups (‘‘Low’’, ‘‘Medium’’, and ‘‘High’’) based on one of the following firm-level characteristics: size, age, inverse of idiosyncratic volatility, analyst coverage, earnings smoothness, accruals quality, and past 12 months' return. Firms in the ‘‘Low’’ group can be generally characterized as having a poorer information environment and being more difficult to value (see, e.g., Francis et al., 2004, 2005a; Zhang, 2006; Bradshaw et al., 2011; Guay et al., 2011). We then compute, for each group separately, the mean and median forecast bias, forecast accuracy, and ERC for the model-based and analysts' forecasts, as well as the average 10-1 realized return spreads associated with the composite model-based and analyst-based ICCs.

In addition to the cross-sectional analysis, we are also interested in the variation of the relative performance over time. We divide the common sample period 1982–2006 (for which both types of forecasts and their associated ICCs are available) into three subperiods of approximately equal length (1982–1990, 1991–1998, and 1999–2006) and compare the performance of the model-based earnings forecasts and ICC with that of analysts' forecasts and ICC for each subperiod.

[33] Due to the availability of the analyst dispersion variable, this regression is restricted to the 1982–2008 period and to firms that are followed by at least two analysts.

Table 8
Relation between composite ICC estimates and future realized returns, subsamples and subperiods.

10-1 realized return spread ($r_{t,t+3}$), composite model-based ICC vs. composite analyst-based ICC

Cross-sectional subsamples		Size		Age		Inverse of idiosyncratic volatility		Analyst coverage		Earnings smoothness		Accruals quality		Past return	
Low	Model	0.1274	*7.78*	0.1565	*5.54*	0.1551	*6.33*	0.1655	*6.76*	0.1082	*3.62*	0.1506	*5.18*	0.1308	*6.92*
	Analysts	0.0000	*0.00*	0.0334	*0.95*	0.0265	*0.74*	0.0329	*1.27*	0.0236	*0.76*	−0.0012	*−0.06*	0.0068	*0.20*
	Difference	0.1274	*4.50*	0.1231	*5.93*	0.1286	*5.91*	0.1326	*4.78*	0.0846	*2.79*	0.1517	*5.53*	0.1240	*5.14*
Medium	Model	0.1768	*6.54*	0.1419	*5.96*	0.0793	*4.17*	0.1465	*5.00*	0.0955	*3.84*	0.0822	*3.33*	0.1075	*4.74*
	Analysts	0.0352	*1.22*	0.0337	*0.91*	0.0264	*0.83*	0.0595	*1.47*	0.0351	*1.54*	0.0339	*1.04*	0.0673	*2.18*
	Difference	0.1416	*6.54*	0.1082	*4.44*	0.0529	*2.74*	0.0870	*3.24*	0.0604	*3.36*	0.0483	*2.02*	0.0402	*1.71*
High	Model	0.1200	*4.48*	0.0573	*2.87*	0.0367	*1.97*	0.0773	*2.67*	0.0837	*4.90*	0.0584	*5.42*	0.1409	*5.37*
	Analysts	0.0757	*2.12*	0.0443	*2.01*	0.0498	*2.31*	0.0496	*1.43*	0.0393	*1.45*	0.0628	*3.27*	0.0926	*3.18*
	Difference	0.0443	*2.71*	0.0130	*0.76*	−0.0131	*−0.54*	0.0277	*1.59*	0.0444	*2.26*	−0.0044	*−0.25*	0.0483	*2.15*
Subperiods		1982–1990		1991–1998		1999–2006				Years around turning points		Other years			
	Model	0.0971	*4.03*	0.0874	*3.92*	0.2192	*7.39*			0.1715	*5.05*	0.1074	*4.59*		
	Analysts	0.0197	*0.37*	0.0099	*0.30*	0.1095	*1.41*			0.1181	*2.34*	−0.0032	*−0.17*		
	Difference	0.0774	*2.25*	0.0774	*5.28*	0.1097	*2.20*			0.0533	*1.96*	0.1107	*6.23*		

This table reports, for different subsamples of firms and different subperiods, the time-series averages of the 10-1 realized return spreads (and their Newey-West *t*-statistics) associated with the composite model-based and analyst-based ICCs as well as the differences between the two return spreads. At the end of June of each year, we sort firms into three groups ("Low", "Medium", and "High") based on one of the following firm-level characteristics: size, age, inverse of idiosyncratic volatility, analyst coverage, earnings smoothness, accruals quality, and past return. *Size* is the natural logarithm of the market equity at the end of June. *Age* is the number of years since the first appearance of the firm (PERMCO) on CRSP. *Idiosyncratic volatility* is the standard deviation of the market model residuals estimated annually at the end of June using the previous 60 monthly returns (24-month minimum). *Analyst coverage* is the number of analysts following the firm as of June. *Earnings smoothness* is earnings volatility divided by operating cash flow volatility estimated using the previous ten years (five-year minimum) of data. *Accruals quality* is computed following Francis et al. (2005a) using a cross-sectional regression of accruals on lagged, current, and future cash flows plus the change in sales revenue and gross property, plant, and equipment (all variables in the regression are scaled by lagged total assets). We estimate the regression each year for each Fama and French (1997) industry with at least 20 observations. Accruals quality is the standard deviation of the firm-level residuals over the previous five years. We multiply earnings smoothness and accruals quality by −1 so that higher values of those variables indicate smoother earnings and higher accruals quality. *Past return* is the cumulative stock return over the past 12 months. In addition to the cross-sectional subsample analysis, we also divide the common sample period (1982–2006) into three subperiods of approximately equal length (1982–1990, 1991–1998, and 1999–2006). We also use the NBER business cycle peaks and troughs to identify periods of economic turning points and separate them out from the normal years.

We also examine the relative performance at economic turning points, as one could argue that analysts are better at identifying shifts in aggregate economic conditions and therefore should outperform the cross-sectional model at such times. We use the NBER business cycle peaks and troughs to identify periods of economic turning points and contrast the relative performance of the model and analysts during those periods with that of normal times.

To conserve space, we do not report the results on forecast bias, forecast accuracy, and ERC (they are available upon request). Those results show that our main findings that the model-based earnings forecasts are less biased, less accurate, and are associated with greater ERCs than analysts' forecasts hold for the majority of the subsamples and subperiods we consider. Moreover, the advantage of the model-based forecasts over analysts' forecasts in terms of forecast bias and ERC is greater for firms with a poorer information environment (smaller, younger firms, firms with higher idiosyncratic volatility, lower analyst coverage, more volatile earnings, lower accruals quality, or lower past returns), for the last subperiod (1999–2006), and during normal times.

Table 8 reports, for different subsamples and subperiods, the average 10-1 realized return spreads (and their time-series Newey-West *t*-statistics) associated with the composite model-based and analyst-based ICCs for the three-year holding period. The table shows that the composite model-based ICC is a more reliable return predictor than the composite analyst-based ICC for almost every cross-sectional subsample we consider. The average realized return spread associated with the composite model-based ICC is positive and significant in every single case. In contrast, the average spread associated with the composite analyst-based ICC, though positive most of the time, is generally much smaller in magnitude and is only significant for the biggest, oldest firms, firms with the lowest idiosyncratic volatility, the highest accruals quality, or the highest past return. The difference in the associated return spreads between the composite model-based and analyst-based ICCs is positive (and almost always statistically significant) for all but two subsamples (for firms with the lowest idiosyncratic volatility or highest accruals quality, the difference is insignificantly negative). Moreover, the difference is always greater for a "Low" subsample than for the corresponding "High" subsample, suggesting that the performance advantage of the model-based ICC over the analyst-based ICC as a return predictor is more pronounced among firms whose information environment is relatively poor. This result complements the aforementioned findings on forecast bias and ERC and points to a direct link between the quality of the underlying earnings forecasts and the reliability of the corresponding ICC estimates. Among firms for which the relative performance of the model-based earnings forecasts over analysts' forecasts is the strongest, the difference in return predictability between the model-based ICC and the analyst-based ICC is thus also the greatest.

In terms of relative performance over time, Table 8 shows that the composite model-based ICC is associated with a positive and significant average realized return spread in every subperiod. On the other hand, the average spread associated with the composite analyst-based ICC is only significant around business cycle turning points. The difference in the associated return spreads between the two ICCs is positive and statistically significant in every subperiod. In addition, the difference is greater for the last subperiod and during normal times.

Taken together, these subsample and subperiod results suggest that for firms that have a poorer information environment (smaller, younger firms, firms with higher idiosyncratic volatility, lower analyst coverage, more volatile earnings, lower accruals quality, or lower past returns) and during normal times, one can benefit the most from using the model-based earnings forecasts and their associated ICC.

6.2. Extensions of the cross-sectional earnings model

We have explored many extensions of our baseline cross-sectional earnings model (Eq. (1)). For example, we have examined alternative models where we include additional earning predictors, such as market value of the firm, book equity, net operating assets, capital expenditures, R&D, and firm age. Many of these variables show up significantly in the regressions. On the other hand, the overall explanatory power of these alternative models is very close to that of the baseline model. In addition, both the performance of the earnings forecasts generated by the alternative models and the performance of the resulting ICC are very similar to those associated with the baseline model. To conserve space, these robustness results are not reported.

One possible concern about our baseline model is that it may potentially be dominated by firms with extreme dollar earnings. To address this concern, we re-estimate the model by scaling earnings and the other level variables using lagged total assets.[34] When we compare the coefficient estimates of the scaled earnings regressions to those of the baseline earnings level regressions, we see that most of the variables retain their signs and significance. In terms of explanatory power, the average regression R^2's are 0.52, 0.34, and 0.25 for the one-, two-, and three-year ahead regressions, respectively, suggesting that a sizable fraction of the variation across firms in (scaled) earnings is captured by the model.[35] The results on forecast bias, forecast accuracy, and ERC are very similar to those based on the earnings level regressions. The model-based earnings forecasts based on the scaled earnings regressions are again less biased, less accurate, and are associated with greater announcement and annual ERCs than analysts' forecasts. Finally, the return predictive power of the ICC based on the scaled earnings regressions is on par with—if not stronger than—that of the ICC based on the earnings level regressions, and much stronger than the predictive power of the analyst-based ICC. The average 10-1 realized return spreads associated with the composite model-based ICC based on the scaled earnings regressions are 10.04%, 13.39%, and 13.61%, respectively, for the one-, two-, and three-year horizons, compared with 10.62%, 11.81%, and 12.03% associated with the composite model-based ICC based on the earnings level regressions, and 3.91%, 4.86%, and 4.53% associated with the composite analyst-based ICC. The difference in the associated return spreads between the composite model-based ICC based on the scaled earnings regressions and the analyst-based ICC is again significant at all horizons. Hence, our key findings are robust to using scaled earnings regressions instead of earnings level regressions to generate earnings forecasts.

We have also examined whether adding the latest consensus analyst forecast to the cross-sectional model could improve the performance of the model-based forecasts and the model-based ICC. Unreported results show that the average coefficient on the analyst forecast variable is economically large (0.4383, 0.3367, and 0.3458 for the one-, two-, and three-year ahead regressions, respectively) and statistically significant (*t*-stats of 8.32, 5.99, and 8.03, respectively) at all three horizons. The introduction of this variable also causes the coefficients on other variables such as lagged earnings to go down in magnitude. However, the average regression R^2 of the extended model (0.87, 0.80, and 0.72 for the one-, two-, and three-year ahead regressions, respectively) is almost identical to that of the baseline model, which suggests that adding analysts' forecasts does not improve the overall explanatory power of the cross-sectional model. The earnings forecasts based on the extended model exhibit higher levels of forecast bias, similar levels of forecast accuracy, and lower levels of ERC than the forecasts produced by the baseline model; they are now equally biased, less accurate, and are associated with lower ERCs than analysts' forecasts. Furthermore, the model-based ICC estimated based on the earnings forecasts from the extended model also produces weaker return predictability results. The average 10-1 realized return spreads associated with the composite ICC based on the extended model are 4.79%, 8.78%, and 9.32% for the one-, two-, and three-year horizons, respectively, which are smaller than the spreads associated with the composite ICC based on the baseline model (10.62%, 11.81%, and 12.03%, respectively). Moreover, the average return spread is not significantly different from that associated with the composite analyst-based ICC (3.91%, 4.86%, and 4.53% for the one-, two-, and three-year horizons, respectively) at the one- and two-year horizons. In sum, our results suggest that incorporating analysts' forecasts in the cross-sectional earnings model offers very little to enhance the performance of the model-based earnings forecasts and the model-based ICC.

[34] The scaling can also create influential observations when total assets are close to zero. To prevent these observations from dominating the regressions, we exclude firms with less than $5 million in assets. We obtain similar results when we use market equity, book equity, sales, or net operating assets as the deflator.

[35] These R^2's are in line with those reported in Hou and van Dijk (2011) but are considerably smaller than the R^2's for the earnings level regressions (0.86, 0.81, and 0.78 for the one-, two-, and three-year ahead regressions), especially over longer horizons. This is not surprising since the scaled regressions overweigh smaller firms for which scaled earnings are more extreme and harder to predict.

6.3. Adjusting for bias in analysts' forecasts

Several recent studies (e.g., Guay et al., 2011; Mohanram and Gode, 2011) demonstrate that the bias in analysts' earnings forecasts is in part predictable. They propose methods to adjust analysts' forecasts for the predictable component of their bias and thus enhance the reliability of the analyst-based ICC. To investigate how their approach affect our results, we follow Mohanram and Gode (2011) to adjust analysts' forecasts and then evaluate the resulting adjusted analysts' forecasts and their associated ICC against the model-based forecasts and the model-based ICC.[36]

In unreported results, we find that after the adjustment, the mean bias of analysts' forecasts is positive (indicating pessimistic forecasts) but not statistically significant at any horizon, which suggests that the adjustment procedure is effective in removing the mean optimistic bias in analysts' forecasts. On the other hand, the median bias of the adjusted analysts' forecasts is considerably larger than the mean bias, and is significant for the one- and three-year ahead forecasts and the weighted forecast and marginally significant for the two-year ahead forecast. Thus, the improved performance of the adjusted analysts' forecasts in terms of mean forecast bias seems to come at the expense of median forecast bias. The mean (median) bias of the model-based forecasts (for the common sample of firm-year observations for which both the model-based forecasts and the adjusted analysts' forecasts are available) is negative (positive) at all horizons and is not significant except for the median bias of the one-year ahead forecast. In addition, the mean (median) bias of the model-based forecasts is almost always smaller in magnitude than the mean (median) bias of the adjusted analysts' forecasts, and the difference in mean (median) bias is significant for the one-year ahead forecast (the three-year ahead forecast and the weighted forecast).

The adjusted analysts' forecasts are on average more accurate than the model-based forecasts, similar to the results for the unadjusted analysts' forecasts. However, only the difference in the mean accuracy between the adjusted analysts' forecasts and the model-based forecasts is significant at all horizons, whereas the difference in the median accuracy between the two types of forecasts is not significant at any horizon. The announcement and annual ERCs associated with the adjusted analysts' forecasts are still lower than those associated with the model-based forecasts, with the differences being economically large and statistically highly significant at all horizons. Thus, the adjustment procedure does not improve the performance of analysts' forecasts relative to the model-based forecasts in terms of ERC.

The ICC based on the adjusted analysts' forecasts shows slightly stronger return predictability than the ICC based on the unadjusted analysts' forecast, but still significantly worse return predictability than the model-based ICC. Over 1983–2006, the average 10-1 realized return spreads associated with the composite ICC based on the adjusted analysts' forecasts are 7.92%, 5.70%, and 4.35% for the one-, two-, and three-year horizons, respectively, which compare favorably to the 4.43%, 4.93%, and 4.36% spreads associated with the composite ICC based on the unadjusted analysts' forecasts, although the difference is not significant at any horizon. On the other hand, they are still considerably smaller than the spreads associated with the composite model-based ICC over the same period (11.47%, 13.21%, and 13.24% for the one-, two-, and three-year horizons, respectively), with the difference being significant at the two- and three-year horizons.

The bottom line is that the evidence on whether the adjustment improves the performance of analysts' forecasts and the analyst-based ICC relative to their model-based counterparts is mixed. An important drawback of the adjustment approach is that the additional data requirement it imposes results in a significantly smaller sample size. The number of firm-year observations for which we can estimate the composite ICC based on the adjusted analysts' forecasts is 49,290, which is about half of the coverage of the composite ICC based on the unadjusted analysts' forecasts (96,974) and less than 30% of the coverage of the composite model-based ICC (172,417).

7. Conclusions

In this paper, we use earnings forecasts from a cross-sectional model instead of analysts' forecasts to proxy for cash flow expectations and estimate the implied cost of capital (ICC) for more than 170,000 firm-year observations over 1968–2008. Our cross-sectional earnings model captures significant variation in future earnings performance across firms using ex ante publicly available information. More importantly, the model produces earnings forecasts that are superior to consensus analyst forecasts in terms of coverage, forecast bias, and earnings response coefficient. We show that the ICC estimated using the model-based earnings forecasts is a more reliable proxy for expected returns than the ICC based on analysts' forecasts.

Our results are robust to different specifications of the cross-sectional earnings model, to adjusting for the predictable component of analysts' forecast bias, and to different methods used to estimate the ICC. In addition, we show that the relative performance of the model-based earnings forecasts and ICC over analysts' forecasts and the analyst-based ICC is more pronounced for firms with a relatively poor information environment, thus highlighting the group of firms for which one can benefit the most from using our model-based earnings forecasts and the associated ICC.

We use our new and improved ICC estimates to re-examine the cross-sectional relation between a broad set of firm-level characteristics and ex ante expected returns. Our results suggest that using the model-based ICC instead of realized returns as a proxy for expected returns has a significant impact on inferences about the cross-section of expected returns.

[36] See Mohanram and Gode (2011) for details of the adjustment procedure.

Our model-based ICC can also be used to investigate a host of other issues in asset pricing, corporate finance, and financial accounting. We leave these issues for future work.

Appendix A

See Table A1.

Table A1
Detailed descriptions of individual ICC estimates.

ICC	Formula and assumptions	Source
GLS	$M_t = B_t + \sum_{\kappa=1}^{11} \frac{E_t[(ROE_{t+\kappa}-R) \times B_{t+\kappa-1}]}{(1+R)^{\kappa}} + \frac{E_t[(ROE_{t+12}-R) \times B_{t+11}]}{R \times (1+R)^{11}}$, where M_t is the market equity in year t, R is the implied cost of capital (ICC), B_t is the book equity, $E_t[]$ denotes market expectations based on information available in year t, and $(ROE_{t+k}-R) \times B_{t+k-1}$ is the residual income in year $t+k$, defined as the difference between the return on book equity and the ICC multiplied by the book equity in the previous year. We estimate the expected ROE in years $t+1$ to $t+3$ using the model-based or analysts' earnings forecasts and book equity determined based on clean surplus accounting ($B_{t+k}=B_{t+k-1}+E_{t+k}-D_{t+k}$, where E_{t+k} is the earnings in year $t+k$, D_{t+k} is the dividend in year $t+k$, computed using the current dividend payout ratio for firms with positive earnings, or using current dividends divided by $0.06 \times$ total assets as an estimate of the payout ratio for firms with negative earnings). After year $t+3$, we assume that the expected ROE mean-reverts to the historical industry median value by year $t+11$, after which point the residual income becomes a perpetuity. Following Gebhardt et al. (2001), we exclude loss firms when calculating the industry median ROE.	Gebhardt et al. (2001)
CT	$M_t = B_t + \sum_{\kappa=1}^{5} \frac{E_t[(ROE_{t+\kappa}-R) \times B_{t+k-1}]}{(1+R)^{\kappa}} + \frac{E_t[(ROE_{t+5}-R) \times B_{t+4}](1+g)}{(R-g) \times (1+R)^5}$, where M_t is the market equity in year t, R is the implied cost of capital (ICC), B_t is the book equity, $E_t[]$ denotes market expectations based on information available in year t, and $(ROE_{t+k}-R) \times B_{t+k-1}$ is the residual income in year $t+k$, defined as the difference between the return on book equity and the ICC multiplied by the book equity in the previous year. We estimate the expected ROE in years $t+1$ to $t+5$ using the model-based or analysts' earnings forecasts and book equity determined based on clean surplus accounting ($B_{t+k}=B_{t+k-1}+E_{t+k}-D_{t+k}$, where E_{t+k} is the earnings in year $t+k$, D_{t+k} is the dividend in year $t+k$, computed using the current dividend payout ratio for firms with positive earnings, or using current dividends divided by $0.06 \times$ total assets as an estimate of the payout ratio for firms with negative earnings). For analysts' forecasts, we impute the four- and five-year ahead forecasts using the long-term growth forecast and the three-year ahead forecast. Following Claus and Thomas (2001), g is set to the current risk-free rate minus 3%.	Claus and Thomas (2001)
OJ	$R = A + \sqrt{A^2 + \frac{E_t[E_{t+1}]}{M_t} \times (g-(\gamma-1))}$, where $A = 0.5\left((\gamma-1) + \frac{E_t[D_{t+1}]}{M_t}\right), \quad g = 0.5\left(\frac{E_t[E_{t+3}]-E_t[E_{t+2}]}{E_t[E_{t+2}]} + \frac{E_t[E_{t+5}]-E_t[E_{t+4}]}{E_t[E_{t+4}]}\right)$, M_t is the market equity in year t, R is the implied cost of capital (ICC), $E_t[]$ denotes market expectations based on information available in year t, E_{t+1} is the earnings in year $t+1$, and D_{t+1} is the dividend in year $t+1$, computed using the current dividend payout ratio for firms with positive earnings, or using current dividends divided by $0.06 \times$ total assets as an estimate of the payout ratio for firms with negative earnings. g is the short-term growth rate. We follow Gode and Mohanram (2003) and use the average of forecasted near-term growth and five-year growth as an estimate of g. γ is the perpetual growth rate in abnormal earnings beyond the forecast horizon. It is set to the current risk-free rate minus 3%.	Ohlson and Juettner-Nauroth (2005)
MPEG	$M_t = \frac{E_t[E_{t+2}] + R \times E_t[D_{t+1}] - E_t[E_{t+1}]}{R^2}$, where M_t is the market equity in year t, R is the implied cost of capital (ICC), $E_t[]$ denotes market expectations based on information available in year t, E_{t+1} and E_{t+2} are the earnings in years $t+1$ and $t+2$, respectively, D_{t+1} is the dividend in year $t+1$, computed using the current dividend payout ratio for firms with positive earnings, or using current dividends divided by $0.06 \times$ total assets as an estimate of the payout ratio for firms with negative earnings.	Easton (2004)
Gordon	$M_t = \frac{E_t[E_{t+1}]}{R}$ where M_t is the market equity in year t, R is the implied cost of capital (ICC), $E_t[]$ denotes market expectations based on information available in year t, and E_{t+1} is the earnings in year $t+1$. This is a special case of the finite-horizon version of the Gordon growth model.	Gordon and Gordon (1997)

This table provides detailed descriptions of the five individual ICC estimates used in this paper. For the model-based ICC, firm valuation is expressed in terms of market equity. The analyst-based ICC is estimated using equivalent formulas expressed on a per share basis.

References

Abarbanell, J., Bushee, B., 1997. Fundamental analysis, future earnings, and stock prices. Journal of Accounting Research 35, 1–24.
Allee, K., 2010. Estimating Cost of Equity Capital with Time-Series Forecasts of Earnings. Working Paper. Michigan State University.
Ang, A., Hodrick, R., Xing, Y., Zhang, X., 2006. The cross-section of volatility and expected returns. Journal of Finance 61, 259–299.
Ang, A., Hodrick, R., Xing, Y., Zhang, X., 2009. High idiosyncratic volatility and low returns: international and further U.S. evidence. Journal of Financial Economics 91, 1–23.
Ashbaugh-Skaife, H., Collins, D., Kinney Jr, W., Lafond, R., 2009. The effect of SOX internal control deficiencies on firm risk and cost of equity. Journal of Accounting Research 47, 1–43.
Ball, R., Brown, P., 1968. An empirical evaluation of accounting income numbers. Journal of Accounting Research 6, 159–178.
Banz, R., 1981. The relationship between return and market value of common stocks. Journal of Financial Economics 9, 3–18.
Bhandari, L., 1988. Debt/equity ratio and expected common stock returns: empirical evidence. Journal of Finance 43, 507–528.
Bharath, S., Shumway, T., 2008. Forecasting default with the Merton distance to default model. Review of Financial Studies 21, 1339–1369.
Blume, M., Friend, I., 1973. A new look at the capital asset pricing model. Journal of Finance 28, 19–33.
Botosan, C., 1997. Disclosure level and the cost of equity capital. The Accounting Review 72, 323–349.
Botosan, C., Plumlee, M., 2002. A re-examination of disclosure level and the expected cost of equity capital. Journal of Accounting Research 40, 21–40.
Botosan, C., Plumlee, M., 2005. Assessing alternative proxies for the expected risk premium. The Accounting Review 80, 21–53.
Botosan, C., Plumlee, M., Wen, H., 2011. The relation between expected returns, realized returns, and firm risk characteristics. Contemporary Accounting Research 28, 1085–1122.
Bradshaw, M., Drake, M., Myers, J., Myers, L., 2011. A Re-Examination of Analysts' Superiority Over Time-Series Forecasts of Annual Earnings. Working Paper. Boston College.
Brown, L., 1993. Earnings forecasting research: its implications for capital markets research. International Journal of Forecasting 9, 295–320.
Brown, L., Hagerman, R., Griffin, P., Zmijewski, M., 1987. An evaluation of alternative proxies for the market's assessment of unexpected earnings. Journal of Accounting and Economics 9, 159–193.
Brown, L., Rozeff, M., 1978. The superiority of analyst forecasts as measures of expectations: evidence from earnings. Journal of Finance 33, 1–16.
Campbell, J., 1991. A variance decomposition for stock returns. Economic Journal 101, 157–179.
Campbell, J., Hilscher, J., Szilagyi, J., 2008. In search of distress risk. Journal of Finance 63, 2899.
Chava, S., Purnanandam, A., 2010. Is default risk negatively related to stock returns? Review of Financial Studies 23, 2523–2559.
Claus, J., Thomas, J., 2001. Equity premia as low as three percent? Evidence from analysts' earnings forecasts for domestic and international stock markets. Journal of Finance 56, 1629–1666.
Collins, D., Kothari, S., Shanken, J., Sloan, R., 1994. Lack of timeliness and noise as explanations for the low contemporaneous return–earnings association. Journal of Accounting and Economics 18, 289–324.
Cooper, M., Gulen, H., Schill, M., 2008. Asset growth and the cross-section of stock returns. Journal of Finance 63, 1609–1651.
Core, J., Guay, W., Verdi, R., 2008. Is accruals quality a priced risk factor? Journal of Accounting and Economics 46, 2–22.
Dhaliwal, D.A., Krull, L., Li, O., Moser, W., 2005. Dividend taxes and implied cost of equity capital. Journal of Accounting Research 43, 675–708.
Dichev, I., 1998. Is the risk of bankruptcy a systematic risk? Journal of Finance 53, 1131–1147.
Diether, K., Malloy, C., Scherbina, A., 2002. Differences of opinion and the cross section of stock returns. Journal of Finance 57, 2113–2141.
Dugar, A., Nathan, S., 1995. The effect of investment banking relationships on financial analysts' earnings forecasts and investment recommendations. Contemporary Accounting Research 12, 131–160.
Easton, P., 2004. PE ratios, PEG ratios, and estimating the implied expected rate of return on equity capital. The Accounting Review 79, 73–95.
Easton, P., 2007. Estimating the cost of capital implied by market prices and accounting data. Foundations and Trends in Accounting 2, 241–364.
Easton, P., Monahan, S., 2005. An evaluation of accounting-based measures of expected returns. The Accounting Review 80, 501–538.
Easton, P., Sommers, G., 2007. Effect of analysts' optimism on estimates of the expected rate of return implied by earnings forecasts. Journal of Accounting Research 45, 983–1015.
Easton, P., Zmijewski, M., 1989. Cross-sectional variation in the stock market response to accounting earnings announcements. Journal of Accounting and Economics 11, 117–141.
Elton, E., 1999. Expected return, realized return, and asset pricing tests. Journal of Finance 54, 1198–1220.
Fama, E., French, K., 1992. The cross-section of expected stock returns. Journal of Finance 47, 427–465.
Fama, E., French, K., 1993. Common risk factors in the returns on stocks and bonds. Journal of Financial Economics 33, 3.
Fama, E., French, K., 1997. Industry costs of equity. Journal of Financial Economics 43, 153–193.
Fama, E., French, K., 2000. Forecasting profitability and earnings. Journal of Business 73, 161–175.
Fama, E., French, K., 2004. New lists: fundamentals and survival rates. Journal of Financial Economics 73, 229–269.
Fama, E., French, K., 2006. Profitability, investment and average returns. Journal of Financial Economics 82, 491–518.
Fama, E., MacBeth, J., 1973. Risk, return, and equilibrium: empirical tests. Journal of Political Economy 81, 607–636.
Francis, J., LaFond, R., Olsson, P., Schipper, K., 2004. Costs of equity and earnings attributes. The Accounting Review 79, 967–1010.
Francis, J., Lafond, R., Olsson, P., Schipper, K., 2005a. The market pricing of accruals quality. Journal of Accounting and Economics 39, 295–327.
Francis, J., Khurana, I., Pereira, R., 2005b. Disclosure incentives and effects on cost of capital around the world. The Accounting Review 80, 1125–1162.
Francis, J., Olsson, P., Oswald, D., 2000. Comparing the accuracy and explainability of dividend, free cash flow, and abnormal earnings equity value estimates. Journal of Accounting Research 38, 45–70.
Francis, J., Philbrick, D., 1993. Analysts' decisions as products of a multi-task environment. Journal of Accounting Research 31, 216–230.
Francis, J., Schipper, K., Vincent, L., 2005c. Earnings and dividend informativeness when cash flow rights are separated from voting rights. Journal of Accounting and Economics 39, 329–360.
Fried, D., Givoly, D., 1982. Financial analysts' forecasts of earnings: a better surrogate for market expectations. Journal of Accounting and Economics 4, 85–107.
Froot, K., Frankel, J., 1989. Forward discount bias: is it an exchange risk premium? Quarterly Journal of Economics 104, 139–161.
Gebhardt, W., Lee, C.M., Swaminathan, B., 2001. Toward an implied cost of capital. Journal of Accounting Research 39, 135–176.
Gordon, J., Gordon, M., 1997. The finite horizon expected return model. Financial Analysts Journal (May/June), 52–61.
Gow, I., Ormazabal, G., Taylor, D., 2010. Correcting for cross-sectional and time-series dependence in accounting research. The Accounting Review 85, 483–512.
Guay, W., Kothari, S., Shu, S., 2011. Properties of implied cost of capital using analysts' forecasts. Australian Journal of Management 36, 125–149.
Hail, L., Leuz, C., 2006. International differences in the cost of equity capital: do legal institutions and securities regulation matter? Journal of Accounting Research 44, 485–531.
Hail, L., Leuz, C., 2009. Cost of capital effects and changes in growth expectations around U.S. cross-listings. Journal of Financial Economics 93, 428–454.
Hayn, C., 1995. The information content of losses. Journal of Accounting and Economics 20, 125–153.
Hirshleifer, D., Hou, K., Teoh, S. The accrual anomaly: risk or mispricing? Management Science, doi:10.1287/mnsc.1100.1290, in press.
Hirshleifer, D., Hou, K., Teoh, S., Zhang, Y., 2004. Do investors overvalue firms with bloated balance sheets? Journal of Accounting and Economics 38, 297–331.

Hong, H., Lim, T., Stein, J., 2000. Bad news travels slowly: size, analyst coverage, and the profitability of momentum strategies. Journal of Finance 55, 265–295.
Hou, K., Loh, R., 2011. Have We Solved the Idiosyncratic Volatility Puzzle? Working Paper. Ohio State University.
Hou, K., Moskowitz, T., 2005. Market frictions, price delay, and the cross-section of expected returns. Review of Financial Studies 18, 981–1020.
Hou, K., Robinson, D., 2006. Industry concentration and average stock returns. Journal of Finance 61, 1927–1956.
Hou, K., van Dijk, M., 2011. Profitability Shocks and the Size Effect in the Cross-section of Expected Stock Returns. Working Paper. Ohio State University.
Hou, K., Zhang, Y., Zhuang, Z., 2011. Understanding the Variation in the Value Relevance of Earnings: A Return Decomposition Analysis. Working Paper. Ohio State University.
Hribar, P., Collins, D., 2002. Errors in estimating accruals: implications for empirical research. Journal of Accounting Research 40, 105–134.
Hribar, P., Jenkins, N., 2004. The effect of accounting restatements on earnings revisions and the estimated cost of capital. Review of Accounting Studies 9, 337–356.
La Porta, R., 1996. Expectations and the cross-section of stock returns. Journal of Finance 51, 1715–1742.
Lakonishok, J., Shleifer, A., Vishny, R., 1994. Contrarian investment, extrapolation, and risk. Journal of Finance 49, 1541–1578.
Lee, C., Ng, D., Swaminathan, B., 2009. Testing international asset pricing models using implied costs of capital. Journal of Financial and Quantitative Analysis 44, 307–335.
Lee, C., So, E., Wang, C., 2010. Evaluating Implied Cost of Capital Estimates. Working Paper. Stanford University.
Lev, B., Ohlson, J., 1982. Market-based empirical research in accounting: a review, interpretation, and extension. Journal of Accounting Research 20, 249–322.
Lin, H.-W., McNichols, M., 1998. Underwriting relationships, analysts' earnings forecasts and investment recommendations. Journal of Accounting and Economics 25, 101–127.
McInnis, J., 2010. Earnings smoothness, average returns, and implied cost of equity capital. The Accounting Review 85, 315–341.
McNichols, M., O'Brien, P., 1997. Self-selection and analyst coverage. Journal of Accounting Research 35, 167–199.
Mohanram, P., Gode, D., 2011. Evaluating Implied Cost of Equity Estimates After Removing Predictable Analyst Forecast Errors. Working Paper. Columbia University.
O'Brien, P., 1988. Analysts' forecasts as earnings expectations. Journal of Accounting and Economics 10, 53–83.
Ohlson, J., Juettner-Nauroth, B., 2005. Expected EPS and EPS growth as determinants of value. Review of Accounting Studies 10, 349–365.
Pástor, L., Sinha, M., Swaminathan, B., 2008. Estimating the intertemporal risk–return tradeoff using the implied cost of capital. Journal of Finance 63, 2859–2897.
Richardson, S., Sloan, R., Soliman, M., Tuna, I., 2006. The implications of accounting distortions and growth for accruals and profitability. The Accounting Review 81, 713–748.
Richardson, S., Tuna, I., Wysocki, P., 2010. Accounting anomalies and fundamental analysis: a review of recent research advances. Journal of Accounting and Economics 50, 410–454.
Sharpe, W., 1978. New evidence on the capital asset pricing model: discussion. Journal of Finance 33, 917–920.
Sloan, R., 1996. Do stock prices fully reflect information in accruals and cash flows about future earnings? The Accounting Review 71, 289–315.
Titman, S., Wei, K.C., Xie, F., 2004. Capital investments and stock returns. Journal of Financial and Quantitative Analysis 39, 677–700.
Wiedman, C., 1996. The relevance of characteristics of the information environment in the selection of a proxy for the market's expectations for earnings: an extension of Brown, Richardson, and Schwager [1987]. Journal of Accounting Research 34, 313–324.
Zhang, X.F., 2006. Information uncertainty and stock returns. Journal of Finance 61, 105–137.