Evaluating Cross-Sectional Forecasting Models for Implied Cost of Capital

by

Kevin Li
University of Toronto
kevin.li@rotman.utoronto.ca

and

Partha Mohanram*
University of Toronto
partha.mohanram@rotman.utoronto.ca

Abstract: The computation of implied cost of capital (ICC) is constrained by the fact that around half of all firms do not have analysts' earnings forecasts. Hou, van Dijk and Zhang (2012, HVZ) present a cross-sectional model to generate forecasts and compute ICC from these forecasts. However, the forecasts from the HVZ model perform worse than those from a naïve random walk model and show anomalous correlations with risk factors. We present two parsimonious alternatives to the HVZ model: the AR model based on the autoregression in earnings and the RIV model based on Feltham and Ohlson (1996). We show that both models outperform the HVZ model in terms of forecast bias, accuracy and earnings response coefficients. Further, the ICC metrics generated from the AR and RIV models outperform those from the HVZ model in terms of correlations with future returns and risk factors. We recommend that future research use these models to generate earnings forecasts.

---

*Corresponding author, Rotman School of Management, University of Toronto, 105 St. George Street, Toronto, Ontario M5S 3E6, Canada. We would like to thank Jacob Thomas, Franco Wong and seminar participants at Erasmus University, University of British Columbia and University of Toronto for helpful comments. All errors are our own.

1. Introduction

Cost of equity plays a central role in valuation, portfolio selection, and capital budgeting. Therefore, measuring and validating cost of equity metrics has been the subject of much research. Inferring cost of equity ex-post from realized returns is problematic because the correlation between expected returns and realized returns is weak (Elton, 1999). Prior research has often documented a weak or even non-existent relation between conventional measures of risk (e.g., β) and realized returns (Fama and French 1992). This has led to the use of implied cost of capital (ICC), which is the discount rate that equates current stock price to present value of expected future dividends.

Prior literature has taken different approaches towards measuring ICC. Gebhardt, Lee and Swaminathan (2001) and Claus and Thomas (2001) use variants of the residual income model to solve for the discount rate that equates price to the sum of book value and the present value of future abnormal earnings. Gode and Mohanram (2003) and Easton (2004) develop proxies based on the abnormal earnings growth model of Ohlson and Juettner-Nauroth (2005). The common feature of all of the approaches to measuring ICC is a reliance on analysts' forecasts. This causes two shortcomings for researchers looking to obtain a reliable proxy for expected returns. First, analyst forecasts are available only for a subset of firms, with almost half of all firms not having analyst coverage in most years. This problem is not trivial because most of the firms without analyst following are typically small and young firms - the kind of firms that would be of the greatest interest to researchers examining issues related to information asymmetry, earnings quality and disclosure where an ICC approach is used most often. Second, an extensive literature has shown that ICC proxies are unreliable showing weak correlations with future returns (Easton and Monahan 2005) and anomalous correlations with risk factors.

A recent paper by Hou, van Dijk and Zhang (2012), henceforth HVZ, offers an interesting approach towards addressing these shortcomings. HVZ run cross-sectional regressions using lagged information to estimate future earnings for horizons of one to five years. The model they use builds on models in Fama and French (2000, 2006) and regresses future earnings on total assets, dividends, earnings and accruals. They use the earnings forecasts from the model to generate ICC estimates based on the approaches in prior literature. HVZ show that their model addresses the shortcomings of relying on analyst forecasts, by providing reliable ICC estimates for a wide cross-section of firms. HVZ show that the model-based ICC generally outperforms the ICC derived from analysts' forecasts. Not surprisingly, the HVZ model has been used in recent research on accounting based valuation (Chang, Landsman and Monahan 2012) and ICC (e.g., Jones and Tuzel 2012; Lee, So and Wang 2011; Patatoukas 2011).

Given the growing attention to the HVZ model, it is imperative to test the model for the following reasons. First, HVZ test only one model and do not benchmark it against other cross-sectional models. Although they show that their ICC estimates are correlated with future returns at the portfolio level, they do not examine the relation between ICC estimates and future returns at the firm level like the prior studies (Gebhardt, Lee and Swaminathan 2001; Gode and Mohanram 2003). In addition, their ICC estimates show many anomalous relations with risk factors, including negative correlations with systematic and idiosyncratic risk.

Second, a recent paper by Gerakos and Gramacy (2013) shows that the HVZ model actually underperforms a naïve random walk model that simply sets future earnings to past earnings. However, a random walk model is impractical for many implied cost of capital metrics that anchor on estimates of short term growth. Further, the level of forecast errors reported in the HVZ model is rather high – the mean absolute error (scaled by price) for one-year-ahead earnings is 0.084 for firms with analyst coverage (Table 3 of HVZ, page 9). If one assumes an average P/E ratio of 12, this represents an absolute error that is on average equal to the estimate of earnings itself. More importantly, the HVZ model generates a much larger forecast errors for firms without analyst coverage where the need for a forecasting model is crucial. Our partition results indicate that the mean scaled absolute forecast error for one-year-ahead earnings for this group generated by the HVZ model is more than twice as large as the error for firms with analyst coverage.

The goal of this paper is to build better cross-sectional models to forecast future earnings. We present and test two parsimonious alternatives to the HVZ model. The first model (AR model) is a simple autoregressive model that forecasts earnings as a function of past earnings, allowing for the differential persistence of profits and losses. The second model (RIV model) is motivated by the residual income valuation models in Ohlson (1995), Feltham and Ohlson (1995, 1996) and incorporates book value and accruals in addition to earnings. We benchmark the HVZ model and our two proposed models against a naïve random walk (RW) model.

We test the HVZ model and the above three alternative models along the following dimensions. We first evaluate the four models on the basis of forecast accuracy and bias. We next look at the earnings response coefficients (ERC), measured as the correlation between forecast surprise and future returns. Finally, we examine the properties of ICC obtained by applying forecasts from these models to the commonly used ICC proxies, in terms of correlations with returns as well as correlations with risk factors.

We find that both the RIV model and the AR model outperform the HVZ model in terms of forecast accuracy, forecast bias and ERC. On average, the forecasts for the whole Compustat population from the RIV model are 28% to 35% more accurate than the forecasts from the HVZ model. The improvement is as large as 42% in small firms and firms without analyst coverage, where model-based forecasts are more relevant and important. On average, the ERCs of the RIV forecasts are twice as large as the ERCs of the HVZ forecasts, indicating that the RIV forecasts better represent market expectations. Consistent with Gramacy and Gerakos (2013), we find that the HVZ model significantly underperforms the naïve random walk model in terms of forecast accuracy, bias and ERC, in the full sample as well as in the subsamples of small firms and firms without analyst coverage. On average, the absolute forecast errors from the RW model are 21% to 28% smaller than the forecast errors from the HVZ model. The fact that the HVZ model cannot outperform the naïve random walk model is quite disconcerting.

We examine the correlation between ICCs derived from the forecasts and future returns along two dimensions – return spreads between extreme portfolios (quintiles) based on the ICC measures, and firm level regressions of future returns on ICC. We find that both the AR model and the RIV model perform better than the HVZ model in terms of portfolio return spreads. Firm-level regressions indicate that the HVZ model produces ICC metrics with the weakest correlation with future realized returns, while the ICC metrics from the RIV model show the strongest correlations with future returns.

We also examine the correlation between the ICC measures and risk factors, consistent with the analysis in Gebhardt, Lee and Swaminathan (2001), Gode and Mohanram (2003) and Botosan and Plumlee (2005). We find that the ICCs based on the RIV model show expected correlations with most risk factors. In contrast, the ICCs based on the HVZ model always show an anomalous negative correlation with systematic risk (β) and an insignificant correlation with idiosyncratic risk and analyst following.

To summarize, we provide two models (RIV and AR) that outperform the HVZ model on multiple dimensions – forecast accuracy, bias, ERC, correlation of ICC proxies with future returns and risk factors. In addition to their superior performance, both these models are grounded in prior theoretical and empirical research in accounting, and are relatively parsimonious. Among the two models, the RIV model performs marginally better on most dimensions. We recommend that future research use the RIV model as the appropriate cross-sectional model to forecast future earnings.

The rest of the paper is organized as follows. In section 2, we discuss the HVZ model and the alternative models developed in this paper. Section 3 discusses the data and empirical execution. Section 4 compares the models on forecast accuracy, bias and ERCs. Section 5 examines the properties of ICC estimates derived from the forecasts. Section 6 discusses the results of sensitivity analyses. Section 7 concludes with implications for future research.

2. The Models

2.1 The HVZ Model

The model developed in HVZ is an extension of the cross-sectional profitability models in Fama and French (2000, 2006), Hou and Robinson (2006), and Hou and van Dijk (2011). The model is specified as:

$$E_{t+\tau} = \alpha_0 + \alpha_1 {*} A_t + \alpha_2 {*} D_t + \alpha_3 {*} DD_t + \alpha_4 {*} E_t + \alpha_5 {*} NegE_t + \alpha_6 {*} AC_t \qquad (1)$$

where $E_{t+\tau}$ is earnings in year t+τ (τ=1 to 5), A_t is total assets, D_t is dividends, DD_t is an indicator variable for dividend paying firms, E_t is earnings, $NegE_t$ is an indicator variable for loss firms and AC_t is working capital accruals. The regression is estimated using the previous ten years of data, ensuring no look-ahead bias (i.e. the regression for one-year-ahead earnings in year t uses data from year t-10 to t-1, or two-year-ahead regression uses data from year t-11 to t-2, etc.). Consistent with HVZ, we estimate the regression at the dollar level with unscaled data. Forecasted earnings for the next five years are estimated by using the coefficients from the above regressions and year t data for each firm. The main advantage of the cross-sectional approach is that it does not impose any survivorship requirement as time series models do. See Appendix A for details of the empirical execution for the HVZ model and the models introduced in this paper.

2.2 The RW Model

We include random walk (RW) model as the naïve benchmark. Although the forecasts from the RW model are not suitable for estimating ICC since they do not allow for growth in earnings, the RW model provides an intuitive benchmark against which to evaluate other earnings forecast models. The RW model does not rely on any parameters. It is specified as:

$$E_{t+\tau} = E_t + \varepsilon \tag{2}$$

2.3 The AR Model

The autoregressive (AR) model is one simple extension of the RW model. It allows for growth in earnings, and hence the forecasts generated by the AR model can be used for ICC estimation. The AR model is specified as:

$$E_{t+\tau} = \beta_0 + \beta_1 {*} NegE_t + \beta_2 {*} E_t + \beta_3 {*} NegE {*} E_t + \varepsilon \tag{3}$$

We include the indicator for negative earnings (NegE) and its interaction term with earnings (NegE*E) to allow for different persistence of profit and loss firms (Li 2011). We estimate the regression using the same approach of HVZ – i.e. we use ten-lagged years of data to estimate the models using all firms with available data and then apply the regression coefficients to firm-specific data to estimate the expected values for each firm. We run the regression at the per-share level by scaling all variables by the number of shares outstanding.

2.4 The RIV Model

One potential drawback of the HVZ model is the reliance on dividends as opposed to earnings and book values. Miller and Modigliani (1961) prove that, ignoring taxes and contracting costs, dividends are irrelevant for asset pricing. As an alternative to the traditional dividend discount valuation models, the residual income valuation (RIV) model derives the relation between price, book value and earnings. The RIV model was developed in early work by Preinreich (1936), Edwards and Bell (1961) and Peasnell (1982) and formalized more recently in a series of papers by Ohlson (1995), Feltham and Ohlson (1995, 1996).

Ohlson (1995) presents a basic model where future residual income depends on current residual income and other information. Feltham and Ohlson (1995) introduce the balance sheet effect of conservatism, which can mechanically increase future residual income because of lower book values. Feltham and Ohlson (1996) further introduce the income statement effect of conservatism through capital expenditures (accruals), which will depress future residual income. Feltham and Ohlson (1996) express future residual income using the following equation (notation simplified):

$$RI_{t+1} = \omega_1 * RI_t + \omega_2 * B_t + \omega_3 * capx_t + \varepsilon \tag{4}$$

where B is book value and RI is residual income. In Eq. (4), ω_1 and ω_2 are expected to be positive and lie between zero and one, while ω_3 is expected to be negative. As the definition of residual income implies that $RI_t = E_t - r*B_{t-1}$ and $RI_{t+1} = E_{t+1} - r*B_t$, we can substitute for RI_t and RI_{t+1} in Eq. (4) and solve for E_{t+1}:

$$E_{t+1} = \omega_1 * E_t + (\omega_2 + r) * B_t + (-\omega_1 * r) * B_{t-1} + \omega_3 * capx_t + \varepsilon \tag{5}$$

Based on Eq. (5), our procedure to estimate future earnings is

$$E_{t+1} = \chi_0 + \chi_1 * E_t + \chi_2 * B_t + \chi_3 * B_{t-1} + \chi_4 * capx_t + \varepsilon \tag{6}$$

In Feltham and Ohlson (1996), capital expenditures refer to all expenditures on assets (not just PP&E as capital expenditures normally pertain to). Accordingly, we set capx to total accruals (TACC from Richardson et al. 2005). We modify Eq. (6) by introducing an interaction term between E_t and a loss dummy (NegE*E). We also remove book value in year t-1 to reduce additional data requirement.[1] The equation we estimate is hence:

$$E_{t+\tau} = \chi_0 + \chi_1 * NegE_t + \chi_2 * E_t + \chi_3 * NegE_t * E_t + \chi_4 * B_t + \chi_5 * TACC_t + \varepsilon \tag{7}$$

We expect the coefficients χ_2 and χ_4 to be positive representing the persistence of earnings, χ_3 to be negative representing the lower persistence of losses (Li 2011), and χ_5 to be negative representing the effects of conservatism. We estimate the regression using the same approach of HVZ.

3. Data and Empirical Execution

3.1 Data

Our estimation sample includes all firms on the Compustat fundamentals annual file up to 2012. We collect stock returns from the CRSP monthly return file and analyst information from the IBES summary file. Appendix A provides the definition of the variables for each model. To

[1] The model including B_{t-1} produces essentially the same results.

minimize the impact of outliers on the regression results, we winsorize all variables each year at the 1st and 99th percentiles.

3.2 Earnings Forecasts for Year t+1 to Year t+5

We follow the methodology in HVZ to estimate the cross-sectional forecast models and the predicted earnings for year t+1 to year t+5. Specifically, for each year between 1969 and 2012, we estimate the three cross-sectional models (HVZ, AR and RIV) using all available observations over the past ten years. For example, if 2000 is our year t, we use data from 1990 to 1999 to estimate the coefficients that will be used to compute the earnings of 2001 (year t+1). Similarly, we use data from 1989 to 1998 to estimate the coefficients that will be used to compute the earnings of 2002 (year t+2). This ensures that the earnings forecasts are strictly out of sample. We estimate each model as of June 30 of each year. To further reduce look-ahead bias, we assume that the financial information for firms with fiscal year ending (FYE) in April to June is not available on June 30. In other words, only the financials of the firms with FYE from April of year t-1 to March of year t are used for estimation of year t. For each firm i and each year t in our sample, we compute earnings forecasts for year t+1 to year t+5 by multiplying the independent variables in year t with the pooled regression coefficients estimated using the previous ten years of data. This method only requires a firm have non-missing values for the independent variables in year t to estimate its future earnings. As a result, the survivorship bias is kept to a minimum. We set the missing value of AC and TACC to zero. However, the results are robust without this requirement.

3.3 Forecast Bias and Accuracy

To evaluate the performance of each model, we first compute forecast bias and forecast accuracy of each model. Consistent with HVZ, forecast bias is the difference between actual earnings and earnings forecasts. We scale the bias by the end-of-June market value of equity if the model forecasts dollar earnings (the HVZ and RW models), or by the end-of-June stock price if the model forecasts earnings per share (the AR and RIV models).[2] Forecast accuracy is defined as the absolute value of forecast bias.

3.4 Estimating the ERC

The second performance measure is the ERCs of the forecasts. We estimate ERCs by regressing the buy-and-hold returns over the next one, two, and three years on the unexpected earnings (i.e. the forecast bias) over the same horizon.[3] The ERCs are cumulative – i.e. the ERC for year t+2 is the regression coefficient of the cumulative buy-and-hold returns for the first two years on the total unexpected earnings for the first two years, and so on. We standardize the unexpected earnings so that they have unit variance each year. As a result, the ERCs are comparable among all model-based forecasts.

3.5 Estimating the ICC Metrics

We use the forecasts from the three cross-sectional models (HVZ, AR and RIV) to estimate implied cost of capital using the four commonly used ICC metrics. We use two ICC metrics based on the abnormal earnings model of Ohlson and Juettner-Nauroth (2005) - the Gode

[2] We estimate the HVZ model at the dollar level as it is specified in their paper. We also perform robustness test by estimating the HVZ model at the per-share level. The inference still holds.
[3] This is called "annual ERC" in HVZ.

and Mohanram (2003) implementation of the full model (ICC_{GM}) and a simplified version based on the PEG or price earnings to growth ratio (ICC_{PEG}). We also use two ICC metrics based on the residual income valuation model – the Gebhardt, Lee and Swaminathan model (ICC_{GLS}) and the Claus and Thomas model (ICC_{CT}). The details of how the four models are empirically estimated are presented in Appendix B. Consistent with the approach commonly used in the literature, we use the average of the ICCs derived from the four individual methods as our ICC metric. To allow for comparison across time, we adjust stock returns and ICCs for the risk-free rate.

4. Comparison of the Forecast Accuracy, Bias and ERC of the Models

4.1 Coefficient Estimates of the Three Cross-sectional Models

Panel A of Table 1 presents the average coefficients and the time-series t-statistics from the HVZ model estimated each year from 1969 to 2012 using the appropriately lagged ten years of data. To conserve space, we only report the results for t+1, t+2, and t+3 earnings regressions (those for t+4 and t+5 regressions are available upon request). The magnitude of the coefficients and the adjusted R^2 are consistent with the results in HVZ. Panel B of Table 1 presents the average coefficients and the corresponding time-series t-statistics from the AR model. The model explains approximately 60%, 44% and 36% of the variations of the EPSs in year t+1, t+2, and t+3, respectively. Although the adjusted R^2 of the HVZ model appears to higher than that of the AR model, this is mainly due to the fact that the HVZ model is estimated at the dollar level. The adjusted R^2 of the HVZ model estimated at the per-share level is 59%, 44% and 36% (untabulated) for t+1, t+2 and t+3 regressions, respectively. Finally, Panel C of Table 1 presents the average coefficients and the corresponding time-series t-statistics from the RIV model. All coefficients have the signs consistent with the theoretical prediction from the residual income valuation model. The three cross-sectional models produce non-missing one-, two-, and three-year-ahead earnings forecasts for 179,362 firm-year observations from 1969 to 2012.

4.2 Forecast Accuracy of the Four Models

To evaluate the performance of the cross-sectional earnings models, we first compare their forecast accuracy and bias. We perform the analysis using the common sample of firm-year observations with non-missing t+1, t+2, and t+3 forecast bias for all models from 1969 to 2008. The time period ends in 2008 because we require non-missing realized earnings in the future three years to calculate forecast bias and accuracy. The common sample includes 119,653 firm-year observations. Table 2 reports the comparison of forecast accuracy.

Panel A of Table 2 reports the time-series averages of the mean and median forecast accuracy for the HVZ, RW, AR, and RIV models in the full sample. Forecast accuracy is the absolute value of forecast bias. A larger number indicates a less accurate earnings forecast. In terms of the mean forecast accuracy, the RIV model produces the most accurate forecasts for all three forecast horizons. The AR model has the second best forecast accuracy, while the HVZ model generates the least accurate forecasts among the four models. For example, the mean accuracy of the RIV model is 0.092 (t=13.56) for one-year-ahead forecasts, while the corresponding mean accuracy of the AR, RW and HVZ models is 0.092 (t=13.63), 0.100 (t=10.57), and 0.127 (t=14.65), respectively. Compared to the forecasts of the HVZ model, the forecasts of the RIV model are on average 28% more accurate. At the one-year-ahead forecast horizon, the RIV model and the AR model have the similar forecast accuracy. However, as the forecast horizon increases, the RIV model produces more accurate forecasts than the AR model.

For example, the mean three-year-ahead forecast accuracy is 0.139 (t=21.48) for the RIV model and 0.146 (t=18.01) for the AR model, with the difference significant at 1% level. The corresponding values for the RW and HVZ models are 0.154 (t=15.65) and 0.215 (t=16.62), respectively. At the three-year-ahead forecast horizon, the RIV model generates forecasts that are on average 35% more accurate than the forecasts from the HVZ model. It is also worth noticing that the forecast accuracy of the RW model is on average 21% to 28% smaller than the HVZ model. The fact that the HVZ model cannot outperform the naïve random walk model is quite disconcerting.

In terms of the median forecast accuracy, the RW model consistently produces the most accurate forecast for all three horizons, although its superiority over the RIV model is only significant at the one-year-ahead horizon. The RIV model has the second best forecast accuracy, while again the HVZ model has the worst forecast accuracy among the four models.

Panel B of Table 2 reports the time-series averages of the mean forecast accuracy in the subsamples partitioned by analyst coverage. A firm is considered as covered by analysts if there is one FY1 consensus forecast on IBES for year t+1. The results show that the mean forecast accuracy is much bigger for firms without analyst coverage, consistent with the presumption that these firms are generally smaller and their earnings are harder to forecast. The AR model and the RIV model continue to outperform the HVZ model in both subsamples and the improvements are more pronounced in the subsample of firms without analyst coverage. For example, for firms without analyst coverage, the differences in mean forecast accuracy between the RIV model and the HVZ model are 0.058 (or 30% improvement) for the one-year-ahead forecasts and 0.124 (or 39% improvement) for the three-year-ahead forecasts. For firms with analyst coverage, the corresponding differences in mean forecast accuracy are 0.010 (or 12% improvement) and 0.016

(or 13% improvement), respectively. The RW model still significantly outperforms the HVZ model in the subsample of firms without analyst coverage. In the subsample of firms with analyst coverage, the HVZ model has better mean forecast accuracy than the RW model. However, the differences are only marginally significant at two-year-ahead forecast horizon.

For the subsample of firms with analyst coverage, researchers have two options to estimate ICC metrics. First, they could use a model-based approach as in HVZ. Second, they could correct the predictable biases in the forecasts as Mohanram and Gode (2013) do, who show that the error correction procedure dramatically improves the performance of ICC metrics. However, for the subsample of firms without analyst coverage, researchers have to use a model-based approach to generate forecasts. Hence, model-based earnings forecasts are much more crucial for firms without analyst coverage. In this important group, the HVZ model significantly underperforms not only the RIV model and the AR model, but also the naïve RW model.

Panel C of Table 2 reports the time-series averages of the mean forecast accuracy in the subsamples partitioned by firm size. Each year, observations are sorted into two equal sized groups based on their end-of-June market value of equity. The results show that in both subsamples the RIV model has the most accurate forecasts. Relative to the HVZ model, the improvement of the RIV model in forecast accuracy is more pronounced in small firms. For example, in small firms the differences in mean forecast accuracy between the RIV model and the HVZ model are 0.069 (or 33% improvement) for the one-year-ahead forecast and 0.150 (or 42% improvement) for the three-year-ahead forecast. The corresponding differences in the large firm are 0.001 (or 2% improvement) and 0.002 (or 3% improvement), respectively. The AR model also outperforms the HVZ model in both subsamples, but the improvements are relatively smaller than those of the RIV model. Finally, the naïve RW model continues to outperform the HVZ model in small firms, whose earnings are more difficult to forecast.

In summary, the results in Table 2 show that both the RIV model and the AR model outperform the HVZ model in terms of forecast accuracy. The improvement is more significant in the groups of firms where model-based forecasts are more relevant and important, i.e., small firms and firms without analyst coverage. In addition, the HVZ model also significantly underperforms the naïve random walk model in the full sample as well as in the subsamples of small firms and firms without analyst coverage.

4.3 Forecast Bias of the Four Models

Table 3 reports the comparison of forecast bias in the common sample of observations. Forecast bias is the difference between the actual earnings and the earnings forecasts, scaled by the end-of-June market value of equity (the HVZ and RW models) or the end-of-June stock price (the AR and RIV models). A negative bias indicates that the forecast is higher than the actual. Panel A of Table 3 reports the time-series averages of the mean and median forecast bias for the four models in the full sample as well as their pair-wise comparisons. Mean forecast biases for the HVZ, AR and RIV models are negative and statistically significant for all forecast horizons. In contrast, mean forecast biases for the RW model are positive and statistically significant for all forecast horizons. This is because the naïve random walk model does not allow for growth in earnings. The magnitude of the forecast bias of the HVZ model is significantly larger than all other models, including the naïve RW model.

Panel B of Table 3 reports the time-series averages of the mean forecast bias in the subsamples partitioned by analyst coverage. The mean forecast bias of the RIV model is generally the smallest in magnitude among all models in both subsamples. However, its one-year-ahead forecast bias has larger magnitude than the RW model for firms with analyst coverage. The HVZ model produces more biased forecast than all other three models for firms without analyst coverage.

Panel C of Table 3 reports the time-series averages of the mean forecast bias in the subsamples partitioned by firm size. For large firms, the mean forecast biases of the AR and RIV models are all statistically insignificant, while the mean forecast biases of the HVZ model are all significantly negative. For small firms, the mean forecast bias of the HVZ, AR and RIV models are all significantly negative, with the forecasts of the HVZ model being the most biased.

In summary, the forecasts from the RIV model generally are the least biased for the whole population as well as in the partitions by analyst coverage and by firm size. In contrast, the forecasts from the HVZ model generally are the most biased, especially for firms without analyst coverage and small firms.

4.4 ERC of Model-based Forecasts

A higher ERC suggests that the market reacts more strongly to the unexpected earnings generated from the model. In other words, the earnings forecasts from the model represent a better approximation of the market expectations. Table 4 reports the time-series averages of the ERCs for all models. ERC is estimated by regressing the buy-and-hold returns over the next one, two, and three years on the unexpected earnings (i.e., the forecast bias) over the same horizon. We standardize the unexpected earnings so that they have unit variance each year. As a result, the ERCs are comparable among all model-based forecasts.

Panel A presents the ERCs for the entire sample. The ERCs for one-, two-, and three-year-ahead forecasts from the RIV model are 0.102 (t=16.11), 0.206 (t=16.27), and 0.326 (t=15.67), respectively, which are the highest among all models. The AR model ranks the second with the corresponding ERCs of 0.102 (t=16.09), 0.200 (t=16.57), and 0.318 (t=16.29), respectively. Surprisingly, even the naïve RW model outperforms the HVZ model for all forecast horizons. For example, the ERC of the three-year-ahead forecast is 0.126 (t=6.44) for the HVZ model and 0.285 (t=8.94) for the RW model, with the difference significant at 1% level.

Panel B of Table 4 reports the time-series averages of the ERCs in the subsamples partitioned by analyst coverage. For firms without analyst coverage, the ERCs for all three forecast horizons from the HVZ model are the lowest among all models, while the ERCs of the RW, AR and RIV models are virtually indistinguishable. For firms with analyst coverage, all four models produce similar ERCs for all horizons, with none of the pair-wise differences being statistically significant.

Panel C of Table 4 reports the time-series averages of the ERCs in the subsamples partitioned by firm size. For small firms, both the AR model and the RIV model produce the highest ERCs, while the HVZ model has the lowest ERCs. In large firms, however, the HVZ model outperforms the other three models, while the RW model is the second best choice.

To summarize, the forecasts from the AR and RIV models represent a better approximation of the market expectations than the forecasts from the HVZ model, both in the full sample and in partitions where model-based forecasts are more important. It appears that with the exception of large firms, the forecasts of the HVZ model generally underperform the forecasts of the naïve random walk model as proxies for the market expectations.

5. Properties of ICC Estimates from the Models

5.1 Relation with Future Returns: Portfolio Tests

In each year, we divide the sample into quintiles based on the ICC metrics. We then compare the equally weighted mean returns to each of the quintiles, focusing on the spreads between the lowest and the highest quintiles. The returns are measured annually for the first three years after portfolio formation, with the compounding period starting four months after the end of the prior fiscal year.[4] To allow for a comparison across time, we subtract the risk-free rate (R_F) from both the annual buy-and-hold returns and the ICC metrics.

Table 5 presents the pooled results of our portfolio tests using annual quintiles. Panel A presents the returns over the future three years for the quintiles formed on the composite ICC metric (average of ICC_{GM}, ICC_{PEG}, ICC_{GLS} and ICC_{CT}) for each forecasting model (HVZ, AR and RIV). Panel B reports the pair-wise comparisons of the return spreads between the lowest and the highest quintiles.

The first column in Panel A provides the mean ICC for each quintile for each of the models. As the results indicate, the mean level of ICC is generally higher for the HVZ model, especially for the higher quintiles. This is potentially related to the higher bias in the HVZ model reported earlier. The HVZ model is more likely to have negative forecast errors – i.e. more likely to generate higher forecasts of earnings which would naturally lead to higher values of ICC.

The first set of rows in Panel A of Table 5 presents the returns and return spreads for the HVZ model. Consistent with their reported results, the return spreads for the three years are

[4] This represents a departure from HVZ, who form calendar time portfolios starting on July 1st. The advantage of our approach is that the financial statement information is equally timely for all observations. The disadvantage is the fact that the compounding period may not be identical for all firms in our sample. As a robustness test, we carry out all tests in a subset of firms with December fiscal year ends (over 60% of the sample) and find virtually identical results.

economically meaningful and statistically significant (4.56%, 6.00% and 3.62%, respectively). Further, the realized returns increase monotonically from the lowest ICC quintile to the highest ICC quintile for the first two years. For the third year, while the return spreads continue to be significantly positive, the returns do not increase monotonically. The returns for the 4th quintile (9.90%) are higher than the returns for the 5th quintile (8.91%).

The next set of rows of Panel A presents the return spreads for the AR model. As the results indicate, the AR model generates higher return spreads. For instance, the return spreads for the AR model are 6.30%, 6.90% and 7.17% for the three years respectively, while the corresponding spreads for the HVZ model are 4.56%, 6.00% and 3.62%., with the difference in spreads being statistically significant in years t+1 and year t+3 (see Table 5, Panel B). Thus, the AR model presents itself as a superior alternative to the HVZ model, consistent with the results for forecast bias, accuracy and ERC.

The final set of rows of Panel A of Table 5 presents the return spreads for the RIV model. The RIV model produces the largest return spreads of the three models for most horizons. For instance, the RIV model produces return spreads of 7.25%, 7.07%, and 6.37% for the next three years. This compares favorably with returns spreads of 4.56%, 6.00% and 3.62% for the HVZ model (differences significant for years t+1 and t+3). However, the RIV model is only mildly superior to the AR model, producing insignificantly larger spreads in year t+1 and year t+2, and insignificantly smaller spreads in year t+3.

5.2 Relation with Future Returns: Firm Level Tests

In addition to the portfolio tests, we perform firm level tests to measure the relation between the ICC metrics generated by the models and future returns. For each year, we estimate cross-sectional univariate regressions with the future returns as the dependent variable and the ICC metric as the independent variable.[5] The benchmark coefficient is "1", where the realized return is on average equal to the ICC proxy. We present the Fama and MacBeth (1973) coefficients and t-statistics in Table 6. Panel A presents the regression results for the ICC metrics derived from each of the three models. Panel B compares the coefficients on the ICC metrics across the models.

The first set of columns of Panel A presents the regression for year t+1. As the results indicate, the ICC metric based on the HVZ model has the lowest correlation with future realized returns with a coefficient of 0.209. In comparison, the ICC metric based on the AR model has a coefficient of 0.584, while the ICC metric based on the RIV model has a coefficient of 0.652. The next set of columns presents the regressions for year t+2 and suggest a similar pattern, with the coefficient on ICC for the HVZ model at 0.192 trailing that for the AR model (0.498) and the RIV model (0.574). Finally, the last set of columns suggests that the pattern persists for year t+3. The ICC metric based on the HVZ model no longer shows any association with realized returns (coefficient 0.087, t-stat 1.04). In contrast, the ICC metrics based on the AR model (coefficient 0.576, t-stat 2.63) and the RIV model (coefficient 0.594, t-stat 2.74) continue to show a strong association with realized returns.

As the comparison on Panel B suggests, the differences in the coefficients on ICC between the HVZ model and the RIV model are statistically significant for all three years, indicating that the ICC estimates derived from the RIV model have a significantly greater association with realized returns. The differences in coefficients between the HVZ model and the

[5] Easton and Monahan (2005) recommend running regressions with the ICC measure and proxies for cash flow news and discount rate news. However, these proxies require forecast revisions, which are not feasible to estimate for cross-sectional models. Hence, we only run univariate regressions.

AR model appear to be economically significant. The mean coefficients of the AR model are more than twice as large as the mean coefficients of the HVZ model for all three years. However, the differences are statistically significant only in year t+3. Finally, the ICC based on the RIV model appears to have slightly stronger correlations with future realized returns than the ICC based on the AR model.[6]

As an alternative to comparing mean coefficients across different sets of regressions, we also run a "horse race" between the ICC metrics by regressing realized returns on all three metrics (labeled ICC_{HVZ}, ICC_{AR} and ICC_{RIV}). The results are presented in Panel C of Table 6. As the results suggest, ICC_{RIV} is the clear winner in this "horse race", with a significant coefficient that approaches the benchmark of 1 for all three years. The coefficients on ICC_{HVZ} and ICC_{AR} are insignificant for all three years.

To summarize, the firm level regressions confirm the results from the portfolio tests. The ICC estimates derived from the AR model and the RIV model show stronger correlations with realized returns than the ICC estimates from the HVZ model. The RIV model appears to perform marginally better than the AR model. Researchers wishing to choose between these two models will have to make a tradeoff between the greater parsimony of the AR model and the slightly stronger results of the RIV model.

[6] A potential concern might be that the lower coefficients on the HVZ model in the return regressions might arise mechanically due to the greater magnitude and greater spread of the ICC estimates generated from the HVZ model. To account for this, we perform the following sensitivity test. We standardize all the ICC measures each year by subtracting the minimum and then dividing by the range (maximum – minimum) for ICC using that method in that year. In other words, we set each ICC = (ICC – min)/(max – min). We then re-estimate the regressions using the standardized ICC measures. We continue to find the weakest relation between ICC from the HVZ model and future returns. For instance, for one-year-ahead returns, the average coefficients on ICC_{HVZ}, ICC_{AR} and ICC_{RIV} from the Fama-MacBeth regressions are 0.183, 0.395, and 0.431, respectively. For two-year-ahead returns, the average coefficients on ICC_{HVZ}, ICC_{AR} and ICC_{RIV} are 0.166, 0.408, and 0.372, respectively. For three-year-ahead returns, the average coefficients on ICC_{HVZ}, ICC_{AR} and ICC_{RIV} are 0.079, 0.390, and 0.338, respectively.

5.3 Relation with Risk Factors

Prior research has evaluated ICC metrics either by evaluating their correlation with realized returns or by analyzing their correlation with risk proxies such as systematic risk, idiosyncratic risk, size, book-to-market, and growth (Gebhardt, Lee and Swaminathan 2001; Gode and Mohanram 2003; Botosan and Plumlee 2005).[7] Our results thus far have shown that the ICC metrics from the AR and RIV models outperform the ICC metrics from the HVZ model as far as the correlation with realized returns is concerned. We now examine the correlation of the ICC metrics with risk factors to ensure that this superior performance is not coming at the expense of anomalous correlations with risk factors.

We use the following risk factors from prior research: (1) Systematic risk (β), calculated using monthly returns over the lagged five years (ensuring that at least 24 observations are available); (2) Firm size (LMCAP), the logarithm of market capitalization at the time of the forecasts; (3) Book-to-market ratio (BM); (4) Idiosyncratic risk (IDIO), the standard deviation of the prior year's monthly returns; (5) Earnings volatility (STDNI), the standard deviation of net income (IBQ) scaled by total assets (ATQ) measured over the previous eight quarters; (6) Leverage (D2A), the ratio of total debt (DLTT+DLC) to total assets (AT); and (7) Analyst following (LFOLLOW), the logarithm of 1+number of analysts following the stock. We expect ICC to be positively related to β, BM, IDIO, STDNI and D2A, and negatively related to LMCAP and LFOLLOW.

We estimate three specifications – the first with only β like the CAPM model, the second with β augmented with size (LMCAP) and book-to-market (BM) like the Fama and French

[7] Easton and Monahan (2010) argue that the latter approach is logically inconsistent as ICC metrics are estimated precisely because of the flaws in conventional measures of risk that often rely on ex-post returns. We present these results to ensure a comparison between our results and those presented in HVZ.

(1992) model, and the final specification with all the proposed risk factors. Regressions are estimated annually and aggregated using the Fama and MacBeth (1973) procedure.

The results are presented in Table 7. At the outset, we note that all proxies correlate strongly in the expected direction with four of the above seven factors – positively as expected with book-to-market (BM), earnings volatility (STDNI), and leverage (D2A) and negatively as expected with size (LMCAP). Our discussion will hence focus on the three remaining risk proxies, β, IDIO, and LFOLLOW, where we find variations among the three forecasting models.

The first set of rows in Table 7 present the regressions for the ICC metric computed from the HVZ model. The results suggest an anomalous negative correlation between ICC_{HVZ} and β in all specifications, and an insignificant correlation with both idiosyncratic risk (IDIO) as well as analyst following (LFOLLOW).

The next set of rows presents the results for the AR model. In the univariate regression, the coefficient on β is positive but marginally insignificant (0.002, t-stat 1.53). In the full specification, we find a strong positive correlation as hypothesized between ICC_{AR} and IDIO (0.212, t-stat 8.39) and a significant negative correlation as hypothesized between ICC_{AR} and LFOLLOW (-0.001, t-stat -4.21). However, the coefficient on β in the full specification is anomalously negative (-0.008, t-stat -8.14).

Finally, the last set of rows presents the results for the RIV model. The coefficient on β is insignificant in the univariate regression (-0.002, t-stat -1.44). In the full specification, we find a strong positive coefficient on IDIO (0.101, t-stat 4.13) but an insignificant negative coefficient on LFOLLOW (-0.000, t-stat -0.25). Similar to the other two measures, the coefficient on β in the full specification is anomalously negative (-0.007, t-stat -7.68).

To summarize, the AR model shows the strongest correlations with risk factors, with six of the seven risk factors (all except β) loading significantly and in the correct direction. The RIV model ranks the second best, with five of the seven risk factors (all except β and LFOLLOW) loading significantly and in the correct direction. For the HVZ model, three risk factors either do not load or load anomalously (β, LFOLLOW and IDIO). The risk regressions hence confirm that the superior performance of the RIV model and the AR model in particular is not coming at the expense of anomalous correlations with risk factors.

6. Sensitivity Analysis

We perform several sensitivity tests to verify the robustness of our results. These tables are not tabulated for brevity, but are discussed below.

A potential problem with the ERC estimation is that the magnitude of ERCs is biased downward in the presence of large forecast errors. Cheong and Thomas (2012) show that ERCs can increase dramatically when observations with extreme forecast errors are deleted. To mitigate the concern that our results of ERC comparison could be driven by some outliers in forecast errors, we truncate at 1%, 5% and 10% on each side of the forecast error distribution for each forecast model. The magnitude of the ERC estimates for all four models increases after eliminating the outliers. However, the ranking of the ERCs as well as the significance of the differences between the ERCs do not change.

The relation between our prediction variables and future earnings could vary not only through time but also across industries. We examine this possibility by estimating regressions by industry and by year, where industry is defined according to the 48 industry classifications in Fama and French (1997). Interestingly, we find that estimating the regressions at the industry-year level actually increase forecast errors for all models. In addition, estimating regressions by industry-year slightly reduces our sample size because certain industries do not have sufficient historical data. Hence, it appears that the parsimonious approach used here as well as in the HVZ paper is preferable.

The evidence in the paper indicates that firm size is an important determinant of the relation between our prediction variables and future earnings. Consequently, we estimate each model by size deciles and year. The size deciles are determined using the end-of-June market value of equity each year. This modification marginally improves the forecast accuracy of all models but has almost no impact on the performance of the ICC metrics. This confirms the validity of the parsimonious approach of running annual cross-sectional regression for the entire population.

One concern that may affect the comparison of the HVZ model with our models, especially the AR and RIV models, is that the HVZ model is estimated at the dollar level, while the AR and RIV models are estimated at the per-share level. We perform robustness test by estimating the HVZ model at the per-share level. We find that the per-share estimation improves HVZ's forecast accuracy and ERC. However, the per-share HVZ model still underperforms the RW, AR and RIV models in terms of forecast accuracy and ERC performance. Hence, the superiority of the AR model and the RIV model is not an artifact of the differences in scaling.

Finally, we use the robust regression technique instead of OLS regression to reduce the impact of outliers on regression coefficients. Robust regression is an iterative procedure that keeps eliminating outliers and re-estimating regressions, until no further outliers are deleted. We find that the robust regression technique marginally improves the mean forecast accuracy of all models. However, the rankings of the models in terms of forecast accuracy and ERC do not change. Further, we find a minimal impact on the properties of the ICC metrics. Again, for reasons of parsimony, we recommend that researchers use a simple OLS regression.

7. Conclusions

Forecasts of future earnings are critical for empirical research in valuation, especially research using implied cost of capital (ICC). Prior research has traditionally used forecasts from analysts, which has restricted the analysis to the subset of covered firms. This has meant that the most interesting firms are often omitted from the analysis. Using time series models to generate forecasts does not satisfactorily address this problem, because these models impose substantial survivorship and age requirements. A recent paper by Hou, van Dijk and Zhang (2012) addresses this problem by using a cross-sectional approach that only requires current information from firms to generate forecasts. Not surprisingly, the HVZ model has been used in recent research on accounting based valuation (Chang, Landsman and Monahan 2012) and ICC (e.g., Jones and Tuzel 2012; Lee, So and Wang 2011; Patatoukas 2011).

Given the widespread adoption of the HVZ model to generate forecasts in lieu of analyst forecasts, it is crucial to evaluate the HVZ model and present alternatives to address its weaknesses. Gramacy and Gerakos (2013) show that the HVZ model performs worse than a naïve random walk model. However, a random walk model is not practical for computing ICC.

In this paper, we present and evaluate two alternatives to the HVZ model, while adopting the cross-sectional forecasting approach in HVZ. Our first model (AR) is a simple autoregressive model which allows for differential persistence of profits and losses. The second model (RIV) is motivated by the residual income valuation model in Feltham and Ohlson (1996), and forecasts future income as a function of current income, current book value of equity and accruals. We test the HVZ model, the above two models and the naïve random walk (RW) model on the basis of their forecast bias, accuracy and earnings response coefficients (ERCs). We also evaluate the ICC estimates generated from the HVZ, AR and RIV models on the basis of their correlation with future returns and risk factors.

We find that both of our models significantly outperform the HVZ model in virtually all the dimensions we examine. Both the AR model and the RIV model generate forecasts that are more accurate and show greater ERCs. These differences are greater in settings where model-based forecasts are likely to be the most useful – for small firms and for firms without analyst coverage. In contrast, the HVZ model performs worse than a naïve random walk model, confirming the results in Gramacy and Gerakos (2013).

In addition, the ICC proxies generated from the AR model and the RIV model show stronger correlations with future returns than the ICC proxies generated from the HVZ model, both at the portfolio level and at the firm level. Lastly, the ICC metrics from the AR model and the RIV model also show more meaningful correlations with suggested risk factors.

The results of our paper have crucial implications for all research where proxies for future expected earnings are required. We recommend that researchers use cross-sectional forecasting models based either on the AR model or the RIV model presented in this paper.

References

Botosan, C. and M. Plumlee. 2005. Assessing Alternative Proxies for the Expected Risk Premium. The Accounting Review 80: 21–53.

Chang, W., Landsman, W. and S. Monahan. 2012. Selecting an Accounting Based Valuation Model. Working paper, INSEAD/University of North Carolina.

Cheong, F. and J. Thomas. 2012. Managerial smoothing of analysts' EPS forecast errors: explanations and implications. Working paper, Rutgers University/ Yale University.

Claus, J., and J. Thomas. 2001. Equity Risk Premium as Low as Three Percent? Evidence From Analysts' Earnings Forecasts for Domestic and International Stocks. Journal of Finance 56: 1629–1666.

Easton, P. 2004. PE Ratios, PEG Ratios, and Estimating the Implied Expected Rate of Return on Equity Capital. The Accounting Review 79: 73–96.

_________ and S. Monahan. 2005. An Evaluation of Accounting-Based Measures of Expected Returns. The Accounting Review 80: 501–38.

_________ and ________. 2010. Evaluating Accounting-based Measures of Expected Returns: Easton and Monahan and Botosan and Plumlee Redux. Working paper, University of Notre Dame/INSEAD.

Edwards, E. and P. Bell. 1961. The Theory of and Measurement of Business Income. University of California Press.

Elton, J. 1999. Expected Return, Realized Return, and Asset Pricing Tests. Journal of Finance 54: 1199–1220.

Fama, E. and K. French. 1992. The Cross-section of Expected Stock Returns. Journal of Finance 47: 427-465.

_________ and ________. 1997. Industry Costs of Equity. Journal of Financial Economics 43: 153-193.

_________ and ________. 2000. Forecasting Profitability and Earnings. Journal of Business 73: 161-175.

_________ and ________. 2006. Profitability, Investment and Average Returns. Journal of Financial Economics 82: 491-518.

________ and J. MacBeth. 1973. Risk, Return, and Equilibrium: Empirical Tests. Journal of Political Economy 71: 607-636.

Feltham, G. and J. Ohlson. 1995. Valuation and Clean Surplus Accounting for Operating and Financial Activities. Contemporary Accounting Research 11 (Spring): 689–732.

——— and ———. 1996. Uncertainty Resolution and the Theory of Depreciation Measurement. Journal of Accounting Research 34: 209–234.

Frankel R., and C. Lee, 1998, Accounting Valuation, Market Expectation, and Cross-Sectional Returns. Journal of Accounting and Economics 25: 283-319.

Gebhardt, W., Lee, C. and B. Swaminathan. 2001. Towards an Implied Cost of Equity. Journal of Accounting Research 39: 135-176.

Gerakos, J. and R. Gramacy. 2013. Regression-based Earnings Forecasts. Working paper, University of Chicago.

Gode, D. and P. Mohanram. 2003. Inferring the Cost of equity Using the Ohlson-Juettner Model. Review of Accounting Studies 8: 399-431.

Hou, K., van Dijk, M., and Y. Zhang. 2012. The Implied Cost of Equity: A New Approach. Journal of Accounting and Economics 53: 504-526.

Hou, K. and D. Robinson. 2006. Industry Concentration and Average Stock Returns. Journal of Finance 61: 1927–1956.

Hou, K. and M. van Dijk. 2011. Profitability Shocks and the Size Effect in the Cross-section of Expected Stock Returns. Working Paper - Ohio State University.

Jones, C. and S. Tuzel . 2012. Inventory Investment and the Cost of Capital. Journal of Financial Economics 107 (3): 557–579.

Lee, C., So, E. and C. Wang. 2011. Evaluating Implied Cost of Capital Estimates. Working paper, Stanford University.

Li, K. 2011. How Well Do Investors Understand Loss Persistence? Review of Accounting Studies 16: 630-667.

Mohanram, P. and D. Gode. 2013. Removing Predictable Analyst Forecast Errors to Improve Implied Cost of Equity Estimates. Review of Accounting Studies (forthcoming).

Miller, M. and F. Modigliani. 1961. Dividend Policy, Growth, and the Valuation of Shares. Journal of Business 34: 235-264.

Ohlson, J. 1995. Earnings, Book Values, and Dividends in Equity Valuation. Contemporary Accounting Research 11: 661-687.

Ohlson, J. and B. Juettner-Nauroth. 2005. Expected EPS and EPS Growth as Determinants of Value. Review of Accounting Studies 10: 349-365.

Patatoukas, P. 2011. Customer-Base Concentration: Implications for Firm Performance and Capital Markets. The Accounting Review 87(2): 363-392.

Peasnell, K. 1982. Some Formal Connections between Economic Values and Yields, and Accounting Numbers. Journal of Business Finance and Accounting 9: 361-381.

Preinreich, G. 1936. The Fair Value and Yield of Common Stock. The Accounting Review 11: 130-140.

Richardson, S, Sloan, R., Soliman, M. and I. Tuna, 2005. Accrual Reliability, Earnings Persistence and Stock Prices. Journal of Accounting and Economics 39: 437-485.

Appendix A: Variable definitions for models used to generate forecasts

HVZ model:

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

Variable	**Definition**	**Xpressfeed variable**
$E_{t+\tau}$	Earnings in year t+τ	ib
A_t	Total asset in year t	at
D_t	Common dividend	dvc
DD_t	A dummy variable that equals 1 for dividend payers and 0 otherwise	
$NegE_t$	A dummy variable that equals 1 for firms with negative earnings and 0 otherwise.	
AC_t	Change in non-cash current assets less change in current liabilities excluding change in short-term debt and change in taxes payable minus depreciation and amortization	Δ(act-che)- Δ(lct-dlc-txp)-dp

AR model:

$$E_{i,t+\tau} = \beta_0 + \beta_1 NegE_{i,t} + \beta_2 E_{i,t} + \beta_3 NegE * E_{i,t} + \varepsilon_{i,t+\tau}$$

Variable	**Definition**	**Xpressfeed variable**
$E_{t+\tau}$	Earnings in year t+τ divided by number of shares outstanding in year t	$ib_{t+\tau}/csho_t$
$NegE_t$	A dummy variable that equals 1 for firms with negative earnings and 0 otherwise.	
$NegE*E_t$	Interaction term of NegE and E	

RIV model:

$$E_{i,t+\tau} = \chi_0 + \chi_1 NegE_{i,t} + \chi_2 E_{i,t} + \chi_3 NegE * E_{i,t} + \chi_4 B_{i,t} + \chi_5 TACC_{i,t} + \varepsilon_{i,t+\tau}$$

Variable	**Definition**	**Xpressfeed variable**
$E_{t+\tau}$	Earnings in year t+τ divided by number of shares outstanding in year t	$ib_{t+\tau}/csho_t$
$NegE_t$	A dummy variable that equals 1 for firms with negative earnings and 0 otherwise.	
$NegE*E_t$	Interaction term of NegE and E	
B_t	Book value of equity divided by number of shares outstanding	$ceq_t/csho_t$
$TACC_t$	Richardson et al. (2005) total accruals, i.e., the sum of the change in WC, the change in NCO, and the change in FIN, divided by number of shares outstanding	WC=(act-che)-(lct-dlc); NCO=(at-act-ivao)-(lt-lct-dltt); FIN=(ivst+ivao)-(dltt+dlc+pstk); All variables deflated by csho

Appendix B: Computing implied cost of capital

The implied cost of equity used in this paper is computed as the average of the four commonly used metrics, ICC_{GM}, ICC_{PEG}, ICC_{GLS} and ICC_{CT}. We briefly describe how these four metrics are computed below.

ICC based on the OJ Model: ICC_{GM} and ICC_{PEG}

Ohlson and Juettner-Nauroth (2005) show that the implied cost of capital can be expressed as:

$$r_e = A + \sqrt{A^2 + \frac{eps_1}{P_0}\left(g_2 - (\gamma - 1)\right)} \text{ where } A = \frac{1}{2}\left((\gamma - 1) + \frac{dps_1}{P_0}\right) \text{ and } g_2 = \frac{(eps_2 - eps_1)}{eps_1}.$$

Gode and Mohanram (2003) make the following assumptions. They set (γ-1) to r_f - 3% where r_f is the risk free rate. In addition, they use the average of short term growth and analysts' long term growth rate (LTG) instead of g_2 to reduce the impact of outliers.

If short term growth ($\frac{eps_2}{eps_1} - 1$) is greater than long term growth rate ($\sqrt[4]{\frac{eps_5}{eps_1}} - 1$), we set g_2 to equal the geometric mean of short term and long term growth rate. If short term growth is less than long term growth, we set g_2 to equal the long term growth rate. Dividends are estimated by calculating current payout for all firms, defined as dividends (DVC) divided by income before extraordinary items (IB) for firms with positive current earnings or dividends divided by 6% of total assets (AT) for firms with negative IB.

In addition, we compute an ICC from a simplified version of the OJ model that ignores dividends and sets ICC to the square root of the inverse of the PEG ratio. We compute ICC_{PEG} as:

$$ICC_{PEG} = \sqrt{\frac{g_2}{\left(PRICE / eps_1\right)}} \text{ where } g_2 \text{ is defined as it is for the } R_{GM} \text{ model}$$

ICC based on the RIV Model: ICC_{GLS} and ICC_{CT}

Gebhardt, Lee and Swaminathan (2001) use the residual income valuation model (RIV) to estimate implied cost of equity. They use EPS estimates for future two years and the expected dividends payout (from historical data) to derive book value and return on equity (ROE) forecasts. Beyond the forecast horizon, they assume that ROE fades to the industry median by year 12. Industry Median ROE is estimated as the median of all ROEs from firms in the same industry defined using the Fama and French (1997) classification over the past five years with positive earnings and non-negative book values, where

ROE is defined as the ratio of net income before extraordinary items (IB) to lagged total common shareholders' equity (CEQ). Abnormal earnings are assumed to remain constant at year 12 levels for perpetuity. The cost of equity is computed numerically by equating current stock price to the sum of the current book value and the present value of future residual earnings – i.e. solving for r in the equation:

$$P_0 = B_o + \sum_{\tau=1}^{12} \frac{(eps_\tau - r * B_{\tau-1})}{(1+r)^\tau} + \frac{(eps_{12} - r * B_{11})}{r(1+r)^{12}}$$

where eps is the forecasted eps (obtained either from explicit forecast or inferred from expected ROE and lagged book value), P_0 is current price per share, B_0 is current book value per share and B_1 through B_{11} are expected future book values per share obtained through the clean surplus relation, setting payout to equal current payout. Current payout is defined as dividends (DVC) divided by income before extraordinary items (IB) for firms with positive current earnings or dividends divided by 6% of total assets (AT) for firms with negative IB. We depart from GLS by using the model forecasts explicitly for years 1 through 5 and then applying ROE convergence.

Claus and Thomas (2001) also use the RIV model to estimate the implied cost of equity. They assume that earnings grow at the analyst's consensus long-term growth rate until year 5, and at the rate of inflation thereafter. The implied cost of equity is estimated numerically by solving the following equation:

$$P_0 = B_o + \sum_{\tau=1}^{5} \frac{(eps_\tau - r * B_{\tau-1})}{(1+r)^\tau} + \frac{(eps_5 - r * B_4) * (1+g)}{(r-g)(1+r)^5}$$

where eps_0 through eps_5 are the forecasted future earnings per share, B_0 is current book value per share and B_1 through B_4 are expected future book values per share. Consistent with Claus and Thomas (2001), g is set to $r_f - 3\%$.

Table 1 Coefficient estimates from the three cross-sectional earnings models, 1969 - 2012.

Panel A: The HVZ Model

	Intercept	A_t	D_t	DD_t	E_t	$NegE_t$	AC_t	Adj. R^2
E_{t+1}	-0.6133	0.0021	0.3148	2.5757	0.7874	1.0792	-0.0449	84%
	(-2.42)	(11.93)	(12.73)	(7.32)	(49.07)	(3.22)	(-6.42)	
E_{t+2}	-0.3585	0.0038	0.4014	3.7143	0.7315	1.8373	-0.0648	78%
	(-1.14)	(12.06)	(9.73)	(7.61)	(28.74)	(4.46)	(-5.21)	
E_{t+3}	0.5986	0.0055	0.4051	4.3166	0.7304	2.3327	-0.0657	74%
	(2.42)	(10.42)	(6.39)	(10.54)	(20.10)	(7.44)	(-4.54)	

Panel B: The AR Model

	Intercept	$NegE_t$	E_t	$NegE*E_t$	Adj. R^2
E_{t+1}	0.1252	-0.2179	0.9363	-0.8080	60%
	(8.45)	(-8.17)	(122.66)	(-27.92)	
E_{t+2}	0.2520	-0.2037	0.9509	-1.0781	44%
	(8.06)	(-5.79)	(95.38)	(-26.19)	
E_{t+3}	0.4108	-0.1980	0.9791	-1.2387	35%
	(8.75)	(-4.19)	(86.83)	(-30.22)	

Panel C: The RIV Model

	Intercept	$NegE_t$	E_t	$NegE*E_t$	B_t	$TACC_t$	Adj. R^2
E_{t+1}	0.0844	-0.1971	0.8608	-0.6277	0.0146	-0.0295	60%
	(5.48)	(-5.56)	(86.95)	(-23.11)	(13.08)	(-8.14)	
E_{t+2}	0.1553	-0.2231	0.8062	-0.9366	0.0292	-0.0514	44%
	(5.10)	(-7.39)	(60.91)	(-9.07)	(18.51)	(-11.71)	
E_{t+3}	0.2626	-0.1915	0.7762	-1.0272	0.0419	-0.0726	36%
	(5.85)	(-3.30)	(47.30)	(-8.13)	(19.62)	(-18.29)	

Each model is estimated annually from 1969 to 2012 using previous ten years of data. The average coefficients and the time-series t-statistics (in parentheses) are reported for the HVZ model (Panel A), the AR model (Panel B), and the RIV model (Panel C). Appendix A provides the definitions of the variables in each model.

Table 2 Forecast accuracy of the cross-sectional earnings models and the random walk model

Panel A: Full sample of 119,653 firm-year observations (1969-2008)

	E_{t+1}		E_{t+2}		E_{t+3}	
	Mean	Median	Mean	Median	Mean	Median
HVZ	0.127***	0.040***	0.172***	0.062***	0.215***	0.082***
	(14.65)	(16.98)	(16.65)	(18.79)	(16.62)	(18.53)
RW	0.100***	0.029***	0.128***	0.046***	0.154***	0.061***
	(10.57)	(22.88)	(13.36)	(28.47)	(15.65)	(26.09)
AR	0.092***	0.031***	0.114***	0.049***	0.146***	0.065***
	(13.63)	(21.85)	(19.36)	(23.97)	(18.01)	(20.67)
RIV	0.092***	0.031***	0.112***	0.047***	0.139***	0.062***
	(13.56)	(22.14)	(20.26)	(26.46)	(21.48)	(22.48)
Comparison						
HVZ-RW	0.027***	0.011***	0.044***	0.016***	0.061***	0.021***
	(5.30)	(7.13)	(4.43)	(7.06)	(4.63)	(7.61)
HVZ-AR	0.035***	0.009***	0.057***	0.014***	0.069***	0.017***
	(7.71)	(5.50)	(7.96)	(6.20)	(6.97)	(5.96)
HVZ-RIV	0.035***	0.009***	0.059***	0.015***	0.076***	0.019***
	(7.42)	(5.63)	(7.79)	(6.44)	(7.66)	(7.01)
RW-AR	0.008*	-0.003***	0.014*	-0.002**	0.008	-0.004*
	(1.91)	(-4.51)	(2.01)	(-2.14)	(0.81)	(-1.86)
RW-RIV	0.008*	-0.002***	0.016**	-0.001	0.015*	-0.002
	(1.84)	(-4.48)	(2.58)	(-1.45)	(1.95)	(-1.04)
AR-RIV	0.000	0.000	0.002	0.001***	0.007***	0.002***
	(-1.04)	(1.56)	(1.31)	(3.15)	(2.75)	(3.55)

Panel B: Partition analysis of mean forecast accuracy by analyst coverage

	No coverage (N=50,242)			With coverage (N=69,411)		
	E_{t+1}	E_{t+2}	E_{t+3}	E_{t+1}	E_{t+2}	E_{t+3}
HVZ	0.194***	0.257***	0.321***	0.086***	0.102***	0.121***
	(11.92)	(15.30)	(16.66)	(9.92)	(16.68)	(24.70)
RW	0.146***	0.183***	0.217***	0.089***	0.111***	0.131***
	(8.92)	(10.03)	(11.35)	(8.39)	(10.83)	(12.57)
AR	0.134***	0.161***	0.208***	0.076***	0.089***	0.106***
	(11.19)	(15.35)	(14.80)	(11.43)	(21.40)	(30.82)
RIV	0.136***	0.161***	0.197***	0.076***	0.089***	0.105***
	(11.06)	(14.83)	(16.78)	(11.40)	(20.50)	(30.25)
Comparison						
HVZ-RW	0.048***	0.074***	0.104***	-0.003	-0.009*	-0.010
	(9.10)	(6.66)	(5.85)	(-1.11)	(-1.87)	(-1.55)
HVZ-AR	0.060***	0.096***	0.112***	0.010***	0.013***	0.015***
	(7.95)	(10.91)	(9.41)	(4.43)	(5.43)	(6.15)
HVZ-RIV	0.058***	0.097***	0.124***	0.010***	0.014***	0.016***
	(7.57)	(11.07)	(10.81)	(4.50)	(5.99)	(7.07)
RW-AR	0.012	0.023*	0.008	0.013***	0.022***	0.025***
	(1.45)	(1.88)	(0.50)	(2.90)	(3.19)	(3.03)
RW-RIV	0.010	0.023**	0.020	0.013***	0.022***	0.026***
	(1.17)	(2.12)	(1.48)	(2.95)	(3.34)	(3.30)
AR-RIV	-0.002***	0.000	0.011*	0.000	0.000	0.001*
	(-3.25)	(0.06)	(2.00)	(1.21)	(0.56)	(1.88)

Table 2 continued

Panel C: Partition analysis of mean forecast accuracy by size

	Small Firms (N=59,819)			Large Firms (N=59,834)		
	E_{t+1}	E_{t+2}	E_{t+3}	E_{t+1}	E_{t+2}	E_{t+3}
HVZ	0.210***	0.283***	0.356***	0.043***	0.060***	0.075***
	(13.96)	(14.92)	(14.76)	(16.60)	(25.65)	(26.19)
RW	0.155***	0.193***	0.229***	0.044***	0.063***	0.079***
	(9.59)	(11.48)	(13.15)	(15.41)	(22.83)	(25.70)
AR	0.141***	0.170***	0.219***	0.042***	0.059***	0.074***
	(12.46)	(17.34)	(15.87)	(18.38)	(26.40)	(25.55)
RIV	0.142***	0.166***	0.205***	0.042***	0.058***	0.073***
	(12.35)	(18.03)	(19.43)	(18.48)	(26.86)	(26.04)
Comparison						
HVZ-RW	0.055***	0.090***	0.127***	-0.001	-0.003**	-0.004**
	(5.54)	(4.73)	(4.92)	(-0.84)	(-2.06)	(-2.37)
HVZ-AR	0.069***	0.114***	0.137***	0.001**	0.001*	0.001
	(7.70)	(7.96)	(6.98)	(2.46)	(1.82)	(1.68)
HVZ-RIV	0.069***	0.117***	0.150***	0.001**	0.002**	0.002***
	(7.40)	(7.73)	(7.58)	(2.33)	(2.54)	(3.03)
RW-AR	0.014*	0.024*	0.010	0.002**	0.004***	0.006***
	(1.84)	(1.89)	(0.57)	(2.17)	(2.73)	(2.97)
RW-RIV	0.013*	0.027**	0.024*	0.002**	0.005***	0.007***
	(1.78)	(2.48)	(1.70)	(2.08)	(3.01)	(3.58)
AR-RIV	-0.001	0.003	0.014**	0.000	0.000***	0.001***
	(-0.90)	(1.19)	(2.62)	(-1.35)	(3.48)	(4.99)

Panel A reports the time-series averages of the mean and median forecast accuracy for the three cross-sectional earnings models and the random walk model, and their pair-wise comparisons. The time-series t-statistics are reported in the parentheses. ***, **, * denote significance at 0.01, 0.05, and 0.10 level, respectively. Forecast accuracy is the absolute value of forecast bias, which is the difference between actual earnings and model-based earnings forecasts scaled by the end-of-June market value of equity (HVZ and RW) or by the end-of-June stock price (AR and RIV). The results are based on the common sample of 119,653 firm-year observations with non-missing t+1, t+2, and t+3 forecast bias from all models. Panel B reports mean forecast accuracy by partition of analyst coverage. A firm is covered by analysts if there is one FY1 consensus forecast on IBES for year t+1. Panel C reports mean forecast accuracy by partition of firm size. Each year, observations are sorted into two equal sized groups based on their end-of-June market value of equity.

Table 3 Forecast bias of the cross-sectional earnings models and the random walk model

Panel A: Full sample of 119,653 firm-year observations (1969-2008)

	E_{t+1}		E_{t+2}		E_{t+3}	
	Mean	Median	Mean	Median	Mean	Median
HVZ	-0.056***	-0.010***	-0.088***	-0.023***	-0.130***	-0.039***
	(-6.10)	(-3.28)	(-6.28)	(-4.86)	(-8.21)	(-6.73)
RW	0.011*	0.008***	0.035***	0.015***	0.046***	0.021***
	(1.86)	(7.04)	(4.00)	(6.71)	(4.45)	(6.49)
AR	-0.019***	0.003**	-0.030***	-0.003	-0.054***	-0.012**
	(-4.41)	(2.04)	(-4.20)	(-0.82)	(-5.09)	(-2.63)
RIV	-0.013***	0.004***	-0.017***	0.000	-0.034***	-0.007*
	(-3.12)	(3.01)	(-2.71)	(0.12)	(-3.85)	(-1.78)
Comparison						
HVZ-RW	-0.067***	-0.018***	-0.123***	-0.038***	-0.176***	-0.061***
	(-8.59)	(-5.57)	(-9.48)	(-7.20)	(-10.56)	(-8.70)
HVZ-AR	-0.037***	-0.013***	-0.058***	-0.021***	-0.076***	-0.027***
	(-5.02)	(-6.10)	(-5.66)	(-6.66)	(-5.94)	(-6.62)
HVZ-RIV	-0.043***	-0.014***	-0.071***	-0.023***	-0.096***	-0.032***
	(-5.62)	(-6.33)	(-6.73)	(-7.00)	(-7.64)	(-7.32)
RW-AR	0.030***	0.005***	0.065***	0.018***	0.100***	0.033***
	(6.49)	(3.58)	(9.11)	(6.17)	(9.80)	(8.06)
RW-RIV	0.024***	0.004***	0.052***	0.015***	0.080***	0.029***
	(5.07)	(2.96)	(8.05)	(5.88)	(9.34)	(7.90)
AR-RIV	-0.006***	-0.001***	-0.013***	-0.003***	-0.020***	-0.005***
	(-7.80)	(-5.72)	(-7.82)	(-5.39)	(-7.79)	(-6.65)

Panel B: Partition analysis of mean forecast bias by analyst coverage

	No coverage (N=50,242)			With coverage (N=69,411)		
	E_{t+1}	E_{t+2}	E_{t+3}	E_{t+1}	E_{t+2}	E_{t+3}
HVZ	-0.072***	-0.128***	-0.214***	-0.020***	-0.021***	-0.032***
	(-4.39)	(-5.96)	(-9.67)	(-3.88)	(-2.93)	(-4.35)
RW	0.023**	0.056***	0.061***	0.004	0.023**	0.032**
	(2.17)	(3.74)	(3.73)	(0.68)	(2.23)	(2.62)
AR	-0.010	-0.033***	-0.082***	-0.021***	-0.019***	-0.027***
	(-1.67)	(-3.05)	(-4.77)	(-4.59)	(-3.48)	(-4.41)
RIV	0.002	-0.006	-0.042***	-0.019***	-0.015**	-0.020***
	(0.37)	(-0.63)	(-3.06)	(-4.23)	(-2.74)	(-3.25)
Comparison						
HVZ-RW	-0.095***	-0.184***	-0.275***	-0.024***	-0.044***	-0.063***
	(-7.05)	(-9.67)	(-12.42)	(-7.49)	(-7.96)	(-8.88)
HVZ-AR	-0.062***	-0.096***	-0.132***	0.001	-0.001	-0.004
	(-4.20)	(-5.97)	(-8.06)	(0.39)	(-0.42)	(-1.24)
HVZ-RIV	-0.075***	-0.122***	-0.172***	-0.001	-0.006*	-0.011***
	(-4.68)	(-6.90)	(-9.98)	(-0.36)	(-2.04)	(-3.71)
RW-AR	0.033***	0.089***	0.143***	0.025***	0.043***	0.059***
	(3.99)	(7.58)	(8.22)	(4.67)	(5.74)	(6.78)
RW-RIV	0.020**	0.062***	0.103***	0.023***	0.038***	0.052***
	(2.34)	(6.03)	(7.62)	(4.23)	(5.22)	(6.20)
AR-RIV	-0.012***	-0.027***	-0.040***	-0.002***	-0.004***	-0.007***
	(-7.00)	(-7.19)	(-7.11)	(-5.45)	(-8.64)	(-10.04)

Table 3 continued

Panel C: Partition analysis of mean forecast bias by size

	Small Firms (N=59,819)			Large Firms (N=59,834)		
	E_{t+1}	E_{t+2}	E_{t+3}	E_{t+1}	E_{t+2}	E_{t+3}
HVZ	-0.106***	-0.168***	-0.250***	-0.007**	-0.009**	-0.010**
	(-6.11)	(-6.36)	(-8.33)	(-2.63)	(-2.37)	(-2.32)
RW	0.017*	0.058***	0.071***	0.004	0.012***	0.021***
	(1.78)	(3.96)	(4.27)	(1.60)	(3.04)	(4.06)
AR	-0.035***	-0.056***	-0.101***	-0.003	-0.005	-0.007
	(-5.20)	(-4.76)	(-5.62)	(-1.30)	(-1.27)	(-1.54)
RIV	-0.023***	-0.030***	-0.063***	-0.003	-0.003	-0.005
	(-3.57)	(-3.02)	(-4.34)	(-1.06)	(-0.97)	(-1.14)
Comparison						
HVZ-RW	-0.123***	-0.225***	-0.321***	-0.011***	-0.021***	-0.031***
	(-8.30)	(-9.24)	(-10.38)	(-10.13)	(-10.02)	(-10.56)
HVZ-AR	-0.071***	-0.112***	-0.149***	-0.003***	-0.004***	-0.003**
	(-4.87)	(-5.60)	(-6.02)	(-5.76)	(-3.98)	(-2.42)
HVZ-RIV	-0.083***	-0.137***	-0.187***	-0.004***	-0.005***	-0.005***
	(-5.48)	(-6.69)	(-7.76)	(-6.13)	(-4.57)	(-3.42)
RW-AR	0.052***	0.113***	0.172***	0.007***	0.017***	0.028***
	(6.29)	(8.94)	(9.50)	(6.86)	(9.05)	(10.73)
RW-RIV	0.040***	0.088***	0.134***	0.007***	0.016***	0.026***
	(4.82)	(7.84)	(9.08)	(6.02)	(8.31)	(10.12)
AR-RIV	-0.012***	-0.025***	-0.038***	-0.001***	-0.001***	-0.002***
	(-7.46)	(-7.51)	(-7.45)	(-3.39)	(-3.94)	(-4.96)

Panel A reports the time-series averages of the mean and median forecast bias for the three cross-sectional earnings models and the random walk model, and their pair-wise comparisons. The time-series t-statistics are reported in the parentheses. ***, **, * denote significance at 0.01, 0.05, and 0.10 level, respectively. Forecast bias is the difference between actual earnings and model-based earnings forecasts scaled by the end-of-June market value of equity (HVZ and RW) or by the end-of-June stock price (AR and RIV). The results are based on the common sample of 119,653 firm-year observations with non-missing t+1, t+2, and t+3 forecast bias from all models. Panel B reports mean forecast bias by partition of analyst coverage. A firm is covered by analysts if there is one FY1 consensus forecast on IBES for year t+1. Panel C reports mean forecast bias by partition of firm size. Each year, observations are sorted into two equal sized groups based on their end-of-June market value of equity.

Table 4 Earnings response coefficient of the model-based earnings forecast, 1969-2008.

Panel A: Full sample of 119,653 firm-year observations

	E_{t+1}	E_{t+2}	E_{t+3}
HVZ	0.060***	0.089***	0.126***
	(6.29)	(5.99)	(6.44)
RW	0.087***	0.172***	0.285***
	(8.87)	(10.94)	(8.94)
AR	0.102***	0.200***	0.318***
	(16.09)	(16.57)	(16.29)
RIV	0.102***	0.206***	0.326***
	(16.11)	(16.27)	(15.67)
Comparison			
HVZ-RW	-0.026***	-0.083***	-0.159***
	(-3.72)	(-4.27)	(-3.69)
HVZ-AR	-0.042***	-0.111***	-0.192***
	(-5.78)	(-7.67)	(-7.47)
HVZ-RIV	-0.042***	-0.117***	-0.200***
	(-5.50)	(-7.13)	(-7.01)
RW-AR	-0.015**	-0.028**	-0.033
	(-2.18)	(-2.31)	(-1.34)
RW-RIV	-0.015**	-0.034***	-0.041*
	(-2.28)	(-2.84)	(-1.91)
AR-RIV	0.000	-0.006	-0.008
	(-0.26)	(-1.60)	(-1.52)

Panel B: Partition analysis of ERC by analyst coverage

	No coverage (N=50,242)			With coverage (N=69,411)		
	E_{t+1}	E_{t+2}	E_{t+3}	E_{t+1}	E_{t+2}	E_{t+3}
HVZ	0.055***	0.061***	0.081***	0.170***	0.319***	0.587***
	(5.85)	(4.57)	(4.47)	(6.63)	(6.91)	(6.43)
RW	0.094***	0.177***	0.281***	0.170***	0.350***	0.577***
	(6.86)	(5.83)	(6.30)	(5.87)	(4.19)	(3.89)
AR	0.100***	0.179***	0.275***	0.153***	0.346***	0.626***
	(12.41)	(12.55)	(12.21)	(8.08)	(6.37)	(6.69)
RIV	0.100***	0.183***	0.280***	0.153***	0.340***	0.602***
	(11.71)	(12.25)	(11.71)	(9.02)	(7.56)	(8.03)
Comparison						
HVZ-RW	-0.039***	-0.117***	-0.200***	0.001	-0.031	0.010
	(-3.30)	(-3.53)	(-3.82)	(0.07)	(-0.63)	(0.13)
HVZ-AR	-0.045***	-0.118***	-0.194***	0.017	-0.028	-0.039
	(-6.66)	(-7.74)	(-7.60)	(0.80)	(-0.84)	(-0.63)
HVZ-RIV	-0.045***	-0.123***	-0.199***	0.017	-0.021	-0.015
	(-6.30)	(-7.23)	(-6.99)	(0.79)	(-0.73)	(-0.27)
RW-AR	-0.006	-0.002	0.006	0.017	0.003	-0.050
	(-0.54)	(-0.07)	(0.16)	(0.63)	(0.05)	(-0.38)
RW-RIV	-0.007	-0.006	0.001	0.017	0.010	-0.025
	(-0.55)	(-0.24)	(0.02)	(0.62)	(0.13)	(-0.20)
AR-RIV	0.000	-0.004	-0.005	0.000	0.006	0.024
	(-0.19)	(-1.21)	(-0.87)	(0.05)	(0.46)	(0.91)

Table 4 continued

Panel C: Partition analysis of ERC by size

	Small Firms (N=59,819)			Large Firms (N=59,834)		
	E_{t+1}	E_{t+2}	E_{t+3}	E_{t+1}	E_{t+2}	E_{t+3}
HVZ	0.056***	0.080***	0.117***	0.403***	0.909***	1.459***
	(6.34)	(6.27)	(8.32)	(11.49)	(11.52)	(11.18)
RW	0.078***	0.149***	0.248***	0.363***	0.740***	1.180***
	(8.26)	(10.21)	(8.00)	(8.03)	(8.80)	(7.46)
AR	0.096***	0.186***	0.298***	0.270***	0.595***	0.998***
	(15.16)	(14.45)	(12.67)	(10.73)	(10.14)	(9.55)
RIV	0.096***	0.191***	0.304***	0.271***	0.590***	0.969***
	(14.82)	(13.92)	(11.93)	(10.87)	(10.50)	(10.29)
Comparison						
HVZ-RW	-0.022***	-0.069***	-0.131***	0.039	0.169**	0.279**
	(-4.06)	(-4.68)	(-3.82)	(1.45)	(2.39)	(2.70)
HVZ-AR	-0.040***	-0.106***	-0.181***	0.133***	0.314***	0.462***
	(-6.19)	(-7.96)	(-7.60)	(4.69)	(4.98)	(4.69)
HVZ-RIV	-0.040***	-0.111***	-0.187***	0.131***	0.320***	0.490***
	(-5.96)	(-7.27)	(-6.97)	(4.62)	(5.05)	(5.16)
RW-AR	-0.018***	-0.037***	-0.050**	0.093**	0.145**	0.183
	(-2.96)	(-3.75)	(-2.48)	(2.44)	(2.20)	(1.37)
RW-RIV	-0.018***	-0.042***	-0.056***	0.092**	0.150**	0.211*
	(-3.01)	(-4.20)	(-3.36)	(2.45)	(2.40)	(1.72)
AR-RIV	0.000	-0.005	-0.006	-0.001	0.005	0.029
	(-0.11)	(-1.44)	(-1.08)	(-0.43)	(0.63)	(1.30)

Panel A reports the time-series averages of the earnings response coefficients (ERC) for the forecasts from the three cross-sectional earnings models and the random walk model, and their pair-wise comparisons. The time-series t-statistics are reported in the parentheses. ***, **, * denote significance at 0.01, 0.05, and 0.10 level, respectively. ERC is estimated by regressing the buy-and-hold returns over the next one, two, and three years on the unexpected earnings (i.e., the forecast bias) over the same horizon. We standardize the unexpected earnings so that they have unit variance each year. The results are based on the common sample of 119,653 firm-year observations with non-missing t+1, t+2, and t+3 forecast bias from all models. Panel B reports mean ERC by partition of analyst coverage. A firm is covered by analysts if there is one FY1 consensus forecast on IBES for year t+1. Panel C reports mean ERC by partition of firm size. Each year, observations are sorted into two equal sized groups based on their end-of-June market value of equity.

Table 5 Return spreads for quintiles of implied cost of capital using model-based forecasts

Panel A: Mean return spreads (%)

Model	Quintile	ICC	$RET_1 - R_F$	$RET_2 - R_F$	$RET_3 - R_F$
HVZ	1	-1.2	4.87	4.36	5.29
	2	1.5	6.37	7.58	7.86
	3	4.4	8.23	8.33	8.21
	4	8.5	8.94	9.50	9.90
	5	21.5	9.43	10.36	8.91
	5 – 1	22.7***	4.56***	6.00***	3.62***
	(t-stat)	(275.44)	(6.98)	(8.80)	(5.39)
AR	1	0.0	5.14	4.74	5.23
	2	1.9	7.01	7.81	8.29
	3	3.3	8.18	8.90	8.70
	4	5.0	9.38	9.59	9.53
	5	11.1	11.45	11.64	12.40
	5 – 1	11.1***	6.30***	6.90***	7.17***
	(t-stat)	(198.04)	(9.92)	(10.08)	(10.07)
RIV	1	-0.5	4.90	4.53	5.28
	2	1.6	6.58	7.28	8.22
	3	3.2	8.00	9.17	8.70
	4	5.0	9.97	10.27	10.69
	5	10.2	12.15	11.60	11.65
	5 – 1	10.7***	7.25***	7.07***	6.37***
	(t-stat)	(220.15)	(11.90)	(10.97)	(9.34)

Panel B: Comparison of mean return spreads (%) across models

	$RET_1 - R_F$	$RET_2 - R_F$	$RET_3 - R_F$
HVZ - AR	-1.75*	-0.90	-3.56***
(t-stat)	(-1.92)	(-0.93)	(-3.64)
HVZ - RIV	-2.69***	-1.07	-2.75***
(t-stat)	(-3.01)	(-1.14)	(-2.87)
AR - RIV	-0.95	-0.18	0.81
(t-stat)	(-1.08)	(-0.19)	(0.82)

Firms are divided into quintiles each year based on the implied cost of capital metric (ICC) computed for each of the three models (i.e., HVZ, AR and RIV). See Appendix A for details of the model estimation and Appendix B for ICC estimation. Panel A presents the pooled equally weighted average of buy-and-hold returns for the first three years after portfolio formation, adjusted for the risk-free rate (RET_1-R_F, RET_2-R_F, and RET_3-R_F, respectively) for quintiles based on ICC as well as the spread between the extreme quintiles. Panel B reports the pair-wise comparisons the spreads. Figures in parentheses represent t-statistics, calculated using a pooled estimate of standard error. ***, **, * denote significance at 0.01, 0.05, and 0.10 level, respectively.

Table 6 Regression of future returns on implied cost of capital using model-based forecasts

Panel A: Univariate regression of future returns on ICC

	RET_1 - R_F			RET_2 - R_F			RET_3 - R_F		
Model	Intercept	ICC	Adj. R^2	Intercept	ICC	Adj. R^2	Intercept	ICC	Adj. R^2
HVZ	0.053*	0.209**	0.96%	0.058**	0.192**	0.76%	0.065***	0.087	0.75%
	(1.90)	(2.27)		(2.04)	(2.16)		(2.22)	(1.04)	
AR	0.043*	0.584***	1.21%	0.054**	0.498***	0.94%	0.053***	0.576***	1.12%
	(1.69)	(2.59)		(2.03)	(2.35)		(2.04)	(2.63)	
RIV	0.046*	0.652***	1.27%	0.055**	0.574***	0.95%	0.054***	0.594***	1.15%
	(1.74)	(3.11)		(1.98)	(2.89)		(1.98)	(2.74)	

Panel B: Comparison of coefficient on ICC across the models

	RET_1 - R_F	RET_2 - R_F	RET_3 - R_F
HVZ - AR	-0.375	-0.306	-0.489
(t-stat)	(-1.54)	(-1.33)	(-2.08)**
HVZ - RIV	-0.443	-0.382	-0.506
(t-stat)	(-1.93)*	(-1.76)*	(-2.18)**
AR - RIV	-0.068	-0.076	-0.017
(t-stat)	(-0.22)	(-0.26)	(-0.06)

Panel C: Regression of future returns on all ICC metrics

	Intercept	ICC_{HVZ}	ICC_{AR}	ICC_{RIV}	Adj. R^2
RET_1 - R_F	0.043*	0.058	-0.239	0.929***	2.34%
	(1.75)	(0.54)	(-0.73)	(3.15)	
RET_2 - R_F	0.049*	0.046	-0.254	0.923***	1.92%
	(1.86)	(0.42)	(-0.73)	(3.53)	
RET_3 - R_F	0.049*	-0.106	-0.197	1.064***	2.15%
	(1.88)	(-0.95)	(-0.47)	(3.02)	

Panel A presents univariate Fama and MacBeth (1973) regressions of future realized returns on metrics of implied cost of capital (ICC) computed for each of the three models. See Appendix A for details of the model estimation and Appendix B for ICC estimation. The dependent variables are the buy-and-hold returns for the first three years after portfolio formation, adjusted for the risk-free rate (RET_1-R_F, RET_2-R_F, and RET_3-R_F, respectively). Panel B reports the pair-wise comparisons of the coefficients on ICC. Figures in parentheses represent t-statistics, calculated using a pooled estimate of standard error. Panel C reports Fama and MacBeth regressions of future realized returns on the three ICC metrics (i.e., ICC_{HVZ}, ICC_{AR} and ICC_{RIV}). ***, **, * denote significance at 0.01, 0.05, and 0.10 level, respectively.

Table 7 Implied cost of capital metrics and risk factors

Metric	Intercept	β (+)	LMCAP (-)	BM (+)	IDIO (+)	STDNI (+)	D2A (+)	LFOLLOW (-)	Adj. R^2
ICC_{HVZ}	0.071^{***}	-0.011^{***}							1.6%
	(11.63)	(-3.44)							
	0.172^{***}	-0.018^{***}	-0.025^{***}	0.047^{***}					53.4%
	(13.33)	(-11.98)	(-13.46)	(17.55)					
	0.163^{***}	-0.018^{***}	-0.025^{***}	0.047^{***}	-0.006	0.021^{***}	0.021^{***}	-0.0002	54.7%
	(12.41)	(-9.00)	(-13.31)	(17.11)	(-0.33)	(5.41)	(5.50)	(-0.53)	
ICC_{AR}	0.041^{***}	0.002							1.0%
	(21.06)	(1.53)							
	0.082^{***}	-0.0002	-0.011^{***}	0.023^{***}					40.5%
	(14.92)	(-0.34)	(-12.11)	(22.83)					
	0.039^{***}	-0.008^{***}	-0.007^{***}	0.025^{***}	0.212^{***}	0.031^{***}	0.023^{***}	-0.0010^{***}	46.4%
	(7.37)	(-8.14)	(-9.51)	(19.59)	(8.39)	(5.41)	(8.91)	(-4.21)	
ICC_{RIV}	0.043^{***}	-0.002							2.2%
	(18.60)	(-1.44)							
	0.050^{***}	-0.003^{***}	-0.007^{***}	0.040^{***}					62.7%
	(8.29)	(-5.89)	(-7.89)	(20.55)					
	0.029^{***}	-0.007^{***}	-0.005^{***}	0.041^{***}	0.101^{***}	0.025^{***}	0.017^{***}	-0.0001	66.0%
	(6.27)	(-7.68)	(-7.70)	(20.95)	(4.13)	(3.06)	(7.20)	(-0.25)	

This table presents firm level regressions of the ICC metrics on the following risk factors: β (systematic risk), LMCAP (size), BM (book-to-market), IDIO (idiosyncratic risk), STDNI (earnings volatility), D2A (leverage) and LFOLLOW (analyst coverage). β is calculated using monthly returns over the lagged five years (ensuring that at least 24 observations are available). LMCAP is the logarithm of market capitalization at the time of the forecasts. IDIO is the standard deviation of the prior year's monthly returns. STDNI is the standard deviation of net income (IBQ) scaled by total assets (ATQ) measured over the previous eight quarters. D2A is the ratio of total debt (DLTT+DLC) to total assets (AT). LFOLLOW is the logarithm of 1+number of analysts following the stock. See Appendix A and B for details of the model estimation and ICC estimation. Regressions are estimated using the Fama and MacBeth (1973) procedure. Figures in parentheses are t-statistics. ***, **, * denote significance at 0.01, 0.05, and 0.10 level, respectively.