Department of Business and Management
Double Degree Program in Corporate Finance
Course of Advanced Corporate Finance

The ICC: A cross-sectional estimation of the Cost of Equity in the Biotech and Pharmaceutical industry

Prof. Pierluigi Murro
SUPERVISOR

Prof. Rosella Santella
CO-SUPERVISOR

Lorenzo Checcacci
ID - 776031
CANDIDATE

Academic Year 2024/2025

Abstract

This study explores the applicability of a cross-sectional earnings model to generate earnings forecasts in the pharmaceutical and biotechnology sector. The earnings forecasts are then used as inputs to calculate firms' cost of equity through the ICC approach. The study finds that a cross-sectional regression based on a limited set of accounting variables is able to provide reliable estimates of future earnings. The inclusion of further predictors to capture sector-specific earnings generation dynamics only marginally improves the performance of the baseline regression. Furthermore, the model-based ICC is positively related to realized returns, representing a reliable proxy for expected returns.

1 Introduction

The goal of this study is to check for the robustness of the cross-sectional estimation of future earnings to calculate a firm's cost of equity, following the approach proposed by Hou et al. (2012). This research evaluates the applicability of the model of Hou et al. (2012) in the pharmaceutical and biotechnology sector, as well as its ability to capture sector-specific dynamics of earnings generation. Furthermore, this study explores the effects of adding R&D expenses to the baseline model as an additional predictor of future earnings, and whether it is possible to improve the accuracy of the estimation of the cost of equity by including factors that capture industry-specific dynamics.

The estimation of the cost of equity capital of firms has long been a central issue in the corporate finance literature. The cost of equity expresses the expected return of a stock or an equity investment, and is a key element to evaluate the risk-return profile of investment alternatives. Expected returns also play a role in valuation, capital budgeting, and portfolio allocation, as well as risk control and performance valuation Hou et al. (2012).

Despite its relevance, to date there is still not a commonly accepted approach to calculate the cost of equity in the corporate finance literature. Most academic studies rely on traditional asset pricing models, such as the CAPM (Sharpe, 1964; Lintner, 1965), which use realized returns as a proxy for unobservable expected returns. Despite their popularity among practitioners, CAPM-related models lack theoretical foundation, as realized returns provide imprecise and noisy estimates of expected returns (Fama and French, 1997).

A more recent stream of literature proposed an alternative approach for the computation of the cost of equity: the implied cost of capital (ICC). The ICC of a firm is the rate of return that equates a firm's stock price to the present value of its expected future cash flows (Hou et al., 2012). This approach is theoretically appealing, as expected returns are directly estimated from current stock prices and cash flows forecasts, instead of relying on noisy and imprecise asset pricing models.

Over the years, the academic literature has analysed the relationship between the ICC and realized returns, exploring different accounting-based techniques to provide more sound estimates of the cost of equity capital. The commonly used approach is to use analysts' earnings forecasts as a proxy for expected earnings, which would then be the input for the ICC calculation. Unfortunately, analysts' estimates have been found to be systematically biased (Francis and Philbrick, 1993) and to suffer from self-selection problems (McNichols and O'Brien, 1997). When included in the ICC calculation, analysts' earnings forecasts yield biased estimates of the cost of equity.

To overcome these limitations, Hou et al. (2012) propose a new approach to estimate the ICC. In particular, they introduce a cross-sectional regression model, which mechanically forecasts future earnings from a set of accounting variables, instead of relying on analysts' estimates. The main advantage of this approach is to provide unbiased forecasts of future earnings, which, in turn, yield reliable expectations of equity returns. Furthermore, it solves analysts' self-selection and coverage issues, while also enhancing the simplicity of the calculation of the cost of equity.

Despite some criticism in terms of model performance (Gerakos and Gramacy, 2013; Li and Mohanram, 2014), the approach proposed by Hou et al. (2012) has been widely used in the ICC literature, due to its simplicity and the quality of its estimates. The authors show that the cross-sectional regression used to forecast future earnings is able to explain up to 86% of the variance of realized earnings. Furthermore, the ICC derived from model-based estimates is positively correlated with realized returns.

This study contributes to the ICC literature by checking the robustness of the model introduced by Hou et al. (2012), as well as exploring potential expansions of the cross-sectional regression, depending on sector-specific dynamics of earnings generation. In particular, this research contributes to the academic literature on the cost of equity estimation in the following ways. First, since the analysis of Hou et al. (2012) is only performed on observations up to 2008, the same methodologies are replicated in an updated timeframe to evaluate whether the main findings also apply to a period of financial turmoil, which includes the 2008–2012 financial crisis as well as the Covid-19 pandemic. Then, this research restricts its analyses to firms in the pharmaceutical and biotechnology sector to study whether the model of Hou et al. (2012) is also applicable to industries with unique earnings generation dynamics. This sector has had a growing impact on the broader economy since the outbreak of the Covid-19 pandemic (Juneja et al., 2024), and its market and investment dynamics are different from those of other economic industries. For these reasons, this sector is an ideal candidate to test whether the model-based ICC provides sound estimates of the cost of equity in industries with specific economic dynamics.

Finally, this study expands the model of Hou et al. (2012) by including R&D expenses among the predictors of future earnings. The idea behind the proposed expansion is that, depending on sample composition, additional independent variables can be added to the baseline model to better capture earnings generation dynamics and improve earnings forecast precision. Since the pharmaceutical and biotechnology sector is heavily reliant on R&D outlays, R&D expenses have been chosen as an additional predictor.

The main findings are as follows. First, the cross-sectional model proposed by Hou et al. (2012) is highly effective in providing unbiased earnings estimates for firms in the pharmaceutical and biotechnology industry in an updated timeframe. Its explanatory power is in line with the original results, ensuring robustness even in industries with peculiar dynamics. Second, the expanded model, including R&D expenses, only marginally improves the explanatory power of the baseline regression, suggesting that the original model is already capable of capturing sector-specific dynamics. Third, the sample composition suggests that the ICC methodology is not applicable to the majority of firms in the sector due to negative expected earnings. Lastly, for firms where ICC is computable, model-based ICCs are positively related to realized returns, supporting the validity of the cost of equity estimates.

This study is structured as follows: Section 2 reviews the literature on ICC; Section 3 describes the data and methodology; Section 4 presents and discusses the results; Section 5 concludes.

2 Literature Review

The estimation of the cost of equity is essential to evaluate the risk-return profile of an investment, as well as for business valuations and capital budgeting decisions. The cost of equity is the rate of return investors require on an equity investment in a firm (Damodaran, 2002), as it is the return that investors can earn by investing in equities of comparable risk (DePamphilis, 2012), representing a relevant benchmark for investment decisions. Theoretically, the cost of equity equals the discount rate of future cash payments to equity holders and therefore it should be positively related to expected future returns.

Despite its relevance in the corporate finance setting, there is not an overall accepted way to compute the cost of equity. Traditionally, the cost of equity is estimated using the Sharpe–Lintner Capital Asset Pricing Model (CAPM), according to which investors are risk averse and choose mean-variance-efficient portfolios (Sharpe, 1964; Lintner, 1965). The CAPM predicts the expected return on equity investments (i.e., the cost of equity) according to the following formula:

$$E(R_i) = R_f + \beta \cdot [E(R_m) - R_f] \tag{1}$$

Where:

- $E(R_i)$ is the expected return on a given equity investment;
- R_f is the rate of return of riskless investments;
- β is the covariance of the equity investment return and the market return over the variance of the market return, and represents the riskiness of the investment;
- $E(R_m) - R_f$ is the risk premium per unit of beta required to invest in risky assets.

Extensive research has been devoted to testing whether the traditional CAPM provides a satisfactory estimate of the cost of equity, highlighting several shortcomings. Fama and French (2004) find that the relation between the beta and the average return is much flatter than the one suggested by the CAPM, advocating for either the failure to completely incorporate an asset's risk or market failures in asset pricing. Several extensions have been implemented in the traditional CAPM, such as the Fama–French three-factor model (Fama and French, 1995) and the Fama–French–Carhart model (Carhart, 1997), but these approaches, despite being widely used, lack a solid theoretical motivation. Fama and French (1997) conclude that the estimates from the traditional CAPM and its extensions are unavoidably imprecise, as they raise problems in defining the correct model specification and estimating risk premia and factor loadings.

Overall, CAPM-related models use past returns as a proxy for expected returns, which is useful considering that expected returns are not observable (Gebhardt et al., 2001). Unfortunately, average realized returns are an imprecise estimate for the equity cost of capital (Fama and French, 1997), as they introduce noise in specifying the correct asset pricing model.

To overcome the deficiencies of the traditional approaches, a stream of literature has started to explore new methodologies for the estimation of the cost of equity. In particular, the cost of equity has been defined as the implied cost of capital (ICC), that is, the expected return that equates a stock's current price to the present value of its expected future free cash flows to equity holders (Li et al., 2013). This approach is theoretically sound, as the cost of equity is, by definition, the internal rate of return (IRR) of future cash payments.

The use of the ICC approach traces back to the research of Botosan (1997), who studied the impact of the disclosure level on the cost of equity. In particular, she argues that, while the traditional CAPM fails to incorporate relevant forward-looking information and introduces noise in the model, the accounting-based valuation models by Edwards and Bell (1961), Ohlson (1995), and Feltham and Ohlson (1995) effectively relate expected earnings to the cost of equity. To confirm the validity of the estimates, the author documents a positive relationship between beta and the ICC and a negative relationship between size and the ICC.

According to Echterling et al. (2015), there are three main accounting-based models to estimate the ICC: the dividend discount model (DDM), the residual income model (RIM), and the abnormal earnings growth model (AEG). The three approaches should theoretically yield identical results, as they are per definition equal and they all consider a constant discount factor of future income streams to derive the equity value, but they differ in the number and nature of inputs required.

Gordon and Gordon (1997) elaborate a two-stage model which starts with a basic DDM approach and considers explicitly forecasted dividends in the short term and a long-term growth rate. According to this method, the equity value is entirely determined by the sum of expected payments to equity holders, discounted to present date. This approach is beneficial due to its simplicity and the low number of inputs required but is highly sensitive to the quality of the related assumptions.

According to the RIM, the market value of equity is equal to the book value of equity, plus the future excess earnings discounted to the valuation date. Claus and Thomas (2001) estimate the cost of equity using the RIM approach, by considering a five-year explicit forecast and a long-term growth rate which is equal to the risk-free rate. Gebhardt et al. (2001) follow a similar logic, but they assume that the abnormal returns decrease over time, so that the ROE of an individual firm gradually converges to the median industry ROE. Overall, the RIM should generally yield more precise estimates than the DDM, as the valuation is anchored to the common shareholders' equity and only a smaller portion of the calculations is sensitive to the assumptions.

The AEG is a variation of the RIM, computing the equity value as the current level of earnings plus the expected variation of abnormal earnings, all discounted to the valuation date. Easton (2004) and Ohlson and Juettner-Nauroth (2005) explore this approach, by developing models that allow for a variation in both the short-term and the long-term variation in abnormal earnings growth. The purpose of this approach is to further refine the ICC calculation, by reducing the sensitivity to initial assumptions.

Over time, the ICC has been widely used and tested in academic research. Gebhardt et al. (2001) find that the firm-level ICC is positively correlated with ex-ante firm characteristics that are known to be associated with five categories of risk, therefore validating the robustness of their results. Moreover, they find a surprisingly low correlation between the ICC and the beta, which is an indication that historical returns are not a good proxy for expected returns. Finally, the authors find that most of the cross-sectional variation of the ICC is explained by four variables: the book-to-market ratio, the dispersion in analysts' forecasts, the long-term growth forecast and the industry risk premium. Claus and Thomas (2001) implement a similar approach, using consensus analysts' forecasts as an input in the ICC calculation. They show that the use of forward-looking information provides a consistently lower estimate of the cost of equity compared to observed returns.

Easton and Monahan (2003) evaluate the reliability of the accounting-based measures used to estimate expected returns within the ICC calculation, by employing a model that includes expected returns, cash flow news and return news. They find that the RIM approach from Gebhardt et al. (2001) is the most precise model in terms of measurement error variance to realized returns, but they highlight how the use of a simpler price-to-forward earnings model yields at least as accurate results as more complicated models. Schröder (2005) extends the examination of ICC estimates to an international setting, by evaluating the relative performance of the different models for European companies. The author finds that, while RIM and DDM approaches yield similar results, the DDM performs better in predicting future stock returns and is more in line with typical asset pricing factors such as the beta, the size or the B/M of a firm. This result is opposite to that of Easton and Monahan (2003) and suggests that the results may or may not be generalizable depending on the sample.

It is worth mentioning that, in the literature, the ICC approach has also been used at the market level. Li et al. (2013) investigate the ability of the ICC to predict future market returns, arguing that the market-level ICC is a less noisy measure than the firm-level ICC. They build a measure of the aggregate ICC, which is the weighted average of the ICC of each firm in the S&P 500 index, computed with a standard DDM. To the aggregate ICC is then subtracted the risk-free rate, to derive the market's implied risk premium. The authors find the aggregate ICC has a strong positive relationship with both in-sample and out-of-sample expected market returns, which is a validation of the usefulness of the ICC approach at the market level.

Hughes et al. (2009) use a theoretical approach to investigate the properties of the ICC, focusing on its relationship with expected returns. They argue that some of the empirical results in the ICC literature may be an artifact, as the ICC naturally differs from expected returns implied by traditional asset pricing theory. Larocque and Lyle (2017) posit that expected returns should be positively correlated with future profitability, as firms that engage in NPV-positive projects should earn higher returns. To estimate the expected returns, four different ICC estimates based on the RIM and the AEG are used. The authors find that only RIM estimates are positively related to current profitability, while none of the estimates is positively related to expected profitability. Moreover, no ICC estimate provides additional information to a simple linear combination of current ROE and B/M of a firm. These results cast doubts on the ability of commonly used ICC models to effectively predict future returns. Penman et al. (2022) join the stream of critics to the ICC, by highlighting models' shortcomings in accounting for long-term growth. First, since the ICC is extracted from observed prices, for a given price the expected long-term growth rate should be consistent with the short-term expected earnings, but commonly used ICC models fail to ensure this condition. Second, in a typical RIM the growth rate of excess earnings is conditional to the cost of equity itself, which creates problems of circularity in the calculation.

One of the most relevant issues in the ICC literature is analysts' bias in earnings forecast estimates. Typical ICC models, in fact, use analysts' earnings forecasts as an input in the calculation of the cost of equity, as they are a proxy for expected earnings. The expected earnings are commonly retrieved from the I/B/E/S database, which provides consensus forecasts. A large body of research documents that analysts tend to be overly optimistic in their forecasts.

Francis and Philbrick (1993) argue that analysts are not always incentivized to provide the most precise earnings forecasts possible. Rather, they typically provide overly optimistic estimates for sell and hold stocks, in order to preserve their professional relationships with the company they are analysing. McNichols and O'Brien (1997) take a slightly different approach, by arguing that analysts do not necessarily report optimistic estimates, but instead self-select the coverage of their analysis to only report on stocks over which they have favourable news. Lin and McNichols (1998) corroborate the hypothesis that conflicts of interest influence analysts' forecasts by documenting that research reports issued by analysts affiliated with a company are significantly more positive than those issued by unaffiliated analysts. Das et al. (1998) and Ke and Yu (2006) align with these findings, by showing respectively that analysts' optimism is more pronounced for firms with less publicly available information and that analysts who issue more optimistic estimates are more likely to access private information from management.

The analysts' bias impact has been widely addressed in the ICC literature. Easton and Sommers (2007) show that the ICC derived from analysts' estimates is upwardly biased by close to 3% in the 1993–2004 period compared to the cost of equity derived from realized earnings. McInnis (2010) examines the relationship between earnings smoothness and the ICC and finds that the link between the two derives from analysts' failure to properly weight prior earnings changes when predicting future earnings changes, which results in systematically positively biased ICC estimates. Guay et al. (2011) report that the failure of some of the most popular models to estimate the ICC to exhibit a positive correlation with realized returns is due to the sluggishness in analysts' forecasts. Analysts, in fact, are slow to adapt their earnings estimates according to relevant information, which causes biased and imprecise results. Mohanram and Gode (2013) build on these findings and show that, adjusting for predictable forecast errors, ICC estimates show a significantly stronger relationship with realized returns.

To overcome the deficiencies of the analyst-based ICC, Hou et al. (2012) introduce a totally new framework for the ICC calculation. The authors generate earnings forecasts using a cross-sectional model, which replaces analysts' estimates as a proxy for expected cash flows. In particular, the expected earnings are mechanically derived from a relatively small and widely available set of accounting variables. These forecasted earnings are then used as an input to compute the ICC, according to the typical DDM, RIM, and AEG methodologies. The results from this paper are striking: the cross-sectional earnings forecasts explain a significant portion of the variation in realized earnings. While these forecasts are slightly less precise than analysts' estimates, they exhibit much lower bias and represent a better proxy for market expectations due to their higher earnings response coefficient (ERC). Moreover, the ICC estimates derived from the model are found to be strong predictors of future realized returns, as a portfolio that goes long on stocks with higher model-based ICC and short on stocks with lower model-based ICC yields significantly higher returns than a portfolio that follows the same strategy using analyst-based ICC. Additionally, the model-based ICC estimates are significantly correlated with firm-specific characteristics known to be risk drivers in the asset pricing literature, such as the book-to-market ratio and firm size.

The study by Hou et al. (2012) is of particular interest in the ICC literature for several reasons. First, as discussed above, it represents a much more reliable and less biased alternative to the traditional analyst-based ICC. Second, it increases coverage, as the model-based ICC can be computed for firms that are not included in analysts' forecasts. Third, the limited number of accounting variables required by the model makes it relatively easy for investors to estimate the cost of equity for a broad set of firms.

Since its introduction, the model by Hou et al. (2012) has been widely used and tested due to its efficiency and innovative approach. Among others, Jones and Tuzel (2013) employ the model from Hou et al. (2012) as an alternative to the traditional RIM to calculate the ICC, finding that the results from both methods are consistent when investigating the relationship between the cost of equity and inventory investments. Adebambo et al. (2024) follow the same approach and use the model from Hou et al. (2012) as a robustness check of traditional ICC methods. Clarkson et al. (2013) use the model introduced by Hou et al. (2012) as a cost of capital proxy to address the limited analyst coverage in the realm of environmental disclosure. This study is, therefore, a relevant demonstration of the benefits generated using the model from Hou et al. (2012) compared to traditional ICC methodologies. As an additional example, Pástor et al. (2022) employ the cross-sectional model from Hou et al. (2012) due to its superior precision in deriving the cost of equity to calculate green returns.

Despite being very popular, the model from Hou et al. (2012) has received some criticism over the years. Gerakos and Gramacy (2013) are among the first to critically evaluate regression-based earnings forecast models used to calculate the ICC. They show that simpler models, such as a naïve random walk (RW) model, lead to more accurate earnings predictions compared to models with a larger set of explanatory variables, such as that of Hou et al. (2012). In particular, the RW model performs better during tumultuous periods, while it is more effective to use a broader set of predictors in more stable market conditions.

Li and Mohanram (2014) perform several analyses to evaluate the performance of the model by Hou et al. (2012). First, they confirm that the model from Hou et al. (2012) underperforms compared to the RW in terms of forecast accuracy, therefore casting doubts on the efficacy of this approach. Then, they compare the performance of this model against an earnings persistence model (EP), which allows for earnings growth, and a typical RIM. Their results are very strong, as they find that the cross-sectional model from Hou et al. (2012) is worse than both alternatives on several dimensions, such as forecast accuracy, forecast bias, and ERC. Similar results are derived when analysing the properties of the ICC estimates obtained from the earnings forecasts in each model. In particular, the ICC estimates from both the EP and the RIM show a more conventional relationship to common risk factors such as the beta and the idiosyncratic risk compared to those of the model of Hou et al. (2012). Moreover, the buy-and-hold returns of a portfolio sorted on the basis of ICC estimates are higher for the EP and the RIM than those of Hou et al. (2012). These results are especially strong for small firms and firms without analyst coverage, which are the very settings where model-based forecasts such as those of Hou et al. (2012) are expected to be most beneficial. Overall, these results cast strong doubts on the efficacy of the model by Hou et al. (2012), as it appears to significantly underperform more simple approaches in the settings where it should offer the greatest advantages.

Hess et al. (2019) evaluate the performance of model-based approaches, such as those of Hou et al. (2012) and Li and Mohanram (2014), in comparison to traditional analyst-based approaches. They find that, despite analysts' earnings estimates being more accurate, model-based ICCs are more reliable at predicting future returns. In particular, a buy-and-hold portfolio sorted on the basis of ICC estimates provides significantly higher return spreads for model-based estimates than analyst-based estimates, indicating a stronger predictive power of future returns. Harris and Wang (2019) align with these findings and show that the model from Hou et al. (2012) is less biased and more efficient, although less accurate than analysts' forecasts. They argue that the model could be further improved by incorporating market information at shorter horizons. Overall, these studies confirm the benefits of using the model-based approach from Hou et al. (2012) to estimate the ICC as opposed to analysts' forecasts.

Schreder and Bilinski (2022) further study the performance of the model from Hou et al. (2012) and the EP and RIM from Li and Mohanram (2014) in an initial public offering (IPO) setting. They show that for firms approaching an IPO—which are commonly characterized by no stock history, no analyst coverage, and consistent losses—all three tested models yield imprecise and consistently biased earnings estimates, resulting in unreliable ICC estimates. On the other hand, a model that combines the characteristics of the one from Hou et al. (2012) and the EP—allowing loss persistence of IPO firms to be factored into the cross-sectional model of Hou et al. (2012)—produces more accurate and reliable estimates in the IPO setting.

Liu et al. (2022) take a new approach within the model-based ICC literature by evaluating the introduction of artificial intelligence (AI) to estimate future earnings. They find that an AI-based model provides more accurate and stable estimates than the model from Hou et al. (2012), while also requiring a smaller pool of data. Simon (2024) confirms that the introduction of machine learning technologies to estimate the ICC results in higher portfolio returns, by both improving accuracy and eliminating systematic distortions of model-based earnings estimates such as those from Hou et al. (2012).

Overall, the study from Hou et al. (2012) is of paramount importance in the ICC literature, having introduced a widely used model to forecast future earnings as an input in the ICC estimation. This model overcomes several shortcomings of the traditional analyst-based ICC, reducing bias, increasing coverage, and allowing economic agents to estimate the cost of equity of firms using only limited and accessible inputs. Over the years, this model has been employed and tested to confirm its performance or propose alternatives. Despite the critiques by other studies, it remains a popular approach in the cost of equity estimation, and it is therefore important to continue testing its properties and performance in different settings.

While the model developed by Hou et al. (2012) has become a relevant approach in the estimation of the ICC, the authors only provide estimates up to 2008. Since then, the broader economy has faced several structural changes on both the financial and economic sides, which may have changed the dynamics of the estimates of the cost of equity and future earnings. From a macroeconomic standpoint, the two main economic downturns were the 2008–2012 subprime financial crisis and the Covid-19 pandemic crisis.

Duarte and Rosa (2015) show how the cost of equity tends to increase in periods of financial turmoil, recession, high inflation or low GDP growth, with a peak in 2012, while it tends to decrease in periods of financial stability. Trunk and Stubelj (2013) confirm that during the 2008–2012 financial crisis the market value of companies tended to decrease more than their fundamental value, which indicates a disproportionate increase in discount rates, and therefore the cost of equity, during times of instability. On the other hand, Kouki and Elkhaldi (2017) argue that the impact of a financial crisis on the cost of equity depends on the degree of international integration of the local financial market. Finally, Pęksyk et al. (2023) show that during the 2008–2012 subprime financial crisis, the cost of equity of major U.S. firms became both higher and more volatile.

Regarding the effects of the Covid-19 pandemic, Bretscher et al. (2020) find that there was a decrease in the cross-sectional equity returns of U.S. firms, partly due to a downward revision of earnings forecasts by analysts. To estimate the effect of Covid-19 on the cost of equity, Ke (2022) employs traditional ICC methodologies such as the RIM and AEG, using analysts' forecasts as a proxy for future earnings. He finds that the cost of equity significantly increased during the Covid-19 pandemic, with firms with greater exposure to Covid-19 experiencing the strongest effects. Pęksyk et al. (2023) confirm that the cost of equity increased during the pandemic due to an initial market overreaction.

Considering the recent evolution of the economy, as well as the disruptions experienced by financial markets, it is therefore relevant to check for the robustness of the model by Hou et al. (2012), evaluating whether the model-based ICC provides reliable cost of equity estimates during financial downturns.

Another relevant evolution pertains to the emergence of a new economy, in which future earnings largely depend on investments in intangible assets (Barth et al., 2023). In particular, there has been an increasing mismatch between revenues and expenses that causes a decline in the value relevance of earnings, which Donelson et al. (2011) and Srivastava (2014) suggest is attributable to economic changes, such as business models focused on intangible outlays like R&D expenses. Barth et al. (2023) indeed find an increase in the value relevance of R&D expenses over the years, while Kanelis and Vorst (2024) report an increasing trend in the risk relevance of R&D expenses. Iqbal et al. (2024) contribute to this stream of literature by exploring the dynamics through which R&D expenses contribute to revenue generation, and therefore earnings, across time and different industries.

R&D expenses are also linked to stock returns, as R&D investments are found to be positively correlated with stock returns. Ali et al. (2012) show that firms with increases in R&D expenses experience positive abnormal returns and that these returns are at least partly attributable to the market's failure to fully recognize R&D's future implications on a timely basis, which is consistent with the market mispricing explanation from prior research (Eberhart et al., 2004). Gu (2016) also reports a positive association between R&D intensity and abnormal returns, but she argues that this relationship is due to the increased risk from R&D outlays. She posits that a firm decides to invest in a project well before its completion and will then assess whether to stop or continue the project depending on whether the expected cash flows exceed a certain threshold. Projects requiring high R&D investments must overcome a higher threshold to create value and are therefore particularly affected by shocks in expected cash flows. For this reason, high R&D investment leads to a higher risk premium and higher expected returns. On the other hand, Donelson and Resutek (2012) argue that, while R&D firms do realize positive excess returns, those are not directly related to R&D investments but are part of the larger value/growth anomaly.

In their study, Hou et al. (2012) consider potential expansions of the model by including additional earnings predictors such as R&D. They report that the inclusion of these variables does not result in a sizable increase in the explanatory power of the model, and therefore retain their original specification. Considering the theoretical framework outlined above, as well as the emergence of an economy based on intangible outlays where R&D expenses are an increasingly important determinant of future earnings, it is relevant to assess whether the inclusion of R&D contributes to increasing the reliability of earnings forecasts. The inclusion of R&D may be particularly relevant for industries with high R&D intensity, therefore providing additional insights on the efficacy of the model.

To refine the scope of the analysis and provide a more relevant contribution to the ICC literature, this study restricts its focus to a single economic sector. Each economic sector, in fact, differs in terms of capital structure, risk profile, and earnings dynamics, and therefore it is relevant to assess the reliability of ICC calculations using the model from Hou et al. (2012) considering how it relates to each industry's individual characteristics. Moreover, since earnings generation dynamics vary across sectors, potential expansions of the baseline model by including additional predictors may exhibit different performances depending on the industry of application.

This study focuses on the pharmaceutical and biotechnology sector. The pharmaceutical industry has had a growing impact on the global economy, contributing a total of $2,295 billion worldwide, of which $755 billion directly, and employing 74.5 million people globally, of which 7.8 million are employed directly (Juneja et al., 2024). This sector has also grown at an annual rate of 5.8% since 2017, making it the fastest-growing industry in 2021 due to its economic contribution during the Covid-19 pandemic. It is expected to reach a market size of $1,470 billion by 2032, driven by increased demand for medical treatments associated with an aging population (Juneja et al., 2024). For these reasons, it is relevant to assess the performance of ICC methodologies in estimating the cost of equity for this sector.

The pharmaceutical sector also heavily relies on R&D outlays. Grabowski et al. (2002) report a steady increase in R&D-to-Sales ratios starting from 1980. While Raghavendra et al. (2012) report a reduction in R&D spending during the 2008–2012 financial crisis, due to both external factors, such as the economic recession, and internal factors, such as reduced profitability, subsequent studies confirm the increasing trend of R&D outlays. DiMasi et al. (2016) find that R&D costs increased every decade from the 1970s to the 2010s, while Juneja et al. (2024) report that R&D expenses have increased at an annual rate of 4% since 2017, with the growth particularly driven by increased demand during the Covid-19 pandemic.

Not only is the pharmaceutical industry heavily reliant on R&D, but these R&D outlays are also a significant factor in revenue generation and market valuation. Iqbal et al. (2024) show that R&D expenses have a relatively long useful life for biotech and pharmaceutical firms. Regarding equity valuation, Hand (2001) finds that the market value of biotech firms is dependent on R&D expenses, while Xu (2006) shows that the volatility of stock prices for biotechnology firms is influenced by their R&D strategy, which represents an important risk factor due to earnings uncertainty. Xu et al. (2007) build on these findings and show that the market value of biotech firms is significantly influenced by the uncertainty of the outcome of R&D expenses, due to a low probability and a long time to commercialization.

The additional risk implied by R&D outcome uncertainty should therefore be reflected in the cost of equity of pharmaceutical and biotechnology firms. While Grabowski et al. (2002) rely on a typical CAPM approach for their study, they acknowledge that traditional asset pricing methodologies may understate the cost of equity of pharmaceutical firms due to their capital structure and the dynamic risk of new R&D projects. Sadorsky and Henriques (2003) develop some variations of the traditional CAPM to factor in the additional risks related to R&D outlays, such as value-at-risk and downside risk, in order to provide more realistic and upwardly corrected cost of equity estimates. Harrington (2012) evaluates the performance of the CAPM and the Fama-French model in the calculation of the cost of equity of biotech and pharmaceutical firms and finds that R&D intensity is positively correlated with the equity cost of capital and that the two models yield materially different estimates. Finally, Bruneo et al. (2024) confirm that biotech firms are indeed riskier than the market due to their reliance on R&D, but they still earn returns in excess of what traditional asset pricing models would predict.

Overall, this diverse framework indicates the relevance of studying the performance of the model-based ICC estimates from Hou et al. (2012) in specific settings. The expansion of the timeframe to the present date is useful to check for the robustness of the approach of Hou et al. (2012), by studying whether cost of equity estimates are in line with recent financial trends. Additionally, while the authors report that the inclusion of additional predictors does not increase the overall explanatory power of the model, it is worth considering industry-specific dynamics to enhance model performance depending on the sector of application. In particular, this study examines the effectiveness of including R&D outlays as an additional predictor in the pharmaceutical and biotechnology sector, which is relevant considering the increasing importance of this industry for the broader economy.

3 Data and Methodology

3.1 Sample Construction

The sample is derived from the intersection of the Compustat North America and CRSP monthly stock databases, accessed through WRDS and merged using the CRSP/Compustat Merged linking table. The sample is limited to securities listed on the NYSE, Amex, and Nasdaq, with `sharecode` 10 or 11, to align with the sample selection criteria of Hou et al. (2012). The observations span the period from 2007 to 2022, in order to extend the temporal scope of the original study by Hou et al. (2012). The sample is restricted to firms with SIC codes 2833, 2834, 2835, 2836, and 8731, reflecting the focus of this study on the pharmaceutical and biotechnology sector. Firms with missing identifiers, accounting values, or returns are excluded from the analysis.

The following variables are derived from Compustat. *Earnings* is defined as income before extraordinary items. *Book equity* corresponds to common stockholders' equity. *Total assets*, *dividends*, and *R&D expense* are also obtained from Compustat. *Accruals* are calculated as the difference between earnings and cash flow from operations.

Monthly CRSP data are aggregated to compute annual returns from July of year t to June of year $t + 1$, consistent with the timing convention in Hou et al. (2012). The *market value of equity* is calculated as the share price multiplied by the number of shares outstanding, also obtained from CRSP.

3.2 Sample Summary Statistics

Table 1 reports summary statistics of the variables included in the dataset as previously described. The sample comprises a total of 5,801 firm-year observations. The variables are defined as follows: A_t is total assets; D_t is dividend payments; DD_t is a dummy variable equal to 1 for dividend payers and 0 otherwise; E_t is earnings; $NegE$ is a dummy variable equal to 1 for firms with negative earnings and 0 otherwise; AC_t denotes accruals; and *R&D* represents research and development expenses.

As can be observed, all variables included in the table are highly skewed, which indicates substantial differences among the firms in the sample. The skewness of the distribution of the total assets variable suggests that the sample is predominantly composed of small firms, while a few large players dominate the market. The earnings variable and the negative earnings dummy show that more than 80% of firms in the sample report negative profitability. This finding is consistent with the business model of the biotech and pharmaceutical sector, where large upfront investments in R&D, patents, and technological implementation are common, causing many firms to be unprofitable during the R&D or testing phase. As will be further discussed, this characteristic imposes limitations on the ICC calculation within the biotech and pharmaceutical industry.

The dividend variable and the dividend payer dummy reveal that very few firms distribute dividends. This observation aligns with the overall lack of profitability and the capital-intensive nature of the sector. Lastly, R&D expenses are also highly skewed, with a small number of firms investing substantially more than others. This heterogeneity in R&D expenditures reflects significant variation across firms, thereby supporting the inclusion of this variable as an additional predictor in the cross-sectional earnings model of Hou et al. (2012).

Table 1: Summary statistics of sample variables

Variable	**Mean**	**1%**	**25%**	**Median**	**75%**	**99%**	**STD**
E_t	154.09	-499.55	-68.68	-27.37	-6.84	6965.92	1446.79
A_t	2305.07	4.16	41.46	114.86	333.22	70283.00	14358.51
D_t	92.63	0.00	0.00	0.00	0.00	3727.94	742.42
DD_t	0.047	0.00	0.00	0.00	0.00	1.00	0.21
NegE_t	0.82	0.00	1.00	1.00	1.00	1.00	0.38
AC_t	-110.36	-3233.20	-25.63	-7.15	-1.34	108.66	835.13
$\text{R\&D}_t$	224.58	0.38	10.93	27.92	70.71	6746.10	1126.85
N				5801			

Table 2 reports the correlation coefficients for the variables in the dataset. As can be seen, there are strong and positive correlations between earnings and total assets, as well as with dividend payments and R&D expenses. Similarly, total assets are also strongly correlated with dividends and R&D outlays. Overall, Table 2 provides a very clear pattern of market dynamics in the pharmaceutical and biotechnology sector, as firms of bigger size, measured by total assets, are able to earn a higher profitability in terms of earnings. Consequently, firms that earn higher profits also have more economic resources to both distribute dividends to shareholders and invest in research and development.

Combining the indications derived from Table 1 and Table 2, the pharmaceutical and biotechnology sector appears to be highly skewed and concentrated. A limited number of large firms are able to report positive earnings and make substantial investments in R&D, while the majority of smaller firms exhibit constrained resources and negative profitability. This pattern highlights the presence of significant heterogeneity within the industry, which may have implications for both the cost of equity estimation and the reliability of earnings forecasts.

Table 2: Correlation matrix of sample variables

	E_t	A_t	D_t	DD_t	Neg E_t	AC_t	R&D$_t$
E_t	1.00						
A_t	0.86	1.00					
D_t	0.85	0.95	1.00				
DD_t	0.49	0.55	0.56	1.00			
Neg E_t	-0.32	-0.30	-0.26	-0.40	1.00		
AC_t	-0.42	-0.73	-0.72	-0.43	0.19	1.00	
R&D$_t$	0.78	0.91	0.91	0.56	-0.30	-0.76	1.00

3.3 Cross-Sectional Earnings Model

To forecast future earnings, this study replicates the model proposed by Hou et al. (2012), which takes the following form:

$$E_{i,t+\tau} = \alpha_0 + \alpha_1 A_{i,t} + \alpha_2 D_{i,t} + \alpha_3 DD_{i,t} + \alpha_4 E_{i,t} + \alpha_5 NegE_{i,t} + \alpha_6 AC_{i,t} + \varepsilon_{i,t+\tau} \quad (2)$$

Where:

- $A_{i,t}$ is total assets,
- $D_{i,t}$ is dividend payments,
- $DD_{i,t}$ is a dummy variable equal to 1 for dividend payers and 0 otherwise,
- $E_{i,t}$ is earnings,
- $NegE_{i,t}$ is a dummy variable equal to 1 for firms with negative earnings and 0 otherwise,
- $AC_{i,t}$ is accruals, computed as the difference between earnings and cash flow from operations.

This cross-sectional model is used to estimate future earnings $E_{i,t+\tau}$ based on current firm-level accounting information. The model is estimated separately for each year using OLS, consistent with the original methodology by Hou et al. (2012).

To evaluate whether R&D expense contributes to the estimation of future earnings in R&D-intensive industries such as the pharmaceutical and biotechnology sector, this study introduces the following expansion of the model:

$$E_{i,t+\tau} = \alpha_0 + \alpha_1 A_{i,t} + \alpha_2 D_{i,t} + \alpha_3 DD_{i,t} + \alpha_4 E_{i,t} + \alpha_5 NegE_{i,t} + \alpha_6 AC_{i,t} + \alpha_7 RD_{i,t} + \varepsilon_{i,t+\tau} \quad (3)$$

Where:

- $RD_{i,t}$ is R&D expense.

This augmented model enables the evaluation of the explanatory power of R&D investments in forecasting future earnings, particularly relevant in sectors characterized by high innovation intensity and intangible-driven value creation.

For each observation in the sample, earnings forecasts up to five years into the future are calculated by estimating cross-sectional earnings regressions over a rolling window of the previous five years ($t-5$ to $t-1$), and using the resulting coefficients to predict earnings for years $t+1$ to $t+5$. This methodology ensures that earnings estimates are out-of-sample and free from look-ahead bias. The use of a five-year, rather than a ten-year, rolling window reduces data requirements, which is more appropriate given the limited scope of the sample compared to the original work by Hou et al. (2012).

3.4 ICC Estimation

The implied cost of capital (ICC) is defined as the internal rate of return that equates the current stock price to the present value of expected future cash flows to equity Hou et al. (2012). As described in the Literature Review section, there are three main accounting-based approaches to derive the ICC: the dividend discount model (DDM), the residual income model (RIM), and the abnormal earnings growth model (AEG). Since each approach has its own strengths and limitations that may influence the final estimates, this study relies on a composite estimation of the ICC, consistent with the methodology of Hou et al. (2012).

Considering the narrower scope of this dataset—both in terms of industry and timeframe—only three individual ICC estimates are employed to construct the composite ICC, rather than the five used by Hou et al. (2012). This trade-off ensures a robust estimation framework while reducing data requirements. It is especially important in this context to avoid overly restrictive assumptions, as ICC estimation is only feasible for firms with positive expected earnings. Since the ICC equates the market value of equity to the present value of future cash flows, negative earnings result in mathematically undefined or economically implausible outcomes in most traditional models. Therefore, ICC estimation is restricted to firms with positive earnings expectations. This constraint is particularly relevant for the current sample, where a substantial number of firms report negative earnings, as discussed earlier.

To mitigate these limitations while maintaining methodological rigor, the study adopts ICC models with fewer required inputs. The three models selected are: the Gordon model (Gordon and Gordon, 1997), representative of the DDM approach; the GLS model (Gebhardt et al., 2001), from the RIM tradition; and the MPEG model (Easton, 2004), representative of the AEG framework. A detailed description of each of these three ICC methodologies follows.

Gordon:

$$M_t = \frac{E_t\left[E_{t+1}\right]}{R} \tag{4}$$

Where:

- M_t is the market equity at year t;
- $E_t[E_{t+1}]$ is expected earnings at time $t+1$;
- R is the ICC.

GLS:

$$M_t = B_t + \sum_{k=1}^{11} \frac{E_t\left[(ROE_{t+k} - R) \times B_{t+k-1}\right]}{(1+R)^k} + \frac{E_t\left[(ROE_{t+12} - R) \times B_{t+11}\right]}{R \times (1+R)^{11}} \tag{5}$$

Where:

- M_t is the market equity at year t;
- B_t is the book equity at year t;
- ROE_{t+k} is the return on equity at year $t+k$. ROEs in $t+1$, $t+2$ and $t+3$ are estimated. After that, the ROE is assumed to mean-revert to the industry median by year $t+11$;
- $(ROE_{t+k} - R) \times B_{t+k-1}$ is the residual income in year $t+k$;
- R is the ICC.

MPEG:

$$M_t = \frac{E_t[E_{t+2}] + R \times E_t[D_{t+1}] - E_t[E_{t+1}]}{R^2} \tag{6}$$

Where:

- M_t is the market equity at year t;
- $E_t[E_{t+1}]$, $E_t[E_{t+2}]$ are the expected earnings at time $t+1$ and $t+2$;
- $E_t[D_{t+1}]$ is the expected dividends at time $t+1$;
- R is the ICC.

The composite ICC for each observation is equal to the mean of the three individual estimates. To maximize the coverage, each observation is only required to have at least one non-missing individual estimate, which is in line with Hou et al. (2012).

4 Main Findings and Results

4.1 Cross-Sectional Earnings Model Coefficients

Table 3 reports the coefficients estimated from the cross-sectional earnings model from Hou et al. (2012) for $E_{(t+1)}$, $E_{(t+2)}$, and $E_{(t+3)}$ estimates. As reported, all the coefficients are statistically significant, which indicates that all the variables included in the model contribute to future earnings estimation. Among the variables, it is important to notice the high and negative coefficient of the negative earnings dummy, which indicates high earnings persistence for the firms in the sample. This consideration, in addition to the skewness of earnings distribution, will impact the ICC calculation that will be performed in the next sections, as firms with negative earnings at time t are projected to have negative earnings in the following periods. The ICC methodology is, on the other hand, economically and mathematically applicable only to firms with positive forecasted future earnings.

Most notably, the adjusted R^2 are respectively 82%, 75%, and 74% for the $t+1$, $t+2$, and $t+3$ time horizons. These values are very similar to those reported by Hou et al. (2012) in their analysis and indicate that the cross-sectional model explains a big portion of the variance of realized earnings.

Overall, Table 3 finds that the model from Hou et al. (2012) performs well in estimating future earnings in this sample, showing that the mechanical approach previously described provides reliable and unbiased forecasts of future earnings in a highly specialized and R&D intensive sector, such as the biotechnology and pharmaceutical one, in an updated timeframe. This analysis, therefore, confirms the robustness of the model of Hou et al. (2012) and suggests that the proposed model's performance is not influenced by sector-specific dynamics, at least in the pharma and biotech industry.

Table 4 reports the coefficients from the cross-sectional earnings model with R&D expansion for E_{t+1}, E_{t+2}, and E_{t+3} estimates. The coefficient of R&D expenses is positive and significant for all time horizons, which means that R&D is a relevant predictor of future earnings. Moreover, the magnitude of the R&D coefficient increases with time horizon, which indicates that R&D outlays have an increased payoff over time, consistent with the long times required for technology and drugs development and testing for the biotechnology and pharmaceutical industry. On the other hand, the inclusion of R&D expenses only marginally increases the adjusted R^2 compared to Hou et al. (2012)'s baseline model.

The results of this analysis are, therefore, twofold: on one side, as expected, the R&D expenditure is a significant determinant of realized earnings in a R&D-intensive sector such as the biotechnology and pharmaceutical one. On the other side, the explored expansion of the model does not yield relevant improvements in earnings

Table 3: Coefficient estimates of baseline cross-sectional earnings model

Variable	E_{t+1}	E_{t+2}	E_{t+3}
A_t	-0.0158^{***}	0.0286^{***}	0.0285^{***}
	(0.0000)	(0.0000)	(0.0000)
D_t	0.2219^{***}	0.4264^{***}	0.5131^{***}
	(0.0000)	(0.0000)	(0.0000)
DD_t	-176.7067^{***}	-216.9669^{***}	-241.4445^{***}
	(0.0008)	(0.0022)	(0.0042)
E_t	0.7856^{***}	0.4035^{***}	0.4061^{***}
	(0.0000)	(0.0000)	(0.0000)
Neg E_t	-69.1595^{***}	-169.0989^{***}	-203.1345^{***}
	(0.0063)	(0.000002)	(0.000001)
AC_t	-0.6536^{***}	-0.2120^{***}	-0.1523^{***}
	(0.0000)	(0.0000)	(0.000001)
Constant	48.0067^{**}	129.1508^{***}	167.6057^{***}
	(0.0397)	(0.0001)	(0.00001)
Observations	4,850	4,007	3,319
R^2	0.8204	0.7524	0.7377
Adj. R^2	0.8202	0.7520	0.7372

forecasting compared to the baseline regression.

Overall, Table 4 confirms the robustness of the model proposed by Hou et al. (2012) and suggests that sector-specific extensions, at least in the biotechnology and pharmaceutical industry, only offer limited enhancements to the performance of the baseline model.

Table 4: Coefficient estimates of expanded cross-sectional earnings model

Variable	E_{t+1}	E_{t+2}	E_{t+3}
A_t	-0.0180^{***}	0.0249^{***}	0.0246^{***}
	(0.0000)	(0.0000)	(0.0000)
D_t	0.1403^{***}	0.2755^{***}	0.2232^{***}
	(0.0012)	(0.00004)	(0.0022)
DD_t	-214.2048^{***}	-298.2554^{***}	-367.8485^{***}
	(0.0005)	(0.00003)	(0.00002)
E_t	0.7683^{***}	0.3757^{***}	0.3473^{***}
	(0.0000)	(0.0000)	(0.0000)
Neg E_t	-59.1015^{**}	-151.1020^{***}	-169.2348^{***}
	(0.0191)	(0.00002)	(0.00004)
AC_t	-0.6004^{***}	-0.1098^{**}	-0.0456
	(0.0000)	(0.0001)	(0.1597)
$R\&D_t$	0.1454^{***}	0.2649^{***}	0.4209^{***}
	(0.0000)	(0.0000)	(0.0000)
Constant	30.2679	97.9175^{***}	112.0259^{***}
	(0.1955)	(0.0022)	(0.0030)
Observations	4,850	4,007	3,319
R^2	0.8220	0.7573	0.7454
Adj. R^2	0.8218	0.7569	0.7449

4.2 Earnings Forecasts Accuracy

Table 5 reports the distribution of the out-of-sample forecasted earnings from $t+1$ to $t+5$, obtained from the 5-year rolling regression. The estimates are derived from the extended model, to make use of the marginal improvements derived from the inclusion of the R&D outlays. As can be seen, the mean of the expected earnings increases monotonically from $t+1$ to $t+5$, which indicates a growing trend in firm profitability. The distribution of the estimates for each time horizon is highly skewed and resembles that of realized earnings in Table 1, as few firms in the sample have positive high forecasted earnings, while the majority have negative forecasted earnings. Furthermore, the estimates indicate that the model incorporates a high level of earnings persistence.

Overall, Table 5 provides preliminary indications on the applicability of the ICC methodology to the sample of this study. As discussed before, the main methodologies to compute the ICC are only applicable to firms with positive expected earnings, but only a small portion of the observations in the sample satisfy these criteria. This sample characteristic, therefore, imposes a severe data restriction to the calculation of the ICC, which will be performed in the next section.

Table 5: Distribution of forecasted earnings

Variable	**Mean**	**1%**	**25%**	**Median**	**75%**	**99%**	**STD**
$E[E_{t+1}]$	132.44	-230.41	-66.87	-50.60	-26.35	6559.07	1236.20
$E[E_{t+2}]$	148.47	-188.22	-70.37	-57.76	-30.19	6479.52	1277.99
$E[E_{t+3}]$	165.03	-236.01	-71.32	-55.73	-19.94	7094.85	1360.02
$E[E_{t+4}]$	177.42	-207.57	-73.94	-59.28	-22.09	7160.40	1461.21
$E[E_{t+5}]$	199.05	-248.11	-85.35	-56.03	-2.72	7767.98	1630.66

Table 6 provides a summary of descriptive statistics of realized and predicted earnings for each time horizon, as well as the forecast accuracy of the model. As can be seen, for both realized earnings and predicted earnings the mean is substantially higher than the median, which is an indication of the skewness of the distribution, as well as the presence of outliers in the right tail of the distribution. Moreover, for both realized and predicted earnings there is a growing trend of the mean over time, which is a further indication of earnings persistence. These findings are in line with those of Tables 1, 3, 4, and 5, as the main conclusions outlined before are confirmed. First, the median firm in the sample has a negative profitability across all time horizons, due to the high upfront and maintenance costs naturally required by the industry dynamics. Second, the market is dominated by some outliers, that earn substantially higher profits than the other players. Third, there is a high level of earnings persistence for both realized and predicted earnings. Fourth, the growing trend of the mean for both realized and predicted earnings is coherent with the increased payoff over time of R&D expenses.

To analyse the accuracy of the proposed model, the mean absolute error (MAE) has been reported in Table 6. The MAE has been chosen as an accuracy measure because it is less sensitive to outliers compared to other commonly used error measures. This consideration is particularly important given the nature of the sample of this study, as the presence of few firms with extreme earnings would have caused other error metrics to be unstable and difficult to interpret. Overall, the MAE is close in magnitude to the mean of realized earnings, which is consistent with the high volatility of both realized and forecasted earnings in all time horizons. Furthermore, the MAE has an increasing trend, which indicates that the accuracy of the model is higher for shorter time horizons and lower for longer time horizons. This finding is aligned with the increase in volatility across time for predicted and forecasted earnings.

Table 6: Comparison of realized and predicted earnings and forecast accuracy

Horizon	**Realized Earnings**			**Predicted Earnings**			**Accuracy**
	Mean	Median	SD	Mean	Median	SD	MAE
t+1	154.09	-27.37	1446.79	132.44	-50.60	1236.20	136.63
t+2	186.88	-26.27	1562.25	148.47	-57.76	1277.99	190.83
t+3	217.03	-24.77	1671.56	165.03	-55.73	1360.02	229.79
t+4	249.52	-23.93	1784.29	177.42	-59.28	1461.21	239.71
t+5	287.94	-22.90	1925.79	199.05	-56.03	1630.66	272.83

4.3 ICC Estimates

To perform the ICC calculation, the Gordon, GLS and MPEG methodologies described in the *ICC Estimation* section have been used. As discussed above, these three approaches are rooted in the ICC literature and are in line with the study from Hou et al. (2012). An important limitation of applying the ICC approach to compute the cost of equity in the pharma and biotech sector is that the calculation only makes economic and mathematical sense for firms with positive forecasted earnings, which is not the case for most of the firms in the dataset. To perform the ICC calculation, therefore, the previously described methodologies are only applied to firms with $E_{t+1} > 0$, which imposes a heavy restriction on the sample. It has to be noted that this limitation is not driven by the application of the model of Hou et al. (2012), neither by the proposed expanded model. Table 3 and Table 4 show, in fact, that both the two cross-sectional models provide reliable estimates of realized future earnings. Rather, the limitation is determined by the ICC approach itself, which is not applicable to loss-making firms. The ICC calculation is performed on a sample of 580 observations.

Table 7 shows the correlation between the three individual ICC estimates and the composite ICC. As can be seen, there is a positive and high correlation between all the individual ICCs, which ranges from 0.47 between Gordon and GLS, to 0.80 between Gordon and MPEG. This finding supports the consistency among the three measures. Interestingly, the composite ICC is almost perfectly correlated with MPEG, which means that the final ICC estimates are mostly driven by their MPEG component, likely because its variability and scale dominate the mean of the individual estimates.

4.4 Relation between ICC Estimates and Realized Returns

To evaluate the quality of the model-based ICC estimates as proxy of expected returns, their relationship with realized returns is examined, following Gebhardt

Table 7: Correlation matrix of ICC estimates

	GLS	Gordon	MPEG	Composite
GLS	1.00			
Gordon	0.47	1.00		
MPEG	0.50	0.80	1.00	
Composite	0.44	0.58	0.99	1.00

et al. (2001), Easton and Monahan (2003), Guay et al. (2011), and Hou et al. (2012). The idea behind this test is that reliable expected returns, proxied by the ICC, should be positively related to future realized returns. To do that, the observations in the sample are sorted by their ICC estimates and then divided in quintiles. Then, the equally weighted buy-and-hold returns for each quintile are calculated for time horizons $t+1$, $t+2$, and $t+3$. Unlike Hou et al. (2012), this study uses quintiles and not deciles to ensure robust results, given the more restricted scope of the analysis. Table 8 reports the results.

As reported in Table 8, there is an overall growing trend in realized returns across the quintiles in all the time horizons, despite some fluctuations. The main takeaway from Table 8 is that the average return spread between extreme quintiles, that is the 5–1 spread, is positive and significant for $t+1$ and $t+2$ time horizons, while for $t+3$, despite being insignificant, it still has the expected positive value. This finding is of paramount importance, as it shows that, in this setting, the model-based ICC has indeed significant predictive power for realized returns. Table 8, therefore, indicates that the cross-sectional approach introduced by Hou et al. (2012), expanded with the inclusion of R&D expenses, manages to provide reliable estimates of expected earnings, which can be used as inputs to derive ICCs that are representative of actual returns.

Table 8: Relationship between composite ICC and realized returns

Quintile	r_{t+1}	p-value	r_{t+2}	p-value	r_{t+3}	p-value
1	0.1102**	0.018	0.0005	0.991	0.0572	0.233
2	0.1683***	0.007	0.0557	0.192	0.0233	0.555
3	0.1166***	0.001	0.1474***	0.000	0.1257***	0.001
4	0.1683***	0.000	0.0467	0.289	0.0935*	0.055
5	0.3687***	0.000	0.1716***	0.007	0.1659**	0.025
5 - 1	0.2585**	0.017	0.1711**	0.025	0.1086	0.213

4.5 Discussion of Results

The goal of this study was to provide practical insights into the applicability of the ICC methodology to calculate the cost of equity of firms in the biotechnology and pharmaceutical sector. To do that, first it has been assessed the efficacy of the cross-sectional model proposed by Hou et al. (2012) to forecast future earnings. Then, it has been explored the possibility to expand the model by including additional variables, namely the R&D expenses, to capture sector-specific dynamics of earnings generation. Finally, the forecasted earnings derived from the model have been used as inputs to calculate the ICC of the observations in the sample, by applying widely used accounting-based models. The main conclusions of the analyses are discussed below.

First, Table 3 confirms the robustness of the cross-sectional model of Hou et al. (2012) in the biotech and pharmaceutical industry in an updated timeframe. The adjusted R^2 reported in Table 3 are, in fact, similar in magnitude to those found by Hou et al. (2012) with a more heterogeneous sample, and more in general, show that the model is able to explain a big portion of the variance of realized earnings. This result represents the main finding of the study, as it confirms that the earnings estimates provided by the cross-sectional model of Hou et al. (2012) are robust in sectors with peculiar market dynamics such as the biotech and pharmaceutical one. It is, therefore, possible to benefit from the high coverage, low bias and low complexity of the model to draw future earnings expectations, while ensuring that the estimates are robust.

Second, Table 4 summarizes the results of including R&D expenses in the cross-sectional model, to incorporate industry-specific dynamics in the estimation of future earnings. On one side, Table 4 shows that R&D expenses are a significant predictor of realized earnings. On the other side, the adjusted R^2 of the expanded model are only marginally higher than those of the model from Hou et al. (2012). Overall, the study finds that, while it is possible to improve the model by including variables that capture sector-specific dynamics, these implementations do not generally provide a sizable increase in the explanatory power of the model, at least in the pharmaceutical and biotech sector. Economic agents such as individual investors or firms' managers may choose the preferred model specification to forecast earnings, facing a trade-off between model performance and complexity.

Third, the ICC estimation casts doubts on the applicability of popular ICC approaches to the biotech and pharmaceutical sector. Table 1 shows that the majority of firms in the sample has negative earnings, while Table 5 and Table 6 find that model-based earnings forecasts are negative for most of the firms over the entirety of the 5-year window. As discussed above, typical ICC methodologies are only applicable for firms with positive expected earnings. This limitation imposes a significant constraint on the possibility to use the ICC as a proxy for the cost of equity in the biotech and pharmaceutical sector. It has to be noted that this limitation is not driven by the previously used cross-sectional model, which provides reliable earnings estimates, but from sample characteristics. Overall, the applicability of the ICC methodology to derive the cost of equity depends on sample-specific or firm-specific characteristics.

Fourth, Table 8 shows that, for the subsample of observations for which it is possible to calculate the ICC, the ICC estimates are indeed a reliable proxy for realized returns. As reported in Table 8, in fact, firms with a higher ICC earn significantly higher returns than those with a lower ICC. This finding reconciles the main results to the ICC literature, by demonstrating that the model-based earnings forecasts provide sound estimates when used as inputs to calculate the ICC. This finding further demonstrates the robustness of the mechanical approaches described before to forecast future profitability. These estimates can, therefore, be confidently used as non-biased expectations of future cash flows within the calculation of the cost of equity.

Overall, these findings confirm the robustness of the future earnings estimates derived from the model by Hou et al. (2012), as well as its applicability to industries with sector-specific dynamics, such as the biotechnology and pharmaceutical one. Despite some general limitations to the applicability of the ICC methodology to the biotechnology and pharmaceutical sector, the ICCs derived from model-based earnings estimates represent a reliable proxy of realized returns.

5 Conclusion

The cost of equity expresses the risk-return profile of an equity investment and represents a key determinant in capital budgeting, firm valuation, and investment decisions. Despite the lack of a universally accepted methodology for calculating the cost of equity, the implied cost of capital (ICC) approach has increasingly gained consensus in the corporate finance literature due to its solid theoretical foundations and relative simplicity.

This study contributes to the ICC literature by examining whether mechanically forecasted earnings can serve as inputs to generate reliable estimates of a firm's cost of equity. Specifically, it assesses the robustness of the cross-sectional earnings model proposed by Hou et al. (2012), evaluating its ability to produce accurate forecasts of future earnings in an updated timeframe and within an industry characterized by distinctive earnings-generation dynamics. To this end, the analysis is restricted to the pharmaceutical and biotechnology sector. In addition, the study explores an extension of Hou et al. (2012) baseline model through the inclusion of R&D expenses, in order to evaluate whether incorporating sector-specific variables can enhance the model's forecasting performance.

The results of this research contribute several important insights to the ICC literature. First, the study confirms the robustness of the model proposed by Hou et al. (2012), which is found to provide reliable estimates of expected earnings in an updated timeframe and within an industry characterized by distinctive earnings generation dynamics, such as the pharmaceutical and biotechnology sector. Second, the proposed model expansion—through the inclusion of R&D expenses—only marginally improves the explanatory power of the model. This suggests that the baseline model by Hou et al. (2012) already captures most of the sector-specific dynamics related to earnings generation. Third, given that the majority of firms in the sample have negative forecasted earnings, the ICC approach cannot be applied to estimate their cost of equity. This finding highlights a crucial limitation of the ICC methodology: its applicability is inherently dependent on firm-specific and sector-specific characteristics. The pharmaceutical and biotechnology sector, dominated by loss-making firms, exemplifies this shortcoming. Fourth, for the subset of firms for which it is feasible to estimate the ICC, the cost of equity is found to be positively related to realized returns. This further supports the reliability of the earnings forecasts produced by the cross-sectional model.

These findings are also relevant for corporate managers, investors, and other economic agents, who can confidently rely on the cross-sectional approach proposed by Hou et al. (2012) to estimate the cost of equity of target firms and inform their investment decisions. Furthermore, economic agents may select the model specification that best suits their objectives, depending on data availability and firm-specific characteristics. Overall, adopting a cross-sectional framework to forecast future earnings—and subsequently using model-based ICC estimates to derive the cost of equity—offers significant advantages in terms of simplicity, theoretical rigor, and unbiased results.

It must be acknowledged that this study presents several limitations concerning its scope, sample composition, and the generalizability of its findings. First, the restricted timeframe of the analysis and the exclusive focus on the pharmaceutical and biotechnology sector may limit the extent to which the results can be extrapolated to other industries or economic contexts. Second, the study explores only one potential model expansion—namely, the inclusion of R&D expenses—so it remains possible that the integration of alternative predictors could have led to more substantial improvements in the performance of the baseline model.

Further research could build on the foundation laid by this study by analysing the performance of the cross-sectional approach proposed by Hou et al. (2012) when applied to industries beyond the pharmaceutical and biotechnology sector. In addition, future work may explore alternative model expansions by incorporating other sector-specific variables that could enhance predictive accuracy. Given the strong potential of the cross-sectional estimation framework developed by Hou et al. (2012), subsequent research should aim to test the robustness of the model across broader contexts and to refine its predictive capabilities. Such efforts would help to provide corporate managers, investors, and financial analysts with an increasingly reliable and precise methodology for estimating the cost of equity.

References

B. Adebambo, R. M. Bowen, S. Malhotra, and P. Zhu. Ceo extraversion and the cost of equity capital. *The Journal of Financial Research*, 2024.

A. Ali, M. Ciftci, and W. M. Cready. Market underestimation of the implications of rd increases for future earnings: The us evidence. *Journal of Business Finance & Accounting*, 39(3-4):289–314, 2012.

M. E. Barth, K. Liu, and C. G. McClure. Evolution in value relevance of accounting information. *The Accounting Review*, 98(1):1–28, 2023.

C. A. Botosan. Disclosure level and the cost of equity capital. *The Accounting Review*, 72:323–349, 1997.

L. Bretscher, A. Hsu, P. Simasek, and A. Tamoni. Covid-19 and the cross-section of equity returns: Impact and transmission. *The Review of Asset Pricing Studies*, 10(4):705–741, 2020.

H. Bruneo, E. Giacomini, G. Iannotta, A. Murthy, and J. Patris. Risk and return in the biotech industry. *International Journal of Productivity and Performance Management*, 2024.

M. M. Carhart. On persistence in mutual fund performance. *Journal of Finance*, 52(1):57–82, 1997.

P. M. Clarkson, X. Fang, Y. Li, and G. Richardson. The relevance of environmental disclosures: Are such disclosures incrementally informative? *Journal of Accounting and Public Policy*, 32(5):410–431, 2013.

J. Claus and J. Thomas. Equity premia as low as three percent? evidence from analysts' earnings forecasts for domestic and international stock markets. *Journal of Finance*, 56:1629–1666, 2001.

A. Damodaran. *Investment Valuation: Tools and Techniques for Determining the Value of Any Asset*. John Wiley & Sons, 2002.

S. Das, C. B. Levine, and K. Sivaramakrishnan. Earnings predictability and bias in analysts' earnings forecasts. *The Accounting Review*, 73(2):277–294, 1998.

D. M. DePamphilis. *Mergers, Acquisitions, and Other Restructuring Activities*. Academic Press, 6th edition, 2012.

J. A. DiMasi, H. G. Grabowski, and R. W. Hansen. Innovation in the pharmaceutical industry: New estimates of r&d costs. *Journal of Health Economics*, 47:20–33, 2016.

D. C. Donelson and R. J. Resutek. The effect of r&d on future returns and earnings forecasts. *Review of Accounting Studies*, 17(4):848–876, 2012.

D. C. Donelson, R. Jennings, and J. McInnis. Changes over time in the revenue-expense relation: Accounting or economics? *The Accounting Review*, 86(3):945–974, 2011.

F. Duarte and C. Rosa. The equity risk premium: A review of models. *Economic Policy Review*, 21(2):39–57, 2015.

P. D. Easton. Pe ratios, peg ratios, and estimating the implied expected rate of return on equity capital. *The Accounting Review*, 79:73–95, 2004.

P. D. Easton and S. J. Monahan. An evaluation of the reliability of accounting-based measures of expected returns: A measurement error perspective. *SSRN Electronic Journal*, 2003.

P. D. Easton and G. A. Sommers. Effect of analysts' optimism on estimates of the expected rate of return implied by earnings forecasts. *Journal of Accounting Research*, 45:983–1015, 2007.

A. C. Eberhart, W. F. Maxwell, and A. R. Siddique. An examination of long-term abnormal stock returns and operating performance following r&d increases. *The Journal of Finance*, 59(2):623–650, 2004.

F. Echterling, B. Eierle, and S. Ketterer. A review of the literature on methods of computing the implied cost of capital. *International Review of Financial Analysis*, 42:235–252, 2015.

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

E. F. Fama and K. R. French. Size and book-to-market factors in earnings and returns. *Journal of Finance*, 50:131–155, 1995.

E. F. Fama and K. R. French. Industry costs of equity. *Journal of Financial Economics*, 43:153–193, 1997.

E. F. Fama and K. R. French. The capital asset pricing model: Theory and evidence. *Journal of Economic Perspectives*, 18(3):25–46, 2004.

G. Feltham and J. Ohlson. Valuation and clean surplus accounting for operating and financial activities. *Contemporary Accounting Research*, 11(2):689–731, 1995.

J. Francis and D. Philbrick. Analysts' decisions as products of a multi-task environment. *Journal of Accounting Research*, 31:216–230, 1993.

W. R. Gebhardt, C. M. C. Lee, and B. Swaminathan. Toward an implied cost of capital. *Journal of Accounting Research*, 39(1):135–176, 2001.

J. Gerakos and R. Gramacy. Regression-based earnings forecasts. *Chicago Booth Research Paper No. 12-26*, 2013.

J. R. Gordon and M. J. Gordon. The finite horizon expected return model. *Financial Analysts Journal*, 53(3):52–61, 1997.

H. Grabowski, J. Vernon, and J. A. DiMasi. Returns on research and development for 1990s new drug introductions. *Pharmacoeconomics*, 20(Suppl. 3):11–29, 2002.

L. Gu. Product market competition, r&d investment, and stock returns. *Journal of Financial Economics*, 119(3):441–455, 2016.

W. Guay, S. P. Kothari, and S. Shu. Properties of implied cost of capital using analysts' forecasts. *Australian Journal of Management*, 36(2):125–149, 2011.

J. R. M. Hand. The market valuation of biotechnology firms and biotechnology r&d. *University of North Carolina Kenan-Flagler Business School*, 2001.

S. E. Harrington. Cost of capital for pharmaceutical, biotechnology, and medical device firms. In P. M. Danzon and S. Nicholson, editors, *The Oxford Handbook of the Economics of the Biopharmaceutical Industry*, pages 75–99. Oxford University Press, 2012.

R. D. F. Harris and P. Wang. Model-based earnings forecasts vs. financial analysts' earnings forecasts. *The British Accounting Review*, 51(4):424–437, 2019.

D. Hess, M. Meuter, and A. Kaul. The performance of mechanical earnings forecasts. *S&P Global Market Intelligence Research Paper Series*, 2019.

K. Hou, M. A. Van Dijk, and Y. Zhang. The implied cost of capital: A new approach. *Journal of Accounting and Economics*, 53(3):504–526, 2012.

J. Hughes, J. Liu, and J. Liu. On the relation between expected returns and implied cost of capital. *Review of Accounting Studies*, 14(2-3):246–279, 2009.

A. Iqbal, S. Rajgopal, A. Srivastava, and R. Zhao. A better estimate of internally generated intangible capital. *Management Science*, 2024. Forthcoming.

C. S. Jones and S. Tuzel. Inventory investment and the cost of capital. *Journal of Financial Economics*, 107(3):557–579, 2013.

M. Juneja, L. Mai, and N. Albu. The economic impact of the global pharmaceutical industry. WifOR Institute & IFPMA, 2024.

D. Kanelis and P. Vorst. Risky business: The risk-relevance of earnings over time. Working Paper, Maastricht University, 2024.

B. Ke and Y. Yu. The effect of issuing biased earnings forecasts on analysts' access to management and survival. *Journal of Accounting Research*, 44(5):965–999, 2006.

Y. Ke. The impact of covid-19 on firms' cost of equity capital: Early evidence from u.s. public firms. *Finance Research Letters*, 46:102242, 2022.

M. Kouki and A. Elkhaldi. Does financial crisis affect the cost of equity estimation? evidence from the tunisian stock exchange. *International Journal of Accounting and Financial Reporting*, 7(2):1–15, 2017.

S. A. Larocque and M. R. Lyle. Implied cost of equity capital estimates as predictors of accounting returns and stock returns. *Journal of Financial Reporting*, 2(1), 2017.

K. Li and P. Mohanram. Evaluating cross-sectional forecasting models for implied cost of capital. *Review of Accounting Studies*, 2014.

Y. Li, D. T. Ng, and B. Swaminathan. Predicting market returns using aggregate implied cost of capital. *Journal of Financial Economics*, 110(3):419–436, 2013.

H.-W. Lin and M. McNichols. Underwriting relationships, analysts' earnings forecasts and investment recommendations. *Journal of Accounting and Economics*, 25:101–127, 1998.

J. Lintner. The valuation of risk assets and the selection of risky investments in stock portfolios and capital budgets. *The Review of Economics and Statistics*, 47 (1):13–37, 1965.

J. Liu, S. Wu, and V. T. Pham. Design of accounting earnings forecasting model based on artificial intelligence. *Scientific Programming*, 2022:Article 8793596, 2022.

J. M. McInnis. Earnings smoothness, average returns, and implied cost of equity capital. *The Accounting Review*, 85(1):1–46, 2010.

M. McNichols and P. O'Brien. Self-selection and analyst coverage. *Journal of Accounting Research*, 35:167–199, 1997.

P. Mohanram and D. Gode. Removing predictable analyst forecast errors to improve implied cost of equity estimates. *Review of Accounting Studies*, 18:443–478, 2013.

J. Ohlson and B. Juettner-Nauroth. Expected eps and eps growth as determinants of value. *Review of Accounting Studies*, 10:349–365, 2005.

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

Ł. Pástor, R. F. Stambaugh, and L. A. Taylor. Dissecting green returns. *Journal of Financial Economics*, 146(2):403–424, 2022.

S. Penman, J. Zhu, and H. Wang. Estimating an equity yield to maturity: Deficiencies of the implied cost of capital (icc) and an alternative, 2022.

M. Pęksyk, K. Śledzik, and B. Bławat. Cost of equity under the circumstances of uncertainties – the example of the us market. *The Annual Journal of Audit and Accounting*, 2023.

M. S. Raghavendra, J. R. Raj, and A. Seetharaman. A study of decrease in r&d spending in the pharmaceutical industry during post-recession. *International Journal of Academic Research Part B*, 4(5):29–47, 2012.

P. Sadorsky and I. Henriques. Risk measures and the cost of equity in the new economy biotechnology industry. *Global Business & Economics Review*, 5(1):37–50, 2003.

M. Schreder and P. Bilinski. A study of the reliability of cross-sectional earnings forecasting models for estimating ipo firms' implied cost of capital. *Accounting Research Journal*, 35(2):219–237, 2022.

D. Schröder. The implied equity risk premium: An evaluation of empirical methods. Technical Report 13/2005, Bonn Graduate School of Economics (BGSE), University of Bonn, 2005.

W. F. Sharpe. Capital asset prices: A theory of market equilibrium under conditions of risk. *Journal of Finance*, 19(3):425–442, 1964.

F. Simon. Model-based earnings forecast accuracy and implied cost of capital portfolio returns, 2024.

A. Srivastava. Why have measures of earnings quality changed over time? *Journal of Accounting and Economics*, 57:196–217, 2014.

A. Trunk and I. Stubelj. The financial-economic crisis and value of equity capital: A case study of slovenian public limited companies 2006–2011. *Expert Systems with Applications*, 40(18):7562–7570, 2013.

B. Xu. R&d strategy and stock price volatility in the biotechnology industry. *Review of Accounting and Finance*, 5(1):59–74, 2006.

B. Xu, M. L. Magnan, and P. E. André. The stock market valuation of r&d information in biotech firms. *Contemporary Accounting Research*, 24(4):1291–1318, 2007.