

Validation of Cost of Equity Metrics

A Study on the Effects of Distribution-Matching and Input Data Adjustments in a Swedish Market Setting

Melinda Nordvall*
22446

Fanny Nyberg**
22436

In this thesis, we look to improve cost of equity validation by further testing of a recently introduced method for statistical adjustments of cost of equity test samples. A resampling method, bin-based distribution-matching, is tested on Swedish data to correct for non-randomness in cost of equity samples and improve testing of the sought-for correlation between cost of equity and realized returns. Using information about returns in the full sample of firms, we reweight observations and construct cost of equity samples that appear randomly drawn from the full sample. We compute four cost of equity metrics, r_{MPEG} , r_{OJN} , r_{GLS} and r_{CT} , and examine their correlation with realized returns before and after resampling. Results show that distribution-matching is feasible in the Swedish market setting and contributes to improve the correlation between cost of equity and realized returns. This improvement is however more modest than previous research would predict. Introducing model-based analyst forecasts to further enhance results shows little or no improvement to the test results, both in combination with distribution-matching and on a stand-alone basis.

Keywords: cost of equity, distribution-matching, resampling, earnings forecasts, expected returns

Tutor: Florian Eugster

Date: December 7, 2015

*22446@student.hhs.se

**22436@student.hhs.se

We would like to express our gratitude to our tutor Florian Eugster for valuable advice

Contents

1. Introduction	1
2. Previous research	4
2.1 Computing cost of equity metrics.....	4
2.2 Validating cost of equity metrics	7
2.3 Adjusting cost of equity input data	8
2.4 Adjusting for non-randomness in cost of equity samples	10
3. Motivation	11
3.1 Area of focus	11
3.2 Research rationale and hypotheses	12
3.3 Practical implications.....	14
4. Methodology	14
4.1 Data and adjustments.....	15
4.2 Computation of implied cost of equity metrics	16
4.3 Resampling with bin-based distribution-matching	19
4.3.1 Defining a reference sample	20
4.3.2 Sorting observations into bins	20
4.3.3 Distribution-matching.....	21
4.4 Additional tests	22
4.4.1 Generating model-based earnings forecasts with the cross-sectional model	22
4.4.2 Introducing an additional implied cost of equity metric.....	24
4.5 Summary of metrics	25
4.6 Association tests with realized returns	25
5. Results	26
5.1 Metrics set 1: Unadjusted metrics (benchmark case)	26
5.1.1 Descriptive statistics for the cost of equity metrics	26
5.1.2 Association tests	27
5.2 Metrics set 2: Distribution-matched metrics.....	29
5.2.1 Results from bin-based distribution-matching.....	29
5.2.2 Association tests	31
5.3 Metrics set 3: Model-based unadjusted metrics	33
5.3.1 Descriptive statistics of the cross-sectional forecast model.....	33
5.3.2 Descriptive statistics of the cost of equity metrics	34
5.3.3 Association tests	36
5.4 Metrics set 4: Model-based distribution-matched metrics	37
5.4.1 Results from bin-based distribution-matching.....	37
5.4.2 Association tests	39
5.5 Summary of results from Metrics sets 1 - 4.....	40
5.6 Results from additional test with $r_{\text{SETTERBERG}}$	41
5.7 Further evaluation.....	42
5.7.1 Tests with absolute values of realized returns	42
5.7.2 Tests excluding years of economic downturn.....	45
6. Limitations of scope	46
7. Conclusion.....	47
8. References	50
APPENDIX	53

1. Introduction

Cost of equity is an important element in company valuation, capital budgeting and portfolio selection. There are numerous methods for computing cost of equity. Similarly, there are many methods used to validate whether the metrics fulfill their purpose of being feasible proxies for expected returns (e.g. Easton, 2004; Ohlson and Juettner-Nauroth, 2005; Gebhardt et al., 2001). Considering cost of equity as a measure of shareholders' expectations, theory predicts that investors require higher returns for holding more risky assets. This implies that high-risk firms are expected to be associated with higher cost of equity than low-risk firms. With this in mind, cost of equity validation begins with examining the association between cost of equity metrics and other proxies for expected returns and/or certain risk measures. For decades, researchers have however been puzzled by the difficulties of obtaining results that align empirical testing with theory (e.g. Easton and Monahan, 2005; Guay et al., 2011). Even though the risk-return concept is widely accepted from a theoretical point of view, researchers continue to obtain equivocal results when testing for the correlation between cost of equity and expected returns, proxied by realized returns. Association tests often indicate weak or even negative correlation, suggesting that pieces of the cost of equity puzzle are flawed or even missing (Easton and Monahan, 2005; Guay et al., 2011; Mohanram and Gode, 2013; Hou et al., 2012; Ecker et al., 2015).

In this thesis, we look to improve cost of equity validation by further testing of a recently introduced method for statistical adjustments of cost of equity test samples. More specifically, we investigate the possibility of correcting for non-randomness in cost of equity samples through a resampling method called bin-based distribution-matching (presented by Ecker et al., 2015). The non-random features of these samples arise as a consequence of data restrictions posed during cost of equity computation. These restrictions result in firms with observable values of cost of equity being a biased selection; often including disproportionately large and profitable firms. Since results obtained from a biased sample cannot be used for conclusions for the whole population without further controls, the generalizability of previous research within cost of equity can be questioned. In other words, if non-randomness in cost of equity samples appears on a systematic level, conclusions from non-corrective studies should be expected to deviate from theoretically predicted results (Ecker et al. 2015). Should this be the case, future research may need to make corrections, such as distribution-matching, in

order to obtain more representative results and align empirical testing with theory.

Previous researchers have taken alternative approaches to tackle the ambiguity in the cost of equity research field. The most common streams propose improvements to implied cost of equity metrics (e.g., Gordon and Gordon, 1997; Claus and Thomas, 2001; Gebhardt et al., 2001; Easton, 2004; Ohlson and Juettner-Nauroth, 2005), alterations of validation methods (Campbell, 1991; Gebhardt et al., 2001; Easton and Monahan, 2005; Botosan et al., 2010; Botosan et al., 2010; Setterberg, 2011; Ecker, 2015) and attempts to improve frequency and accuracy of input data from analyst forecasts (Hou et al., 2012; Jones and Tuzel, 2012; Patatoukas, 2011; Li and Mohanram, 2014). Within all these areas, methods that to some extent improve validation results have been introduced, though none have persistently shown superior to other methods. A majority of previous research is concentrated to the U.S. market (e.g. Hou et al., 2012; Ecker et al., 2015), potentially because of the large sample sizes and availability of data. Through shifting focus from earlier methods to the role of non-randomness in cost of equity samples, we however find it interesting to run tests in a smaller, less documented market setting. Distribution-matching as such challenges the common assumption that larger sample sizes by definition increases statistical reliability, by suggesting that results could be more generalizable when reducing the sample size (Ecker et al., 2015). This could indeed be beneficial for small markets, and we thereby investigate whether additional practical issues arise with a smaller starting sample, or whether the resampling model does succeed to improve cost of equity validation also for small markets.

In this study, we test four cost of equity metrics; r_{MPEG} , r_{OJN} , r_{GLS} and r_{CT} , which origin from Easton (2004), Ohlson and Juettner-Nauroth (2005), Gebhardt et al. (2001) and Claus and Thomas (2001). The tests are based on a dataset of Swedish listed firms in the years 1995-2014. With distribution-matching, we use information on the marginal distribution of a full realized returns sample to redistribute observations in the cost of equity subsamples. The intention is to construct cost of equity samples that appear to be randomly drawn from a full population of firms. By doing so, we aim to further validate bin-based distribution-matching, as presented by Ecker et al. (2015), as a means to improve cost of equity validation. We hypothesize that the procedure will be less efficient in the Swedish market setting, due to the small sample size. In addition, we investigate whether combining bin-based distribution-matching with other models may further improve the correlation between cost of equity and realized returns, and thus make up for the potentially modest results arising from the small

market setting. The additional models tested are a cross-sectional model used to generate earnings forecasts (as opposed to using analyst followings) and a fifth cost of equity metric, $r_{\text{SETTERBERG}}$, particularly suitable for the Swedish market. In different sets we compute i) unadjusted metrics and ii) metrics based on modelled earnings forecasts. On these sets, we run the distribution-matching procedure. In addition, we compute the fifth metric, $r_{\text{SETTERBERG}}$ to be compared with our original four implied cost of equity metrics.

Results based on our benchmark case of unadjusted metrics show that none of the four metrics can present a positive and significant correlation with realized returns, which is the relation predicted by theory. Instead, findings are in line with previous empirical research (Ecker et al., 2015; Easton and Monahan, 2005; Guay et al., 2011) and suggest weak or negative correlations. Studying the distribution of realized returns in the full population sample (our reference sample) versus the returns in each respective cost of equity subsample, we notice differences. Through distribution-matching, we align the distributions and construct cost of equity subsamples that appear to be randomly drawn from the full population of firms (as opposed to our original cost of equity samples). Having done so, we find positive correlations between realized returns and the cost of equity metrics r_{GLS} and r_{CT} , as predicted by theory. The positive effects from distribution-matching are thereby evident, though not as substantial as in the original U.S. setting.

Our attempt to combine distribution-matching with other suggested methods from the cost of equity research field turn out less effective. Results show that performing distribution-matching on cost of equity samples based on model-based earnings forecasts instead of actual analyst followings does not improve the correlation results. Lack of improvement is also evident when we test the Setterberg (2011) metric, $r_{\text{SETTERBERG}}$. Based on these results, we cannot accept nor reject our hypothesis that the effects of distribution-matching can be enhanced by combining it with other suggested models. We can however note that the matter is definitely complex in a small market due to the smaller sample sizes.

The thesis is divided as follows; first, a section devoted to previous literature where we describe research on cost of equity metrics, cost of equity validation and attempts to improve the test setting and input data. Second, we will provide a motivation for the chosen methodology and subject in relation to previous research, and also state our hypotheses. Third, our methodological procedure is described followed by a presentation of the results of

the tests performed. Fourth, after stating the effects of the limitations of scope, we summarize our conclusions and the main findings of the thesis, accompanied by input for future research.

2. Previous research

In this section, an overarching presentation of previous research within the cost of equity field is given. Since the field has been of high relevance for a long period of time, it is under no means possible to cover all aspects and previous contributions. We therefore exclude parts of the field that do not directly connect to the chosen direction in this thesis. Firstly, we provide a background to cost of equity computation and introduce the metrics that are most frequently used. Secondly, we present research regarding cost of equity validation. Thirdly we outline suggested adjustments to input data and handling of analyst forecasts. Lastly, we cover the newest addition; adjustments for non-randomness in samples.

2.1 Computing cost of equity metrics

Cost of equity refers to the return that stockholders require for a company. More specifically, it is the compensation that the market demands in exchange for owning an asset and bearing the risk of ownership. Financial theory predicts that an investor owning a risky asset expects to be compensated with a higher return. As stated by Hou and van Dijk (2011), previous research has almost exclusively relied on ex post realized returns to measure ex ante expected return. A stream of researchers has however argued that realized return is not at all an optimal proxy for expected return (e.g., Blume and Friend, 1973; Elton, 1999; and Vuolteenaho, 2002). Elton (1999) mean that realized returns can be decomposed into an expected and an unexpected part, where the latter is assumed to have a mean value of zero when using realized return as a direct proxy. The author argues that this assumption does not hold and that unexpected returns correlate across firms and time, causing average realized returns to be a biased measure when chosen to represent expected returns. Vuolteenaho (2002) further claims that unexpected return is not only existent, but actually driving firm-specific realized returns, giving more reason to expect noisiness in a realized return proxy.

To find better proxies for expected returns, recent accounting and finance studies (e.g. Gordon and Gordon, 1997; Claus and Thomas, 2001; Gebhardt et al., 2001; Easton, 2004; Ohlson and Juettner-Nauroth, 2005) propose estimation through the *implied cost of capital approach*. This refers to backing out the cost of equity from some kind of valuation model, i.e. solve for the

discount rate that sets the modelling value of a share equal to the share price.¹ Currently, there are three types of models mainly used to arrive at the implied cost of equity. These are the dividend discount model (Botosan, 1997 and Brav et al., 2005), residual income valuation models based on Ohlson (1995) (O’Hanlon and Steele, 2000; Gebhardt et al., 2001; Claus and Thomas, 2001; Baginski and Wahlen, 2003), and the Ohlson and Juettner-Nauroth model (Gode and Mohanram, 2003 and Easton, 2004). The following table outlines the four most frequently used cost of equity computation methods:

Easton (2004)
$r_{MPEG} = A + \sqrt{A^2 + \frac{eps_2 - eps_1}{P_0}} \quad \text{where } A = dps_1/(2P_0)$
Ohlson & Juettner-Nauroth (2005)
$r_{OJN} = A + \sqrt{A^2 + \frac{eps_1}{P_0} * \frac{eps_2 - eps_1}{eps_1} - (\gamma - 1)} \quad \text{where } A = \left((\gamma - 1) + \frac{dps_1}{P_0} \right) \frac{1}{2}$
Gebhardt et al. (2001)
$P_0 = bps_0 + \sum_{t=1}^{12} (1 + r_{GLS})^{-t} \left((ROE_t - r_{GLS}) bps_{t-1} \right) + (1 + r_{GLS})^{-12} (r_{GLS}^{-1}) \left((ROE_{ss} - r_{GLS}) bps_{12} \right)$
Claus & Thomas (2001)
$P_0 = bps_0 + \sum_{t=1}^5 (1 + r_{CT})^{-t} \left((ROE_t - r_{CT}) bps_{t-1} \right) + (R - g)(1 + r_{CT})^{-5} (r_{CT}^{-1}) \left((ROE_{t+5} - r_{CT})(1 + g) bps_5 \right)$

¹ This opposed to for example CAPM or the Fama French models, which would be deemed more conventional though criticized for lack of accuracy (e.g. Fama and French, 1997).

The metrics are all related, but differ in the length of forecast period as well as in handling terminal values and growth. The Ohlson and Juettner-Nauroth (2005) model, r_{OJN} , relates a firm's price per share to next year's expected earnings per share (eps), short term growth in eps, long-term growth in eps and cost of equity capital. The model is based on the assumption that the present value of dividends per share (dps) determines price, but does not use dividend values beyond the first period or specify how these may evolve over time. The focus is on capitalized forward earnings as well as an adjustment factor for future growth in abnormal earnings, calculated as cum-dividend earnings minus earnings for the previous year.

Easton (2004) provides a similar cost of equity metric, the modified PEG ratio (r_{MPEG}), which originates from Ohlson and Juettner-Nauroth's earlier work. Also relating back to abnormal earnings, the model is based on earnings and earnings growth. Easton (2004) compares the metric with the expected return obtained by the original PEG ratio, which is the price-to-earnings ratio divided by forecasted growth in earnings. Evidence is provided for r_{MPEG} being closely related to the Abnormal Earnings Growth model, with differences only occurring in cases of abnormal growth in the first and second year of the forecast period. The difference between r_{MPEG} and r_{OJN} is that Easton does not incorporate long-term growth in eps, only short-term.

Gebhardt et al. (2001) make use of a residual income valuation (RIV) model to derive the implied cost of capital. The forecast period is twelve years before a terminal value is estimated. For the first three years of the forecast period, analysts' forecasts of earnings are used. For the remaining years, future earnings are estimated based on the assumption that return on equity (ROE) from year three fades towards the steady state ROE^2 , which is industry specific and based on historical mean values.

Claus and Thomas (2001) take a similar, but somewhat simpler approach to the RIV model. Analogous to Gebhardt et al. (2001), analyst forecasts of earnings are used for the first three years of the forecast period. However, instead of a twelve year time frame, the model is built upon a five year forecast period before steady state is assumed and the terminal value calculated. Four and five year forward-looking earnings are estimated using a growth rate and analyst forecasts available for the third year. The long-term growth rate is assumed to

² Earnings forecasts are for the majority of firms only available for three years ahead.

converge towards the inflation rate, in such a manner that the steady state growth rate is equal to the inflation and assumed not to vary between industries.

Cost of equity approaches emerging from abnormal earnings based models (such as r_{OJN} and r_{MPEG}) have been praised for the small computational burden required (e.g. Easton, 2004; Ohlson and Juettner-Nauroth, 2005). Ohlson (2005) also underlines another advantage by pointing out that the models are not reliant upon the clean surplus relation, a theoretical assumption arguably not always reasonable to hold in practice. However, critique has also arisen with respect to lacking preciseness and accuracy of the metrics. The RIV based metrics have on the contrary been promoted as more precise, but impose a heavy computational burden and are not as frequently used in practice. Pros and cons aside, the four metrics (r_{MPEG} , r_{OJN} , r_{GLS} and r_{CT}) are commonly accepted and, compared to peers, the metrics most frequently used in cost of equity research (e.g. Ecker et al., 2015; Easton and Monahan, 2005; Hou et al., 2012). Despite the fact that these metrics remain “best practice” for computation, all cost of equity metrics are consistently difficult to validate through empirical testing; these metrics being no exception.

2.2 Validating cost of equity metrics

One of the most widely used cost of equity validation methods is to run association tests between the implied cost of equity metrics and realized returns, studying correlation as well as the slope coefficients from a simple regression (Gebhardt et al., 2001; Gode and Monahanram, 2003; Guay et al., 2005 etc.). If a positive correlation can be found, the metrics are considered feasible proxies for expected returns. This is in turn dependent on realized returns also being a good proxy for expected returns. As previously stated, the latter is not necessarily the case. To create a more suitable proxy, some researchers suggest adjustments to the simple return regression, such as controlling for cash flow news and expected return news (Campbell, 1991; Easton and Monahan, 2005; Botosan et al., 2011).

As an alternative to the realized return approach, Fama and French (1997) tested the correlation between cost of equity and conventional risk measures such as beta, idiosyncratic risk, size and the book-to-market ratio. Variations of this approach are still common practice (Gebhardt et al., 2001; Botosan and Plumlee, 2005; Botosan et al., 2011; Setterberg, 2011; Ecker, 2015), often as complementary testing to the association tests with realized returns.

Researchers look for a negative association with market cap and positive correlation with beta and book-to-market ratios.

The different validation procedures have had varying degrees of success. Easton and Monahan (2005) make a thorough investigation of four cost of equity metrics to determine their feasibility as proxies for expected returns. Association tests show that the majority of the metrics lack the expected correlation with realized returns. A naïve measure of expected return (the inverse of price to forward earnings) contains no more, and often less, measurement error than the sophisticated implied cost of equity metrics. Botosan et al. (2011) are in some areas more successful; after redefining control variables from Easton and Monahan (2005) they find a positive correlation between realized returns and implied cost of equity metrics, and also find significant correlations with risk measures. On the contrary, Guay et al. (2011) find results aligned with Easton and Monahan (2005) and show that implied cost of equity metrics as well as estimates from the Fama and French three factor model are uncorrelated with future annual and monthly returns.

Neither the transition to implied cost of equity metrics nor the alterations of validation methods seem to have succeeded in persistently proving positive correlations between expected return proxies. As an effect of this, other research streams starts off in accepting a certain type of metric and a specific validation method, but focus adjustments on other components in the cost of equity puzzle.

2.3 Adjusting cost of equity input data

One of these components is input data, of which the level of quality has been a topic for discussion. Most of the data used for cost of equity computation are based on reported figures from annual reports and further assumptions about how these values will evolve over time. The exception is data collected from analyst forecasts, which are thought to represent the market's expectation of future earnings. Guay et al. (2011), Easton and Monahan (2005) and Hou et al. (2012) all question the accuracy of these forecasts, meaning that flawed and biased figures may be at least partly responsible for the ambiguous results within the cost of equity research field. Easton and Monahan (2005) find that accurate forecasts generate stronger correlations with realized returns, suggesting that the forecast quality play a central role in cost of equity reliability. Analysis indicates that the errors in the tested cost of equity metrics

are not random, but arise as a result of forecasts being sluggish with respect to information in past stock returns. More specifically, for analyst forecasts to be a feasible proxy for market expectations, analysts are required to quickly incorporate stock price changes into their forecasts. If this is not done sufficiently, the consequent sluggishness introduces errors in the cost of capital estimates; errors that correlate with past security price performance. Guay et al. (2011) focus on long-term forecasts and find that these are more affected by sluggishness, as given large weight in the computation of implied cost of equity metrics.

Hou et al. (2012) also criticize the reliability of analyst forecasts, and introduce a model for generating replacement values. Based on historical accounting data, they compile a cross-sectional model that generates proxies for expected earnings. The model allows them to use estimated earnings forecasts as input data in cost of equity calculations instead of using values from analyst forecasts. The method is anchored in previous research (e.g., Fama and French, 2000, 2006; Hou and Robinson, 2006; Hou and van Dijk, 2012) showing that cross-sectional models are able to explain a large fraction of the variation in expected profitability across firms. Results show that the model-based earnings forecasts are less accurate than analyst forecasts, but leads to better coverage and a higher earnings response coefficient.

Another approach that suggests alternative inputs to analyst forecasts is the fundamental RIV metric ($r_{\text{SETTERBERG}}$), developed by Setterberg (2011). It builds upon the metric by Gebhardt et al. (2001), r_{GLS} , and uses an extended RIV model as starting point when backing out the cost of equity. Through letting firm specific mean values of historical ROE determine future earnings forecasts, rather than analysts' earnings forecasts, the data constraints are loosened and more observations can be included. Setterberg (2011) finds that using the fundamental RIV approach when deriving cost of equity improves the correlation between the metrics and realized returns as well as risk measures.

Summarizing the research stream related to input data and analysts' forecasts, several approaches have been developed to get around the potential issue of biased input data. Hou et al.'s (2012) cross-sectional model and the Setterberg (2011) RIV metric have proven that avoiding analyst forecasts may provide for a better validation basis of cost of equity metrics.

2.4 Adjusting for non-randomness in cost of equity samples

Fairly unrelated to other previous research streams, Ecker et al. (2015) introduce a new area of adjustment, focusing on the accuracy of statistical practice. Instead of assessing the potential issue of analyst forecast quality, they problematize the sample of firms that have been given forecasts in the first place. According to Ecker et al. (2015), analyst followings are assigned after certain selection criteria; one example being size, since larger companies tend to be of higher interest to the public. On top of these criteria, implied cost of equity metrics pose other requirements such as positive earnings growth (Easton 2004; Ohlson-Juettner 2005) and positive value of absolute earnings (Easton 2004; Ohlson-Juettner 2005; Gebhardt 2001; Claus and Thomas 2001). Combining these criteria, a seemingly “randomly drawn” cost of equity sample may consist of firms that are on average larger and more profitable than the whole population of firms.

Providing evidence that a randomly drawn subsample of U.S. firms is in fact biased in the above described way, Ecker et al. (2015) evaluate the possibility of correcting for non-random sampling through a resampling method called distribution-matching. Using information about the marginal distribution of the dependent variable (realized returns) in a reference sample as well as in cost of equity subsamples, they resample observations from the non-random subsamples until a point where it appears as if samples are random draws from the full population. This procedure is performed in two different formats; the KS (Kolmogorov-Smirnov) based distribution-matching and the bin-based distribution-matching, both tested and proven efficient. The latter is a simplified version of the former, providing slightly less accurate results but imposing less heavy computational burden. Five cost of equity measures are tested; Claus and Thomas (2001), Gebhardt et al. (2001), Easton (2004), Ohlson and Juettner-Nauroth (2005) and a naïve model using Value Line dividends forecasts. When performing benchmark association tests based on the original (non-random) subsamples, all metrics show negative or insignificant correlations with realized returns. After resampling with distribution-matching however, all five metrics show positive correlation with realized returns. This indicates that random samples can be constructed, and more importantly, that samples of cost of equity may be in need of resampling in order to give a justified picture of expected returns.

Non-randomness in samples is not a new feature per se, but new in this particular setting. Ecker et al. (2015) discuss how distribution-matching relates to other resampling and missing

data approaches, mainly multiple imputations and Heckman-type selection models. They show that distribution matching leads to better results in terms of improving the correlation between cost of equity and realized returns.³

3. Motivation

In this section, previous research areas are tied together to motivate the direction of our thesis. An initial statement of our primary and secondary aim is followed by further motivation of the rationale behind our chosen research area. We also state hypotheses and practical implications for our chosen direction.

3.1 Area of focus

Based on the current state of research in the cost of equity field in general, and the problem of validating cost of equity metrics in particular, we focus on further investigating the latest addition to the literature. When applying distribution-matching to cost of equity samples in the U.S. market, Ecker et al. (2015) provide a new perspective to the research field. Through reaching particularly good results and presenting the model as a standalone solution for better cost of equity validation, Ecker et al. (2015) distance non-random sampling corrections from other progress made in previous cost of equity research.⁴ In this thesis, we aim to test the bin-based distribution-matching method in a new setting (the Swedish market) as well as challenge its stand-alone status through combining it with other selected models. We intend to answer whether distribution-matching is a sufficient stand-alone procedure or if it may reach further potential through incorporation with other methods.

Primary aim: Further validate the bin-based distribution-matching model (as presented by Ecker et al., 2015) as a means to improve cost of equity validation, through testing the model in new setting.

Secondary aim: Investigate whether combining bin-based distribution-matching with other efficient models may improve test results in a small market setting.

³ Further information on how these models relate can be found in Ecker et al. (2015).

⁴ Ecker et al. (2015) do however state that such a distinction may not be necessary over time.

3.2 Research rationale and hypotheses

Outlining previous research, it is clear that the cost of equity field is well-documented, however characterized by uncertainty regarding best practices and procedures. Much research is conducted with the same ambition; to improve validation methodology for cost of equity. In other words, the fundamental issue is that cost of equity metrics are difficult to validate empirically.

Several interconnected research areas have developed, all concentrating on identifying and correcting features that obstruct cost of equity testing. We have identified three main research areas that focus on i) improvement of cost of equity metrics, ii) alternation of the methods for validation and iii) adjustments for flaws in input data. Separate from these streams, Ecker et al. (2015) find that resampling through distribution-matching is in itself sufficient to enable validation of cost of equity metrics. This rather controversial position naturally calls for validation and brings us to our first aim: to further validate the bin-based distribution-matching model (as presented by Ecker et al. 2015) as a means to improve cost of equity validation.

There are several possible approaches for validating the bin-based distribution-matching method. For example, rerunning the original tests in the same (or in a very similar) setting would be an effective way to confirm previous results. However, this would not provide much room for development. We therefore choose to accept the U.S. market results, and instead test whether the model holds in a differently characterized market: the Swedish market. The small market setting strains the statement made by Ecker et al. (2015); that reducing sample size may actually increase the generalizability of results. This notion contradicts the standard approach of running tests using the largest possible number of observations, and is motivated by the increased strength in a distribution-matched sample, which is supposed to outweigh the negative effects from sample reduction. Although intuitively promising for smaller markets, this statement also questions whether the method is feasible in a setting where the sample size is substantially smaller to begin with. Since the resampling model imposes further sample reduction to an already small market, we predict that distribution-matching may not be as efficient when applied to the Swedish market.

Hypothesis I: *Bin-based distribution-matching as a means to improve the correlation between cost of equity and realized returns will be less efficient in a Swedish market setting.*

Predicting that a small market setting could introduce additional problems to the Ecker et al. (2015) bin-based model, we choose to investigate whether other areas of cost of equity research can contribute to improved results. By relying on unadjusted input data, Ecker et al. (2015) disregard the progress made within the research stream about input data quality. Since we expect the Swedish market to pose additional challenges to the model, we choose the opposite direction and incorporate such adjustments. This leads us to our secondary aim: to investigate whether combining bin-based distribution-matching with other efficient models may improve results in the Swedish market.

For this purpose, we concentrate on two additional methods, deliberately choosing approaches that have indicated success in other settings and that we believe could facilitate distribution-matching on the Swedish market. The first method is the cross-sectional model for generating earnings forecasts presented by Hou et al. (2012). Using model-based forecasts instead of analyst followings, the intention is to produce better proxies for expected earnings before performing the distribution-matching procedure. Moreover, the model is designed to increase the number and dispersion of observations, which could prove helpful when resampling a small number of observations. Though the model has been criticized for various reasons⁵, several researchers have reached uplifting results using it and thereby incentivizing us to incorporate it in our test setting⁶. In addition, we choose to include the cost of equity metric presented by Setterberg (2011). It was originally developed with Swedish data as basis and imposes less data restrictions than comparable metrics; two features that we predict will be beneficial with regards to our chosen market setting. We foresee that combining these alternate methods with bin-based distribution-matching is feasible, and that the positive effects from model-based input data and less data restrictions may improve final results.

Hypothesis II: *Bin-based distribution-matching can be successfully combined with other methods for improving the validation of cost of equity metrics in a Swedish market setting. Modeling earnings forecasts through the Hou et al. (2012) cross-sectional model and/or*

⁵ For example, Li and Mohanram (2014) criticize the cross-sectional earnings model developed by Hou et al. (2012) by pointing out that the model was not benchmarked to other cross-sectional models and has not been tested on firm level (although correlations have been proved to be stronger on portfolio level). They also mean that the Hou et al. (2012) model show somewhat odd correlations with certain risk factors. Apart from Li and Mohanram (2014), Gerakos and Gramacy (2013) show that the cross-sectional model underperforms a naïve random walk model that sets future earnings to past earnings. Li and Mohanram (2014) suggest two alternative cross-sectional models that in quantitative testing are shown to perform better in terms of forecast biases, accuracy, earnings response coefficients and correlations. The EP model forecasts earnings as a function of past earnings, whereas the RI model is based on residual income valuation.

⁶ See for example Jones and Tuzel (2012), Lee et al. (2011) and Patatoukas (2011).

computing implied cost of equity following Setterberg (2011) before performing the distribution-matching procedure will improve the correlation between cost of equity and realized returns.

3.3 Practical implications

With the aspiration to meet our aims, Ecker et al. (2015) is our primary most important paper to set up our methodological framework. In addition, the procedures described in Claus and Thomas (2001), Gebhardt et al. (2001), Easton (2004), Ohlson and Juettner-Nauroth (2005) will be followed in order to compute implied cost of equity metrics. These methods constitute the most recognized and most frequently used implied cost of equity metrics in current research.⁷ All four metrics are included in order to allow for robustness checks of results. We will further follow the procedure presented by Hou et al. (2012) to generate model-based earnings forecasts, and the method described by Setterberg (2011) to compute a final implied cost of equity metric.⁸

4. Methodology

In the following section, our methodological procedure will be described. Firstly, we introduce the data and the adjustments that have been made in order to make the set suitable for testing. Secondly, we outline our method and assumptions for computing cost of equity metrics. Then we proceed to the actual models; distribution-matching and the additional tests performed through cross-sectional estimation of earnings forecasts as well as the Setterberg (2011) metric. Lastly, we describe the simple association tests between cost of equity and realized returns.

⁷ Since our focus lies upon improving the *validation* of cost of equity metrics rather than the metrics themselves, we do not intend to question the metrics per se. Instead we make the assumption that they are sufficient measures of implied cost of equity.

⁸ The chosen methods naturally imply that there are several areas that we have chosen not to investigate. As previous literature indicates, there are several alternative methods and models that could have been considered in order to investigate our secondary aim (see for example Li and Mohanram 2014). Unfortunately, the computational efforts for performing distribution-matching on the four standard cost of equity metrics, estimating earnings forecasts with the cross-sectional model and computing the additional metric by Setterberg (2011) before running the final association tests pose a heavy workload alone. Also, the introduction of more additional tests would both obstruct the interpretation of our results and possibility to draw valid conclusions. Enlightening the most prominent alternative methods in the previous literature section, we aim to increase the transparency of our thesis and hopefully provide an explanation for unexpected test results. It also provides a foundation for potential improvement for further research.

4.1 Data and adjustments

The dataset includes 852 unique firms, all which have been listed in Sweden sometime during the chosen period 1995-2014. In order to maximize the sample size we include not only Nasdaq OMX Stockholm Exchange, but also second-degree listings such as Aktietorget and First North. Despite this, our sample is significantly smaller than almost any sample looked upon in the U.S., as a direct effect of the Swedish market being smaller than the U.S. market. As of 2014, Sweden had just over 300 companies listed on the Stockholm Stock Exchange, whereas the corresponding number in the U.S. approached 5000. With this in mind, we refer to our sample as “small”.

All calculations are made on an annual basis, with yearly observations rather than monthly. Other tests within the field vary between using firm-year-observations (Setterberg, 2011; Gebhardt, 2001; Claus and Thomas, 2001; Easton, 2004; Easton and Monahan, 2005) and firm-month-observations (e.g. Ecker et al., 2015), with the majority using the first alternative. Monthly data provides the general benefit of more observations (increasing sample size and preciseness), but has not proven unanimously superior in association tests on cost of equity. Therefore, we choose simplicity and stick to yearly observations.

We investigate the years 1995-2014, which hence constitute the timespan for the final association tests. Since ten years of historical data is necessary for the cross-sectional earnings forecast model, we have been collecting data on Swedish companies from 1985-2014. The time span is set to include as many years and thereby observations as possible, while still maintaining a sufficiently high quality of data. Data from the early 80's and back is heavily characterized by missing values and incomplete information. The Swedish market is thus not as well-documented as the U.S. market, which is investigated in many previous studies (Easton and Monahan, 2005; Easton, 2004; Ohlson and Juettner-Nauroth, 2005; Gebhardt, 2001; Hou et al., 2012). Setterberg (2011) ran tests on the Swedish market and chose to collect data manually, with the intention to increase the quality of data. However, this approach decreases the number of firms observed. In the trade-offs between sample size, quality of data and limitations of scope (time), we deem the 1995-2014 time span a sensible choice. We have also chosen not to search for and introduce additional databases than our main source in order to minimize overlap and potential mismatching.

Our main data source is Datastream, and necessary modifications have been made to refine the data set to minimize the risk of errors and obsolete observations affecting our results. Dead (de-listed) observations have been eliminated (after de-listing point in time) to rule out out-of-date information. Furthermore, several companies have different share types (A, B, C) causing them to appear more than once with the same information on all variables except for price. To be consistent in the sample, we aimed to keep the observation based on the share type most comparable to common stock, and thus in most cases kept the B-share.

Observations that lack information in any of the variables needed for a particular measure (e.g. missing earnings forecasts, dividend per share, reported book value etc.) are disregarded, since missing values would result in misleading implied cost of equity measures. In line with Easton and Monahan (2005) we delete firms whose fiscal year does not end in December (less than 5% of the original sample). This enables a consistent approach in calculation of the measures and is unlikely to infer with the data, except for decreasing the sample size slightly.

4.2 Computation of implied cost of equity metrics

We calculate four cost of equity metrics; r_{MPEG} (Easton, 2004), r_{OJN} (Ohlson and Juettner-Nauroth, 2005), r_{GLS} (Gebhardt et al., 2001) and r_{CT} (Claus and Thomas, 2001). All metrics are calculated on a yearly basis for the years 1995 to 2014. The procedure and assumptions made to compute each cost of equity metric are presented in this section. If no alternative source is stated, assumptions originate from the main metric papers.

Easton (2004)

$$r_{MPEG} = A + \sqrt{A^2 + \frac{eps_2 - eps_1}{P_0}} \text{ where } A = dps_1 / (2P_0)$$

- eps_1 and eps_2 refer to earnings per share and are measured as of December 31 year -1.
- dps_1 is the forecasted dividend per share for year 1. Consistent with Setterberg (2011) we use dps_{-1} , i.e. the actual dividend from the previous year, as a proxy for dividend forecasts, as these are difficult to obtain for Swedish companies.
- P_0 is the price at the valuation point in time.

Ohlson & Juettner-Nauroth (2005)

$$r_{OJN} = A + \sqrt{A^2 + \frac{eps_1}{P_0} * \frac{eps_2 - eps_1}{eps_1} - (Y - 1)} \text{ where } A = \left((Y - 1) + \frac{dps_1}{P_0} \right) \frac{1}{2}$$

- eps_1 and eps_2 are measured as of December 31 year -1.
- dps_1 is the forecasted dividend per share for year 1. Consistent with Setterberg (2011) we use dps_{-1} , i.e. the actual dividend from the previous year, as a proxy for dividend forecasts, as these are difficult to obtain for Swedish companies.
- P_0 is the price at the valuation point in time.
- $(Y - 1)$ is the perpetual growth rate in abnormal earnings beyond the forecast horizon. Following Gode and Mohanram (2003), Setterberg (2011) and Hou et al. (2012), we aimed to set the growth equal to the risk-free rate minus three percent. However, since Sweden has experienced low rates in recent years we obtain negative numbers for several years. We therefore alter the approach, and set the growth equal to the risk-free rate minus the average rate of inflation. For the years that still obtain negative values, we set the growth to zero. We use the year-end Swedish repo rate as the measure of risk-free rate.

Gebhardt et al. (2001)

$$P_0 = bps_0 + \sum_{t=1}^{12} (1 + r_{GLS})^{-t} ((ROE_t - r_{GLS})bps_{t-1}) + (1 + r_{GLS})^{-12} (r_{GLS}^{-1}) ((ROE_{ss} - r_{GLS})bps_{12})$$

- bps_0 is the book value of equity per share (ex-dividend) at the end of year -1, i.e. at the valuation point in time.
- bps_1 to bps_{12} are calculated under the assumption of the clean surplus relation using $bps_t = bps_{t-1} + (eps_t * (1 - dividend\ payout\ ratio))$
- P_0 is the price at the valuation point in time.
- eps_1, eps_2 and eps_3 are measured as of December 31 year -1.
- ROE_t is the return on equity. For years 1-3, values for ROE_t are measured as eps_1, eps_2 and eps_3 divided by the beginning of period book value per share.
- From year four to year twelve, ROE_t is assumed to fade linearly towards the steady state ROE.
- The steady state value for ROE (ROE_{ss}) is assumed to equal the historical median ROE in the industry that each respective firm operates in. Median industry ROE is calculated using rolling three-year windows, requiring at least three observations to calculate the median. We choose to include firms with negative earnings in the estimation of industry median ROE, consistent with Botosan and Plumlee (2005) and Setterberg (2011).
- The payout ratio, assumed constant, is used to forecast future book values of equity. Estimating the payout ratio is done through dividing the dividend per share with the earnings per share for year $t=-1$ i.e. $\frac{dps_{-1}}{eps_{-1}}$. Any negative values of the payout ratio are set equal to zero, and correspondingly, any values above one is set to one.

Claus & Thomas (2001)

$$P_0 = bps_0 + \sum_{t=1}^5 (1 + r_{CT})^{-t} ((ROE_t - r_{CT}) bps_{t-1}) + (R - g)(1 + r_{CT})^{-5} (r_{CT}^{-1}) ((ROE_{t+5} - r_{CT})(1 + g) bps_5)$$

- bps_0 is the book value of equity per share (ex-dividend) at the end of year -1, i.e. at the valuation point in time.
- eps_1 , eps_2 and eps_3 are measured as of December 31 year -1.
- eps_4 and eps_5 are computed using the long-term growth forecast and the three-year ahead forecast. Following Claus and Thomas (2001), long-term growth is set to the risk-free rate minus three percent.
- P_0 is the price at the valuation point in time.
- Consistent with Claus and Thomas (2001), we estimate the expected ROE in years t+1 to t+5 using the model-based or analysts' earnings forecasts and book equity determined based on clean surplus accounting ($B_{t+k} = B_{t+k-1} + E_{t+k} - D_{t+k}$, where E_{t+k} is the earnings in year t+k, D_{t+k} is the dividend in year t+k, computed using the current dividend payout ratio).

Following the definitions specified above, we compute one set of unadjusted cost of equity metrics, including no further modifications or assumptions. This set serves as our benchmark in determining whether resampling observations through distribution-matching results in an improved correlation between cost of equity metrics and realized returns.

4.3 Resampling with bin-based distribution-matching

After having produced our unadjusted metrics set, we perform bin-based resampling through distribution-matching, analogous to Ecker et al. (2015). This procedure consists of three main parts. Firstly, we define a reference sample to determine the desired distribution of subsamples. Secondly, we sort return observations in the reference sample and each separate cost of equity sample into bins. Lastly, we perform distribution-matching, i.e. resample cost of equity observations in such a manner that the distributions of the corresponding returns mimic the distribution in the reference sample. Each step is described in more detail below.

4.3.1 Defining a reference sample

A reference sample can be thought of as a “full population” sample, lingering upon the idea that any random subsample drawn from that reference sample should follow the same distribution with regards to the dependent variable. In a cost of equity setting, the data available is more or less complete with regards to realized returns. Similar to Ecker et al. (2015), we thus define our reference sample as all listed firms that have presented realized returns for any year in the chosen time span 1995 to 2014. Per definition, all companies with realized returns also have a cost of equity. In reality however, lack of information restricts the number of firms for which an implied cost of equity metric may be computed. Data on realized returns are available for nearly all listed Swedish companies whilst the cost of equity can be calculated for only 34-50 % (depending on the selected metric).

4.3.2 Sorting observations into bins

Having defined our reference sample, we turn to our set of unadjusted cost of equity metrics (containing calculations of r_{MPEG} , r_{OJN} , r_{GLS} and r_{CT}). We start by displaying the distribution (with regards to realized returns) for the reference sample as well as for each individual cost of equity sample. This is done through sorting observations into pre-defined bins (strata) of 500 basis points⁹. This gives a clear picture of what proportion of observations that can be found within each specified return interval. As exemplified in the table below, each cost of equity sample bin has a corresponding reference bin in the full return sample, which could either be of equal proportion (i.e. not in need of resampling) or different proportion.

Example

Return bin (interval)	Cost of Equity Sample	Reference Sample
Bin 1 (-0.05-0.00)	3%	3%
Bin 2 (0.00-0.05)	1%	5%
Bin 3 (0.05-0.10)	6%	2%
Bin 4 (0.10-0.15)	5%	5%

The table is compiled for illustrational purposes, and has nothing to do with our actual bin results.

⁹ Ecker et al. (2015) choose to define strata of 100 basis points. Increased bin width (to 500 basis points) could result in a loss of precision in the distribution-matching process. Instead of matching the proportion of observations for each percentage of realized returns, our construction implies that firms with realized returns in the range of five percentage points are bundled in the same bin. The Swedish market being small however makes a strata definition of 100 basis points impractical since many bins end up with only one or a few observations. The adjustment to 500 basis points per bin is necessary to meet the restrictions posed by a smaller market. Ecker et al. (2015) admits that their definition of strata is rather ad-hoc, which we deem justifies our decision to adjust the bin size for our chosen setting.

4.3.3 Distribution-matching

All cost of equity bins that contain different proportions of return observations compared to their corresponding reference bins are in need of resampling for the distributions to match. In such cases, we aim to arrive at a point where the proportion of observations in each bin equals the proportion in its corresponding reference sample bin. For example, if one of the bin is shown to contain 5% of the reference sample observations and only 1% of the cost of equity subsample observations in a given year (see bin 2 in the example table above), we resample that particular cost of equity bin to increase its proportion to 5% of the sample size in that year.

As the resampling is made on the basis of proportions, the desired final sample size can be freely specified. In our case, we set the resampled subsamples to equal the original cost of equity sample sizes, and thereby aim for a different number of observations for each respective year and metric. This procedure implies heavier computational burden than the alternative of deciding upon a sample size fixed over time. However, it also leads to a redistributed sample more similar to its original counterpart, naturally in all aspects except the distribution of realized returns. Referring to the purpose of the procedure (to randomize the subsample), fixing the sample size over the years would distinguish the subsample from the population of firms rather than aligning them. This we deem inappropriate.

Multiplying the desired sample size with the bin proportions gives the number of observations that have to be drawn out of each cost of equity bin in the resampling procedure. Observations are drawn randomly from the original cost of equity samples with replacement, which leads to some observations appearing several times in the resampled subsample, while other observations may just as likely be excluded. The randomized drawing of observations is repeatedly performed until we have effectively reweighted observations and reached the desired proportions for each respective bin.

Through resampling observations according to the procedure above, we create our second set of metrics. The values of specific cost of equity estimates have not changed, but each value has been given a different weight in the new subsample. Distribution-matching is performed for all four cost of equity metrics; r_{MPEG} , r_{OJN} , r_{GLS} and r_{CT} .

4.4 Additional tests

Having resampled the unadjusted cost of equity metrics we start a new computation procedure aiming to replace analyst forecasts with better input and potentially improve results. We examine two different ways of doing this; cross-sectional estimation of earnings and the Setterberg cost of equity metric.

4.4.1 Generating model-based earnings forecasts with the cross-sectional model

Following Hou et al. (2012) we use their cross-sectional model to generate proxies for expected earnings. This means that for each year, we estimate the following pooled cross-sectional regression using the previous ten years of data:

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

Where

$E_{i,t+\tau}$	denotes the earnings of firm i in year $t+\tau$ ($\tau = 1$ to 5)
$A_{i,t}$	is total assets
$D_{i,t}$	is the dividend payment
$DD_{i,t}$	is a dummy variable that equals 1 for dividend payers and 0 otherwise
$NegE_{i,t}$	is a dummy variable that equals 1 for firms with negative earnings and 0 otherwise
$AC_{i,t}$	is a measure to account for accruals

Since ten years of historical data is used to estimate the coefficients for each year, the data necessary for our first cost of equity year (1995) is collected from 1985 and onwards. When downloading the variables, we aimed to follow the definitions from Hou et al. (2012) as closely as possible. Since our data is obtained from another database, deviations do however occur. Each deviation from the original procedure has been designed to make sure that we follow a comparable procedure.

For *Earnings* we use net earnings before extraordinary items and preferred dividends from Datastream, from where we also have obtained *Total assets* and *Dividend payments*. *Accruals* is a computed variable and thereby gives most room for interpretation. Hou et al. (2012) have chosen a split between two different methods; the balance sheet method (where accruals have been set to equal the change in non-cash current assets less the change in current liabilities

excluding the change in short-term debt and the change in taxes payable minus depreciation and amortization expense) and the cash flow statement method (where accruals are the difference between earnings and cash flows from operations). The transition between the methods occurs in 1988, as cash flow statement information was not available for U.S. firms before that time. Transferring this into our Swedish setting we notice that even though cash flow statements are available from 1988 for Swedish firms in Datastream, extremely few values have been registered for the early years. Therefore, we choose to continue with the balance sheet method until 1994, which is when cash flow statements are available for a more substantial part of the firms.

Moreover, we choose to slightly adjust the balance sheet method (used for the years 1985-1994), with the sole purpose of producing a sufficient number of accruals observations. The adjustment means completely excluding the change in taxes payable, since this number is available for such a limited number of firms. In general, computing accruals on an overall level (as proposed by Hou et al., 2012) is quite arbitrary. To calculate a representative value, an interpretation of each company's annual report would be necessary to assess the value of the accruals. Within the scope of time for this thesis, we however accept the proxy used by Hou et al. (2012) even though we question its accuracy. Since *Accruals* is imprecise in itself we do not require numbers on all variables (i.e. non-cash current assets, current liabilities, short-term debt and depreciation and amortization expense) to compute a number, but settle with the information available.

Having chosen and defined the variables, we run the cross-sectional regressions and obtain coefficients for each year and forecast period (one, two and three years). These are used to estimate firm-specific earnings forecasts. After constructing these estimates for earnings forecasts, we input the numbers as analyst followings.

Having replaced analyst followings, we follow the same metrics computation procedure as when arriving at the unadjusted cost of equity metrics (see section 4.2). We compute the same four metrics; r_{MPEG} , r_{OIN} , r_{GLS} and r_{CT} . This set of metrics is studied separately as well as further adjusted through distribution-matching. We thus repeat the resampling procedure for our new metrics set based on model based earnings. Again, this means specific cost of equity estimates stay the same, but are given new weights in the manually created subsamples.

4.4.2 Introducing an additional implied cost of equity metric

As an additional method to avoid the use of analyst forecasts, we compute $r_{\text{SETTERBERG}}$; an adjusted cost of equity metric designed by Setterberg (2011). The metric builds upon Gebhardt et al (2001), and uses an extended RIV model as starting point when backing out cost of equity. Below follows the formula and our assumptions.

Setterberg (2011):

$$P_0 = bps_0 + \sum_{t=1}^{12} (1 + r_{RIV})^{-t} ((ROE_t - r_{RIV}) bps_{t-1}) + (1 + r_{RIV})^{-12} \left(\left(\frac{V_T}{BV(E)_T} - 1 \right) * bps_{12} \right)$$

- bps_0 is the book value of equity per share (ex-dividend) at the end of year -1, i.e. at the valuation point in time.
- For years 1 to 3, ROE_t is set to the firm-specific historical mean ROE, calculated using rolling three-year windows based on years -3 to -1. Consistent with Setterberg (2011) we require at least two firm-year ROE observations to obtain a mean value.
- In year 4 to 12, the ROE is assumed to linearly fade towards the steady state value.
- dpf is the dividend payout fraction, measured as dps_{-1} divided by bps_{-2} , i.e. the dividend per share for year -1 divided by the equity book value per share at the end of year -2. The value of dpf is assumed to be constant for years 1-3.
- bps_t for years 1 to 3 is calculated using the following expression: $bps_t = bps_{t-1} + bps_{t-1} * (ROE_t - dpf)$. From year 4 to year T, bps_t is forecasted assuming that the growth rate in owners' equity in year 3 fades linearly towards the steady state value, g_{ss} .
- g_{ss} , is assumed to be 2%.
- ROE_{ss} is assumed to be a function of the cost of equity capital, the steady state growth and the accounting measurement bias for the firm, according to the following expression:

$$ROE_{ss} = r + \left(\frac{V_T}{BV(E)_T} - 1 \right) (r - g_{ss})$$

- $\left(\frac{V_T}{BV(E)_T} - 1 \right)$ is the permanent measurement bias (PMB) of owner's equity at time T, a measurement of accounting conservatism. Values of PMB are industry specific, and collected from Runsten (1998).

The model differs from the other metrics in a number of aspects. Firstly, it generates earnings forecasts through using firm specific mean values of historical ROE, rather than analysts' earnings forecasts. This implies fewer data constraints and leads to a larger number of cost of equity observations. Secondly, it differs from Gebhardt et al. (2001) on how to assess dividend input values. Setterberg (2011) relates dividends to owners' equity rather than earnings, with the intuition that owners' equity is less volatile and thus provides a more stable impression of dividends. It is argued that stable dividends are more suitable to base forecasts on. Moreover, rather than assuming that the payout ratio is constant up to steady state, the growth in owner's equity is allowed to fade towards the growth rate in steady state, which prevents a large jump in dividends prior to entering the steady state. In addition to all this, the metric was developed specifically for the Swedish market setting. We consider the metric to be particularly suitable for small markets, since it does not exclude observations due to for instance negative earnings pattern (which the other measures by definition do).

An unadjusted Setterberg (2011) metric is computed according to the specifications above. Afterwards, we follow the same resampling process as for the other metrics. Since it does not require any analyst forecasts, the cross-sectional model is not applicable to combine it with.

4.5 Summary of metrics

Metrics set 1: Unadjusted cost of equity metrics

Metrics set 2: Distribution-matched observations, based on Metrics set 1

Metrics set 3: Cost of equity metrics based on model-based earnings

Metrics set 4: Distribution-matched observations, based on Metrics set 3

Additional: One unadjusted and one distribution-matched version of $r_{\text{SETTERBERG}}$

4.6 Association tests with realized returns

When all sets of metrics have been computed, we run association tests between the cost of equity metrics and realized returns. In a simple regression of realized (excess) returns and each of the cost of equity metrics (separately), we estimate the slope coefficients. We use the same regression framework as Ecker et al. (2015):

$$R_{i,t} - R_{f,t} = \delta_0 + \delta_{1,t} \text{Cof}E_t + \varepsilon_{i,t}$$

Worth mentioning is that this is the most conservative type of validation test, not including any further control variables. Researchers such as Botosan et al. (2011) suggest a much more extensive regression framework, controlling for size, leverage, beta and other factors. Similar to Ecker et al. (2015) we acknowledge the contribution and potential need for such additional variables in a validation process however choose to only look at the most conservative equation above. This provides a view of the overall effects of distribution-matching. Including further control variables in the regression is instead a suggestion for further research.

As an additional measure the Pearson correlation coefficient is calculated. The comparison between these coefficients constitutes the basis for our results. Following the argumentation of Ecker et al. (2015), Pearson's correlation coefficient is relevant for interpretation in addition to the regression coefficient since it is not affected by changes in the relative standard deviations of the two variables. The correlation coefficients are thus not affected by dispersion or changes in the number of observations, but rather isolate the effect from the distribution-matching itself.

5. Results

5.1 Metrics set 1: Unadjusted metrics (benchmark case)

5.1.1 Descriptive statistics for the cost of equity metrics

The first metrics set consists of the four different metrics (r_{MPEG} , r_{OJN} , r_{GLS} and r_{CT}) without further adjustments. The set is referred to as our benchmark case. Descriptive statistics show that cost of equity values vary between the metrics, with mean values ranging from 14.02% (r_{GLS}) to 17.58% (r_{OJN}). The median values are somewhat lower, 12.83% and 16.75% respectively, and standard deviations range from 7.32% to 12.57%. Note that the descriptive statistics are presented before association tests, i.e. before list-wise deletion of observations.

Table 1 – Descriptive statistics for cost of equity metrics (benchmark case)

	Realized Returns	r_{MPEG}	r_{OJN}	r_{GLS}	r_{CT}
Obs.	6952	2818	2477	2352	2334
Mean	0.0778	0.1506	0.1758	0.1402	0.1619
Std. Dev.	0.5699	0.0826	0.0844	0.0732	0.0980
Skewness	1.5197	2.4393	2.0366	4.3827	2.5118
Kurtosis	9.7484	15.0279	13.3990	38.6137	14.5217
Min	-0.9980	0.0034	0.0139	0.0072	0.0066
P5	0.7425	0.0596	0.0654	0.0671	0.0500
Q1	0.2847	0.0996	0.1229	0.1022	0.1056
Median	0.0397	0.1319	0.1675	0.1283	0.1421
Q3	0.3515	0.1808	0.2101	0.1602	0.1926
P95	0.9994	0.3021	0.3212	0.2421	0.3425
Max	5.2715	0.9529	0.9527	0.9977	0.9906

The table shows descriptive statistics for all five cost of equity metrics in their original design are presented, and descriptive statistics for realized returns are presented for comparison. Cost of equity ranges from values close to 0 % up to a maximum of 99.06 %. Outliers (larger than 100 %) are dropped from the sample. On realized returns, values range from -99.8% to 527% with a theoretical range from -100% to infinity. Median value of realized returns is 3.97 % indicating that an average firm can be expected to have slight positive, but close to zero, realized returns.

As predicted, the number of observations available varies substantially between the different metrics. The models based on abnormal earnings (r_{MPEG} and r_{OJN}) impose fewer data restrictions since they require less information. This leads to a slightly higher number of obtained observations, especially noticeable in r_{MPEG}. In comparison, the metrics subject for most restrictions (the RIV based metrics: r_{GLS} and r_{CT}) can only be obtained for 2253 and 2334 observations respectively. In comparison to the reference sample of realized returns, cost of equity numbers can be obtained for 34%-40% of the observations hence implying that data restrictions determine the appearance of our cost of equity subsample.

5.1.2 Association tests

Running association tests with realized returns and the unadjusted cost of equity metrics provides results in line with previous literature. Three of the metrics (r_{MPEG}, r_{OJN} and r_{GLS}) show a negative slope coefficient when regressed against realized returns, where all except r_{GLS} are significant. The last metric (r_{CT}) show a positive correlation with realized returns, although not significant.

Table 2 – Association tests between cost of equity and realized returns (benchmark case)

	r_{MPEG}	r_{OJN}	r_{GLS}	r_{CT}
Regression coefficient	-1.5969	-1.2234	-0.1751	0.0048
Pearson's correlation coefficient	-0.2486	-0.1990	-0.0215	0.0008
p-value	0.0000	0.0000	0.3635	0.9725
t-stat	-12.03	-8.94	-0.91	0.03
Obs.	2431	2104	2003	2000

The table provides a summary of association tests between realized returns and implied cost of equity metrics in their original design, i.e. the benchmark case. The association tests have been run for all four metrics (r_{MPEG} , r_{OJN} , r_{GLS} and r_{CT}). We present the regression slope coefficients, Pearson's correlation coefficient as well as the p-value and t-stat for each metric. The association is negative in three out of four metrics. The Claus & Thomas metric (r_{CT}) is the only one that shows a slight positive, however not significant, association to realized returns.

Testing Pearson's correlation between realized returns and each respective metric, we obtain coefficients ranging from -0.2486 to 0.0008, where all metrics with a negative correlation coefficient had a corresponding negative slope coefficient in the regression. These findings correspond to what a portion of other researchers have found in previous studies (Easton and Monahan, 2005; Guay et al., 2011; Hou et al., 2012). Notable in our case as well as for theirs is that results are the opposite of what is predicted by theory. As previously mentioned, we have excluded outliers above and below certain points to ensure that our results are not driven by extreme values. In addition to this, we include a complementary association test based on leverage (see Appendix for full overview). Theory would predict higher leverage to be associated with higher risk (see e.g. Modigliani and Miller, 1957), and thus we look for the same correlation pattern as between cost of equity and realized returns. Our results support our first regression and again contradict what theory predicts from well-behaving cost of equity metrics. This is an outcome crucial for the overall rationale of our thesis and we believe the results to be attributable to non-randomness in our cost of equity samples, following the argumentation by Ecker et al. (2015).

Conclusion Metrics set 1: Association tests with unadjusted cost of equity metrics and realized returns show weak or negative slope coefficients and correlation coefficients, which is in line with previous literature but contradicts common risk-return intuition.

5.2 Metrics set 2: Distribution-matched metrics

5.2.1 Results from bin-based distribution-matching

The second set of metrics contains the samples of the four cost of equity metrics after distribution-matching. In order to confirm the need of distribution-matching as well as the effect of the bin-based method, the descriptive statistics of the samples with regards to realized returns are considered (see Table 3).

Comparing the original cost of equity sample (benchmark case) with the reference sample, it appears that the subsamples have noticeably smaller standard deviation, skewness and kurtosis. The characteristics of our sample thus suggest that there are non-random features present, which distinguishes the distribution of our unadjusted cost of equity samples from the distribution of our reference sample. We deem this supportive of our conclusion from the first metric set; that the negative correlation between cost of equity and realized return may be attributed to non-aligned distributions. The descriptive statistics are in line with findings from the U.S. market setting presented by Ecker et al. (2015), though the differences in our Swedish data set are considerably smaller. The effects from the distribution-matching are however similar; the kurtosis and skewness of the distribution-matched sample have been increased hence aligned with the values of the reference sample. The standard deviation is also improved, although still notably smaller than what we were aiming for.

Table 3 - Descriptive statistics of realized returns, before and after resampling (r_{GLS})

	Obs.	Mean	Std. Deviation	Skewness	Kurtosis
Reference sample	6952	0.0778	0.5699	1.5197	9.7484
Actual cost of equity sample	2003	0.2114	0.4868	1.3486	8.2942
Distribution-matched cost of equity sample	1722	0.1355	0.5068	1.5015	9.1203

The table numerically illustrates the effect of distribution-matching, by outlining specific statistics (mean, standard deviation, skewness and kurtosis) on realized returns for the reference sample, the actual cost of equity sample and the distribution-matched cost of equity sample. Statistics show that skewness, kurtosis and standard deviation is notably lower in the actual cost of equity sample compared to the reference sample. After the distribution-matching process it can be seen that the distribution-matched sample has features more aligned with the reference sample. This example is based on the r_{GLS} measure and corresponding descriptive statistics for the other metrics are presented in Appendix.

Across the metrics, the distribution-matching has been performed with varying success with regards to matching of skewness, kurtosis and standard deviation with the levels in the reference sample. A potential source for these different degrees of success is the practical issue that arises in the distribution-matching process when the reference sample has observations in a certain bin that the cost of equity sample completely lacks. The lack of dispersion in the cost of equity sample thus leads to observations not being resampled for particular years, as there are no observations to draw from.¹⁰ This would have a negligible effect in a large sample, but due to our small sample size it is evident that it affects the overall distribution-matching results. Referring to the weak improvement of the standard deviation, this is one plausible explanation.¹¹ Moreover, these occurrences support our decision to use a larger bin width than originally used by Ecker et al. (2015). The fact that we experience problems arising from lack of dispersion with our current bin sizes indicates that using narrower bins would impose even worse practical issues.

The described effects from the resampling procedure are visualized in Figure 1, where the reference sample, the original cost of equity sample and the distribution-matched sample from the metric r_{GLS} are plotted. As can be seen in the figure, the post distribution-matching observations follow a pattern more similar to the distribution in the reference sample. This, in combination with similar descriptive measures (skewness, kurtosis and, to some extent, standard deviation) of the two samples, indicate that the resampling procedure was successful in terms of aligning the distributions and create a sample that appears more randomly drawn from the dataset. The characteristics of the cost of equity samples are slightly changed after the distribution-matching. Most notable is the significant decrease in observations, which ranges from 23.8 % (r_{MPEG}) to 27.0 % (r_{GLS}).¹²

¹⁰ Consider the fact that the cost of equity sample in general contains a proportionally larger amount of large and profitable firms. The smaller dispersion with regards to realized returns will, at least for some time periods, give rise to intervals (bins) of realized returns in the reference sample that are unrepresented in the original cost of equity sample. As an illustrating example; in year 2010, the reference sample contains 3.53% of the observations in the bin for realized returns between -55% and -50%. When resampling for r_{GLS} is performed, it is not possible to draw any observations in this particular bin since there are no observations with observable values for cost of equity with a corresponding realized return between -50% and -55%. The bin in question will therefore be set to 0% in the distribution-matched sample, and hence the distributions cannot per definition be completely matched.

¹¹ Further description of the distribution-matching results can be found in the Appendix, where the distributions for all metrics are plotted in their original and resampled appearance.

¹² Full descriptive statistics for Metrics set 2 can be found in the Appendix.

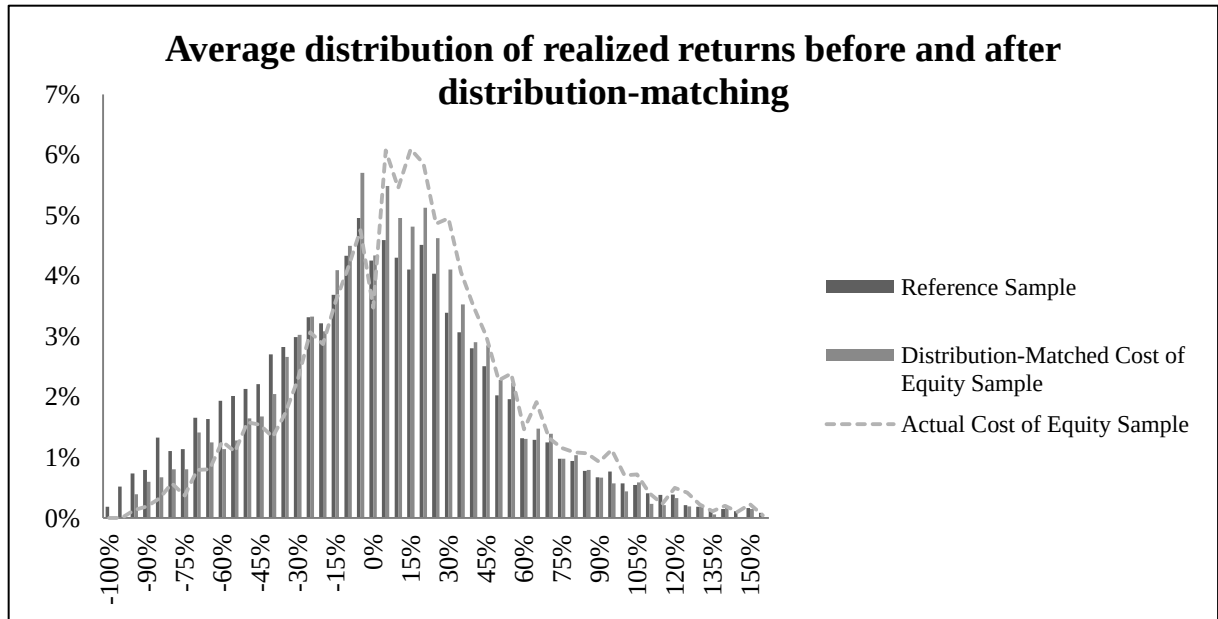


Figure 1: the difference between reference sample, actual cost of equity sample and distribution-matched sample is graphed to visualize the effects of the resampling process. The density (empirical distribution) of realized returns for each sample is plotted. For the purpose of this graph, the data is truncated at +160%. This example is based on the r_{GLS} measure (corresponding graphs for the other metrics are presented in Appendix II).

5.2.2 Association tests

Running the association tests between the resampled cost of equity metrics and realized returns, we notice several differences in the results. The overall effect from the resampling is positive but more modest than in Ecker et al. (2015)'s corresponding tests on the U.S. market. Positive correlation is obtained from two measures: r_{GLS} and r_{CT} . Compared to the benchmark case, r_{GLS} has changed from negative but not significant results (Pearson's correlation of -0.0215 with t-stat -0.91) to indicating a positive, significant correlation (Pearson's correlation of 0.1133 and t-stat 4.72). The r_{CT} remains insignificantly positive, however with increased size of the coefficients (Pearson's correlation changing from 0.0008 to 0.0255) and significance (t-stat rising from 0.03 to 1.05). The two simpler AEG based measures, r_{MPEG} and r_{OJN} , show negligible change and remain significantly negatively correlated to realized returns.

Table 4 - Association tests between cost of equity and realized returns, after distribution-matching (Metrics set 2)

	r_{MPEG}	r_{OJN}	r_{GLS}	r_{CT}
Regression coefficient	-1.6396	-1.2749	0.8510	0.1489
Pearson's correlation coefficient	-0.2455	-0.1937	0.1133	0.0255
p-value	0.0000	0.0000	0.0000	0.2920
t-stat	-10.40	-8.44	4.72	1.05
Obs.	1696	1830	1722	1717

The table provides a summary of association tests between realized returns and implied cost of equity metrics after distribution-matching. The association tests have been run for all four metrics (r_{MPEG} , r_{OJN} , r_{GLS} and r_{CT}). We present the regression slope coefficients, Pearson's correlation coefficient as well as the p-value and t-stat for each metric. r_{GLS} shows the most improvement; the association with realized returns is now positive and significant.

The results indicate that the resampling procedure on its own has significantly less effect in this test than in its original U.S. setting. The positive correlation which was proven for all five metrics in the Ecker et al. (2015) study is only found for two measures in this setting. Notable is that the Gebhardt metric (r_{GLS}) is the implied cost of equity metric that in previous literature has been proving the overall strongest result, regardless of which adjustment method that has been used. This is also the metric that responded best in terms of correlation to realized returns after the distribution-matching process in our tests.

Conclusion Metrics set 2: Descriptive statistics show that the distribution-matching procedure was successful but less precise in terms of aligning cost of equity return distributions with the reference sample distribution. The method has a positive impact on the association between realized returns and cost of equity, but with only r_{GLS} and r_{CT} showing positive correlation coefficients the effect is noticeably more modest than in its original design.

5.3 Metrics set 3: Model-based unadjusted metrics

5.3.1 Descriptive statistics of the cross-sectional forecast model

Following the model set up by Hou et al. (2012) we run a cross-sectional regression to arrive at 1-5Y earnings forecasts. Table 5 shows mean and median values as well as standard deviation and quartiles for all variables included in the cross-sectional regression. Comparing the numbers with the data set in the original model by Hou et al. (2012), we find that the mean values are fairly similar to their data base. An exception is that we have a substantially higher mean value of assets, which could be attributed to the fact that a large part of the dataset is made up of Swedish large cap firms (due to more default data on smaller firms).

Table 5 – Descriptive statistics for variables in cross-sectional earnings model

	E	A	D	DD	NegE	AC
Obs.	4943	5001	4743	4743	4943	4143
Mean	473	26534	192	0.6	0.3	-276
Std. Dev.	1868	129991	556	0.5	0.4	5364
Skewness	4.3	10	5.4	-0.6	1.0	32
Kurtosis	71.5	135	40.33	1.3	2.1	1628
Min	-21264	0.17	0.00	0.0	0.0	-38153
P5	-187	39	0.00	0.0	0.0	-2732
Q1	-2.67	290	0.00	0.0	0.0	-203
Median	32	1470	9.7	1.0	0.0	-21
Q3	281	9124	106	1.0	1.0	4.00
P95	2884	77192	1007	1.0	1.0	647
Max	32222	2484968	6369	1.0	1.0	272312

The table shows descriptive statistics for the six variables in the cross-sectional regression are presented. The variables are Earnings (*E*), Total Assets (*A*), Total Dividends (*D*), a dummy taking the value 1 if the company pays dividends (*DD*), an earnings dummy taking the value 1 if earnings are negative (*NegE*) and Accruals (*AC*). Up until 1994, accruals have been calculated with the balance sheet methods, i.e. set equal to the change in non-cash current assets less the change in current liabilities excluding the change in short-term debt minus depreciation and amortization expense. After 1994, accruals have been calculated with the cash flow statement method, i.e. set equal to the difference between earnings and cash flows from operations. We present the time-series averages of the cross-sectional mean, median, standard deviation, and selected percentiles. All variables except *DD* and *NegE* are expressed in MSEK.

Table 6 shows the average regression coefficients obtained from the cross-sectional model. All coefficients have a positive effect on future earnings except for the short-term (1-2Y) effect of negative earnings and accruals. The adjusted R^2 naturally decreases with the forecast horizon, and the cross-sectional model hence have the lowest explanatory power for future earnings five years ahead (35.07%). For one year earnings forecasts, the explanatory power is sufficiently higher (49.32%).

Table 6 – Regression results from the cross-sectional earnings model

	Intercept	E	A	D	DD	NegE	AC	Adj. R ²
1Y Earnings forecast	42.1444	0.4738	0.0056	0.6865	75.8695	-29.0186	-0.0357	0.4932
2Y Earnings forecast	34.0393	0.1684	0.0082	1.2266	158.3267	50.0776	-0.0390	0.3886
3Y Earnings forecast	52.4635	0.0350	0.0106	1.8295	155.2865	17.8147	0.0005	0.3802
4Y Earnings forecast	62.1863	0.0029	0.0103	2.3374	146.3056	21.1956	0.0025	0.3614
5Y Earnings forecast	46.1380	0.2186	0.0165	1.5116	191.1280	31.6135	0.0046	0.3507

The table shows average coefficients from the pooled regressions estimated for each year between 1995 and 2013 using the previous 10 years of data. The variables are Earnings (*E*), Total Assets (*A*), Total Dividends (*D*), a dummy taking the value 1 if the company pays dividends (*DD*), an earnings dummy taking the value 1 if earnings are negative (*NegE*) and Accruals (*AC*). Up until 1994, accruals have been calculated with the balance sheet methods, i.e. set equal to the change in non-cash current assets less the change in current liabilities excluding the change in short-term debt minus depreciation and amortization expense. After 1994, accruals have been calculated with the cash flow statement method, i.e. set equal to the difference between earnings and cash flows from operations.

5.3.2 Descriptive statistics of the cost of equity metrics

The third set of metrics contains the four cost of equity metrics, (r_{MPEG} , r_{OJN} , r_{GLS} , r_{CT}), with interchanged earnings forecasts values, and will be referred to as the model-based benchmark case. As compared to the original base case, the characteristics of the cost of equity metrics change distinctly when model-based earnings forecasts are used. The mean values have sufficiently increased, and range between 30.05% (r_{GLS}) and 38.07% (r_{MPEG}) (compare to 14.02% and 17.58% respectively in the base case). Also the median values, percentiles and standard deviations follow a similar pattern of increased values. As an example, the standard deviations are moved from base case levels of 7.32%-12.57% to a new level of 23.34%-26.17%.

Table 7 - Descriptive statistics for cost of equity metrics (based on Metrics set 3)

	r_{MPEG}	r_{OJN}	r_{GLS}	r_{CT}
Obs.	2108	2159	2304	1753
Mean	0.3807	0.3792	0.3005	0.3381
Std. Dev.	0.2448	0.2409	0.2334	0.2617
Skewness	0.6495	0.7051	1.1937	0.7848
Kurtosis	2.4861	2.5506	3.4642	2.5910
Min	0.0061	0.0096	0.0044	0.0035
P5	0.0668	0.0809	0.0698	0.0180
Q1	0.1825	0.1804	0.1249	0.1272
Median	0.3263	0.3187	0.2087	0.2630
Q3	0.5424	0.5374	0.4207	0.5066
P95	0.8583	0.8568	0.8061	0.8779
Max	0.9993	0.9993	0.9966	0.9999

The table shows descriptive statistics for all four cost of equity metrics adjusted with model-based earnings forecasts, and descriptive statistics for realized returns are presented for comparison. Cost of equity ranges from values close to 0 % up to a maximum of 99.99 %. Outliers (larger than 100 %) are dropped from the sample.

The seemingly extreme features of the cost of equity metrics obtained in the third set are to some extent in line with the findings of Hou et al. (2012) who presents significantly higher mean values for model-based cost of equity measures than its analyst-based equivalents. For example, the average mean value (all tested metrics included) increases from 8.52% to 14.92% when analyst-based forecasts are replaced with model-based forecasts. This supports the notion that model-based earnings are less precise than analyst forecasts. In our setting, the values are even more extreme (for example, the mean value of r_{GLS} more than doubles; from 14.02% to 30.05%).

A potential explanation is that the model in our case is calibrated using Swedish historical data, a set of observations that is substantially smaller than the U.S. historical data used in its original construction. The smaller amount of observations would be a feasible explanation to the rather high estimated earnings-forecasts. For instance, the mean value of A is notably higher in our case due to the smaller Swedish sample (including a greater proportion of large firms due to data limitations). Expanding the data set to include some other similar markets, or even to a full European level, could thereby lead to a more refined calibration of the model and more sensible model-based earnings forecasts. However, since it is important that the data set used for calibration of the model is representative for the sample for which the model is used to generate earnings forecasts, such a solution would have to be further evaluated.

Applying the model to a smaller market than in its original format is inevitably not efficient without further adjustments.

In terms of coverage we do not obtain more observations per year, which is opposite to Hou et al. (2012). We deem this closely related to the lack of historical data for some of the variables (e.g. AC and D), causing many observations to drop out since earnings forecasts cannot be modeled without this information.¹³

5.3.3 Association tests

Running the association tests with realized returns show a slight improvement compared to the unadjusted metrics in the benchmark case (Metrics set 1). The regression and correlation coefficients of r_{MPEG} and r_{OJN} remain significantly negative and are slightly closer to zero. r_{GLS} switches signs and shows a clear positive, however insignificant correlation with realized returns. r_{CT} on the other hand actually turns negative (formerly positive) but remains insignificant.

Table 8 - Association tests between cost of equity and realized returns (Metrics set 3)

	r_{MPEG}	r_{OJN}	r_{GLS}	r_{CT}
Regression coefficient	-0.2575	-0.2709	0.0963	-0.0103
Pearson's correlation coefficient	-0.1183	-0.1245	0.0418	-0.0049
p-value	0.0000	0.0000	0.1089	0.8730
t-stat	-4.81	-5.14	1.60	-0.16
Obs.	2108	2159	2304	1753

The table provides a summary of association tests between realized returns and implied cost of equity metrics after computing cost of equity based on the earnings forecasts generated by the cross-sectional model. The association tests have been run for all four metrics (r_{MPEG} , r_{OJN} , r_{GLS} and r_{CT}). We present the regression slope coefficients, Pearson's correlation coefficient as well as the p-value and t-stat for each metric. The results show a slight improvement compared to the benchmark case (Metrics set 1), with r_{GLS} showing most improvement.

Referring back to previously obtained results, it is notable that r_{GLS} remains the metric responding most positively to changes. The abnormal earnings based models (r_{MPEG} and r_{OJN}) remain negatively correlated with realized returns after both adjustments, whereas the r_{CT} measure shows ambiguous results.

¹³ The observant will notice that r_{CT} is suffering the most significant loss of observations (dropping to 1753 from its original 2000) as opposed to the closely related metric r_{GLS} . This difference is related to the interpretation of the metrics; for the computation of r_{CT} , model-based forecasts are used for forecast year 1-5, and for the r_{GLS} measure model-based forecasts are used for forecast year 1-3. Since the Hou et al. (2012) model is based on forward-looking data to generate model-based earnings forecasts, the numbers required can be obtained for 1995-2010 for r_{CT} and for 1995-2012 for r_{GLS} thereby explaining the significant difference in sample size between the two metrics.

Conclusion Metrics set 3: Generating cross-sectional earnings heavily boost cost of equity numbers; more than what seems reasonable. Association tests with realized return (before resampling) show a slight positive improvement, but the effect is less apparent than the effect from distribution-matching

5.4 Metrics set 4: Model-based distribution-matched metrics

5.4.1 Results from bin-based distribution-matching

In the fourth set of metrics, we perform the resampling process on the cost of equity samples computed with model-based earnings forecasts. The resampled subsamples show a distribution in realized returns more similar to the reference sample. The skewness and kurtosis in the model-based benchmark case samples are significantly higher (except for r_{GLS}) than in the reference sample (e.g. 1.75 and 10.11 compared to 1.51 and 9.74 for r_{GLS}), which is successfully adjusted for in the distribution-matched samples. The mean values are close to unaffected by the resampling process, whereas the standard deviations decrease after distribution-matching. These effects are in line with Ecker et al. (2015).

Table 9 - Descriptive statistics of realized returns, before and after resampling (r_{GLS})

	Obs.	Mean	Std. Deviation	Skewness	Kurtosis
Reference sample	6952	0.0776	0.5699	1.5197	9.7484
Actual cost of equity sample	2304	0.2297	0.5462	1.7526	10.1189
Distribution-matched cost of equity sample	2047	0.1610	0.5686	1.6368	8.8672

The table numerically illustrates the effect of distribution-matching, by outlining specific statistics (mean, standard deviation, skewness and kurtosis) on realized returns for the reference sample, the actual cost of equity sample and the distribution-matched cost of equity sample. This table focuses on results from Metrics set 3, i.e. cost of equity metrics based on modelled earnings forecasts. Statistics show that skewness, kurtosis and standard deviation is notably higher in the actual cost of equity sample compared to the reference sample. After the distribution-matching process it can be seen that the distribution-matched sample has features more aligned with the reference sample. This example is based on the r_{GLS} measure. Corresponding descriptive tables for the other metrics can be found in Appendix.

Compared to the distribution-matching of the original base case (Metrics set 2) the results of the resampling process are slightly better in the sample with model-based analyst forecasts. As predicted, the use of model-based analyst forecasts results in a wider dispersion of cost of equity observations. This diminishes the problem of cost of equity bins that potentially lack

observations, which was described earlier. In the case with model-based forecasts, the greater dispersion of cost of equity values increase the accuracy of the average distribution-matching since a higher portion of the reference sample bins can be matched with a corresponding cost of equity bin. For comparison, the average drop-out due to empty bins in the original distribution-matched case is 12.7%, and the corresponding number in the model-based distribution-matched case is 10.7%, implying the amount of dropped observations is decreased by 10.3% when using model-based analyst forecasts.

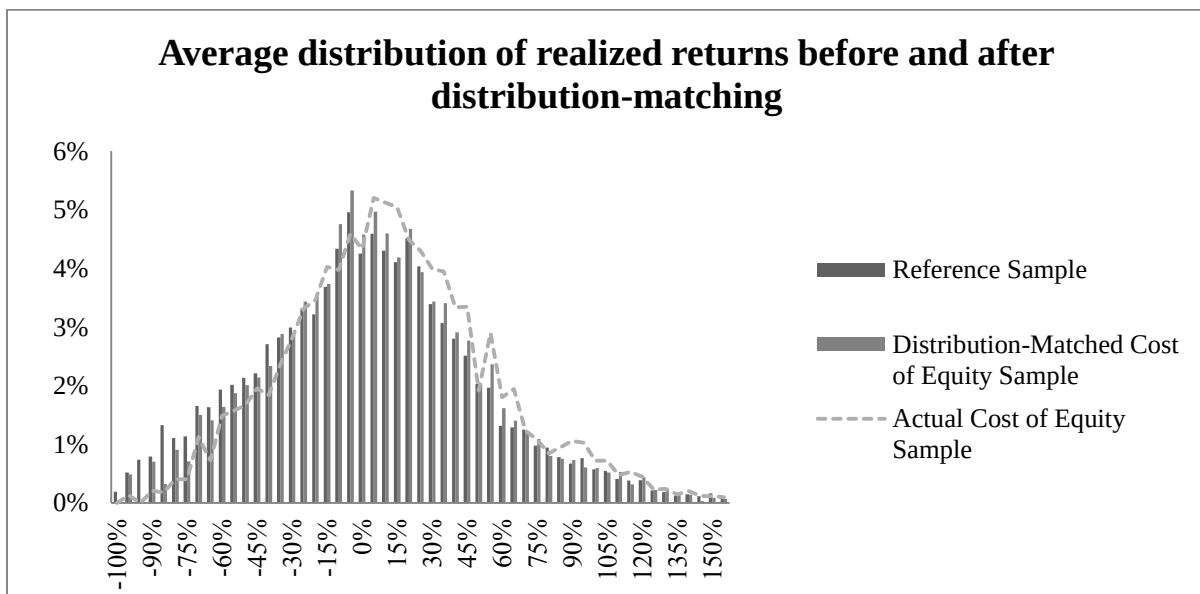


Figure 2: the difference between reference sample, actual cost of equity sample and distribution-matched sample is graphed to visualize the effects of the resampling process. The density (empirical distribution) of realized returns for each sample is plotted. For the purpose of this graph, the data is truncated at +160%. This example is based on the r_{GLS} measure (corresponding graphs for the other metrics are presented in Appendix).

The similar-looking distributions as well as matching of standard deviation indicates that the resampling procedure was successful in terms of aligning the distributions and create a sample that appears randomly drawn from the data set; even more so than in the Metrics set 2. The features of the samples with regards to the distribution of cost of equity show little change after the resampling procedure. The mean values remain high (30.82%-38.15%) and the standard deviations as well as the median values show no significant change. Most significant is the reduction in number of observations caused by the resampling, which caused the sample sizes to drop with 19.5% (r_{GLS}) to 28.9% (r_{CT}).¹⁴

¹⁴ Full descriptive statistics for Metrics set 4 are available in the Appendix.

5.4.2 Association tests

Running association tests on the model-based cost of equity sample provides results fairly similar to the previous tests. Compared to the model-based base case (Metrics set 2), we see improved correlation with realized returns among the RIV based metrics, r_{GLS} and r_{CT} which both show a positive correlation and regression coefficients. Similar to Metrics set 2, the r_{GLS} measure alone is positive and significant (Pearson's correlation 6.98% and t-stat 2.78). The abnormal earnings growth based models, r_{MPEG} and r_{OJN} , remain significantly negatively correlated to realized returns. Additionally, we notice that r_{CT} is showing a slight positive correlation in this setting, though it has been ambiguous and not significant throughout all test settings.

Table 10 - Association tests between cost of equity and realized returns (Metrics set 4)

	r_{MPEG}	r_{OJN}	r_{GLS}	r_{CT}
Regression coefficient	-0.2449	-0.3670	0.1615	0.0633
Pearson's correlation coefficient	-0.1624	-0.1636	0.0698	0.0304
p-value	0.0000	0.0000	0.0056	0.3460
t-stat	-4.21	-6.72	2.78	0.94
Obs.	1896	1931	2047	1493

The table provides a summary of association tests between realized returns and implied cost of equity metrics after distribution-matching of the cost of equity metrics based on the earnings forecasts generated by the cross-sectional model. The association tests have been run for all four metrics (r_{MPEG} , r_{OJN} , r_{GLS} and r_{CT}). We present the regression slope coefficients, Pearson's correlation coefficient as well as the p-value and t-stat for each metric. The results are similar to the distribution-matching of the unadjusted metrics in terms of coefficient signs, but are weaker in terms of size and significance levels.

The overall effect on the association tests from the combined use of model-based earnings forecasts and distribution-matching is weak. An interesting notation is that despite the fact that model-based forecasts increase sample dispersion and thereby the effectiveness of the distribution-matching, the final effect on the correlation with realized returns is non-existing. We thus see the same effect from distribution-matching as when resampling unadjusted metrics. This may be an effect of the less precise earnings forecasts and thereby possibly inconsistent cost of equity metrics. The dispersion of observations gave us the opportunity to better match the distributions; however, if cost of equity metrics are flawed to begin with, the resampling procedure will not improve the realized return association.

Conclusions Metrics sets 4: Resampling cost of equity observations from model-based earnings forecasts allows for more precise matching of distributions. However, running association tests show no further improvement of the correlations between cost of equity and realized returns. In fact, the strongest correlation observed between a cost of equity metric and realized returns is found in Metrics Set 2; resampled, analyst-based r_{GLS} show a significant correlation coefficient of 11.33%.

5.5 Summary of results from Metrics sets 1 - 4

Conclusion Metrics set 1: Association tests with unadjusted cost of equity metrics and realized returns show weak or negative slope coefficients and correlation coefficients, which is in line with previous literature but contradicts common risk-return intuition.

Conclusion Metrics set 2: Descriptive statistics show that the distribution-matching procedure was successful but less precise in terms of aligning cost of equity return distributions with the reference sample distribution. The method has a positive impact on the association between realized returns and cost of equity, but with only r_{GLS} and r_{CT} showing positive correlation coefficients the effect is noticeably more modest than in its original design.

Conclusion Metrics set 3: Generating cross-sectional earnings heavily boost cost of equity numbers; more than what seems reasonable. Association tests with realized return (before resampling) show a slight positive improvement, but the effect is less apparent than the effect from distribution-matching.

Conclusions Metrics sets 4: Resampling cost of equity observations from model-based earnings forecasts allows for more precise matching of distributions. However, running association tests show no further improvement of the correlations between cost of equity and realized returns. In fact, the strongest correlation observed between a cost of equity metric and realized returns is found in Metrics Set 2; resampled, analyst-based r_{GLS} show a significant correlation coefficient of 11.33%.

5.6 Results from additional test with $r_{\text{SETTERBERG}}$

The results of the Setterberg (2011) metric, before and after resampling are presented below. We include r_{GLS} (our most successful metric) for comparison.

Table 11 - Association tests between cost of equity and realized returns

	$r_{\text{SETTERBERG}}$	$r_{\text{SETTERBERG}}$ (resampled)	r_{GLS}	r_{GLS} (resampled)
Regression coefficient	-0.8148	-0.7341	-0.1751	0.8510
Pearson's correlation coefficient	-0.2014	-0.1840	-0.0215	0.1133
p-value	0.0000	0.0000	0.3635	0.0000
t-stat	-10.99	-9.85	-0.91	4.72
Obs.	3103	2800	2003	1722

The table provides a summary of association tests between the Setterberg cost of equity metric, $r_{\text{SETTERBERG}}$, and realized returns, both before and after distribution-matching. The table also includes r_{GLS} as a comparable metric. Results show that distribution-matching was a lot less efficient for the Setterberg metric in terms of improving the correlation with realized returns.

Notable is that the metric $r_{\text{SETTERBERG}}$ shows results similar to r_{MPEG} and r_{OJN} , and far from the positive results of the closely related r_{GLS} metric. The less restrictive data requirements did in its original study (Setterberg, 2011) prove effective to enhance the positive correlation with realized returns. We find the opposite results; neither in original nor resampled version a positive correlation occurs. The few data restrictions allow us to compute the metric for a larger number of firms and arrive at the implied cost of equity for 3103 observations during the time period. However, when regressed with realized returns, the slope coefficient is notably negative. This association remains after resampling as well.

Since including the metric under the pretenses that it was originally developed for a Swedish market setting, the results are somewhat disappointing. However, when Setterberg (2011) computed the metric, the data collection process was a lot more extensive than what we have carried out. Our data can thus be expected to be of less quality and potentially not suitable for the model. Additionally, it can be argued that industry specifications by Runsten (1998) are outdated. Since we have from the start aimed to keep an as large sample size as possible, we

have sorted companies into industries by a simple conversion of the pre-set industries in Datastream to the industries specified in Runsten (1998). This simplified procedure could also potentially affect the outcome of the metric negatively.

5.7 Further evaluation

Elaborating on the results from the different association tests, some main observations can be concluded: i) the resampling procedure does improve correlation between cost of equity and realized returns but only by a fraction of what was shown in the U.S. setting, ii) exchanging analyst forecasts to model-based forecasts enhances the accuracy of distribution-matching, but does not namely affect the association between cost of equity and realized returns, iii) r_{GLS} is the only measure that responds with positive, significant correlation to the use of distribution-matching and/or model-based earnings forecasts, and iv) the RIV based models (r_{GLS} and r_{CT}) perform better than the abnormal earnings growth models (r_{MPEG} and r_{OJN}) throughout all tests. In order to further understand the fundament of these results, we have made some additional tests.

5.7.1 Tests with absolute values of realized returns

The first additional test is performed to shed light on the differences in performance between the AEG based models (r_{MPEG} and r_{OJN}) and the RIV models (r_{GLS} and r_{CT}). The concluded differences in outcome are surprising, since we expect all proxies for expected returns to show similar results and characteristics when tested. Commenting on the robustness of results of the distribution-matching model, it is worrying that the effects are not strong across all metrics. To our knowledge, r_{MPEG} and r_{OJN} have been tested with varying results in previous research, though we do not find an immediate explanation to their poor performance in our test setting. One potential source of the differences is that r_{MPEG} and r_{OJN} features larger sample sizes as a consequence of lesser data restraints than the RIV models, and thereby contain additional information. If this selection of additional observations is characterized by less beneficial features for correlation results, the association test results would be affected negatively.

To elaborate further on this, we have made additional tests on absolute values of realized returns. We believe that such tests could add intuitive understanding to our results, even though such tests are not common practice in within the research field. Relating this to previous research, we know that realized return is often questioned with regards to its feasibility as a proxy for expected returns (Easton and Monahan, 2005; Botosan et al., 2011;

Vuolteenaho, 2002). Proxying expected returns with realized returns builds upon the assumption that the unexpected part of realized return has a mean value of zero, which we deem unlikely by the appearance of our dataset. If the proxies used are not sufficient, the association tests would show weak results regardless of whether distribution-matching is used or not.

Our intuition for using absolute values of return is the following: a low-risk firm would have a low cost of equity with investors expecting moderate returns. On the contrary, a high-risk firm would have a high cost of equity, with investors expecting a higher return in exchange for bearing the risk. A high-risk firm is characterized by more uncertainty than a low-risk peer, and the outcome of realized return will to a larger extent vary between highly positive (driven by expected return) or highly negative (driven by unexpected return). A high-risk firm will thus have a high cost of equity, but also be more likely to provide either high *or* low realized returns due to the uncertainty. Following this logic, we introduce absolute values of realized returns as a simple adjustment to examine whether they might provide a more fair view of expected returns. Inspired by additional testing in Ecker et al. (2013) we run association tests based on absolute values¹⁵, testing whether it could strengthen the relation between cost of equity and realized returns. This interpretation of expected returns thus allows firms with low return and high cost of equity to strengthen the association with realized returns instead of reducing it.

¹⁵ Ecker et al. (2013) split the cost of equity samples into two parts and run association tests on observations with positive returns and negative returns separately, finding a positive correlation to realized returns in the positive subsample and a corresponding negative correlation with returns in the negative subsample.

Table 12 - Association tests on Metrics set 1, absolute values

	r_{MPEG}	r_{OJN}	r_{GLS}	r_{CT}	$r_{SETTERBERG}$
Regression coefficient	-0.1052	-0.1282	0.1317	0.0530	0.0637
Pearson's correlation coefficient	-0.0219	-0.0276	0.0761	0.0683	0.0568
p-value	0.3030	0.2240	0.0010	0.0040	0.0020
t-stat	-1.03	-1.22	3.24	2.90	3.05
Obs.	2431	2104	2003	2000	3103

Association tests - Metrics set 1: The table provides a summary of association tests between realized returns and implied cost of equity *before* distribution-matching. The difference between this table and others, is that tests are made on absolute values of returns, assuming that these are better proxies for expected returns. A significantly negative return would indicate a high-risk company, and from an investor perspective a hope for a high return as compensation.

Table 13 - Association tests on Metrics set 2, absolute values

	r_{MPEG}	r_{OJN}	r_{GLS}	r_{CT}	$r_{SETTERBERG}$
Regression coefficient	0.2816	0.0434	1.8793	0.7623	0.2037
Pearson's correlation coefficient	0.0667	0.0089	0.3370	0.1796	0.0724
p-value	0.0020	0.7050	0.0000	0.0000	0.0000
t-stat	3.10	0.38	14.82	7.54	3.8200
Obs.	1696	1830	1722	1717	2800

Association tests - Metrics set 2: The table provides a summary of association tests between realized returns and implied cost of equity *after* distribution-matching. The difference between this table and others is that tests are made on absolute values of returns, assuming that these are better proxies for expected returns. A significantly negative return would indicate a high-risk company, and from an investor perspective a hope for a high return as compensation.

The association tests on absolute values are run for the metrics with analyst-based earnings forecasts (Metrics set 1 and Metrics set 2)¹⁶, including the additional metric $r_{SETTERBERG}$. The results are positive both in the benchmark case and, even more so, in the distribution-matched case. In the benchmark case, an overall improvement is presented. The AEG based metrics (r_{MPEG} and r_{OJN}) still have negative correlation and regression coefficients, however the remaining metrics are positive and significant. In the distribution-matched samples, the

¹⁶ Find the corresponding regression for Metrics set 4 in Appendix.

positive effects are even more substantial where all five metrics have a positive association with realized returns (four out of five metrics showing significant results).

Interestingly, we obtain positive coefficients for the distribution-matched r_{MPEG} and r_{OJ} , which has not been seen in any other test. This supports our notion that earlier results of r_{MPEG} and r_{OJN} have been driven by the associations between *negative* returns and cost of equity, rather than positive returns. This also seems to be the case for the $r_{\text{SETTERBERG}}$ metric, possibly indicating that the fraction of negative returns increase with the number of observations.

Testing with absolute values is useful, since it allows us to disregard the zero-assumption of unexpected returns that is otherwise imposed in our original test setting. It however implies other assumptions about the level of high versus low risk. Using absolute values indicate that realized returns (positive or negative) close to zero are regarded as “low”, whereas realized returns (positive or negative) going towards infinity are regarded as “high”.¹⁷ Although this may be intuitive over time, it remains clear that the level for what the market regard as “low” versus “high” returns changes between individual years. For instance, in years of economic downturn, a value of -20% in realized return could be normal for a low-risk firm, indicating a low theoretical cost of equity value. On the contrary, the same return level for the same firm in another year could be seen as a “high” loss, hence suddenly attributed to a high-risk firm.

5.7.2 Tests excluding years of economic downturn

Considering the issue of non-adjustments for yearly variations, it is not only evident in our tests with absolute values but could also be an obstruction of the original tests made in section 5.1 to 5.4. Since the association tests are run on pooled observations, tests are performed disregarding the fact that a high return one year can be considerably different from a high return another year. In order to test whether years of unusual levels might obstruct the relationship between cost of equity and realized returns, we perform additional tests where years of particular economic downturn is excluded from the sample. In particular, we run new association tests on Metrics set 2, excluding the years 2001 and 2008 that were characterized with untypically low company returns on the Swedish market. The overall returns during those years are significantly lower than average, where 78.9 % versus 99.4 % of all observations have returns below zero. This can be compared to the average of 41.6 % observations with negative values of realized returns in the whole sample. The results are shown in the tables below.

¹⁷ “High” returns in this sense are expected to be accompanied by a high cost of equity and “low” returns are expected to be accompanied by a low cost of equity.

Table 14 - Association tests on Metrics set 2, excluding year 2001 and 2008

	r_{MPEG}	r_{OJN}	r_{GLS}	r_{CT}
Regression coefficient	-2.0253	-1.3007	1.4109	0.4030
Pearson's correlation coefficient	-0.3054	-0.1921	0.1930	0.0707
p-value	0.0000	0.0000	0.0000	0.0050
t-stat	-14.10	-8.04	7.82	2.81

Association tests - Metrics set 2: The table provides a summary of association tests between realized returns and implied cost of equity after distribution-matching. The difference between this table and others, is that the years 2001 and 2008 have been excluded, since these were deemed to have particularly negative impacts on overall association results. These years coincide with years of financial difficulties on the Swedish stock market. Since the years 2001 and 2008 were excluded after the resampling process, the outcome for other years have not been affected.

The results from the above regressions are partially positive; excluding the years 2001 and 2008 has a substantial positive effect on the two RIV based metrics (r_{GLS} and r_{CT}). Both the regression coefficients and the Pearson's correlation coefficients are now positive and more significant as compared to the results from the full sample tests in section 5.2. This can be interpreted as an indication that years with unusual patterns of values, such as during economic downturns, dampen the effect from the average years and hence affect the results obtained from the original association tests. Relating back to previous research, realized returns in 2001 and 2008 could on this basis be argued to be driven by unexpected returns rather than expected returns and thereby prove weaker correlation to cost of equity.

6. Limitations of scope

By choosing to focus only on distribution-matching, we automatically make assumptions regarding the quality and accuracy of the downloaded data. As outlined in previous research, earlier contributions within the field focus on improving other areas such as altering metrics, change the realized return regression or make adjustments to analyst forecasts. We do not discuss if there might be better ways to arrive at the implied cost of equity hence, if the metrics are inadequate, we are making adjustments to something that is incorrect to begin with. The same goes for the association tests with realized returns; Easton and Monahan (2005) have shown that controlling for cash flow news and return news generates results more in line with theory. Instead of incorporating this, we follow Ecker et al. (2015) and thus

assume that realized return is a feasible proxy for expected returns, again potentially searching for positive correlation in something that is not supposed to correlate in the first place. We believe that to fully come to terms with aligning empirical results with theory within the cost of equity field, alterations would have to be made within all areas. However, the scope of such work is too broad for this thesis.

Another limiting feature of this study is the chosen bin size in the resampling process. In its original performance, the bin-based method was performed on U.S. sample by sorting the data into bins of 100 basis points, whereas we have chosen larger bins of 500 basis points. Consequently our resampling is less precise, with observations being grouped on a broader level. As observations are drawn randomly with replacement, each firm observation stands the same chance to be drawn in each round. With larger bins, we risk overestimating or underestimating the effects of resampling, by drawing extreme observation within each bin several times. We considered this a trade-off between preciseness and overall functionality, since it makes little sense to have almost as many bins as we have observations.

In our work we have also chosen to equate distribution-matching with the bin-based method, even though it was originally presented by Ecker et al. (2015) as a simplified alternative to the KS (Kolmogorov-Smirnov) based distribution-matching. The KS method was criticized for being complicated and time consuming, however resampling with that method did lead to better results. Since we have chosen to use the simpler procedure, we cannot make any conclusions of distribution-matching as phenomenon but only the bin-based method as composed in our thesis.

7. Conclusion

This study investigates the possibility of improving cost of equity validation through correcting for non-randomness in samples with bin-based distribution-matching. The non-random features of the samples are assumed to occur as a result of data restrictions in cost of equity settings, resulting in a biased sample of firms obtaining cost of equity calculations. We test four main cost of equity metrics; r_{MPEG} , r_{OJN} , r_{GLS} and r_{CT} , originating from Easton (2004), Ohlson and Juettner-Nauroth (2005), Gebhardt et al (2001) and Claus and Thomas (2001), and run tests based on a dataset of Swedish listed firms in the years 1995-2014. In different turns we compute i), unadjusted metrics and ii), metrics based on modelled earnings forecasts,

and run the bin-based distribution-matching procedure on both sets. We also compute a fifth metric, $r_{\text{SETTERBERG}}$, as an additional test and run regressions on absolute values of returns.

Our primary aim is to validate the bin-based distribution-matching method presented by Ecker et al. (2015), as a means to improve cost of equity validation. Our secondary aim is to investigate whether combining bin-based distribution-matching with other models (a cross-sectional model for generating earnings forecasts by Hou et al., 2012 and a RIV based metric by Setterberg, 2011), may further improve the correlation between cost of equity and realized returns. Our initial hypothesis was that introducing resampling in a smaller market setting could reduce the effect from the bin-based distribution-matching as compared to its original setting. We further hypothesized that distribution-matching is not necessarily optimal as a stand-alone procedure, and that results obtained in a small market setting could potentially be enhanced if combined with other models from the cost of equity research field.

Results based on our benchmark case of unadjusted metrics show that none of the four metrics can present a positive and significant correlation with realized returns, which would be the relation predicted by theory. Instead, findings are in line with previous empirical research (Ecker et al., 2015; Easton and Monahan, 2005; Guay et al., 2011) and suggest weak or even negative correlations in association tests. Studying the distribution of the returns in our reference sample versus the returns in each respective cost of equity subsample, we notice differences. Through distribution-matching, we align the distributions and construct cost of equity subsamples that appear to be randomly drawn from the full population of firms (as opposed to our actual cost of equity samples). Results point to positive correlations between realized returns and the RIV based cost of equity measures (r_{GLS} and r_{CT}). To allow for further interpretation of results, we rerun our initial tests using absolute values of return rather than actual returns. After resampling, we find reliably positive correlations between absolute returns and four of the cost of equity metrics. Overall, the improvements of distribution-matching are evident, but not as apparent as in a U.S. setting, which confirms our initial hypothesis regarding small markets.

Our attempt to combine distribution-matching with other suggested methods from the cost of equity research field turned out to be less successful. Results show that distribution-matching performed on cost of equity samples based on modelled earnings forecasts rather than actual analyst followings does not improve the correlation results in this setting. However, it is

difficult to state whether this is attributed to distribution-matching being most feasible as a stand-alone method, or if the underlying data and number of observations were not sufficient enough to fulfill a satisfactory calibration of the cross-sectional model. The latter seems likely, considering the generation of extreme earnings forecasts from the model and, thereby, heavily exaggerated cost of equity numbers. Lack of improvement also shows in the Setterberg (2011) metric, $r_{\text{SETTERBERG}}$. Similar to the cross-sectional model, we question whether we have had enough quality of data to make a fair assessment. Based on these results, we cannot accept nor reject our hypothesis that the effects of distribution-matching can be enhanced by combining it with other suggested models.

On a broader view, our findings to some extent confirm Ecker et al. (2015)'s notion that researchers might benefit from more closely examining distribution alignments and correct for deviations between a cost of equity sample and the full returns sample. Corrective actions, such as distribution-matching, may improve generalizability of results. Additionally, our findings suggest that researchers should be more careful in a small market setting, as models and methods developed for larger sample sizes may generate more modest results or simply not hold. In this study, distribution-matching is shown to be a helpful tool for a Swedish data set, but not at all as efficient as in a larger market.

For future research we suggest further investigation of distribution-matching in combination with other improvement methods. Our somewhat odd results from cross-sectional earnings forecasts open up for a similar methodology on greater scale, for example on the U.S. or overall European market. If more reasonable results can be obtained from the model, a more fair view of its feasibility in combination with distribution-matching could be given. Furthermore, our study opts for a conservative approach by running regressions based on realized returns. We see a possibility of extending the regression model and include control variables in accordance with for example Botosan and Plumlee (2011). Lastly, we deem our additional tests with absolute values of realized returns to be an area for potential future development. If some variation of absolute returns (before or after distribution-matching) could be argued to give a representative picture of expected returns (as our simplified test suggests), it might propose that some cost of equity metrics used in accounting and finance have greater construct validity than previous empirical results indicate.

8. References

- Baginski, S.P. & Wahlen, J.M. 2003, "Residual income risk, intrinsic values, and share prices", *The Accounting Review*, vol. 78, no. 1, pp. 327-351.
- Blume, M.E. & Friend, I. 1973, "A NEW LOOK AT THE CAPITAL ASSET PRICING MODEL", *The Journal of Finance*, vol. 28, no. 1, pp. 19.
- Brav, A., Lehavy, R., & Michaely, R. 2005. "Using expectations to test asset pricing models". *Financial Management*, 34, 31–64.
- Botosan, C.A., 1997, "Disclosure level and the cost of equity capital". *The Accounting Review*, 72, 323–349.
- Botosan, C.A. & Plumlee, M.A. 2005, "Assessing Alternative Proxies for the Expected Risk Premium". *The Accounting Review*, vol. 80, no. 1, pp. 21-53.
- Botosan, C.A., Plumlee, M.A. & Wen, H. 2011, "The Relation between Expected Returns, Realized Returns, and Firm Risk Characteristics". *Contemporary Accounting Research*, vol. 28, no. 4, pp. 1085.
- Campbell, J. Y., 1991, "A Variance Decomposition of Stock Returns". *The Economic Journal*, Vol. 101, pp 157-179.
- Claus, J. & Thomas, J. 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.D. 2004, "PE Ratios, PEG Ratios, and Estimating the Implied Expected Rate of Return on Equity Capital", *The Accounting Review*, vol. 79, no. 1, pp. 73-95.
- Easton, P.D. & Monahan, S.J. 2005, "An Evaluation of Accounting-Based Measures of Expected Returns", *The Accounting Review*, vol. 80, no. 2, pp. 501-538
- Ecker, F., Francis, J., Olsson, P., Schipper, K. 2013, "Associations Between Realized Returns and Risk Proxies Using Non-Random Samples". Working paper, Duke University.
- Ecker, F., Francis, J., Olsson, P., Schipper, K. 2015, "Non-Random Sampling and Association Tests on Realized Returns and Risk Proxies". Working paper, Duke University.
- Elton, J. 1999. "Expected return, realized return, and asset pricing tests". *Journal of Finance*, 54, 1199–1220
- Fama, E.F. & French, D.R. 1997, "Industry costs of equity", *Journal of Financial Economics*, vol. 43, no. 2, pp. 153-193.
- Fama, E., French, K., 2000. "Forecasting profitability and earnings". *Journal of Business* 73, 161–175.

- Fama, E., French, K., 2006. "Profitability, investment and average returns". *Journal of Financial Economics* 82, 491–518.
- Gebhardt, W.R., Lee, C.M.C. & Swaminathan, B. 2001, "Toward an implied cost of capital", *Journal of Accounting Research*, vol. 39, no. 1, pp. 135-176.
- Gerakos, J., & Gramacy, R., 2013. "*Regression-based earnings forecasts*". Working paper, University of Chicago.
- Gode, D., and Mohanram, P., 2003, "Inferring the cost of capital estimates using the Ohlson-Juettner model". *Review of Accounting Studies* 8: 399–431.
- Gordon, J., and Gordon, M., 1997. "The finite horizon expected return model". *Financial Analysts Journal* 53(3): 52–61.
- Guay, W., Kothari, S.P. & Shu, S. 2011, "Properties of implied cost of capital using analysts' forecasts", *Australian Journal of Management*, vol. 36, no. 2, pp. 125-149.
- Hartwig, F. 2012, "The Use Of Capital Budgeting And Cost Of Capital Estimation Methods In Swedish-Listed Companies", *Journal of Applied Business Research*, vol. 28, no. 6, pp. 1451-1476.
- Hou, K., Robinson, D., 2006. "Industry concentration and average stock returns". *Journal of Finance* 61, 1927–1956.
- Hou, K., van Dijk, M., 2011. "Profitability Shocks and the Size Effect in the Cross-section of Expected Stock Returns". Working Paper. Ohio State University
- Hou, K., van Dijk, M.,A. & Zhang, Y. 2012. "The implied cost of capital: A new approach", *Journal of Accounting & Economics*, vol. 53, no. 3, pp. 504.
- Jones, C., & Tuzel, S., 2012. "Inventory investment and the cost of capital". *Journal of Financial Economics*, 107(3), 557–579.
- Lee, C., So, E., & Wang, C. 2011. "*Evaluating implied cost of capital estimates*". Working paper, Stanford University.
- Li, K.K. & Mohanram, P. 2014. "Evaluating cross-sectional forecasting models for implied cost of capital". *Review of Accounting Studies*, vol. 19, no. 3, pp. 1152-1185.
- Lys, T., & Sohn, S. 1990. "The association between revisions of financial analysts' earnings forecasts and security-price changes". *Journal of Accounting and Economics*, 13, pp. 341–363
- Modigliani, F. & Miller, M. 1957. "The cost of capital, corporate finance and the theory of investment". *The American Economic Review*, vol. XLVIII, no. 3, pp. 261-297
- Mohanram, P. & Gode, D. 2013. "Removing predictable analyst forecast errors to improve implied cost of equity estimates", *Review of Accounting Studies*, vol. 18, no. 2, pp. 443-478

- O'Hanlon, J., & Steele, A. 2000. "Estimating the equity risk premium using accounting fundamentals". *Journal of Business Finance and Accounting*, 27, 1051–1084.
- Ohlson, J.A. 1995. "Earnings, book values, and dividends in equity valuation". *Contemporary Accounting Research*, 11, 661–687.
- Ohlson, J.A. 2005. "On Accounting-Based Valuation Formulae*". *Review of Accounting Studies*, vol. 10, no. 2-3, pp. 323-347.
- Ohlson, J.A. & Juettner-Nauroth, B. 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, 363–392.
- Setterberg, H., 2011, "The pricing of earnings: essays on the post-earnings announcement drift and earnings quality risk". Dissertation, Stockholm School of Economics.
- Vuolteenaho, T. 2002. "What drives firm-level stock returns?". *Journal of Finance*, 57, 233–264

APPENDIX

Appendix 1 – Metrics set 1: Association tests between cost of equity and leverage

	r_{MPEG}	r_{OJN}	r_{GLS}	r_{CT}
Regression coefficient	-11.6601	-3.8300	-0.6169	-1.5463
Pearson's correlation coefficient	-0.0759	-0.0490	-0.0194	-0.0617
p-value	0.0000	0.0200	0.3620	0.0040
t-stat	-3.8300	-2.3200	-0.9100	-2.8900
Obs.	2431	2104	2003	2000

The table provides a summary of association tests between leverage (as opposed to realized returns) and implied cost of equity in the metrics original design. Leverage is long-term debt scaled by book value of equity. The association is negative in all four metrics, even though we would expect leverage to be associated with higher risk and hence a higher cost of equity.

Appendix 2 – Metrics set 2: Descriptive statistics of cost of equity metrics after resampling

	r_{MPEG}	r_{OJN}	r_{GLS}	r_{CT}
Obs.	2148	1830	1716	1708
Mean	0.1559	0.1744	0.1390	0.1555
Std. Dev.	0.0895	0.0784	0.0682	0.0854
Skewness	2.6379	1.3648	4.9564	1.9763
Kurtosis	15.8990	8.5836	53.1715	10.4736
Min	0.0034	0.0139	0.0179	0.0078
P5	0.0591	0.0617	0.0673	0.0507
Q1	0.1010	0.1225	0.1036	0.1051
Median	0.1370	0.1695	0.1297	0.1392
Q3	0.1839	0.2106	0.1596	0.1861
P95	0.3194	0.3197	0.2383	0.3149
Max	0.9529	0.7082	0.9977	0.7276

The table presents descriptive statistics for Metrics set 2 (containing our four cost of equity metrics after distribution-matching). We show a selection of statistics, including number of observations, mean, standard deviation, skewness, kurtosis, max- min- and median values as well as quartiles for each metric. The number of observations in each sample is reduced from its original size by an average of 25%. The mean values and standard deviation have remained fairly stable (+/- 0.5%) as compared to the unadjusted benchmark case. The maximum cost of equity values represented in r_{OJN} and r_{CT} have decreased significantly as a consequence of the resampling.

Appendix 3 – Metrics set 2: Descriptive statistics for realized returns, before and after resampling (r_{MPEG})

	Obs.	Mean	Std. Deviation	Skewness	Kurtosis
Reference sample	6952	0.0777	0.5699	1.5196	9.7484
Actual cost of equity sample	2431	0.1851	0.5111	1.7105	11.9662
Distribution-matched cost of equity sample	1696	0.1295	0.5182	1.3942	8.6734

The table numerically illustrates the effect of distribution-matching, by outlining specific statistics (mean, standard deviation, skewness and kurtosis) on realized returns for the reference sample, the actual cost of equity sample and the distribution-matched cost of equity sample. Statistics show that skewness, kurtosis and standard deviation is notably higher in the actual cost of equity sample compared to the reference sample. After the distribution-matching process it can be seen that the distribution-matched sample has features more aligned with the reference sample. This example is based on the r_{MPEG} measure.

Appendix 4 – Metrics set 2: Descriptive statistics for realized returns, before and after resampling (r_{OJN})

	Obs.	Mean	Std. Deviation	Skewness	Kurtosis
Reference sample	6952	0.0777	0.5699	1.5196	9.7484
Actual cost of equity sample	2104	0.1937	0.5002	1.6648	12.5553
Distribution-matched cost of equity sample	1830	0.1442	0.5148	1.5551	9.7236

The table numerically illustrates the effect of distribution-matching, by outlining specific statistics (mean, standard deviation, skewness and kurtosis) on realized returns for the reference sample, the actual cost of equity sample and the distribution-matched cost of equity sample. Statistics show that skewness, kurtosis and standard deviation is notably higher in the actual cost of equity sample compared to the reference sample. After the distribution-matching process it can be seen that the distribution-matched sample has features more aligned with the reference sample. This example is based on the r_{OJN} measure.

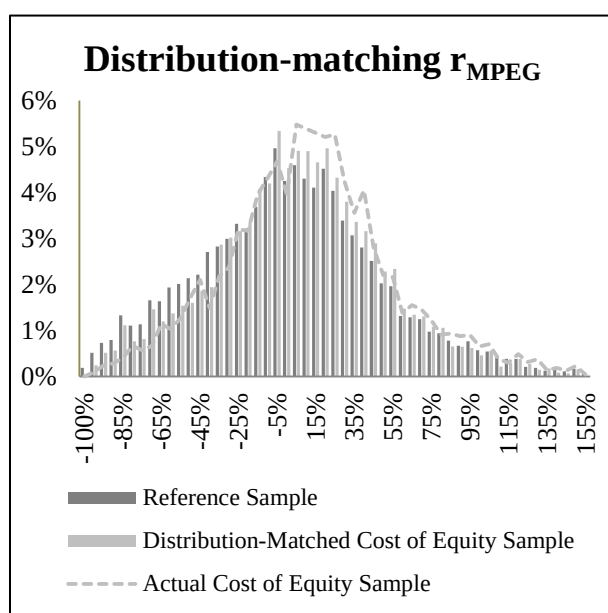
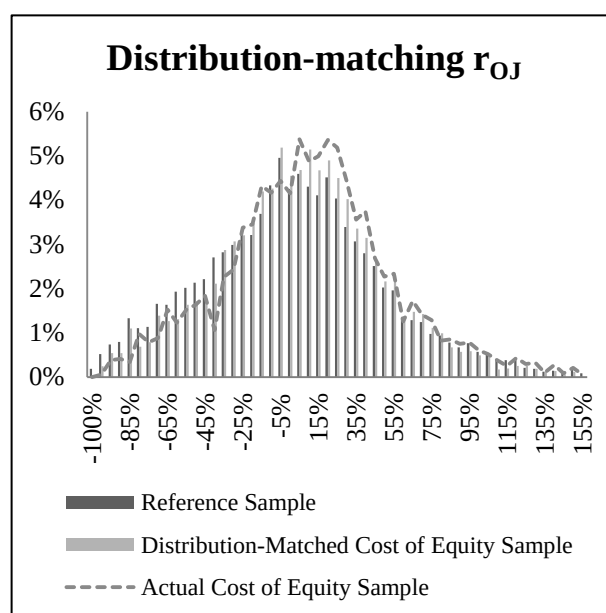
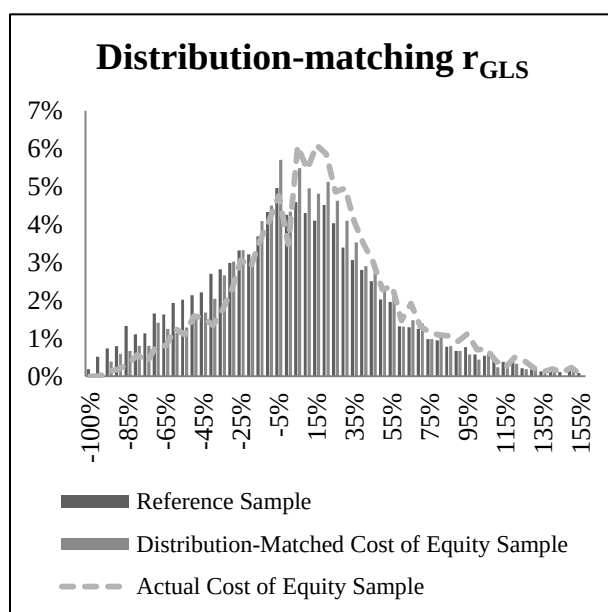
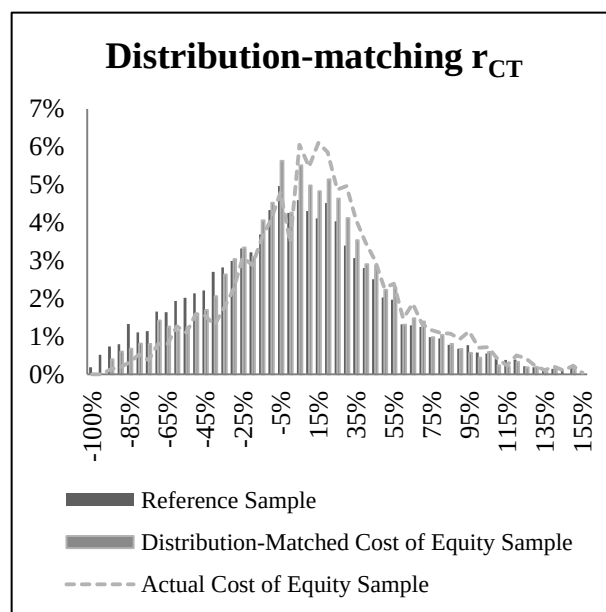
Appendix 5 – Metrics set 2: Descriptive statistics for realized returns, before and after resampling (r_{CT})

	Obs.	Mean	Std. Deviation	Skewness	Kurtosis
Reference sample	6952	0.0777	0.5699	1.5196	9.7484
Actual cost of equity sample	2000	0.2138	0.4913	1.4357	8.8108
Distribution-matched cost of equity sample	1717	0.1386	0.5118	1.6181	10.1367

The table numerically illustrates the effect of distribution-matching, by outlining specific statistics (mean, standard deviation, skewness and kurtosis) on realized returns for the reference sample, the actual cost of equity sample and the distribution-matched cost of equity sample. Statistics show that skewness, kurtosis and standard deviation is notably lower in the actual cost of equity sample compared to the reference sample. After the distribution-matching process it can be seen that the distribution-matched sample has features more aligned with the reference sample. This example is based on the r_{CT} measure.

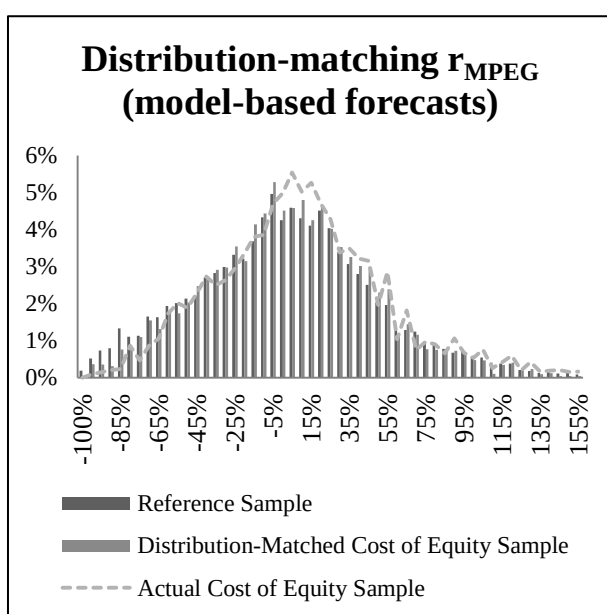
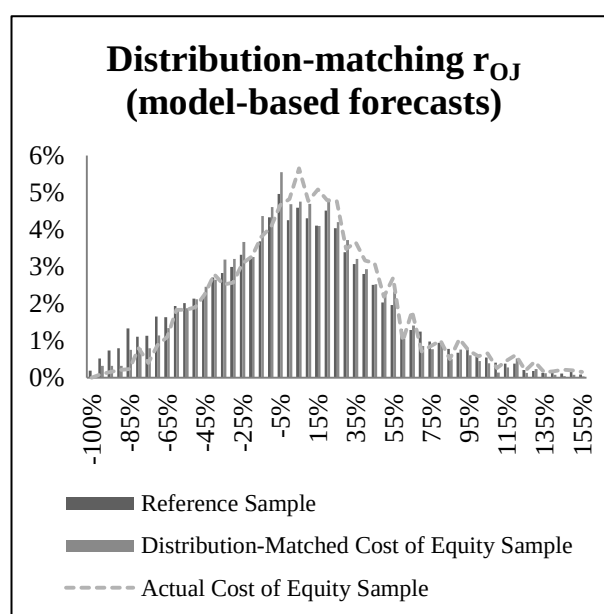
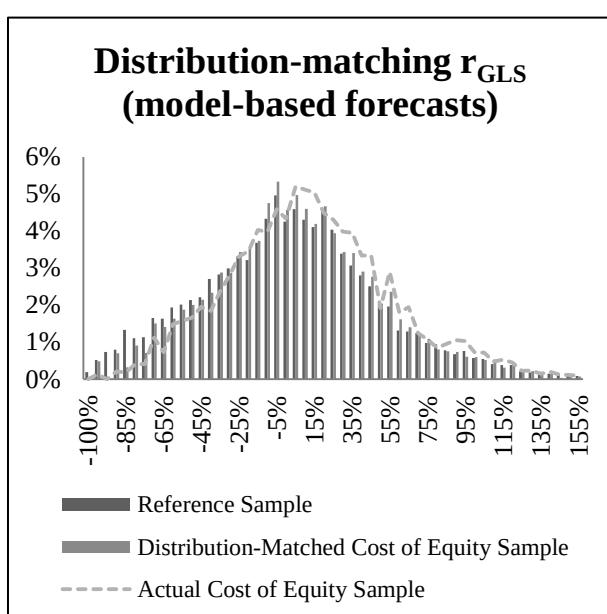
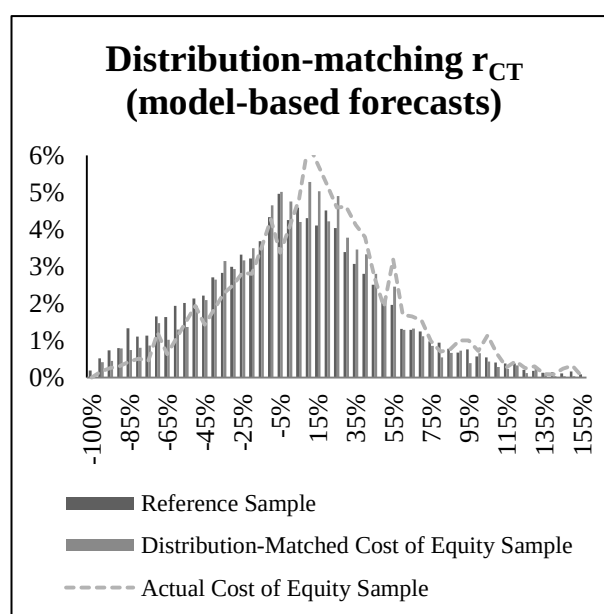
Appendix 6 – Metrics set 2: Overview of distribution-matching results for the cost of equity metrics (based on the benchmark case of unadjusted metrics)

The following figures illustrate the effects of distribution matching for r_{MPEG} , r_{OJ} , r_{GLS} and r_{CT} . Note that r_{GLS} is the same as the one presented in the results section. Each figure demonstrates three distributions; one of the reference sample, one of the actual cost of equity sample and one of the cost of equity sample after distribution-matching. Each distribution refers to observations of returns and show the proportion of observations that can be found in each return interval. In other words, each curve plots the density (empirical distribution) of realized returns for each sample. For the purpose of this graph, the data is truncated at +160%.



Appendix 7 – Metrics set 4: Overview of distribution-matching results for the cost of equity metrics (based on modelled earnings)

The following figures illustrate the effects of distribution-matching for the cost of equity metrics r_{MPEG} , r_{OJ} , r_{GLS} and r_{CT} ; all based on modelled earnings. Note that r_{GLS} is the same as the one presented in the results section. Each figure demonstrates three distributions; one of the reference sample, one of the actual cost of equity sample and one of the cost of equity sample after distribution-matching. Each distribution refers to observations of returns and show the proportion of observations that can be found in each return interval. In other words, each curve plots the density (empirical distribution) of realized returns for each sample. For the purpose of this graph, the data is truncated at +160%.



Appendix 8 – Metrics set 4: Descriptive statistics for cost of equity metrics after resampling observations based on the cross-sectional earnings forecasts

	r_{MPEG}	r_{OJN}	r_{GLS}	r_{CT}
Obs.	1896	1931	2047	1493
Mean	0.3761	0.3815	0.3082	0.3436
Std. Dev.	0.2427	0.2423	0.2327	0.2602
Skewness	0.6870	0.6976	1.1871	0.6819
Kurtosis	2.5503	2.5156	3.4279	2.4161
Min	0.0061	0.0194	0.0044	0.0053
P5	0.0647	0.0834	0.0769	0.0174
Q1	0.1825	0.1828	0.1355	0.1354
Median	0.3207	0.3227	0.2153	0.2626
Q3	0.5319	0.5376	0.4174	0.5380
P95	0.8541	0.8587	0.8244	0.8551
Max	0.9993	0.9986	0.9943	1.0000

The table presents descriptive statistics for Metrics set 4, containing our four cost of equity metrics after distribution-matching of the cost of equity observations computed based on the cross-sectional earnings forecasts rather than direct analyst forecasts. We show a selection of statistics, including number of observations, mean, standard deviation, skewness, kurtosis, max- min- and median values as well as quartiles for each metric. The mean values and standard deviation have remain fairly stable (+/- 0.5%) as compared to Metrics set 3.

Appendix 9 – Metrics set 4: Average distribution of realized returns, before and after distribution-matching of the cost of equity observations based on modelled earnings forecasts (r_{MPEG})

	Obs.	Mean	Std. Deviation	Skewness	Kurtosis
Reference sample	6952	0.0778	0.5699	1.5197	9.7484
Actual cost of equity sample	2108	0.1738	0.5474	1.7973	10.6522
Distribution-matched cost of equity sample	1896	0.1209	0.5463	1.6059	9.6192

The table numerically illustrates the effect of distribution-matching, by outlining specific statistics (mean, standard deviation, skewness and kurtosis) on realized returns for the reference sample, the actual cost of equity sample and the distribution-matched cost of equity sample. This table focuses on results from Metrics set 3, i.e. cost of equity metrics based on modelled earnings forecasts. Statistics show that skewness, kurtosis and standard deviation is notably higher in the actual cost of equity sample compared to the reference sample. After the distribution-matching process it can be seen that the distribution-matched sample has features more aligned with the reference sample. This example is based on the r_{MPEG} measure.

Appendix 10 – Metrics set 4: Average distribution of realized returns, before and after distribution-matching of the cost of equity observations based on modelled earnings forecasts (r_{OJN})

	Obs.	Mean	Std. Deviation	Skewness	Kurtosis
Reference sample	6952	0.0778	0.5699	1.5197	9.7484
Actual cost of equity sample	2159	0.1860	0.5409	1.7978	10.8024
Distribution-matched cost of equity sample	1931	0.1319	0.5409	1.5691	8.9106

The table numerically illustrates the effect of distribution-matching, by outlining specific statistics (mean, standard deviation, skewness and kurtosis) on realized returns for the reference sample, the actual cost of equity sample and the distribution-matched cost of equity sample. This table focuses on results from Metrics set 3, i.e. cost of equity metrics based on modelled earnings forecasts. Statistics show that skewness, kurtosis and standard deviation is notably higher in the actual cost of equity sample compared to the reference sample. After the distribution-matching process it can be seen that the distribution-matched sample has features more aligned with the reference sample. This example is based on the r_{OJN} measure.

Appendix 11 – Metrics set 4: Average distribution of realized returns, before and after distribution-matching of the cost of equity observations based on modelled earnings forecasts (r_{CT})

	Obs.	Mean	Std. Deviation	Skewness	Kurtosis
Reference sample	6952	0.0777	0.5699	1.5196	9.7484
Actual cost of equity sample	1753	0.2280	0.5465	1.4991	9.3618
Distribution-matched cost of equity sample	1493	0.1358	0.5690	1.7020	10.9094

The table numerically illustrates the effect of distribution-matching, by outlining specific statistics (mean, standard deviation, skewness and kurtosis) on realized returns for the reference sample, the actual cost of equity sample and the distribution-matched cost of equity sample. This table focuses on results from Metrics set 3, i.e. cost of equity metrics based on modelled earnings forecasts. Statistics show that skewness, kurtosis and standard deviation is notably lower in the actual cost of equity sample compared to the reference sample. After the distribution-matching process it can be seen that the distribution-matched sample has features more aligned with the reference sample. This example is based on the r_{CT} measure.

Appendix 12 – The Setterberg metric: descriptive statistics before and after resampling

	$r_{\text{SETTERBERG}}$	$r_{\text{SETTERBERG}}$ (resampled)
Obs.	3478	2769
Mean	0.1624	0.1682
Std. Dev.	0.1257	0.1327
Skewness	1.8659	2.0091
Kurtosis	9.0117	9.8616
Min	0.0000	0.0001
P5	0.0182	0.0188
Q1	0.0728	0.0746
Median	0.1373	0.1437
Q3	0.2152	0.2178
P95	0.3900	0.4012
Max	0.9788	0.9788

The table shows descriptive statistics for the Setterberg cost of equity metric, before and after resampling. We include number of observations, mean, standard deviation, skewness, kurtosis, min, max, median and selected percentiles. The number of observations has been reduced by 20%. The mean values and standard deviation have remained fairly stable.

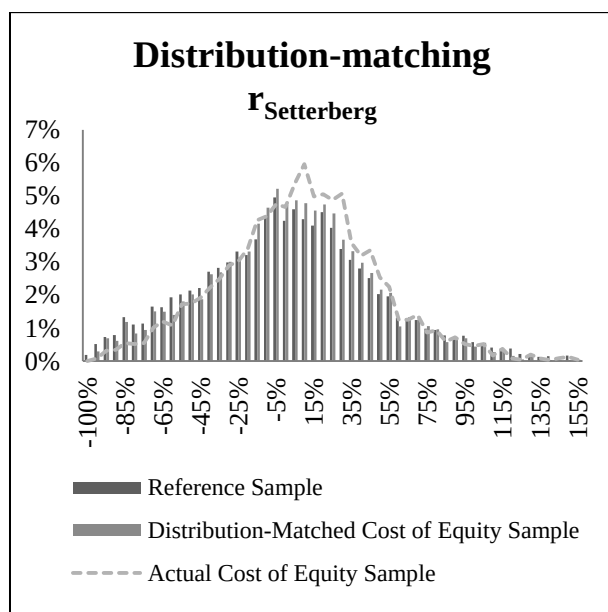
Appendix 13 – Setterberg metric: Descriptive statistics for realized returns, before and after resampling ($r_{\text{SETTERBERG}}$)

	Obs.	Mean	Std. Deviation	Skewness	Kurtosis
Reference sample	6952	0.077757	0.569939	1.51969	9.7484
Actual cost of equity sample	3103	0.0970	0.4621	1.0395	7.5858
Distribution-matched cost of equity sample	2800	0.0629	0.5287	1.3580	8.6380

The table numerically illustrates the effect of distribution-matching, by outlining specific statistics (mean, standard deviation, skewness and kurtosis) on realized returns for the reference sample, the actual cost of equity sample and the distribution-matched cost of equity sample. Statistics show that skewness, kurtosis and standard deviation is notably lower in the actual cost of equity sample compared to the reference sample. After the distribution-matching process it can be seen that the distribution-matched sample has features more aligned with the reference sample. This example is based on the $r_{\text{SETTERBERG}}$ measure.

Appendix 14 – The Setterberg metric: descriptive statistics before and after resampling

The following figure illustrates the effects of distribution-matching for $r_{\text{SETTERBERG}}$. Three distributions are shown; one of the reference sample, one of the actual cost of equity sample and one of the cost of equity sample after distribution-matching. Each distribution refers to observations of returns and show the proportion of observations that can be found in each return interval. In other words, each curve in the graph plots the density (empirical distribution) of realized returns for each sample. For the purpose of this graph, the data is truncated at +160%.



Appendix 15 – Metrics set 4: Association tests for cost of equity metrics based on absolute values

	r_{MPEG}	r_{OJN}	r_{GLS}	r_{CT}
Regression coefficient	-0.0587	-0.0749	0.1125	0.0762
Pearson's correlation coefficient	-0.0507	-0.0440	0.0672	0.0505
p-value	0.1950	0.0750	0.0080	0.1180
t-stat	-1.30	-1.78	2.67	1.57
Obs.	1896	1931	2047	1493

Association tests - Metrics set 4: The table provides a summary of association tests between realized returns and implied cost of equity *after* computation of cross-sectional earnings forecasts and distribution-matching. The difference between this table and others, is that tests are made on absolute values of returns, assuming that these are better proxies for expected returns. A significantly negative return would indicate a high-risk company, and from an investor perspective a hope for a high return as compensation.

Appendix 16 – Metrics set 4: Association tests for cost of equity metrics, excluding year 2001 and 2008

	r_{MPEG}	r_{OJN}	r_{GLS}	r_{CT}
Regression coefficient	-0.2373	-0.3180	0.1729	0.1540
Pearson's correlation coefficient	-0.1578	-0.1443	0.0774	0.0793
p-value	0.0000	0.0000	0.0040	0.0240
t-stat	-3.96	-5.65	2.91	2.27

Association tests - Metrics set 4: The table provides a summary of association tests between realized returns and implied cost of equity after inputting cross-sectional estimates of earnings forecasts and distribution-matching. The difference between this table and others is that the years 2001 and 2008 have been excluded, since these were deemed to have particularly negative impacts on overall association results. These years coincide with years of financial difficulties on the Swedish stock market. Since the years 2001 and 2008 were excluded after the resampling process, the outcome for other years have not been affected.