Exploring Multi-factor Models as a cross-check on allowed returns at PR24

Report prepared for Water UK

November 2022

Contents

1 Important notice

This Report (the Report) has been prepared by KPMG LLP ('KPMG', 'we' or 'our') for Water UK on the basis of an engagement contract dated 8 September 2022 between Water UK and KPMG (the "Engagement Contract"). Water UK commissioned the work regarding Ofwat's proposals relating to cost of equity cross-checks in the Draft Methodology for the PR24 price control. The agreed scope of work is included in Appendix 1 of this Report. Water UK should note that our findings do not constitute recommendations as to whether or not Water UK should proceed with any particular course of action.

This Report is for the benefit of Water UK only. It has not been designed to be of benefit to anyone except Water UK. In preparing this Report we have not taken into account the interests, needs or circumstances of anyone apart from Water UK, even though we may have been aware that others might read this Report. We have prepared this Report for the benefit of Water UK alone.

This Report is not suitable to be relied on by any party wishing to acquire rights against KPMG (other than Water UK) for any purpose or in any context. Any party other than Water UK that obtains access to this Report or a copy and chooses to rely on this Report (or any part of it) does so at its own risk. To the fullest extent permitted by law, KPMG does not assume any responsibility or liability in respect of our work or this Report to any party other than Water UK.

In particular, and without limiting the general statement above, since we have prepared this Report for the benefit of Water UK alone, this Report has not been prepared for the benefit of any other person or organisation who might have an interest in the matters discussed in this Report, including for example water companies or regulatory bodies.

Information in this Report is based upon publicly available information and reflects prevailing conditions as of the date of the Report, all of which are accordingly subject to change. Although we endeavour to provide accurate and timely information, there can be no guarantee that such information is accurate as of the date it is received or that it will continue to be accurate in the future. Information sources and source limitations are set out in the Report. We have satisfied ourselves, where possible, that the information presented in this Report is consistent with the information sources used, but we have not sought to establish the reliability or accuracy of the information sources by reference to other evidence. We relied upon and assumed without independent verification, the accuracy and completeness of information available from public and third-party sources. KPMG does not accept any responsibility for the underlying data used in this report.

The findings expressed in this Report are (subject to the foregoing) those of KPMG and do not necessarily align with those of Water UK.

For the avoidance of doubt, it is Water UK's sole responsibility to decide what should be included in their response or submission to Ofwat. KPMG has not made any decisions for Water UK or assumed any responsibility in respect of what Water UK decides, or has decided to, include in its response or submission.

2 Executive summary

On 7 July 2022 Ofwat published its Draft Methodology (DM) for the next water price control (PR24), which sets allowed revenues for the five-year period to 31 March 2030. In its DM Ofwat outlined its proposed approach to cross-checking the cost of equity (CoE) implied by its estimate based on the Capital Asset Pricing Model (CAPM).

Ofwat has proposed the use of Market-Asset Ratio (MAR) evidence to cross-check the CAPM-derived CoE. The use of cross-checks such as MARs to inform the allowed CoE aims to derive more robust and more precise estimates and to avoid solely relying on CAPM given its widely recognised shortcomings, for example, that it relies on a number of simplifying assumptions,[1] is an imprecise[2] model and has limited power to explain observed returns.

The identification of empirical shortcomings in the CAPM has led the academic community to test whether additional variables (business or asset characteristics)—the additional 'risk factors'—could improve the explanatory power of the CAPM, i.e. better fit the observed returns. This resulted in the genesis of multi-factor asset pricing models (MFMs),[3] which link the return on an asset based on its exposure to the market risk factor (which underpins the CAPM) *and* a set of additional systematic factors. These additional factors explain more closely higher required returns for some assets and lower for other assets. Empirical analysis has found that MFMs can have *stronger* explanatory power and empirical performance than the CAPM, i.e. better explain observed returns.

This Report was commissioned by Water UK to explore whether relevant financial literature, regulatory principles and empirical analysis could support the use of MFMs as an alternative, robust cross-check for setting allowed returns at PR24 and to estimate required returns for UK regulated utilities, water companies in particular, using MFMs.

The evaluation of MFMs in this Report as a potential cross-check is structured as follows:

- Introduction to MFMs and identification of primary factors and leading models in academic research as potential candidates for use as the cross-checks for setting allowed returns at PR24
- Consideration of the rationale for using MFMs as a cross-check
- Commentary on the theoretical and empirical foundations and relative performance of leading MFMs
- Empirical analysis and statistical testing of selected MFM(s) based on UK data
- Assessment of the potential implications of MFM evidence for allowed CoE at PR24

2.1 Introduction to MFMs, primary factors and leading models based on the latest academic research

The CAPM is the standard asset pricing model applied by practitioners and has been traditionally set out in the majority of finance textbooks for the valuation of securities. It is used by all UK regulators as the primary methodology for setting the allowed CoE for price controls, reflecting its simplicity, straightforward interpretation, and relative ease of use.

The primary limitation of the CAPM is that it relies solely on one risk factor and assumes that all systematic risk that is relevant for pricing can be captured by the market factor (i.e. the combination of the CAPM-beta and market risk premium). Insofar as this assumption does not hold, the CAPM may

1 UKRN (2022), Guidance for regulators on the methodology for setting the cost of capital – consultation, p11 and Ofwat (2021), PR24 and beyond: Discussion paper on risk and return, p20
2 CMA (2021), RIIO2 Final Determination, Volume 2A: Joined Grounds: Cost of equity, para 5.718
3 MFMs are based on the same core underlying principle as the CAPM, i.e. that there is a direct relationship between risk and required returns. MFMs can be perceived as effectively augmenting the CAPM with additional explanatory factors.

not price in all systematic risks which are relevant to explain observed returns. This is confirmed in practice: empirically, the CAPM does not explain the significant variation in observed returns, which means that there must be other factors[4] that are relevant for pricing of risk and for deriving the correct risk premium.

Since the early 1980s academic research explored empirical shortcomings of the CAPM and factors that could potentially add to the explanation of the cross-section[5] of returns provided by the market factor. Building on previous research, Fama and French (1992, 1993)[6] published seminal analysis which confirmed empirical contradictions in the central predictions of the CAPM. The results of their analysis questioned the usefulness and power of market betas and hence the CAPM in explaining observed returns.

Since Fama and French (1993), a large number of studies have tried to explain returns and, in particular, anomalies[7] in cross-sectional returns unexplained by CAPM and other models by applying various additional factors. Over time academic research has converged on small number of factors to augment the market factor to derive better asset pricing models.

There are now four additional factors (over and above the market factor as applied in the CAPM) which are widely used in the latest MFMs based on business characteristics that drive the overall risk premium:

- The size factor, which captures the risk premium associated with small stocks relative to big stocks;
- The investment factor captures the risk premium associated with stocks with a low level of investment in total assets relative to stocks with a high level of investment in total assets;
- The profitability factor, which captures the risk premium associated with stocks with high profitability relative to stocks with low profitability; and
- The value factor, which captures the risk premium associated with 'value' stocks (high book-to-market value) relative to 'growth' stocks (low book-to-market value).

These additional factors are used in the leading MFMs to explain observed returns, achieve the 'best fit' based on observed returns and build a model that estimates expected returns for a given asset more robustly.

These factors have been found to be empirically robust – for example in the US – to explain observed returns and have strong theoretical justifications, albeit there is no single universally accepted theory behind each factor.

As a result of the ongoing development and refinement of MFMs in academic research, Hou et al's q-factor model (2015)[8] and Fama and French's five-factor model (FF5F) (2015)[9] have emerged as two of the leading MFMs. The two papers, which established these models each have several thousand academic citations from subsequent work, which underscores their importance in academic research and hence their consideration as relevant models to derive cross-checks to CAPM in this Report.

Hou et al (2017)[10] compared the performance of several asset pricing models – including the CAPM, the Fama-French three-factor model (FF3F), the FF5F and the q-factor model[11] – in explaining stock returns observed on the New York Stock Exchange (NYSE). They found that the FF5F and the q-factor model have better empirical performance than the alternatives, with the q-factor model having

4 A variable or business characteristic with which asset returns are correlated.
5 I.e. across a sample of stocks over the same period of time.
6 Fama, E. F., & French, K. R. (1992). The cross-section of expected stock returns. the Journal of Finance, 47(2), 427-465; Fama, E. F., & French, K. R. (1993). Common risk factors in the returns on stocks and bonds. Journal of financial economics, 33(1), 3-56.
7 Asset pricing anomalies are patterns in average stock returns that are not explained asset pricing models.
8 Hou, K., Xue, C., & Zhang, L. (2015). Digesting anomalies: An investment approach. The Review of Financial Studies, 28(3), 650-705.
9 Fama, E. F., and K. R. French. "A Five-Factor Asset Pricing Model." Journal of Financial Economics,116 (2015), 1–22.
10 Hou, K., Xue, C., & Zhang, L. (2017). A comparison of new factor models (Working Paper No. 2015-03-05). Columbus, OH: Fisher College of Business.
11 Their analysis also covered Carhart's (1997) four-factor and Pastor and Stambaugh's (2003) models.

the best performance in explaining the momentum anomalies and the FF5F in the value-versus-growth anomalies.

Overall, there is now significant consensus on the set of factors to be included in MFMs. Following the development of these more complete factor models, they have become the primary sophisticated tool used to explain observed returns. This means that MFMs should represent a useful input into the cross-checks for PR24 given their potential to improve performance of the CAPM.

2.2 The rationale for using MFMs as a cross-check for PR24

A more granular, nuanced[12] view of risk provided by MFMs relative to the CAPM

CAPM and MFMs both have the same starting point, namely stocks' observed return. MFMs explain observed returns with reference to multiple explanatory factors that, by design, are expected to provide a more granular view of and better captures the systematic risk associated with individual stocks than a single factor model like the CAPM, which relies on a simplifying assumption that all risks relevant to pricing can be captured by the single market factor.

The table below sets out a comparison between CAPM and MFMs across a number of dimensions. Both models rely on the same basic methodology and theoretical underpinning and use risk-free rate, Total market returns and market beta in their calculations of expected return. The difference is that the MFMs include three or four additional factors.

Table 1: Comparison of MFMs to CAPM

	CAPM	MFMs
Methodology for estimating returns	The asset pricing formula is the same for both models[1] apart from the number of factors (explanatory variables) used	
Theoretical underpinning	Based on the same core underlying principle that there is a direct relationship between risk factors and required returns	
Data requirements	Returns data	Returns and accounting data
Risk free rate	Incorporated in both models (using index linked gilt)	
Total market return	Incorporated in both models (reflects the outturn total returns on a representative sample of UK equities)	
Market beta	Incorporated in both models (derived by regressing returns on market risk premium)[2]	
Other factor premia and beta[3,4]	None	Three or four others (derived by regressing returns market risk premium and three / four other factors)

Source: KPMG analysis
1 CAPM model: $R_{it} - R_{ft} = \beta_{Mkt,i}(R_{Mt} - R_{ft})$; q-factor model: $R_{it} - R_{ft} = \beta_{Mkt,i}(R_{Mt} - R_{ft}) + \beta_{size,i}Size_t + \beta_{I/A,i}I/A_t + \beta_{ROE,i}ROE_t$
2 The CAPM-market beta is calculated by regressing a stock's return on the market risk premium, whereas for the q-factor model the market beta is calculated concurrently with the three other factor betas by regressing a stock's return on the set of four factor premia.
3 Similar to the market risk premium that predicts that the market portfolio of equities has higher return than the risk-free bond due to its exposure to market risk, other factor premia explain the higher returns observed for stock portfolios based on size, profitability and investment factors.
4 Similar to the market beta that measures the sensitivity of a stock's return to market risk, the other factor betas measure the sensitivity of a stock's return to size, profitability and investment risk factors.

The empirical evidence indicates that market factor in the CAPM is on its own insufficient to explain observed returns for stocks, implying that the CAPM places too much weight on a single factor. This shortcoming of the CAPM is widely acknowledged in academia and by practitioners. For example, Goldman Sachs Investment Management notes that:

12 Using Multifactor Models (cfainstitute.org)

"…since the CAPM's introduction, ample empirical evidence ad theoretical evidence has surfaced to suggest the ***world is substantially more complex than single-factor models can allow****…* ***The premium associated with market risk is not the only dependable source of return as long-term returns also derive from a number of other global risks****"* and *"In these* [multi-factor] *models, every factor, such as size or value, reflects a distinct risk. Practitioners have also developed factor models for risk, which focus solely on explaining assets' volatility and co-movement."*

As both the CAPM and MFMs start from observed returns, the use of additional explanatory factors does not suggest stocks have *more* risk. The risk of a stock is directly reflected in its observed returns and is therefore unrelated to the choice of the model. The choice of the model (i.e. the type and the number of explanatory factors) only changes the power of the model to *explain* risk and hence estimate expected returns.

Empirical performance of MFMs relative to the CAPM

The leading MFMs in academic research are underpinned by a combination of economic theory and empirical research.

The guiding principle behind MFMs is to capture the drivers of systematic risk which explain observed returns. The additional factors included in MFMs have been established based on robust theoretical principles and justification, exploratory analysis of returns and empirical testing of the statistical significance of additional factors to explain risk. As a result, MFMs provide a more nuanced, granular view of risk, better fit the empirical returns data and have stronger explanatory power.

The latest MFMs have been proven to be statistically robust and to materially improve on the empirical performance of the CAPM[13] based on US data[14]. In general, the explanatory power of MFMs has improved over time as MFMs have developed.

Use of MFMs by academics and practitioners and in regulatory precedent

Academics have long used MFMs as the mainstream model for explaining observed returns which is recognised by both standard corporate finance textbooks and academic papers[15].

Practitioners are also increasingly moving away from and supplementing the CAPM with analysis based on MFMs. In particular, large assets managers, including those who have historically invested in regulated utilities, now use MFMs extensively to manage their portfolios[16].

A review of academic literature, corporate finance textbooks and practitioners' asset pricing methodologies indicates that MFMs are increasingly prevalent as asset pricing models to measure risk and improve on the empirical performance of the CAPM.

UKRN and the CMA have also recognised the stronger power of MFMs compared to the CAPM. Whilst MFMs have been considered in the past[17] by UK regulators (Ofwat, CAA, Ofgem, Ofcom) as a tool which could be used to estimate regulatory CoE, regulatory analysis of MFMs was predominantly concentrated in the early 2000s and has not been substantively revisited thereafter as MFMs have developed. In particular, MFMs have undergone a process of development and refinement which has been informed by a long series of academic studies by several authors over the 1980s-2000s. The leading MFMs which are now favoured in academia include a broadly common set of factors and are significantly more empirically and theoretically robust than the earlier MFMs.

---

13 For example, Fama and French noted in 2004 that "unfortunately, the empirical record of the model is poor – poor enough to invalidate the way it is used in applications. The CAPM's empirical problems may reflect theoretical failings, the result of many simplifying assumptions. But they may also be caused by difficulties in implementing valid tests of the model". Eugene F. Fama and Kenneth R. French (2004), The Capital Asset Pricing Model: Theory and Evidence

14 For example, Fama and French (1993, 1996, 2015), Hou et al (2015), Green, J., Hand, J. R., & Zhang, X. F. (2017). The characteristics that provide independent information about average US monthly stock returns. The Review of Financial Studies, 30(12), 4389-4436.

15 For example, McKinsey & Company, Valuation: Measuring and Managing the Value of Companies notes that "*given the strength of Fama and French's empirical results, the academic community now measures risk with a model commonly known as the Fama-French three-factor model*".

16 For example, Blackrock notes that "*Now – why do factors work? Extensive research, including that of Nobel prize winners, has proven that certain factors have driven returns for decades*." Separately, Goldman Sachs Asset Management manages a suite of portfolios which use MFMs to forecast returns on securities, for example, the CORE Equity Portfolio and ActiveBeta ETFs

17 For example, as part of PR04, PR09 in water, Q5 appeal in aviation, TPCR4 in energy.

2.3 Theoretical and empirical foundations and relative performance of leading MFMs

The q-factor model and the FF5F contain a common set of factors and have strong theoretical underpinnings and empirical performance, although there is evidence the former outperforms the latter in head-to-head tests.

Based on US data, both models significantly improve on the performance of both the CAPM and the FF3F (which has been the only MFM considered in UK regulation). This Report considers both models and carries out empirical analysis of each model based on UK data.

Overview of basis in asset pricing theory

Hou et al (2019, 2021)[18] apply an '*investment-based'* model to estimate returns, whereas Fama and French (2018)[19] adopt a *'consumption-based'* model. As highlighted by Zhang[20], the basic philosophy of investment-based models is to price risky assets from the perspective of their suppliers (firms), as opposed to their buyers (investors). Conversely, consumption-based models price risky assets from the perspective of their buyers.

Both models complement one another as together they explain observed returns from the perspectives of supply and demand for risky assets. *Factors and genesis*

Both the q-factor model and the FF5F arrive at a common set of factors to explain observed returns (market, size, investment and profitability)[21]. The FF5F additionally includes a value factor. This consensus has represented significant progress in the development of MFMs.[22]

For both models, the factors have combined theoretical and empirical foundations:

- In both cases, the investment and profitability factors (and the value factor) are justified based on the economic theories underpinning the respective models.
- In both cases, the market and size factors are incorporated in the models based on empirical evidence that these factors help to explain observed returns, however they have in the past been tied strongly to economic theories.

Theoretical foundation

The theoretical foundation for the q-factor model is the Net Present Value (NPV) rule of Corporate Finance and the FF5F is based on the Dividend Discount Model (DDM) of Valuation Theory. The NPV rule links a firm's real investment decisions to its expected return, whilst DDM links the market value of a firm's equity to its expected return.

Ultimately, both models have similar theoretical underpinnings for the investment and profitability factors (and value factor). Although at face value the theoretical underpinnings for these factors may appear somewhat different, Zhang notes in practice these are "*(virtually) identical"*[23] *and* the basis for the q-factor model can be mathematically translated into DDM which underpins the FF5F[24][25].

18 Hou, K., Mo, H., Xue, C., & Zhang, L. (2019). Which factors?. Review of Finance, 23(1), 1-35; Hou, K., Mo, H., Xue, C., & Zhang, L. (2021). An augmented q-factor model with expected growth. Review of Finance, 25(1), 1-41.
19 Fama, E. F., & French, K. R. (2018). Choosing factors. Journal of financial economics, 128(2), 234-252.
20 Zhang, Lu, q-Factors and Investment CAPM (December 3, 2019). Fisher College of Business Working Paper No. 2019-03-030, Charles A. Dice Working Paper No. 2019-30.
21 However, the factors are in all cases constructed differently as explained in Appendix 4.
22 Cochrane, J. H. (2011). Presidential address: Discount rates. The Journal of finance, 66(4), 1047-1108, p.16
23 Zhang Lu (December 2016), Factor Wars, Tsinghua Financial Review 37, 101-104.
24 Specifically the Gordon Growth Model, refer to section 1.2.2 of Hou et al (2015)
25 Some academics consider that both FF5F and q-factor model can ultimately be seen as derived from the DDM.

Empirical foundation

Both models include the market factor based on the same rationale, i.e. that market factor is included to capture common variation in returns over time (time-series variation) while accounting for variation in returns across stocks (cross-sectional variation) with other factors.

Hou et al include the size factor as size is a popular investment style for mutual funds. Hou et al note that while the size factor helps the q-factor model fit the observed returns across size deciles, its incremental effect in capturing known anomalies in returns is minimal. As a result, the authors conclude that the size factor plays only a secondary role in the q-factor model, whereas the investment and profitability factors are more prominent.

Fama and French recognise that whilst the size factor is not explicitly linked to their extension of DDM, it is observed to help forecast returns based on empirical evidence. They consider the size factor must therefore implicitly improve the predictive power of profitability and investment or explain variation in observed returns over different holding periods[26].

A comparison of the q-factor model and the FF5F across the dimensions above is set out in Table 2.

Table 2: Comparison of the q-factor model and the FF5F

	q-factor	FF5F
Basis in asset pricing	Investment-based	Consumption-based
Theoretical foundation	NPV rule of Corporate Finance	Dividend Discount Model of Valuation Theory
Genesis	Combination of theoretical and empirical[27]	
Number of factors	4	5
Market factor	✓	
Size factor	✓	
Value factor	×	✓
Profitability factor	✓	
Investment factor	✓	

Source: KPMG analysis

Whilst both the q-factor model and the FF5F are leading MFMs in the academic literature, there is evidence that the former is a stronger and more robust model relative to the latter.

The q-factor model outperforms the FF5F on statistical tests

The stronger empirical performance of the q-factor model relative to FF5F (and other models such as Carhart's four-factor model and Fama-French six-factor model) has been demonstrated by the authors of the q-factor model (Hou et al 2017, 2019) as well as several independent comparisons of the q-factor and FF5F models (Barillas and Shanken (2015), Stambaugh and Yuan (2016), Green et al (2017)[28]).

26 Fama and French focus on short-term returns (specifically, one-month returns) in their study. They consider that expected returns for a given stock could differ under different holding periods i.e., for the same stock, short-term returns could differ from long-term returns.
27 The starting point for investment and profitability (and value) factors is finance theory, whereas for the size and market factors the starting point was empirical.
28 Green, J., Hand, J. R., & Zhang, X. F. (2017). The characteristics that provide independent information about average US monthly stock returns. The Review of Financial Studies, 30(12), 4389-4436.

The value factor in the FF5F is redundant and without the value factor, the FF5F reduces to a 'noisy' variant of the q-factor model

Zhang (2017) shows that the value factor included in the FF5F is redundant because it can be seen as another manifestation of the investment factor based on economic theory. Moreover, Fama and French (2015) recognise that empirically the value factor is mostly absorbed by other FF5F factors. Zhang (2017) considers that without the value factor, the FF5F reduces to a 'noisy' variant of the q-factor model.

The q-factor model has fewer factors than, and is therefore preferable to, the FF5F

The q-factor model has fewer factors than the FF5F. This reduces the scope for distortive cross-correlations amongst the factors and means the model is simpler to adopt and apply in practice.

2.4 Empirical analysis and statistical testing of chosen MFM(s) using UK data

The Report calibrates and evaluates the performance of the q-factor and FF5F models relative to the CAPM and one another based on UK data[29].

The calibration of MFMs requires the construction of each of the factors included in the model. This requires the estimation of factor loadings and factor premia. Factor loadings are regression coefficients which represent the sensitivity of a stock's return to each risk factor included in the model (e.g. beta), whereas factor premia represent the additional return that is expected for taking on the associated risk (e.g. market risk premium).

The assumptions and methodologies for calibrating the models and evaluating results are consistent with those adopted by Ofwat and in academic literature. In particular, the construction of factors follows the approaches used by Fama and French and Hou et al. The selection of comparators, estimation and averaging windows and other assumptions required to derive the CoE are aligned to Ofwat's PR24 DM.

In order to assess whether MFMs warrant inclusion in the cross-checks for PR24 and what weight they might merit, the Report tests the performance of the models for (1) all companies used to calibrate the model and (2) for regulated utility stocks. In combination these tests cover both the factor premia and factor loadings used to derive the CoE estimates under the CAPM and MFMs. The tests find that the q-factor model has better empirical performance than the CAPM based on UK data and warrants inclusion as a cross-check for PR24. By contrast, the FF5F does not have stronger explanatory power than the CAPM and has not been considered further in this Report.

Statistical tests for q-factor and FF5F models

A two-stage statistical testing is undertaken to evaluate the empirical performance of the models for all companies used to calibrate the model. The statistical tests deployed – the factor spanning[30] and Gibbons-Ross-Shanken (GRS)[31] tests – are the standard tests applied in the academic literature to assess the statistical robustness of asset pricing models. Both of these tests allow for the assessment of the performance of different models on a relative basis.

The factor spanning test is the most important test that MFMs need to pass. The test examines whether additional risk factors add to the explanation of observed returns provided by an existing model. The test is binary; either additional factors add to the explanation of returns (a pass), or they

29 I.e. financial and accounting data for the largest 250 firms on London Stock Exchange excluding financials and real estate for each year during 1981 - 2022. The list of firms includes companies that have since de-listed or have gone bankrupt to avoid survivorship bias. Survivorship bias results from the use of a dataset that consists of survivors over a period, not the full set of companies that were listed. As the characteristics of survivors are likely to differ systematically from those who have delisted, the results will be biased.

30 Factor spanning regressions are a means to test if an explanatory factor can be explained by a combination of other explanatory factors. Spanning tests are performed by regressing returns of one factor against the returns of all other factors and analysing the intercepts from that regression.

31 Gibbons, M. R., Ross, S. A., & Shanken, J. (1989). A test of the efficiency of a given portfolio. Econometrica: Journal of the Econometric Society, 1121-1152.

do not (a fail). As such, the spanning test directly assesses whether one model is superior to (or subsumes) another in explaining observed returns. In consequence the factor spanning test is applied as the first-stage test and the GRS test as the second stage, with only the model(s) that pass the first stage taken forward to the second.

The results of the spanning test revealed (1) the q-factor model subsumes the CAPM (and the FF5F) (passes); and (2) the FF5F does not subsume the CAPM (fails). As such, only the q-factor model proceeds to the next stage of statistical testing, the GRS test. The FF5F failure of the spanning test could be explained by the following:

- The value factor in the FF5F appears to be redundant and inclusion of this factor may add 'noise'. Zhang (2017) shows that the value factor can be seen as another manifestation of the investment factor which could add 'noise' to the model.
- The profitability factor in the FF5F may capture a hidden investment effect[32] which could result in cross-correlations amongst factors and dampen the model's explanatory power.
- The profitability factor in the FF5F uses a different definition of profit to the well-established Peasnell model and thus may have weaker explanatory power for observed returns.

The GRS test[33] is widely adopted in testing MFMs.[34] It indicates whether an asset pricing model could explain the observed returns of all the tested portfolios. The test regresses the portfolio returns on factor premia for each portfolio separately. If the intercept terms of all the tested portfolios are jointly indistinguishable from zero, the model passes GRS test. The test is binary in the sense that a model could either pass or fail the test, but to assess the performance of different models on a relative basis, the next question would be how much of the variation in observed returns could be explained by the model. Here, the adjusted R^2[35] is a useful indicator.

Both the CAPM and q-factor fail the GRS test, which indicates neither model could describe the observed returns of all the tested portfolios. In particular, the portfolios that failed the test are mainly small-size portfolios. This has relatively limited impact on the usefulness of these models for PR24 as listed regulated utilities are large companies. After excluding the ten small-size portfolios, both the CAPM and q-factor pass GRS test.

Whilst both models pass the test after excluding small portfolios, the Report also considers the relative performance of the two models, measured by adjusted R^2. For all 25 portfolios, the q-factor model has consistently higher explanatory power than the CAPM. On average, the adjusted R^2 of the q-factor model is 64.5%, meaning that 64.5% variability of portfolio returns could be explained by the q-factor model, while for the CAPM, the adjusted R^2 is 45.2%.

In summary, the q-factor model based on UK data passes the spanning and performs better on GRS tests – as evidenced by the materially adjusted R^2 across all test portfolios relative to CAPM – and so is empirically robust in addition to having strong theoretical underpinning.

Overall the Report tests the performance on q-factor using an approach that is consistent with that applied in academic research where MFMs are evaluated as potential replacements for the CAPM as the primary methodology for estimation of returns. As a result, the bar applied in this Report to MFM evidence as a potential cross-check is significantly greater than the MAR cross-check included in the PR24 DM. All else equal this suggests that MFM evidence should be considered to be a primary cross-check. This implies in cross-checking the CAPM that weight should be attached to evidence implied by the q-factor model over and above other cross-checks.

---

32 This is because the FF5F measure of profitability divides in-year profit by contemporaneous book equity which, relative to the q-factor approach for calculating this value, incorporates an extra measure of the investment factor (difference between contemporaneous assets and one-year lagged assets).
33 Gibbons, M. R., Ross, S. A., & Shanken, J. (1989). A test of the efficiency of a given portfolio. Econometrica: Journal of the Econometric Society, 1121-1152.
34 For example, Gregory et al (2013), Hou et al (2015, 2017, 2019) and Fama and French (2015, 2017)
35 Adjusted R2 is a corrected goodness-of-fit (model accuracy) measure for linear models. It identifies the percentage of variance in the target field that is explained by the input or inputs.

2.5 Implications of MFM evidence for allowed CoE at PR24

To assess the impact and implications of the evidence from the q-factor model for the allowed CoE at PR24, the Report compares CoE estimates derived using the CAPM and the q-factor model as at each date for the two pure play water comparators (United Utilities and Severn Trent) across all estimation windows. The resulting differentials between the two models are set out in the table below.

A positive differential means that the CoE derived using the q-factor model exceeds that derived using the CAPM whereas a negative differential means that CAPM-derived CoE exceeds that derived using the q-factor model. The results set out in the table below show that the differentials across all estimation windows and both cut-off dates are positive These point to a conclusion that the proxies proposed by Ofwat for the UK water sector, have a higher systematic risk exposure than that implied by the CAPM. All else equal, this implies that where the CAPM is used to set allowed CoE in a price control setting, the model is likely to under-estimate systematic risk exposure, and this should be taken into account in setting the point estimate for CoE and cross-checking returns implied by CAPM.

Table 3: CoE differentials between estimates derived using the CAPM and q-factor

Cut-off date	Estimation window	Water portfolio
March 31, 2022	10-year	0.47%
	5-year	0.39%
	2-year	0.52%
February 28, 2020	10-year	1.73%
	5-year	2.20%
	2-year	2.96%

Source: KPMG analysis
Note: The estimates presented for each estimation window represent the spot rate.

Empirical analysis based on UK data finds that the CoE derived using the q-factor model is 0.39 – 2.96% higher than that derived using the CAPM.

The evidence considered in this Report imply that in selecting a point estimate for CoE based on the CAPM that weight should be attached to evidence implied by the q-factor model over and above other cross-checks. In practice MFM evidence indicates that the point estimate for the allowed CoE for PR24 should be 0.39 – 2.96% higher than the mid-point of the CAPM-derived CoE range. There is material variance between the differentials across the two cut-off dates. This is primarily driven by the structural break associated with Covid, which has resulted in a marked ‘flight to safety’ effect. Excluding data from the Covid period the implied differential is 1.73% - 2.96%.

The variance in returns implied by the two models can be viewed in the context of the extensive academic research which explored empirical shortcomings and contradictions of the CAPM, which has limited power to explain observed returns (which ultimately led to the genesis of MFMs). The q-factor model has been shown to have stronger empirical performance than CAPM based on UK data, and the variances set out in the table above should be considered in this context.

2.6 Comparison of MFMs to alternative cross-checks

The cross-checks adopted for PR19 (broker forecasts and MAR) and proposed for PR24 (MAR only) have weaker explanatory power than the CAPM and cannot improve upon the CAPM-derived CoE. In contrast MFMs are empirically more robust based on US data, are adopted by academics as a primary methodology for estimation of returns and are supported by economic theory. All else equal this suggests that MFMs could act as a *primary* cross-check on CAPM-implied returns.

If there is a risk that CAPM could over- or under-state returns, there is a requirement for a robust model to sense-check CAPM-derived returns and alternative models such as MFMs can help to select for example a point estimate within a range. It is possible (given that MFMs can improve on the

explanatory power of CAPM) that MFM models could ultimately replace CAPM as the primary methodology for setting returns, however this would represent a material change to the regulatory approach to returns estimation and is not proposed in this Report.

The table below sets out an assessment of MFMs as a cross-check against robust criteria (the basis for each criterion is set out in Appendix 7). The same assessment is carried out for the MAR cross-check which Ofwat proposes to apply at PR24 (based on the DM) on a basis consistent with KPMG's *Use of market-to-asset ratios (MARs) as a cross-check in the context of regulatory price controls* Report.

Table 4: A high level assessment of potential cross-checks against criteria

Criterion	MFMs	MAR
Transparent		
Targeted		
Objective		
Incentive		
Consistent		

Source: KPMG analysis
Note: Green indicates that the cross-check meets the criterion well; Amber that it partially does so; and Red that it does not do so.

Overall, the assessment of the MFM asset pricing models against robust criteria indicates that estimates derived using the preferred MFM model offer robust estimates of returns and hence a good cross-check on the estimates of allowed returns. The MFM model:

- Is based on a methodology with stronger explanatory power than the primary methodology for estimation of returns (CAPM);
- Is based on a transparent methodology and can be grounded in established academic literature;
- Is a targeted cross-check on and unbiased estimator of CoE as the output of the MFM analysis is an estimate of CoE;
- Capture a more granular view of risk for water companies; and
- Is the standard approach adopted by academics for estimation of expected returns.

By comparison, MAR as the primary cross-check currently under consideration for PR24 does not represent a comparably robust cross-check. This is because:

- There are many unknowns in the determination of a company's value, and the calculated MAR cannot be solely attributed to a difference between investors' assumed return of equity and the allowed return. As a result, the MAR cross-check is not targeted.
- Academic literature and research are generally clear that MAR cannot be used to observe CoE without controlling for all other factors which influence companies' values. This is not possible in the case of MARs as the factors may not be quantifiable and controllable.
- The application of MARs by the regulator to revise the allowed CoE (whether up or down) could cause regulated companies to behave in ways which were not intended and not in the best interests of consumers.
- In case of transaction MAR, there are additional issues with transparency and objectivity of the cross-check. The input information to derive traded MARs is generally not publicly available (especially for private companies) and MARs are also by nature biased towards the winning bidder's aims, assumptions, and strategic advantages.

3 Context and scope

3.1 Context

On 7 July 2022 Ofwat published the Draft Methodology (DM) for the next water price control (PR24) which sets allowed revenues for the five-year period to 31 March 2030. In its DM Ofwat outlined its proposed approach to cross-checking the cost of equity (CoE) implied by its estimate based on the Capital Asset Pricing Model (CAPM).

This section provides an overview of Ofwat's proposed approach to cross-checking the CAPM-derived CoE as well as wider regulatory considerations around application of cross-checks. It also sets out the rationale for exploring other asset pricing models, such as multifactor models (MFMs) as an alternative approach to cross-checking allowed returns implied by the CAPM.

3.1.1 The role of cross-checks in setting allowed returns

The CAPM is the primary methodology adopted in utility regulation to estimate the CoE, however, it has widely accepted shortcomings:

- The CMA in its determination for RIIO GD&T2 noted that: *"CAPM is an imperfect and imprecise tool – but that it is broadly regarded as the best model on which to base an estimate of the cost of equity for a regulatory price control."*[36]
- Ofwat in its PR24 discussion paper on risk and return notes that: *"The CAPM is a relatively* simple *model, with inputs that are readily obtainable, and which is familiar to stakeholders. Nevertheless, the CAPM represents a simplification of the real world and comes with known limitations."*[37]
- The recent UKRN consultation on guidance for regulators on the methodology for setting the cost of capital notes that: *"Regulators should continue to use the capital asset pricing model (CAPM) as their primary approach for estimating the cost of equity"* while recognising *"CAPM is a relatively simple model…"* [38]

In this context, regulators including Ofwat are considering other sources of market-based evidence to cross-check the CAPM-derived CoE to determine an estimate that is appropriate 'in the round'. Regulators have positioned the role of cross-checks as follows:

- Ofwat in its PR24 DM states that: *"Our proposed implementation of the CAPM…is reliant on significantly backwards-looking data, particularly on TMR, where we propose to capture over 120 years of historical evidence. One implication of this approach may be an allowed return which is slow to adapt to changing market conditions. Because our objective is to set an allowed return aligned with investors' expectations over 2025-30, it is therefore important to cross-check our CAPM-derived estimates against estimates from alternative approaches underpinned by more recent and forward-looking data."*[39]
- Ofwat also notes that: *"For our point estimate we propose that we would ordinarily use the midpoint of this CAPM-derived plausible range. We consider there should be a high evidential bar for moving away from this central estimate, limited to evidence from our market cross-checks. We expect that any adjustment would be modest and would in any case lie within the endpoints of our CAPM-derived cost of equity stated range."*[40]
- Ofwat in its PR24 discussion paper also commented that: "*We propose to estimate the base return on equity with most weight placed on the capital asset pricing model (CAPM). Since this is likely to produce a range of potential values, we see merit in using appropriate cross-checks. For*

36 CMA (2021), RIIO2 Final Determination, Volume 2A: Joined Grounds: Cost of equity, para 5.718
37 Ofwat (2021), PR24 and beyond: Discussion paper on risk and return, p20
38 UKRN (2022), Guidance for regulators on the methodology for setting the cost of capital – consultation, p10-11
39 Ofwat (2022), PR24 Draft Methodology, Appendix 11 – Allowed return on capital, p24
40 Ibid, p25

example, we propose to look at evidence from market-to-asset ratios while recognising that, for any cross-check, a degree of judgement will be required" and *"…the CAPM represents a simplification of the real world and comes with known limitations. Therefore, in addition to using the CAPM, we propose to use cross-checks outside the CAPM framework to gain assurance over CAPM estimates."*[41]

- UKRN in its recent consultation notes that: "*Since the CAPM is just one model of expected returns, market benchmarks…provide a sense-check on the CAPM point estimate when such market data are available"* and *"as available cross-checks themselves may be uncertain and reliant on assumptions, there should be a high evidential bar to deviating from the mid-point of the* [CAPM] *cost of equity range".*[42]
- As part of the RIIO GD&T2 appeals the CMA considered that the role of cross-checks is to assess whether the CAPM-implied returns appear materially miscalibrated relative to market-based evidence[43]. In this context, the CMA commented that: *"the ultimate requirement should be to ensure that the overall cost of equity allowance is sufficient to attract investors and allow companies to finance their activities"* and *"market-based cross-checks can help with this process."*[44]
- The CMA also concluded that no individual cross-check proposed by Ofgem could be presumed to be effective in assessing whether an estimate of the CoE is correct as otherwise any such cross-check would be capable of *replacing* the CAPM as the primary methodology for estimating returns[45].
- The CMA in its PR19 re-determination noted that cross-checks of the point estimate for CoE – in particular financeability – are valuable given that CAPM could be used to derive a wide range of potential estimates for the CoE[46]. Further, the CMA considered that *"arguments for picking a point estimate higher than the midpoint include…to take into account a cross-check on market data and financeability ratios."*[47]

Overall recent regulatory determinations have emphasised the importance of cross- or sense-checking returns implied by the CAPM with reference to alternative market benchmarks, whilst not ultimately challenging the role of the CAPM as the primary methodology to underpin setting allowed returns. In this context the Report explores whether MFMs could provide robust, additional evidence to sense-check returns implied by the CAPM for water companies.

3.1.2 Ofwat's cross-checks for PR19 and PR24

At PR19 Ofwat applied cross-checks to the CAPM-derived CoE and has proposed to maintain this step for PR24 CoE estimation. Ofwat has generally considered that its cross-checks, in particular the Market-to-Asset Ratio (MAR), counter the risk that the cost of equity has been set too low[48].

At PR19 Ofwat incorporated two cross-checks, namely, the MAR and broker forecasts. At the PR19 appeal the CMA evaluated both cross-checks in detail and additionally considered financeability as a cross-check. The CMA concluded that Ofwat's cross-checks "*were insufficiently robust to change the choice of point estimate…*[49]*"* whilst *"financeability should be a valuable cross-check"*[50].

41 Ofwat (2021), PR24 and beyond: Discussion paper on risk and return, p3 and p20
42 UKRN (2022), Guidance for regulators on the methodology for setting the cost of capital – consultation, p23-24
43 CMA (2021), RIIO2 Final Determination, Volume 2A: Joined Grounds: Cost of equity, para 5.718
44 Ibid, para 5.723
45 Ibid, para 5.718
46 CMA (2021), PR19 Final Determination, para 9.1378
47 Ibid, para 9.1240
48 Ibid, para 9.1346
49 Ibid, para 12.48
50 Ibid, para 9.1383

MAR and broker forecasts at PR19 and RIIO2 appeals

The MAR cross-check featured prominently during both the recent PR19 and RIIO2 appeals. The CMA concluded in both appeals that: *"it is difficult to use MARs to accurately infer small adjustments to a CAPM-based estimate of the cost of equity".*[51] The CMA recognised the difficulties in correctly interpretating MAR data as a result of several factors:

- Traded premiums for listed stocks tend to be lower than premiums from transactions;
- Discounted deals are unlikely to make it to market; and
- There are a wide range of assumptions that go in to bid prices, there is also significant evidence in relation to the 'winner's curse' i.e. bid prices rely on assumptions which exceed underlying economic value[52].

In its PR19 re-determination the CMA remained *"cautious about using market prices to determine the point estimate for the cost of equity*[53]*"* and did not give MAR analysis significant weight in coming to a final view on the point estimate for CoE[54].

As in its assessment of the MAR, at PR19 the CMA considered that *"caution is warranted when interpreting broker forecasts of the cost of equity in relation to utility companies"*[55]. During the RIIO2 appeals, the CMA did not find the cross-check[56] to be "wrong".

Role of financeability in CMA PR19 decision

The CMA *"disagree*[d] *with Ofwat's submission that the need to maintain credit metrics can never be part of the WACC assessment"*[57]. In particular, the CMA considered the use of credit ratios provides a check on whether the CoE appears to be broadly consistent with the credit ratings assumed throughout the determination.

Cross-checks to apply at PR24 and importance of development of robust cross-checks

For PR24, Ofwat has only proposed to consider MAR evidence as a cross-check and not the broker forecasts. In doing so Ofwat appears to have reflected, in part, the CMA's views on cross-checks from the PR19 appeals, although it is not clear whether and on what basis Ofwat considers that direct inferences about CoE can be drawn from MAR[58]. Further, it has not included the financeability cross-check, noting that *"we do not see the financeability assessment as a test for whether an individual component of the price control package, such as the allowed return…is reasonable.*[59]*"*

As a result, an important question for PR24 is whether there are more robust, alternative cross-checks which could be deployed to refine the CAPM-derived CoE range or inform the selection of the point estimate for the CoE.

3.1.3 MFMs as a potential alternative cross-check to the CAPM-derived CoE

The CAPM is a single factor pricing model which relies only on the market factor to explain observed returns. The simplicity of the CAPM, straightforward interpretation and the ease of calculation made it widely popular among academics and practitioners. However, academic research has over time identified a number of empirical shortcomings in the CAPM to explain observed returns.

The logical next step was to test whether additional variables (business or asset characteristics)—the additional 'risk factors'—improve the explanatory power of the CAPM, i.e. better fit the market data.

---

51 CMA (2021), RIIO2 Final Determination, Volume 2A: Joined Grounds: Cost of equity, para 5.681. Also see CMA (2021), PR19 Final Determination, para 9.1358
52 Ibid, para 5.706
53 CMA (2021), PR19 Final Determination, para 9.1358
54 Ibid, para 9.1362
55 Ibid, para 9.1365
56 Investment manager TMR forecasts. Ofgem combined these TMR forecasts with estimates of other CAPM parameters to cross-check the CoE.
57 CMA (2021), PR19 Final Determination, para 9.1378
58 Ofwat (2022), PR24 Draft Methodology, Appendix 11 – Allowed return on capital, p25
59 Ofwat (2021), PR24 and beyond: Discussion paper on risk and return, p53

This led to the genesis of MFMs. MFMs have been used as the preferred asset pricing models in academia for almost thirty years and are being increasingly relied upon by practitioners.

MFMs link the return on an asset based on its exposure to the market risk factor (which underpins the CAPM) *and* a set of additional systematic factors. Empirical analysis has found that these additional factors can provide additional explanatory power relative to using only the market factor to explain observed returns, which is also supported strongly by economic theory. These additional factors explain more closely higher required returns for some assets and lower for other assets. As a result, these MFMs generally have *stronger* explanatory power and empirical performance than the CAPM, i.e. better explain observed returns.

MFMs are based on the same core underlying principle as the CAPM, i.e. that there is a direct relationship between risk and required returns. MFMs can be perceived as effectively *augmenting* and *extending* the CAPM with additional explanatory factors.

Based on the above, this Report explores whether MFMs with stronger explanatory power than the CAPM could potentially represent a robust cross-check for the CAPM-derived CoE for regulatory price setting.

3.2 Scope and structure of this Report

This Report was commissioned by Water UK to explore whether relevant financial literature, regulatory principles and empirical analysis could support the use of MFMs as an alternative, robust cross-check for setting allowed returns at PR24. This Report considers MFMs as potential cross-check in three steps:

- First, it introduces MFMs as an asset pricing tool, sets out their development and evolution, considers the rationale for these asset pricing models as cross-checks on the CAPM-derived CoE and discusses two of the leading MFMs in academic research as potential candidates for inclusion in the cross-checks for PR24 (Section 4).
- Second, it sets out the approach and methodology for empirical analysis based on two of the leading MFMs using UK data, covering data collection, model calibration and statistical robustness testing based on the approaches followed in academic research (Section 5); and
- Third, it sets out discussion and interpretation of the results from the MFM analysis and potential implications for the allowed CoE at PR24 (Section 6).

4 MFMs as a potential cross-check for CoE

This section introduces MFMs as an asset pricing tool, sets out their development and evolution, provides a brief overview of the relevant financial literature on MFMs, considers the rationale for MFMs as cross-checks on the CAPM-derived CoE and discusses leading MFMs based on the latest academic research.

4.1 Development and evolution of MFMs

The CAPM is a single factor asset pricing model (based on the market factor) and is a widely applied approach for estimating allowed returns in a regulatory setting. There is now substantial academic evidence which indicates that the CAPM has empirical shortcomings and does not have strong power to explain observed returns.

In response to the weak empirical performance of the CAPM, academic research has sought to identify and test additional factors which could potentially better explain observed returns. This is the genesis of MFMs. These models have represented the benchmark pricing model in academic literature for almost three decades.

MFMs are reflect a substantive body of academic research. There is also significant consensus on the factors to be included in these models which is the result of ongoing development and refinement that dates back to the 1980s.

The CAPM is the standard asset pricing model applied by practitioners and traditionally set out in the majority of finance textbooks for the valuation of securities. The CAPM assumes a linear relationship between risk and return and estimates the return required by investors for a particular stock based on its exposure to *systematic* risk. Systematic risk captures the volatility of a stock's return relative to the market as a whole and is measured by beta. The remuneration required for each unit of beta is represented by the excess return on the market portfolio, i.e. the market risk premium. The required return of a stock is derived as the product of the market risk premium and the stock's beta combined with the risk-free rate.

The CAPM has long shaped the way academics and practitioners estimate expected returns. It is used by all UK regulators as the primary methodology for setting the allowed CoE for price controls, reflecting its simplicity, straightforward interpretation and ease of use. The primary limitation of the CAPM is that it relies solely on one risk factor, effectively assuming that all the systematic risk that is relevant for pricing can be captured by the market factor (i.e. the combination of the CAPM-beta and market risk premium). Insofar as this assumption does not hold, empirically the CAPM may not price in all risks which are relevant to explain observed returns. The bias arises as a result of the omission from the model of other variables (i.e. factors) that could help explain observed returns (omitted variable bias[60]).

In this context the performance of the CAPM has been extensively challenged in various research papers which have revealed limitations in its explanatory power. For example as early as 1976, Stephen Ross developed the arbitrage pricing theory (APT)[61], which explains the returns of a stock based on its exposure to several systematic risk factors. Similar to the CAPM, APT assumes a linear relationship between risk and return. However, unlike the CAPM, which effectively collapses all

60 Omitted-variable bias occurs when a statistical model leaves out one or more relevant explanatory variables. The bias results in the model attributing the effect of the missing variables to those that were included.
61 Ross, S., 1976a. The arbitrage theory of capital asset pricing. Journal of Economic Theory 13, 341–60.

systematic risks into a single risk factor, APT does not specify the number and nature of the underlying risk factor(s).

During the 1980s[62] academic research explored empirical shortcomings of the CAPM and factors that could potentially add to the explanation of the cross-section[63] of average returns provided by the market factor. This included exploration of, for example, size, value, leverage, earnings / price ratios as potential explanatory factors. A discussion of factors used in leading MFMs is set out in Section 4.3.1.

Building on previous research, Fama and French (1992, 1993)[64] published seminal analysis which confirmed empirical contradictions of the central predictions of the CAPM. The results of their analysis questioned the usefulness and power of market betas in explaining observed returns. The authors found that in the long run, small stocks have generated higher returns than large stocks and value stocks have generated higher returns than growth stocks, albeit they contain more risk.

As a result, the authors introduced additional size and value factors, arguing that the combination of the market factor and the two additional factors could help to explain patterns observed in stock returns that were not explained by the CAPM. This model, Fama-French three-factor model, (FF3F) represented a benchmark asset pricing model for stock valuation for roughly three decades. Other academics have extended the FF3F, for example, Carhart (1997)[65] augmented the FF3F with a momentum factor to form a four-factor model. As discussed below the momentum factor continues to be subject to debate.

As discussed in Harvey, Liu, and Zhu (2016)[66], since Fama and French (1993), a large number of studies have tried to explain returns and, in particular, the anomalies[67] in cross-sectional returns unexplained by CAPM and other models by applying various factors. As a result of this research, a number of new factors emerged. However, over time academic research has converged on small number of factors to derive better asset pricing models on which the literature is now reaching an agreement. For example, the q-factor and Fama-French five-factor models discussed below include a common set of factors.

Hou et al (2012[68], 2015)[69] proposed a four-factor model, which includes market, size, investment, and profitability factors. This model is dubbed the *q-factor* model because Hou et al base their model on corporate investment decisions, for which the neoclassical *q theory*[70] offers a benchmark. The authors noted that the q-factor model goes a long way toward explaining many anomalies that the FF3F could not, such as anomalies related to earnings surprise, financial distress, investment and return on equity (ROE).

After Hou et al put forward the q-factor model[71], Fama and French (2015)[72] separately proposed a five-factor model (FF5F) where, besides market factor, the dominant factors contributing to observed returns were identified as size, value, profitability, and investment factors. The authors found that these factors combined had better explanatory power than the previous three-factor model.

Fama and French (2018)[73] augmented the FF5F with the momentum factor to improve its predictive performance (FF6F). The momentum factor was established by Jegadeesh and Titman (1993)[74] based on the observation that stocks which have performed well (poorly) in the past continue to

---

62 Fama, E. F., & French, K. R. (1992). The cross-section of expected stock returns. the Journal of Finance, 47(2), 427-465 providers an overview of the research during 1980s
63 I.e. across a sample of stocks over the same period of time.
64 Fama, E. F., & French, K. R. (1992). The cross-section of expected stock returns. the Journal of Finance, 47(2), 427-465; Fama, E. F., & French, K. R. (1993). Common risk factors in the returns on stocks and bonds. Journal of financial economics, 33(1), 3-56.
65 Carhart, Mark M. "On Persistence in Mutual Fund Performance." The Journal of Finance, vol. 52, no. 1, 1997, pp. 57–82
66 Harvey, C. R., Liu, Y., & Zhu, H. (2016). … and the cross-section of expected returns. The Review of Financial Studies, 29(1), 5-68.
67 An asset pricing anomaly is a statistically significant difference between the realized average returns associated with certain characteristics of securities, or on portfolios of securities formed on the basis of those characteristics, and the returns that are predicted by a particular asset pricing model.
68 An earlier version of the 2015 paper in the next footnote was published as a National Bureau of Economic Research working paper 18435
69 Hou, K., Xue, C., & Zhang, L. (2015). Digesting anomalies: An investment approach. The Review of Financial Studies, 28(3), 650-705.
70 Introduced by Brainard and Tobin (1968) and Tobin (1969)
71 According to Zhang, the q-factor model predates the FF5F by three to six years. Table 1, Factors War (2016), Lu Zhang (this is an English translation of an article published in Tsinghua Financial Review 37, 101-104, in Chinese).
72 Fama, E. F., and K. R. French. "A Five-Factor Asset Pricing Model." Journal of Financial Economics,116 (2015), 1–22.
73 Eugene F. Fama, Kenneth R. French, Choosing factors, Journal of Financial Economics, Volume 128, Issue 2, 2018, Pages 234-252
74 Jegadeesh, Narasimhan, and Sheridan Titman. "Returns to Buying Winners and Selling Losers: Implications for Stock Market Efficiency." The Journal of Finance, vol. 48, no. 1, 1993, pp. 65–91.

perform well (poorly). Barroso and Santa Clara (2015)[75] argue that the addition of the momentum factor in the FF6F to capture risk is controversial. In addition, Hou et al (2019)[76] find that the FF6F is subsumed by the q-factor model in head-to-head factor spanning tests[77] i.e. the FF6F does not provide any additional explanatory power relative to the q-factor model. For these reasons, and as there is limited theoretical justification for this factor – the momentum factor is often omitted in empirical studies of observed returns and is not considered further in this Report.

Overall, MFMs have developed significantly over time and there is now substantive consensus on the relevant set of factors to be included in MFMs. Following the development of the MFMs, they have become the primary sophisticated tool used by academics as well as certain practitioners to explain observed returns. This means that MFMs should represent a useful input into the cross-checks for PR24 given their potential to improve performance of the CAPM.

4.2 Rationale for using MFMs as a cross-check

MFMs represent the most logical cross-check for the CAPM and should in principle be added to Ofwat's cross-checks for PR24. This is because they provide a more granular view of risk than the CAPM, improve upon the empirical performance of the CAPM and are widely used in academia and by practitioners as the more robust asset pricing model.

The additional factors included in MFMs have been established based on robust theoretical principles, exploratory analysis of observed returns and empirical testing of their statistical significance.

The cross-checks adopted for PR19 and proposed for PR24 have weaker explanatory power than the CAPM and cannot improve upon the CAPM-derived CoE. In contrast, MFMs are a superior asset pricing model to the CAPM.

This section considers the rationale for adopting MFMs in a regulatory setting to cross-check allowed returns, focussing on:

- A more granular and nuanced[78] view of risk provided by MFMs relative to the CAPM
- Empirical performance of MFMs relative to the CAPM
- Use of MFMs by academics and practitioners
- Regulatory precedent on MFMs

A more granular and nuanced view of risk provided by MFMs relative to the CAPM

There are a number of overarching similarities between MFMs and the CAPM. The table below sets out a comparison between the CAPM and MFMs across a number of dimensions.

The CAPM and MFMs both have the same starting point, namely stocks' observed return and rely on the same basic methodology and theoretical underpinning and use risk free rate, total market returns and market beta in their calculations of expected return. The difference is that the MFMs typically include three or four additional factors to estimate the risk premium, namely, size, investment, profitability and value.

75 Barroso, P., & Santa-Clara, P. (2015). Momentum has its moments. Journal of Financial Economics, 116(1), 111-120.
76 Hou, K., Mo, H., Xue, C., & Zhang, L. (2019). Which factors?. Review of Finance, 23(1), 1-35.
77 Factor spanning regressions are a means to test if an explanatory factor can be explained by a combination of other explanatory factors. Spanning tests are performed by regressing returns of one factor against the returns of all other factors and analysing the intercepts from that regression.
78 Using Multifactor Models (cfainstitute.org)

Table 5: Comparison of MFMs to the CAPM

	CAPM	MFMs
Methodology for estimating returns	The asset pricing formula is the same for both models[1] apart from the number of factors (explanatory variables) used	
Theoretical underpinning	Based on the same core underlying principle as CAPM i.e. that there is a direct relationship between risk factors and required returns	
Data requirements	Returns data	Returns and accounting data
Risk free rate	Incorporated in both models (using index linked gilt)	
Total market return	Incorporated in both models (reflects the outturn total returns on a representative sample of UK equities)	
Market beta	Incorporated in both models (derived by regressing returns on market risk premium)[2]	
Other factor premia and beta[3,4]	None	Three or four others (derived by regressing returns market risk premium and three / four other factors)

Source: KPMG analysis
1 CAPM model: $R_{it} - R_{ft} = \beta_{Mkt,i}(R_{Mt-} R_{ft})$; q-factor model: $R_{it} - R_{ft} = \beta_{Mkt,i}(R_{Mt-} R_{ft}) + \beta_{size,i}Size_t + \beta_{I/A,i}I/A_t + \beta_{ROE,i}ROE_t$
2 The CAPM-market beta is calculated by regressing a stock's return on the market risk premium, whereas for the q-factor model the market beta is calculated concurrently with the three other factor betas by regressing a stock's return on the set of four factor premia.
3 Similar to the market risk premium that predicts that the market portfolio of equities has higher return than the risk-free bond due to its exposure to market risk, other factor premia explain the higher returns observed for stock portfolios based on size, profitability and investment factors.
4 Similar to the market beta that measures the sensitivity of a stock's return to market risk, the other factor betas measure the sensitivity of a stock's return to size, profitability and investment risk factors.

As a result MFMs can be seen as an extension of the CAPM with a number of key methodological assumptions, data requirements and approaches to parameters *the same* across both models. However, the key difference – namely the inclusion of additional factors in MFMs – by design, are expected to provide a more granular view of and better captures the risk associated with individual stocks than a simplified single factor model like the CAPM.

The empirical evidence – primarily based on data from the US – shows that the market factor in the CAPM is on its own insufficient to explain observed returns for stocks, implying the CAPM places too much weight on a single factor. This shortcoming of the CAPM is widely acknowledged, for example, Goldman Sachs Investment Management notes that:

"…since the CAPM's introduction, ample empirical evidence ad theoretical evidence has surfaced to suggest the world is substantially more complex than single-factor models can allow… The premium associated with market risk is not the only dependable source of return as long-term returns also derive from a number of other global risks" and *"In these* [multi-factor] *models, every factor, such as size or value, reflects a distinct risk. Practitioners have also developed factor models for risk, which focus solely on explaining assets' volatility and co-movement."*

As both the CAPM and MFMs start from observed returns, the use of additional explanatory factors does not suggest stocks have more risk. The risk of a stock is directly reflected in its observed returns and is therefore unrelated to the choice of the model. The choice of the model (i.e. the type and the number of explanatory factors) only changes the power to explain that risk.

The additional factors in MFMs, like the market factor, are systematic in nature as they relate to market-wide, non-diversifiable risks. More precisely, the additional factors are proxies for (directly unobservable) macroeconomic risk exposures. For example, in the case of the size factor, there are macroeconomic risks that affect stocks differently depending on their size.

The additional factors relate to common attributes across all stocks such as size. The underlying risk with the size factor stems from how a stock's reaction to macroeconomic risks is affected by size. Stocks that are small (relative to the average stock in the market) are expected to perform worse than large stocks in adverse macroeconomic conditions. Given that exposure to macroeconomic risks is common across all stocks, the interaction of such risks with size cannot be diversified away. The size

beta indicates the sensitivity of a stock's return to the size factor i.e., how closely it behaves like a small stock. The same principle applies to other factors.

By comparison, the underlying risk for the market factor is the market's reaction to macroeconomic risks. It is well understood that this underlying risk cannot be diversified away. The market beta indicates the sensitivity of a stock's return to the market factor i.e., how closely it follows the behaviour of the market portfolio.

Empirical performance of MFMs relative to the CAPM

The leading MFMs in academic research are underpinned by a combination of economic theory and empirical research. The relevant economic theories are discussed in detail in Section 4.3.2.

The guiding principle behind MFMs is to capture the drivers of systematic risk which explain observed returns. The additional factors included in MFMs have been established based on robust theoretical principles and justification, exploratory analysis of observed returns and empirical testing of their statistical significance. In combination, this and the fact that MFMs provide a more nuanced, granular view of risk mean that they better fit the empirical returns data and have stronger explanatory power.

The latest MFMs have been proven to be statistically robust and shown to materially improve upon the empirical performance of the CAPM[79] based on US data[80]. In general, the explanatory power of MFMs has improved over time as MFMs have developed.

Use of MFMs by academics and practitioners

Academics have long used MFMs as the mainstream model for explaining observed returns which is recognised by both standard corporate finance textbooks and academic papers:

- *"Given the strength of Fama and French's empirical results, the academic community now measures risk with a model commonly known as the Fama-French three-factor model."* (McKinsey & Company, Valuation: Measuring and Managing the Value of Companies)
- *"Fama and French (1993) spurred widespread use of three factors, motivated by violations of the single-factor CAPM related to firm size and value versus-growth measures."* (Stambaugh and Yuan, 2016)

In addition, practitioners are increasingly moving away from and supplementing CAPM with analysis based on MFMs. In particular, large asset managers, including those who have historically invested in regulated utilities, now use MFMs extensively to manage their portfolios. These asset managers have developed their own proprietary MFMs which are inspired by but are not precise reflections of the leading models in the academic literature.

This Report does not consider use of these proprietary MFMs as it is not clear how they have been specified and as they have not been widely tested, unlike those in the academic literature. However, the key inference from the popular use of such models in industry is that practitioners are no longer relying solely on the CAPM and instead on models which rely on multiple factors to price risk.

Examples of large asset managers which use and cite MFMs include:

- "*Now – why do factors work? Extensive research, including that of Nobel prize winners, has proven that certain factors have driven returns for decades*." (Blackrock)
- Goldman Sachs Asset Management manages a suite of portfolios which use MFMs to forecast returns on securities, for example, the CORE Equity Portfolio and ActiveBeta ETFs

79 For example, Fama and French commented in 2004 that "unfortunately, the empirical record of the model is poor – poor enough to invalidate the way it is used in applications. The CAPM's empirical problems may reflect theoretical failings, the result of many simplifying assumptions. But they may also be caused by difficulties in implementing valid tests of the model". Eugene F. Fama and Kenneth R. French (2004), The Capital Asset Pricing Model: Theory and Evidence

80 For example, Fama and French (1993, 1996, 2015), Hou et al (2015), Green, J., Hand, J. R., & Zhang, X. F. (2017). The characteristics that provide independent information about average US monthly stock returns. The Review of Financial Studies, 30(12), 4389-4436.

- "*Our stock selection is driven by our proprietary multifactor model… By combining these factors, the model generates an overall score for each company… which is a forecast of the investment's expected relative return*" (Hermes Investment Management)
- "*For over 40 years MSCI…has researched factors to determine their effects on long term equity performance. MSCI has developed Factor Indexes and Factor Models in consultation with the world's largest investors...*" (MSCI)
- *"We're measuring and monitoring our own portfolios' factors… The goal: avoiding unintentional factor biases—unintended tilts toward a factor—in which a portfolio's exposures are significantly greater (or less) than expected"* and *"Sharp moves within asset classes highlight the need for enhanced risk analysis and management through factor analysis"* (JP Morgan Asset Management)
- Similarly a standard textbook for investment and financial professionals notes that: *"Multifactor models have come to dominate investment practice, having demonstrated their value in helping asset managers and asset owners address practical tasks in measuring and controlling risk"* (CFA Institute, CFA Program Level II, Portfolio Management and Wealth Planning)

Finally, evidence from MFMs has been considered by regulators in different sectors and jurisdictions. For example, the Federal Reserve Bank of New York (which plays a leadership role in monetary policy, financial supervision and the payments system) in its analysis of banks' cost of capital comments that "***Empirically, there are a myriad of approaches for measuring the cost of capital.*** *In this section,* ***we investigate the robustness of our findings by estimating several alternative measures of the cost of capital.*** *Table 8 repeats the key difference-in-differences specifications for banks and Top banks for* ***alternative cost of capital measures including three-factor estimates from the Fama and French (1993) model, five-factor estimates that incorporate additional interest rate and term spread factors***".[81]

A review of academic literature, corporate finance textbooks and practitioners' asset pricing methodologies indicates that MFMs are increasingly prevalent as asset pricing model(s) to measure risk and improve on the empirical performance of the CAPM.

Regulatory precedent on MFMs

MFMs have been considered in the past[82] by UK regulators (Ofwat, CAA, Ofgem, Ofcom) as a tool which could be used to estimate regulatory CoE. However, regulatory analysis of MFMs was predominantly concentrated in the early 2000s and has not been substantively revisited thereafter as MFMs have developed.

These previous analyses focussed on the FF3F which was established around thirty years ago (1993). MFMs have moved on significantly since then. In particular, they have undergone a process of development and refinement which has been informed by a long series of academic studies by several authors over the 1980s-2000s. The leading MFMs which are now favoured in academia include a broadly common set of factors and are significantly more empirically and theoretically robust than the FF3F.

Regulatory analysis also largely focused on MFMs as potential replacements for the CAPM as the primary approach to set allowed CoE. By contrast, this Report assesses the merits of MFMs as a *cross-check*, to be used in conjunction with the CAPM to underpin regulatory revenue setting. This suggests that a different bar is appropriate for evaluating the role of MFMs than the one applied during earlier analyses. The application of a different bar to cross-check evidence relative to the primary methodology for estimating returns is consistent with the approach followed by Ofwat and the CMA. Notwithstanding this different bar, it is important to note that MFMs are the only cross-check which has been proven to *improve* the empirical performance of the CAPM. All else equal this signals that MFM evidence is more relevant and reliable for the purposes of cross-checking the CAPM-implied CoE than other cross-checks.

---

81 Federal Reserve Bank of New York, Evaluating Regulatory Reform: Banks' Cost of Capital and Lending (July 2020), p. 29
82 For example, as part of PR04, PR09 in water, Q5 appeal in aviation, TPCR4 in energy.

Sections 4.3, 5 and Appendix 6 consider the issues explored as part of past regulatory analyses of MFMs and comment on the extent to which they remain relevant based on the current academic research and empirical results of the analysis put forward in this Report.

More recently, UKRN and CMA have recognised the stronger power of MFMs compared to the CAPM:

- NATS CMA appeal: "*our understanding is that multi-factor models have been rejected for use by regulators not because they are wrong – academic evidence suggests they are better in explaining actual returns to investors. They have been rejected because they are hard to populate in practice. The most popular multi-factor model, the Fama-French model, has been considered by some regulators for use but it was concluded that it was not feasible to populate the model*"[83]
- 2018 Wright et al paper for UKRN: "*similarities between stocks with stocks with certain types of characteristics mean that multifactor models inevitably are able to explain greater proportion of the cross–section of realised returns better than CAPM can" and "multifactor models can provide a helpful cross-check on standard techniques for estimating CAPM betas*"[84]

Separately, in the ED2 Draft Determination Ofgem commented that the FF5F is a relatively recent addition to the literature (~2015) and noted that MFMs were explicitly considered in the UKRN 2018 report, which concluded that CAPM is the best available model, despite numerous caveats[85]. First, the UKRN report acknowledged that MFMs have better explanatory power and can serve as a useful cross-check on CAPM-implied returns. Second, consistent with the recommendations of UKRN, this Report explores whether MFM evidence could supplement rather than replace the CAPM as the primary methodology for setting allowed CoE. Third, although the FF5F and the q-factor model are relatively recent additions to the MFM literature, they have the best empirical performance and strong theoretical foundations as discussed in Section 4.3.2.

4.3 Leading MFMs in academic research

Since the early 1980s academic literature has progressively identified additional factors that, together with the market factor already captured within the CAPM, explain observed returns. There are now four additional factors which are widely used in MFMs (size, investment, profitability and value). This consensus has represented significant progress in the development of MFMs and addresses the criticism levied on MFMs as part of previous regulatory analysis that there was judgement involved in the selection of risk factors. The convergence to a common set of factors in academic literature implies that there is now less scope for judgment involved.

The q-factor model and the FF5F have emerged as two of the leading MFMs in academic research. Both models contain a common set of factors and have strong theoretical underpinnings and empirical performance based on US data, although there is evidence the former outperforms the latter in head-to-head tests.

The factors included in both of these models are empirically robust based on US data and have several compelling economic interpretations, in addition to the ones put forward by Hou et al and Fama and French. No interpretation has gained universal acceptance amongst academics.

Based on US data, both of these models significantly improve upon the performance of both the CAPM and the FF3F which has been the only MFM considered in UK regulation. As a result, this Report considers both models and carries out empirical analysis of both models using UK data.

This section explains each of the primary factors employed in the latest academic research and provides an overview of two of the leading MFMs in the latest academic research, in terms of their theoretical underpinning and empirical performance.

83 NATS (En Route) Plc/CAA Regulatory Appeal Final report, Appendix D: Technical note on betas and gearing, para 26.
84 Wright, S., Burns, P., Mason, R., & Pickford, D. (2018). Estimating the cost of capital for implementation of price controls by UK Regulators, p.29
85 Ofgem (2022), RIIO-ED2 Draft Determinations – Finance Annex, Appendix 3, Consultancy report E6

4.3.1 Intuitive explanation of each factor employed in the latest academic research

Since the early 1980s academic literature has progressively identified additional factors that, together with the market factor already captured within the CAPM, explain observed returns. There are now four additional factors which are widely used in the latest MFMs (size, investment, profitability and value), including the q-factor model and the FF5F. The FF5F includes market, size, investment, profitability and value factors whereas q-factor includes market, size, investment and profitability factors. All of these factors have strong theoretical justification, albeit there is no universally accepted theory behind each factor.

Each factor captures the risk premium between stocks with characteristics associated with higher risk/return and lower risk/return. For example, the size factor captures the risk premium associated with small stocks relative to big stocks, given small stocks are associated with higher risk/return. The concept of higher risk for small stocks is supported by economic theory whilst higher return for small stocks has been proven empirically.

This section sets out the basis for the economic interpretation of each factor employed in the latest academic research, i.e. market, size, investment, profitability and value factors.

Market factor

The market factor (or market excess return) captures the risk premium associated with a market index relative to the risk-free rate, as applied in the CAPM.

The market factor is based on widely accepted theory about risk and return which dates back to Markowitz's portfolio theory (1959)[86]. The implication of this theory is that stocks with higher market betas pay off less when markets decline, i.e., they have higher systematic risk. Investors thus require higher expected returns to hold these stocks.

Whilst Fama and French's earlier studies (1992) found anomalies when relying solely on the market factor to predict returns, their later analysis (1993) suggests the market factor is nonetheless required in MFMs to explain why stock returns are above the risk-free rate. Put another way, the market factor is retained in MFMs to capture common variation in returns over time whilst additional factors account for variation in returns across stocks. This is also consistent with Hou et al's analysis.

Size factor

The size factor captures the risk premium associated with small stocks relative to big stocks.

Fama and French (1993) find that small stocks had a prolonged earnings depression following the recession in 1980-1982 and did not participate in the boom of the middle and late 1980s. By contrast big stocks avoided this long earnings depression. This implies small stocks are more sensitive to macroeconomic conditions, and are therefore riskier, than big stocks.

Another popular explanation for the size premium, first investigated by Stoll and Whaley (1983)[87], is based on liquidity. Big stocks are generally *more* liquid and investors are willing to trade off returns for liquidity, in consequence the returns for big stocks are *lower* than small stocks.

Investment and profitability factors

The investment factor captures the risk premium associated with stocks with a low level of investment in total assets relative to stocks with a high level of investment in total assets.

The profitability factor captures the risk premium associated with stocks with high profitability relative to stocks with low profitability.

---

86 Markowitz, H(1959), Portfolio Selection: Efficient Diversification of Investments, Cowles Foundation Monograph #16 (Wiley, New York)
87 Stoll, H. R., & Whaley, R. E. (1983). Transaction costs and the small firm effect. Journal of Financial Economics, 12(1), 57-79.

Fama and French extend the Dividend Discount Model (DDM) to justify the inclusion of the investment and profitability factors[88]. This is explained in greater detail in Appendix 2. Fama and French's extension of DDM effectively states that the market value of equity is equal to *(profitability – investment) / discount rate*. For a given level of profitability, high investment stocks should earn lower returns than low investment stocks. Similarly, for a given level of investment, high profitability stocks should earn higher returns than low profitability stocks.

The investment CAPM forms the basis of the q-factor model (discussed in greater detail in Section 4.3.2) and rationalises the inclusion of the investment and profitability factors.

Intuitively, the investment CAPM represents the point at which the marginal benefit of investment equals the marginal cost of investment, for a firm. Formulaically, the investment CAPM says *expected profitability / discount rate* is equal to *investment costs*. Here, *expected profitability / discount rate* reflects the marginal benefit of investment and *investment costs* reflects the marginal cost of investment. Investment costs are an increasing function of investment.

The investment CAPM can be restated as: *discount rate = expected profitability / investment costs.* It is clear that, for a given level of expected profitability, high investment stocks (with high investment costs) should earn lower returns than low investment stocks. Similarly, for a given a level of investment, high expected profitability stocks should earn higher returns than low expected profitability stocks.

Value factor

The value factor captures the risk premium associated with 'value' stocks (high book-to-market value) relative to 'growth' stocks (low book-to-market value).

Zhang (2002)[89] observes that in recessions, value stocks are burdened with more unproductive capital and find it more difficult to disinvest than growth stocks. This is because (1) companies face higher costs in cutting than in expanding capital, i.e., costly reversibility[90], which prevents value stocks from disinvesting, and (2) discount rates are higher in recessions, which increases the hurdle rate for investment and in turn intensifies incentives for value stocks to disinvest. This propagates the effect of costly reversibility. These factors deprive value stocks of flexibility to disinvest, and in turn to smooth dividends, in recessions. As a result, assets in place are perceived to be riskier than growth options.

Fama and French (1993) suggest low book-to-market ratios are associated with growth i.e., persistently high earnings on book equity (or high returns on capital) that result in high stock prices. High book-to-market ratios are associated with financial distress i.e., persistently low earnings on book equity that result in low stock prices. This implies that value stocks will perform poorly in a recession and hence can be seen as riskier than growth stocks.

4.3.2 Overview of leading MFMs

In academic research, Hou et al's q-factor model (2015) and Fama and French's FF5F (2015) have emerged as two of the leading MFMs. The two papers which established these models have, more than 2000 and 6000 citations respectively, which confirms their relevance for the estimation of returns.

Both the q-factor model and the FF5F have been shown to have strong empirical performance based on US data. Hou et al (2017)[91] compared the performance of several asset pricing models – including the CAPM, FF3F, FF5F and the q-factor model[92] – in explaining stock returns observed on the New York Stock Exchange. They found that the FF5F and the q-factor model have better empirical

88 DDM states the market value of equity of a stock is the present value of its future dividend cashflows.
89 Zhang, Lu, The Value Premium (November 13, 2002). Simon School of Business Working Paper No. FR 02-19
90 Costly reversibility means that firms face higher costs in cutting than in expanding capital stocks.
91 Hou, K., Xue, C., & Zhang, L. (2017). A comparison of new factor models (Working Paper No. 2015-03-05). Columbus, OH: Fisher College of Business.
92 Their analysis also covered Carhart's (1997) four-factor and Pastor and Stambaugh's (2003) models.

performance than alternative models, with the q-factor model having the best performance in explaining the momentum anomalies and the FF5F in the value-versus-growth anomalies.

Hou et al (2019, 2021[93]) apply an '*investment-based'* model to estimate returns, whereas Fama and French (2018) adopt a *'consumption-based'* model. As highlighted by Zhang[94], the basic philosophy of investment-based models is to price risky assets from the perspective of their suppliers (firms), as opposed to their buyers (investors). Conversely, consumption-based models price risky assets from the perspective of their buyers.

Consumption-based models represent the traditional approach for estimating returns, for instance, the CAPM is an example of a consumption-based model[95]. Empirical studies have revealed that traditional consumption-based models are not able explain anomalies in observed returns[96]. As a result, academics such as Zhang[97] consider these anomalies are primarily due to the relationship between firms' characteristics and returns, which has led to the genesis of *investment-based* models.

Consumption-based models focus on the demand of risky assets (i.e., the investors) whereas investment-based models focus on the supply of risky assets (i.e., the firms). Both models complement one another as together they explain observed returns from the perspectives of supply and demand.

q-factor model

The q-factor model can be seen as an empirical implementation of the investment CAPM. The investment CAPM prices risky assets from the perspective of firms. It connects expected returns solely to firms' characteristics using an investment model.

The investment CAPM is underpinned by the neoclassical q theory of investment which is an important economic theory for explaining firms' investment behaviour. The investment CAPM builds on Cochrane's research (1991)[98] which applied a q theory-based investment model to study asset prices and found that a firm's investment returns should equal its stock returns under certain conditions.[99] The derivation of the investment CAPM includes expressing stock returns using firm-level accounting measures which is supported by accounting and valuation literature such as Peasnell (1982)[100].

The investment CAPM is effectively a restatement of the NPV rule in Corporate Finance[101]. The NPV rule says that a firm should invest in a given project only if its NPV[102] is greater than or equal to zero[103]. In the case of multiple projects, a firm should start with the project with the highest NPV and keep investing until the NPV of a new project is nil. Where the NPV of a new project is nil, this means its *expected profitability / discount rate* is equal to its *investment costs*. In other words, this project (i.e. the last project in which a firm should invest) represents the point at which the marginal benefit of investment is equal to the marginal cost of investment. The investment CAPM takes the NPV rule and translates it into asset pricing theory by rewriting the rule as: *discount rate = expected profitability / investment costs*.

The investment CAPM implies that investment and profitability factors are relevant for explaining observed returns. This finding is supported by a vast empirical literature in finance and accounting on the investment premium and profitability premium[104].

---

93 K., Mo, H., Xue, C., & Zhang, L. (2021). An augmented q-factor model with expected growth. Review of Finance, 25(1), 1-41.
94 Zhang, Lu, q-Factors and Investment CAPM (December 3, 2019). Fisher College of Business Working Paper No. 2019-03-030, Charles A. Dice Working Paper No. 2019-30.
95 For clarity whilst CAPM is an example of a consumption-based model, there is also (separately) a consumption-based CAPM.
96 Ibid.
97 Ibid.
98 Cochrane, J. H. (1991). Production-based asset pricing and the link between stock returns and economic fluctuations. Journal of Finance, 46, 209-237.
99 Restoy and Rockinger (1994) prove that the equality between the investment and stock returns holds under more general conditions.
100 Peasnell, K. V. (1982). Some formal connections between economic values and yields and accounting numbers. Journal of Business Finance & Accounting, 9(3), 361-381.
101 Zhang, Lu, q-Factors and Investment CAPM (December 3, 2019). Fisher College of Business Working Paper No. 2019-03-030, Charles A. Dice Working Paper No. 2019-30.
102 I.e. the discounted value of its future cash flows net of investment costs today.
103 I.e. where the benefit of investment is greater than or equal to the cost of investment.
104 Zhang, L. (2017). The investment CAPM. European Financial Management, 23(4), 545-603.

Hou et al add size and market factors to the investment and profitability factors implied by the investment CAPM to form the q-factor model. The rationale for the inclusion of these factors is as follows:

- The size factor is included as size is a popular investment style for mutual funds. Hou et al note that while the size factor helps the q-factor model fit the observed returns across size deciles, its incremental effect in capturing known anomalies in returns is minimal. As a result, the authors conclude that the size factor plays only a secondary role in the q-factor model, whereas the investment and profitability factors are more prominent.
- The market factor is included to capture common variation in returns over time (time-series variation) while accounting for variation in returns across stocks (cross-sectional variation) with the q-factors (size, investment and profitability factors).

The value factor is not included in the q-factor model:

(1) Zhang (2017) mathematically restates the investment CAPM as investment costs = market value of equity / book value of equity. It follows that high investment stocks (with high investment costs) tend to have high valuations (high market-to-book value or low book-to-market value).
(2) High valuations are a characteristic of growth stocks which are associated with lower returns.
(3) High investment stocks are associated with lower returns.
(4) Together (1)-(3) imply that growth stocks with high valuations invest more[105] and thus earn lower returns. This is consistent with the direction of the risk premia for both the value and investment factors.

As such, Zhang considers the value factor duplicates the investment factor and could distort results where the investment factor is included.

FF5F

The FF5F is based on the fundamentals of DDM which has widely accepted theoretical justification based on Williams (1938)[106] and is widely used by market practitioners. Regulators such as Ofwat and Ofgem have considered DDM analysis to inform their calibration of total market returns[107].

Fama and French build on DDM in two ways (explained in greater detail in Appendix 2):

- First, they express the cashflow-based model in accounting terms by re-writing dividend cashflows as earnings less investment[108]. This is supported by a vast literature on accounting and valuation, such as Peasnell (1982). Peasnell, for example, shows the economic value of a stock can be derived by discounting accounting measures of profit.
- Second, they divide through the DDM equation by book value of equity.

For Fama and French, their extension of DDM implies there are three 'natural' components for estimating returns, namely, the value, profitability, and investment factors. They form the five-factor model by combining these three factors with the size and market factors:

- Fama and French recognise that whilst the size factor is not explicitly linked to their extension of DDM, it is observed to help forecast returns based on empirical evidence. They consider the size factor must therefore implicitly improve the predictive power of profitability and investment or explain variation in returns over different holding periods[109].

---

105 Intuitively, growth stocks with high valuations should invest more (i.e., be high investment stocks) in line with q theory. Q theory says firms with high valuations have, under constant returns to scale of investment , high ratios of market value of investment / replacement cost of investment. Accordingly, firms with high valuations will invest more given that the value of investment is worth more than it costs to replace. This reinforces the overlap between the investment and value factor
106 Williams, J.B. (1938) The Theory of Investment Value. Harvard University Press, Cambridge, MA
107 Ofgem (2018), RIIO2 Sector Specific Methodology Annex: Finance
108 Based on "clean surplus" accounting. In "clean surplus" accounting all valuation changes in book value (e.g. depreciation and revaluation) must flow through the P&L account.
109 Fama and French focus on short-term returns (specifically, one-month returns) in their study. They consider that expected returns for a given stock could differ under different holding periods i.e., for the same stock, short-term returns could differ from long-term returns.

- Fama and French previously noted in their paper for the FF3F that the market factor is required to capture common variation in returns whilst additional factors are left to explain differences in returns across stocks. This is consistent with Hou et al's view.

Comparison of the q-factor model and the FF5F

The two models are similar in terms of included factors, genesis and theoretical underpinnings.

Both the q-factor model and the FF5F arrive at a common set of factors to explain returns (market, size, investment and profitability)[110]. The FF5F additionally includes a value factor. This consensus has represented significant progress in the development of MFMs. This development also addresses the criticism levied on MFMs as part of previous regulatory analysis[111], which commented that there was judgement involved in the selection of risk factors. The convergence to a common set of factors in academic literature implies that there is now less scope for judgment.

For both models, the included factors have a combined theoretical and empirical genesis:

- In both cases, the investment and profitability factors (and the value factor) are justified based on the economic theories underpinning the respective models.
- In both cases, the market and size factors are incorporated in the models based on empirical evidence that these factors help to explain returns, however they have in the past been tied strongly to economic theories.

Both models have similar theoretical underpinnings for the investment and profitability factors (and value factor). Although at face value the theoretical underpinnings for these factors may appear different, Zhang notes these are in fact "*(virtually) identical"*[112]. Hou et al (2015) show that the investment CAPM (which forms the basis of the q-factor model) can be mathematically translated into DDM which underpins the FF5F[113][114].

The models have different bases in asset pricing (investment- vs consumption-based) and have strong, although somewhat, different empirical performance. Whilst both the q-factor model and the FF5F are leading MFMs in the academic literature, there is evidence that the former is a stronger and more robust model.

The q-factor model outperforms the FF5F on statistical tests based on US data

- Hou et al (2017) show that the q-factor model has greater explanatory power than the FF5F. Namely, the q-factor model materially explains the value, profitability, and investment FF5F factors whereas the FF5F cannot explain the investment and profitability q-factors.
- Zhang (2017)[115] points to several independent empirical comparisons of both models which broadly corroborate the results of Hou et al (2017). For example, Barillas and Shanken (2015) develop a new test procedure that allows model comparison and find results in favour of the q-factor model. The analysis from Stambaugh and Yuan (2016) finds that (1) the q-factor model explains the FF5F factors in time series regressions, but the FF5F cannot explain the q-factors and (2) the q-factor model outperforms the FF5F in explaining a wide array of anomalies in the cross section of returns.
- Green et al (2017)[116] show that the q-factor model best captures the independent determinants of average returns, relative to Carhart's four-factor model and the FF5F. This means the q-factor model yields the fewest incrementally significant factors beyond those specified by the model itself.

---

110 However, the factors are in all cases constructed differently as explained in Appendix 4.
111 See Appendix 6 for greater detail
112 Zhang Lu (December 2016), Factor Wars, Tsinghua Financial Review 37, 101-104.
113 Specifically the Gordon Growth Model, refer to section 1.2.2 of Hou et al (2015)
114 Some academics consider that both FF5F and q-factor model can ultimately be seen as derived from the DDM.
115 Zhang, L. (2017). The investment CAPM. European Financial Management, 23(4), 545-603.
116 Green, J., Hand, J. R., & Zhang, X. F. (2017). The characteristics that provide independent information about average US monthly stock returns. The Review of Financial Studies, 30(12), 4389-4436.

- Hou et al (2019) run further statistical tests which reaffirm and build on their previous results, demonstrating the q-factor model is more empirically robust than the FF5F. They show that q-factor model largely subsumes the FF5F (2015) and the FF6F (2018).

The value factor in the FF5F is redundant and without the value factor, the FF5F reduces to a 'noisy' variant of the q-factor model

- Zhang (2017) shows that the value factor included in the FF5F is redundant because it can be seen as another manifestation of the investment factor based on economic theory. This is discussed in greater detail above. Moreover, Fama and French (2015) recognise that empirically the value factor is mostly absorbed by other FF5F factors.
- Zhang (2017) considers that without the value factor, the FF5F reduces to a 'noisy' variant of the q-factor model.

The q-factor model has fewer factors than, and therefore may be preferable to, the FF5F

- The q-factor model has fewer factors than the FF5F. This reduces the scope for distortive cross-correlations amongst the factors and means the model is simpler to adopt and apply in practice.

Overall, the q-factor model and the FF5F have emerged as two of the leading MFMs in academic literature, with evidence that the former is a more robust model with stronger empirical performance. Based on US data, both of these models significantly improve upon the performance of both the CAPM and the FF3F which has been the only MFM considered in UK regulation. Furthermore, both of these models have stronger theoretical underpinnings than previous variants of MFMs[117]. As a result, this Report takes both models forward for more detailed explanation of included factors and empirical analysis using UK data.

117 See for example, Competition Commission (2003), The Competition Commission's report on the charges made by mobile operators for terminating calls - 18 February 2003

5 Approach and methodology for MFM analysis

The Report considers whether any of the leading MFMs are better at estimating the required returns for UK regulated water companies than the CAPM based on empirical analysis. The key models considered are the q-factor model and the FF5F which are the most relevant options in this context. The steps to calibrate and evaluate the performance of the MFMs relative to the CAPM and one another are as follows:

- Collect the data required to calibrate the models for a sufficiently representative sample of the UK stock market (Section 5.1)
- Construct the factors included in each of the q-factor model and the FF5F (Section 5.2)
- Carry out statistical tests and examine regression statistics to determine (1) whether the models are superior (in terms of explanatory power) to the CAPM and (2) which of the two models is superior to the other, based on the UK data (Section 5.3)
- Estimate and compare the CoE under the CAPM and under those MFMs which are superior to the CAPM based on statistical tests and regression statistics (Section 6)

The analysis is undertaken using daily returns data which is consistent with Ofwat's proposed approach to estimating betas for PR24 and also addresses an area for improvement identified in earlier analyses of MFMs in a regulatory setting[118], i.e. to explore the analysis on a daily rather than a monthly basis.

5.1 Data collection

MFMs are calibrated based on accounting, market value and returns data. This Report relies on the period March 1981[119] to March 2022 (with the start date consistent with Gregory et al, 2013[120]) as this provides a sufficiently long window from which to draw a robust sample size and an appropriate investment horizon to establish persistent patterns in observed returns.

The collection of data for MFM analysis is based on three steps:

- Obtain a long list of listed UK stocks for each year of the analysis.
- For each company in the long list, collate data to enable the filtration to exclude companies with non-viable market and book values, foreign companies, financial services and real estate companies and AIM listings (as discussed below).
- Source for all the remaining companies the returns, market value and accounting data required for the construction of factors.

Step 1

Datastream and Bloomberg were initially explored as potential sources of the long list of companies. However, it was not possible to obtain a long list from these sources for each relevant year. As a result, the long list is instead sourced from the London Share Price Database (LSPD). LSPD provides a comprehensive list of all listed UK stocks in a given year, including companies that have since de-listed and / or gone bankrupt.

118 Europe Economics (2007), CAA's price control reference for Heathrow and Gatwick airports, 2008-2013 Supporting paper II Cost of capital – analysis of responses to CAA's initial proposals. Smithers & Co (2006), Report on the Cost of Capital provided to Ofgem
119 More specifically, March 1981 for the accounting data and October 1981 for the returns data.
120 Alan Gregory, Rajesh Tharyan and Angela Christidis, Journal of Business Finance & Accounting, 40(1) & (2), 172–214, January/February 2013

De-listed stocks are included in the dataset to avoid survivorship bias. Survivorship bias results from the use of a dataset that consists of survivors over a period, not the full set of companies that were listed. As the characteristics of survivors are likely to differ systematically from those who have delisted, the results will be biased.

Consistent with Gregory et al (2013), the list of listed stocks is downloaded as at the end of September each year over 1981-2021.[121]

Step 2

Data is required for filtration to exclude companies with non-viable market and book values, foreign companies, financial services and real estate companies and AIM listings. As a result, data is required for companies on the long list from LSPD. Both market value and returns information can be obtained from LSPD. As LSPD does not include accounting information, Datastream and Bloomberg were used to source this information based on SEDOL numbers from LSPD.

The process for collection of data identified significant gaps in the Datastream dataset pre-2000, primarily related to companies which have since de-listed. The resulting sample, after removing stocks with incomplete data, was too small to generate statistically significant results.

Comparison to the dataset used for Gregory et al (2013) which partially relied on DataStream indicated that older data required for filtration may have been deleted from the platform. At the same time, there appeared to be new data on the platform which likely resulted from DataStream's merger with Eikon and data sharing across the two providers.

As a result, the following approach was followed in order to arrive at a robust dataset for filtration:

- For companies from the LSPD list that had the data required for filtration available on Datastream and / or Bloomberg, these sources were used to apply the required filters.
- The filtered list of companies derived from Datastream and Bloomberg was supplemented by the filtered dataset used in Gregory et al (2013). The period up to 2005 relies on the Gregory et al (2013) dataset, the period between 2005 – 2010 relies on a combined dataset and 2010 onwards only the new dataset is used. The data collection methods behind the Gregory et al (2013) dataset (October 1980 to December 2010) are set out in their paper. An overview of key sources is set out below:
  - Market values and sector tags are obtained from LSPD.
  - Accounting data is primarily sourced from DataStream[122], with missing values filled in with data from: Thomson One Banker; tailored Hemscott data (from the Gregory et al 2011 study of directors' trading) obtained by subscription; and hand collected data on bankrupt firms from Christidis and Gregory (2010).
  - By combining several data sources Gregory et al filled potential data gaps in the data available from DataStream.

The filtration criteria are as follows:

- Financial firms are excluded in line with Fama and French (1992). Similarly, Hou et al (2015) note that financial firms are excluded from their analysis. Fama and French (1992) state the following logic for excluding financials: "*We exclude financial firms because the high leverage that is normal for these firms probably does not have the same meaning as for nonfinancial firms, where high leverage more likely indicates distress*".
- Subsequently, Fama and French (1993) makes clear that "*REITS, ADRs and units of beneficial interest are also excluded"*. On this basis and consistent with Gregory et al (2013), real estate firms are excluded in the analysis in this Report.

121 Fama and French use June t Mcap matched with Dec t-1 financial year ends because the majority of US companies have a December financial year end. By contrast, the most common financial year end in the UK is March. As a result, Gregory et al match Sept t Mcap with March t financial year end data.
122 The historical accounting data which now appears to have been deleted from Datastream was still present when Gregory et al were developing their analysis.

- Additional reasons why financials and real estate are excluded are as follows:
  - First, in banking the treatment of customer deposits and loans means that "book-to-market" has a very different interpretation in banking compared to industrial and commercial companies.
  - Second, in financials, profits are mainly driven by interest rate spreads.
  - Third, in banks, real estate and mutual funds, book to market ratios will tend to cluster around unity.
- AIM listed stocks are excluded following Gregory et al (2013). The rationale is that AIM stocks have not historically been viewed as investible by many fund managers (due to e.g., high failure rates based on findings from Gregory et al (2010)[123] and poorer standards of reporting) and the premise behind the factor analysis is to build factors to price stocks in the investible universe.
- Foreign firms are excluded following Gregory et al (2013). This is consistent with Ofwat's approach of focusing solely on UK listed utility stocks to estimate beta at PR19.
- Stocks with missing or negative book values and market values are excluded to ensure a complete and robust dataset. Fama and French (1992) and Hou et al (2015) similarly exclude firms with negative book-to-market.

Appendix 3 sets out a detailed comparison of sample formation approaches across Fama-French, Hou et al and this Report. Overall, the approaches established by these papers have been followed to the extent possible given the differences between the US and UK markets. Where it was not possible to align the approach exactly due to differences between the two markets, the Report has preserved the principles of the US approach whilst adjusting for UK specific circumstances. The appendix only covers sample formation approaches as in all other respects (such as factor formation), the Report has maintained full consistency with Fama-French and Hou et al.

The filtered list collated in this manner for each year was taken forward to the next stage which compiled all the data required to construct the factors, i.e. the third step described above.

Step 3

Some of the accounting information required to construct the factors was not readily available in Datastream, Bloomberg or in the Gregory et al (2013) dataset as this analysis was undertaken before the FF5F and q-factor models were established. For clarity, whilst Gregory et al (2013) dataset includes the data required for filtration and to construct the FF3F, it does not include accounting data required to construct the profitability and investment factors.

This necessitated an additional source of accounting data to enable the construction of these additional factors (relative to the FF3F). The following approach was followed:

- Archived data from previous academic studies, which had accounting data available between 1980s and 2005, was identified as an additional data source.
- 2005 onwards accounting data was sourced from Datastream or Bloomberg which have significantly better data availability post 2000.

123 Gregory, Alan and Guermat, Cherif and Al-Shawawrah, Fawaz, UK IPOs: Long Run Returns, Behavioural Timing and Pseudo Timing (2009-11). Journal of Business Finance & Accounting, Vol. 37, Issue 5-6, pp. 612-647, June/July 2010

Following these data collection steps a robust dataset comprising the largest 250 stocks filtered in accordance with the requirements of the analysis was compiled. This is the largest electronically available dataset of returns and accounting data. The robustness of the resulting dataset is underpinned by:

- The exploration of multiple reputable data sources.
- The long period and the large number of companies covered by the time series of data.
- The cross-check of resulting factor premia on UK data against the factor premia from the US market over the same period. There is broad consistency between these factors, although the UK factors are more conservative overall.

The data downloaded and associated factors are set out in Table 6 and Table 7 below.

Table 6: Full dataset for q-factor

Data downloaded	Type of data	Relevant factor
Net income before extraordinary items	Accounting	Return-on-equity (profitability) factor
Book value of equity		Return-on-equity (profitability) factor
Total assets (year t and year t-1)		Investment-to-asset (investment) factor
Market value of equity	Market	Size factor
Total return index (daily) of all stocks passing filtration criteria		Factor risk premium (return-on-equity, investment-to-asset, and size factors)
Total return index (daily) of FTSE All-Share		Market factor
90-day Treasury bill rate (daily)		Market factor

Source: KPMG analysis

Table 7: Full dataset for FF5F

Data downloaded	Type of data	Relevant factor
Revenue	Accounting	Profitability factor
Total operating expense		Profitability factor
Net interest expense		Profitability factor
Book value of equity		Value factor, profitability factor
Total assets (year t and year t-1)		Investment factor
Market value of equity	Market	Value factor, size factor
Total return index (daily) of all stocks passing filtration criteria		Factor risk premium (profitability, size, value, and investment factors)
Total return index (daily) of FTSE All-Share		Market factor
90-day Treasury bill rate (daily)		Market factor

Source: KPMG analysis

Market value data from September is matched with accounting data from March in each year, leaving a 6-month gap between the two following Fama and French (2015). This ensures accounting information is readily available before the market reflects the information into the pricing of stocks, thereby avoiding look ahead bias[124] in the results.

124 Look ahead bias occurs when an analysis relies on information that was not yet available during the period being analysed.

5.2 Factor construction

The calibration of MFMs requires the construction of the factors included in each model. This requires the estimation of factor loadings and factor premia. Factor loadings are regression coefficients which represent the sensitivity of a stock's return to each risk factor included in the model (e.g. beta), and factor premia represent the additional return that is expected for taking on the associated risk (e.g. market risk premium).

The subsequent sections discuss the approach to estimating each of the additional factors included in the q-factor model and the FF5F. The market risk premium is estimated in the same manner for both the CAPM and MFMs. The calculation of market betas for MFMs follows a broadly similar logic to the CAPM but is not directly comparable to the CAPM in practice. The CAPM-market beta is calculated by regressing a stock's return on the market risk premium, whereas for MFMs the market beta is calculated concurrently with the other factor loadings by regressing a stock's return on a set of factor premia. Given the logical similarities in the calibration of the market risk factor between the CAPM and MFMs, this factor is not separately discussed below.

For reference, the MFMs are calibrated based on daily returns (to maintain consistency with the estimation of PR24 beta based on daily data) whilst the statistical tests (set out in Section 5.3) are carried out based on monthly returns (to maintain consistency with the approaches established in academic literature),

5.2.1 q-factor model factor construction

Additional factors included in the q-factor model are constructed based on Hou et al (2015) approach. In order to construct factors, each year stocks – comprised of the largest 250 firms[125] – need to be sorted into groups three times based on size, investment-to-asset (I/A) and return-on-equity (ROE) factors. Table 8 below outlines how each factor is measured and stocks are sorted.

Table 8: Measurement and sorting of size, I/A and ROE factors

Factor	Measurement	Approach to sorting
Size	Market value as of September $year_t$	Stocks are ranked and sorted into *2* groups (small and large) based on 50th percentile of market value.
Investment-to-asset (I/A)	(Total assets as of March $year_t$ – total assets as of March $year_{t-1}$) / total assets as of March $year_{t-1}$	Stocks are ranked and sorted into *3* groups (low, mid and high I/A) using breakpoints at the 30th and 70th percentile.
Return-on-equity (ROE)	Net income before extraordinary income as of March $year_t$ / book value of equity as of March $year_{t-1}$	Stocks are ranked and sorted into *3* groups (low, mid and high ROE) using breakpoints at the 30th and 70th percentile.

Source: KPMG analysis

In total, the intersections of 2 size, 3 I/A, and 3 ROE groups creates 18 (2×3×3) portfolios[126]. These 18 portfolios are then aggregated into 6 categories: large and small size, high and low ROE, and high and low I/A. This is illustrated in Table 9.

125 As noted in Gregory et al (2013), an important issue in the London Stock Exchange is that there is a large "tail" of small and illiquid stocks that are not part of the tradable universe of the major institutional investors. Therefore, researchers in the UK recognise the importance of this by applying a breakpoint in forming the factors (Gregory et al 2001, 2003, 2013; Gregory and Michou 2009; Dimson, Nagel and Quigley 2003).

126 Stocks are sorted into groups based on a single characteristic (e.g., size) whereas portfolios are formed based on an intersection of two (FF5F) or three (q-factor) characteristics.

Table 9: Components of size, I/A and ROE factors

Size factor	I/A factor	ROE factor
9 small size portfolios (intersection of small size × 3 I/A groups x 3 investment groups)	6 low I/A portfolios (intersection of low I/A x 2 size groups x 3 ROE groups)	6 low ROE portfolios (intersection of high ROE x 2 size groups x 3 I/A groups)
9 large size portfolios (intersection of large size × 3 I/A groups x 3 investment groups)	6 high I/A portfolios (intersection of high I/A x 2 size groups x 3 ROE groups)	6 high ROE portfolios (intersection of high ROE x 2 size groups x 3 I/A groups)

Source: KPMG analysis

The portfolios are constructed and rebalanced annually (October) based on the annual market value and accounting data described in Table 6. Within each year (October – September), the composition of stocks within each portfolio remains unchanged.

The return of a portfolio is the average return of all stocks within the portfolio. Factor premia are calculated by taking the simple average of the portfolio returns as specified below:

- Size factor = (sum of the daily returns of the 9 small size portfolios - sum of the daily returns of the 9 large size portfolios) / 9
- I/A factor = (sum of the daily returns of the 6 low I/A portfolios - sum of the daily returns of the 6 high I/A portfolios) / 6
- ROE factor = (sum of the daily returns of the 6 high ROE portfolios - sum of the daily returns of the 6 low ROE portfolios) / 6

After daily factor premia are constructed, factor loadings ($\beta_{Mkt,i}$, $\beta_{size,i}$, $\beta_{I/A,i}$, $\beta_{ROE,i}$) are calculated by regressing the daily excess return of a stock ($R_{it} - R_{ft}$) on the daily factor premia ($R_{Mt-} R_{ft}$, $Size_t$, I/A_t, and ROE_t) over a specific time horizon. The factor loadings are the coefficients of the regression.

The regression model employed is in line with Hou et al (2015):

$$R_{it} - R_{ft} = \alpha_i + \beta_{Mkt,i}(R_{Mt-} R_{ft}) + \beta_{size,i}Size_t + \beta_{I/A,i}I/A_t + \beta_{ROE,i}ROE_t + e_{it}$$

where

- a_i is the intercept term
- e_{it} is the difference between the expected excess return from the model and the actual excess return for the stock. It has an expectation of zero and represents the risk of the stock that is unrelated to any of the factors

5.2.2 FF5F model factor construction

Additional factors included in the FF5F are constructed based on Fama and French (2015) approach. This approach follows the same general principles as outlined for the q-factor model with differences relating to, for example, the number of sorts (reflecting the number of factors) and the way factors are measured. Appendix 4 outlines the differences in factor construction for the two models.

In the case of the FF5F, stocks are sorted four times based on size, value, profitability, and investment factors. Table 10 below outlines how each factor is measured and how the stocks are sorted.

Table 10: Measurement and sorting of size, value, profitability, and investment factors

Factor	Measurement	Approach to sorting
Size (Small minus Big: SMB)	Market value as of September $year_t$	Stocks are ranked and sorted into *2* group (small and big) based on 50th percentile of market value.
Value (High minus Low: HML)	Book-to-market is book value of equity as of March $year_t$ / market value as of September $year_t$	Stocks are ranked and sorted into *3* groups (low, neutral, and high) using breakpoints at the 30th and 70th percentile.
Profitability (Robust minus Weak: RMW)	Operating profit[127] as of March $year_t$ / book value of equity as of March $year_t$	Stocks are ranked and sorted into *3* groups (weak, neutral, and robust) using breakpoints at the 30th and 70th percentile.
Investment (Conservative minus Aggressive: CMA)	(Total assets as of March $year_t$ – total assets as of March $year_{t-1}$) / total assets as of March $year_{t-1}$	Stocks are ranked and sorted into *3* groups (conservative, neutral and aggressive) using breakpoints at the 30th and 70th percentile.

Source: KPMG analysis

The portfolios for the FF5F are formed as follows:

- Portfolios for the value factor are formed by the intersection of the 3 value groups (high, neutral, low) and 2 size groups (small, big), which gives 6 (3 × 2) portfolios for the value factor.
- Similarly, portfolios for profitability and investment factors are formed by the intersection of their own 3 groups and 2 size groups which gives 6 (3 × 2) portfolios for each factor.
- As a result, 18 portfolios are formed based on value, profitability, and investment factors.
- The portfolios for the size factor represent a reallocation of these 18 portfolios into 9 small size portfolios and 9 big size portfolios (i.e. no new portfolios need to be formed for size).
- A subset of the 18 portfolios is used to construct each of the HML, RMW and CMA factors. For example, for the HML factor, the high value portfolios and low value portfolios are used whereas neutral are not. This is illustrated in Table 11.

Table 11: Components of SMB, HML, RMW and CMA factors

SMB	HML	RMW	CMA
9 small size portfolios (intersection of small size with 3 value groups, with 3 profitability groups, and with 3 investment groups)	2 high book-to-market portfolios (intersection of high book-to-market x 2 size groups)	2 robust profitability portfolios (intersection of robust profitability x 2 size groups)	2 conservative investment portfolios (intersection of conservative investment x 2 size groups)
9 big size portfolios (intersection of big size with 3 value groups, with 3 profitability groups, and with 3 investment groups)	2 low book-to-market portfolios (intersection of low book-to-market x 2 size groups)	2 weak profitability portfolios (intersection of weak profitability x 2 size groups)	2 aggressive investment portfolios (intersection of aggressive investment x 2 size groups)

Source: KPMG analysis

127 Operating profit is measured by (revenue – total operating expense – net interest expense).

The return of a portfolio is the average returns of all stocks within the portfolio. Factor premia are calculated by taking the simple average of the portfolio returns as specified below:

- SMB factor = (sum of the daily returns of the 9 small size portfolios - sum of the daily returns of the 9 big size portfolios) / 9
- HML factor = (sum of the daily returns of the 2 high book-to-market portfolios - sum of the daily returns of the 2 low book-to-market portfolios) / 2
- RMW factor = (sum of the daily returns of the 2 robust profitability portfolios - sum of the daily returns of the 2 weak profitability portfolios) / 2
- CMA factor = (sum of the daily returns of the 2 conservative investment portfolios - sum of the daily returns of the 2 aggressive investment portfolios) / 2

After daily factor premia are constructed, factor loadings (b_i, s_i, h_i, r_i, c_i) are calculated by regressing the daily excess return of a stock ($R_{it} - R_{ft}$) on the daily factor premia ($R_{Mt-} R_{ft}$, SMB_t, HML_t, RMW_t and CMA_t) over a specific time horizon. The factor loadings are the coefficients of the regression.

The regression model employed is in line with Fama and French (2015):

$$R_{it} - R_{ft} = a_i + b_i(R_{Mt-} R_{ft}) + s_i SMB_t + h_i HML_t + r_i RMW_t + c_i CMA_t + e_{it}$$

5.3 Overview and results of statistical tests

A two-stage statistical testing of the robustness of the q-factor, FF3F and FF5F models is undertaken to evaluate the empirical performance of the models relative to the CAPM based on UK data and whether they warrant inclusion in the cross-checks for PR24. The statistical tests deployed – factor spanning and Gibbons-Ross-Shanken (GRS) tests – are widely recognised by the academic literature. The factor spanning test is applied as the first-stage test and the GRS test as the second stage, with only the model(s) that pass the first stage taken forward to the second.

The q-factor model passes the spanning test whilst the FF5F fails. This may be driven by (1) the inclusion of the redundant value factor in the FF5F which may add 'noise' to the model, (2) cross-correlations amongst factors and a weaker explanatory power due to the hidden investment effect[128], and (3) divergence from the definition of profit in the well-established Peasnell model which may result in a weaker explanatory power for observed returns.

As a result, only the q-factor model is taken forward to stage-two testing where the q-factor model is found to perform better on the GRS test than the CAPM.

The q-factor model based on UK data passes the spanning test and performs better on GRS tests – as evidenced by the materially adjusted R^2 across all test portfolios relative to the CAPM[129] – and so has stronger explanatory power than the CAPM. As a result, the model warrants inclusion in the cross-checks for PR24.

In academic literature, all asset pricing models are examined for their validity and usefulness before they are used to draw conclusions regarding the drivers of observed returns or to estimate expected returns. In order to assess whether MFMs warrant inclusion in the cross-checks for PR24, it is essential to evaluate their explanatory power relative to the CAPM and to determine whether the models are subsumed by the CAPM in explaining observed returns.

This section applies statistical tests widely recognised in empirical asset pricing literature[130] to the two MFMs which are candidates for the inclusion in the PR24 cross-checks, the q-factor model and the

128 This is because the FF5F measure of profitability divides in-year profit by contemporaneous book equity which, relative to the q-factor approach for calculating this value, incorporates an extra measure of the investment factor (difference between contemporaneous assets and one-year lagged assets).
129 Adjusted R2 is a corrected goodness-of-fit (model accuracy) measure for linear models. It identifies the percentage of variance in the target field that is explained by the input or inputs.
130 Fama & French (2015); Fama & French (2017); Hou, Xue & Zhang (2015), Hou, Xue & Zhang (2017); Hou, Mo, Xue & Zhang (2019)

FF5F, based on UK market data. In combination the two tests cover the factor premia and the factor loadings for the *test portfolios* described in Sections 5.2.1 and 5.2.2. In this Report, the factor premia are calculated based on the data for all companies used to calibrate the model whereas the factor loadings can be calculated for either (1) based on all companies or (2) specific stocks. In this section, the factor loadings are calculated based on all companies. The Report also tests the statistical significance of factor loadings for *regulated utilities comparators* in Section 5.4.2.

Statistical testing is undertaken in two stages (1) factor spanning test and (2) GRS test and only the model(s) that pass the first stage taken forward to the second. Note that the tests are conducted on the monthly return data on a basis consistent with the academic literature.

5.3.1 Factor spanning test

The factor spanning test is the first and most important test that MFMs need to pass. It is adopted by researchers to enable comparison of the performance of different asset pricing models (e.g. Hou et al (2017, 2019), Fama and French (2017)[131], Foye (2018)[132], Rugwiro and Choi (2019)[133]).

The test examines whether additional risk factors add to the explanation of observed returns provided by an existing model. The spanning test can directly assess whether one model is superior to (or subsumes) another model. Passing the spanning test means that the model being examined provides additional explanatory power to the other model.

The spanning test is conducted by regressing the additional factors[134], either individually or jointly, on existing factors based on monthly data[135]. The test can be interpreted as follows:

- If the intercept of the regression is statistically different from zero (at a 5% significance level), the additional factor(s) add to the explanation of observed returns over the sample period i.e., a pass; and
- If the intercept of the regression is not statistically different from zero (at a 5% significance level), the additional factor(s) are subsumed by those in the existing model and can be considered redundant i.e., a fail.

In addition to testing whether additional factors contribute to the explanatory power of an existing model, the test can also be used to compare two models that employ different set of factors (e.g. q-factor vs. FF5F). This requires a joint test of factors. Individual tests of factors do not apply in this case given the factors are different instead of additive to existing factors (as is the case for FF5F vs FF3F).

In the case of the FF5F, the test is performed in two stages. First, the FF3F is compared with the CAPM to examine the explanatory power of SMB and HML. Then, the FF5F is compared with the FF3F to examine the explanatory power of the two additional factors, CMA and RMW.

Table 12 below provides a detailed comparison among asset pricing models.

131 Fama, E. F., & French, K. R. (2017). International tests of a five-factor asset pricing model. Journal of financial Economics, 123(3), 441-463.
132 Foye, J. (2018). A comprehensive test of the Fama-French five-factor model in emerging markets. Emerging Markets Review, 37, 199-222.
133 Rugwiro, S., & Choi, S. B. (2019). Re-examination of Fama–French Models in the Korean Stock Market. Asia-Pacific Financial Markets, 26(1), 23-45.
134 For example, we could test whether SMB factor provides additional explanation on return compared to market risk premium factor by running $SMB_t = a_i + b_i(R_{Mt} - R_{ft}) + e_{it}$. If the intercept term a_i is statistically different from zero, it means SMB adds to the explanation of returns above and beyond the market risk factor
135 For statistical testing the Report uses monthly data but for implementation it follows regulatory precedent and uses daily data.

Table 12: Factor spanning tests on MFMs

	FF3F vs. the CAPM	FF5F vs. FF3F	q-factor vs. the CAPM	q-factor vs. FF3F	q-factor vs. FF5F
Additional factors	SMB, HML	RMW, CMA	ROE, investment, size	ROE, investment, size	ROE, investment, size
Existing factors	$R_M - R_f$	$R_M - R_f$, SMB, HML	$R_M - R_f$	$R_M - R_f$, SMB, HML	$R_M - R_f$, SMB, HML, RMW, CMA
Pass / fail	Individually: Fail (SMB, HML) Jointly: Fail	Individually: Fail (RMW, CMA) Jointly: Fail	Individually: Pass (ROE, investment), Fail (size) Jointly: Pass	Jointly: Pass (individual test does not apply)	Jointly: Pass (individual test does not apply)
Implication	SMB and HML are redundant. FF3F does not subsume the CAPM	RMW and CMA are redundant. Neither FF3F nor FF5F subsume the CAPM	q-factor model subsumes the CAPM[136]	q-factor model subsumes FF3F	q-factor model subsumes FF5F

Source: KPMG analysis

The results of the spanning test revealed that the q-factor model subsumes the CAPM and the FF5F:

- The additional three factors in the q-factor model provide the model with greater explanatory power than the CAPM for the UK market returns.
- The q-factor model jointly provides additional explanations on observed returns compared to the FF3F and FF5F, suggesting that the q-factor model subsumes the two MFMs. This is consistent with what has been tested by Hou et al (2019) that the q-factor model largely subsumes the FF5F, using the US market data[137].

As a result, the q-factor proceeds to the next stage of statistical testing.

By contrast, the FF5F failed the factor spanning test, which indicates that the additional four factors included in the model can be considered, both individually and jointly, redundant based on empirical analysis. This means that the FF5F based on UK data is not robust, cannot improve upon the explanatory power of the CAPM and does not warrant inclusion in the cross-checks for PR24. As a result, the FF5F does not proceed to the next stage of statistical testing.

Whilst both the q-factor model and the FF5F are leading MFMs in the academic literature, the FF5F failure of the spanning test could be explained by the following:

1. The value factor in the FF5F is redundant and inclusion of the factor may add 'noise'

The FF5F incorporates the value factor. Zhang (2017) shows that the value factor can be seen as another manifestation of the investment factor and is therefore redundant in the FF5F. As a result, the inclusion of the value factor in the FF5F could lead to distortive cross-corelations amongst the FF5F factors without providing any additional benefit to the model's explanatory power. Put another way, the value factor may simply add 'noise' to the model.

136 Whilst the size factor fails individually, as two of the three additional factors pass individually and the model passes the joint test, the overall result is a pass.
137 Hou, K., Mo, H., Xue, C., & Zhang, L. (2019). Which factors? Review of Finance, 23(1), 1-35.

2. The profitability factor in the FF5F may capture a hidden investment effect which could result in cross-correlations amongst factors and dampen the model's explanatory power

The FF5F measure of profitability divides in-year profit by contemporaneous book equity. Assuming stocks are debt-free for simplicity, then book equity would equal (total) assets and the FF5F measure of profitability would become in-year profit / contemporaneous assets.

It is economically more logical to use one-year lagged assets as opposed to contemporaneous assets as intuitively, in-year profit is generated by one-year lagged assets. Indeed, the q-factor measure of profitability is based on one-year lagged assets[138].

Given the use of one-year lagged assets seems more suitable, the FF5F measure of profitability may contain a 'hidden' investment effect. This is because the difference between contemporaneous assets and one-year lagged assets (i.e., asset growth) is exactly the measure of the investment factor. As a result, there could be distortive cross-correlations between the FF5F profitability and investment factors which could dampen the model's explanatory power.

For clarity, the expectation is that the value factor is completely explained by the investment factor whereas the profitability factor is partially explained by the investment factor.

3. The profitability factor in the FF5F uses a different definition of profit to the well-established Peasnell model and thus may have weaker explanatory power for observed returns

Some academics have posited that the structure behind the FF5F could be seen as another variant of the accounting model described by Peasnell (1982). Peasnell derives a discount model which computes the economic valuation of firms using accounting measures of profit, provided the accounting is clean surplus[139].

The FF5F measure of profitability uses a definition of in-year profit that is close to a pre-tax operating profit definition. This is quite different to the clean surplus definition of profit assumed by Peasnell which is instead closer to the bottom line.

The Peasnell model has long been established and is widely recognised in academic literature. Thus, by adopting a different definition of in-year profit to Peasnell, the FF5F profitability factor may have weaker explanatory power for observed returns.

5.3.2 GRS test

The GRS test[140] is widely adopted in testing MFMs.[141] It indicates whether an asset pricing model could explain the observed returns of all the test portfolios. The test regresses the portfolio returns on factor premia for each portfolio separately. If the intercept terms of all the tested portfolios are jointly indistinguishable from zero, the model passes GRS test. The test is binary in the sense that a model could either pass or fail the test, but to assess the performance of different models on a relative basis, the next question would be how much of the variation in observed returns could be explained by the model. Here, the adjusted R^2 is a useful indicator.

Failing the GRS test is not uncommon. As illustrated by Fama and French (2015) and Hou et al (2015), both the CAPM and MFMs fail the GRS test based on US market data. However, this does not invalidate the asset pricing models. As noted in Fama and French (2015), asset pricing models are simplified propositions about returns: *"We are less interested in whether competing models are rejected than in their relative performance, which we judge using GRS and other statistics. We want to identify the model that is the best (but imperfect) story for average returns on portfolios formed in different ways."*[142]

138 It uses one-year lagged book equity but in the case of no debt, this is equal to one-year lagged (total) assets.
139 In "clean surplus" accounting all valuation changes in book value (e.g. depreciation and revaluation) must flow through the P&L account.
140 Gibbons, M. R., Ross, S. A., & Shanken, J. (1989). A test of the efficiency of a given portfolio. Econometrica: Journal of the Econometric Society, 1121-1152.
141 For example, Gregory et al (2013), Hou et al (2015, 2017, 2019) and Fama and French (2015, 2017)
142 Page 10, Fama, E. F., & French, K. R. (2015). A five-factor asset pricing model. Journal of financial economics, 116(1), 1-22.

Following Fama and French (2015), test portfolios are constructed based upon size and book-to-market (B/M) ratio. Stocks are allocated to five size groups (Small to Big) and allocated independently to five B/M groups (Low to High). The intersections of the two sorts produce 25 Size-B/M portfolios. GRS test then runs an asset pricing model on each portfolio, and tests whether the expected values of all 25 intercepts are zero. Both the CAPM and q-factor fail the GRS test, which indicates neither model could describe the observed returns of all the tested portfolios. In particular, the portfolios that failed the test are mainly small-size portfolios. This has relatively limited impact on the usefulness of these models for PR24 as listed regulated utilities are large companies. After excluding the ten small-size portfolios, both the CAPM and q-factor pass GRS test.

Whilst both models pass the test after excluding small portfolios, the Report also considers the relative performance of the two models, measured by adjusted R^2. For all 25 portfolios, the q-factor model has consistently higher explanatory power than the CAPM. On average, the adjusted R^2 of the q-factor model is 64.5%, meaning that 64.5% variability of portfolio returns could be explained by the

Overall the Report tests the performance of the q-factor model using an approach that is consistent with that applied in academic research where MFMs are evaluated as potential replacements for the CAPM as the primary methodology for estimation of returns. As a result, the bar applied in this Report to MFM evidence as a potential cross-check is significantly greater than the MAR cross-check included in the PR24 DM. All else equal this suggests that MFM evidence should be considered to be a primary cross-check. This implies in cross-checking the CAPM that weight should be attached to evidence implied by the q-factor model over and above other cross-checks.

5.4 Application of the q-factor model to UK regulated utilities

Estimation of CoE for regulated water companies using the q-factor model and the CAPM requires loadings and premia for each of the factors. Statistical analysis of the q-factor loadings for Severn Trent / United Utilities (SVT/UUW) value-weighted portfolio indicates that these factors are individually significant in most cases and jointly significant in all cases. This suggests that additional factors included in the q-factor model are useful in explaining water company returns, individually and jointly.

In order to assess the implications of the evidence from the q-factor model for required CoE for PR24, the Report compares CoE estimates derived using the q-factor model and the CAPM under different and cut-off dates for the value-weighted pure play water portfolio comprised of SVT/UUW comparators. Pennon (PNN) is not included in this portfolio as there is insufficient pure play data as at the two cut off dates used in the Report. Notably, Ofwat has proposed to review whether to include PNN data in beta estimation in the final methodology and has noted that reflecting this data would not be straightforward due to difficulties in accounting for cash holdings from the disposal of Viridor with gearing.

Estimation of CoE for regulated utilities using the q-factor model and the CAPM requires loadings and premia for each of the factors.

- Annualised factor premia for CoE estimation are derived as set out in Section 5.2 based on the full dataset employed. Consistent with other statistical tests applied in this Report, the statistical significance of the premia is evaluated on a monthly basis in Appendix 5.
- Factor loadings for regulated utilities are estimated based on the regression model outlined in Section 5.2 and their statistical significance is evaluated in Section 5.4.2.[143]

This section sets out the approach and assumptions underpinning the derivation of CoE using the q-factor model and the CAPM and presents the factor loadings and associated significance levels for each cut-off date and averaging window for each of the comparators.

143 i.e. regress daily excess returns ($R_{it} - R_{ft}$) on daily factors ($R_{Mt} - R_{ft}, Size_t, I/A_t$ and ROE_t) to estimate the factor loadings.

5.4.1 The approach for estimating the CoE for regulated utility comparators

The first step in estimating the CoE is to derive the excess returns (i.e. CoE less risk-free rate) from each of the models.

Table 13 below provides detailed specification of assumptions underpinning the estimation of CoE for regulated utilities and associated rationale.

Table 13: Assumptions underpinning calibration of factor loadings and premia and the estimation of CoE for regulated utilities

	Assumptions	Rationale
Assumptions underpinning the calibration of factor loadings and premia		
Estimation window	2-, 5- and 10-year	Consistent with estimation windows set out in PR24 DM
Cut-off dates	March 31, 2022	The latest reporting date at which all the financial and accounting data is available to construct daily factors
	February 28, 2020	Consistent with the pre-Covid cut-off used by the CMA at PR19[144]. Used to assess the relative differences between q-factor and the CAPM-derived CoE before Covid and the Russia-Ukraine war
Averaging windows	Spot	Ofwat has not signalled in the DM that it will use other averaging windows apart from spot
Data frequency	Daily	Consistent with Ofwat's approach in the PR24 DM
Dependent variables for the regression	Excess returns of SVT/UUW portfolio	Consistent with comparators adopted in the PR24 DM
Independent variables for the regression (CAPM)	Market risk premium[145]	According to the CAPM specification
Independent variables for the regression (q-factor)	Market risk premium, investment, ROE	According to q-factor specification
Assumptions underpinning the estimation of CoE for regulated utilities		
Factor loadings	Calculated based on the assumptions above	See above
Factor premia	For market risk premium, the CMA's total market return (TMR) of 6.81% is adopted in conjunction with the risk-free rate estimates set out below. This applies to both CAPM and q-factor	Consistent with the CMA's point estimate for PR19
	For the premia for other factors, average annualised factor premia of the longest horizon available are used (average annualised factor premia from the first period Oct	Consistent with the approach deployed by the regulators and the CMA when deriving the TMR i.e. the

144 CMA (2021), PR19 Final Determination, para 9.468
145 The market risk premium used to calibrate factor loadings and premia uses a TMR derived based on the data collected for the MFM analysis and uses a risk-free rate based on the 90-day Gilt rate. This is consistent across CAPM and q-factor.

	Assumptions	Rationale
	1981/Sept 1982 to the last period Oct 2021/March 2022)	use of data since 1900 to derive TMR in order to ensure robustness
Inflation (to deflate nominal factor premia)	An average of (1) CPI series used unadjusted for the latest CPI and CPIH back-cast and (2) RPI series adjusted for 30 bps 'formula effect' and 90 bps RPI-CPIH wedge (Oct 1981 to March 2022 horizon)	Consistent with the inflation series used in CMA's PR19 TMR
Gearing basis	Book value of net debt	Book value of net debt is adopted for simplicity
Debt beta (relevant for the market factor only)	0.075	Consistent with the CMA's point estimate for PR19
Risk-free rate (for calculating CoE)	-0.97% (March 31, 2022 cut-off)	Consistent with the CMA's PR19 approach for setting the point estimate for the risk-free rate[146]
	-0.97% (February 28, 2020 cut-off)	

Source: KPMG analysis

The excess return of a stock is calculated using the product of the factor premia and factor loadings. Factor loadings are estimated as at each cut-off date, whereas the factor premia are derived based on the longest run of data is applied in each case[147]. As the actual gearing for comparators differs from notional gearing, the Harris-Pringle formula[148] is adopted to de- and re-lever factor loadings. The excess returns thus derived are then added to a risk-free rate estimate calculated using the CMA's PR19 methodology to arrive at a CoE estimate.

5.4.2 Factor loadings and associated significance levels for the regulated utility comparators

Section 5.3 statistically tests the factor loadings for all the stocks via the test portfolios used to construct the factors. This section assesses the statistical significance of the factor loadings for the regulated utility comparators specifically.

The tables below present factor loadings and the associated significance levels across the three averaging windows and cut-off dates included in the analysis. Where a factor loading is statistically significant[149][150], this means that it has statistical power to explain observed returns and actively contributes to the explanatory power of the model. The tables also present the joint significance[151] of the factors i.e. whether they jointly have statistical power in explaining observed returns.

The following observations emerge for the SVT/UWW portfolio based on the results presented in the tables:

- For the March 2022, cut-off date, three out of four factors in the q-factor model (market risk premium (MRP), size, and ROE) individually exhibit significant explanatory power at the 1% significance level. In addition, all the factors, jointly, are statistically significant at the 1% level and jointly contribute to the explanatory power of the model.

146 I.e. a 6-month average of (1) 20-year ILD rate and (2) an average of AAA Non-Gilts 10-15 years and AAA Non-Gilts 10+ years indices
147 In the case of market risk premium, the CMA's PR19 assumptions of TMR and risk-free rate are used.
148 This is the formula used by regulators to de-lever and re-levered raw equity betas to derive the notional beta. The Harris-Pringle formula. See 'Risk·Adjusted Discount Rates-Extensions from the Average·Risk Case', (1985), Robert S. Harris and John J. Pringle
149 Significance level defines the strength of evidence in probabilistic terms. Typical significance levels used in statistics are 1%, 5% and 10% with 1% being the strongest evidence in probabilistic terms.
150 Statistical significance is assessed based on p-values. P-value represents the probability that that a factor beta is statistically indifferent from zero (which is the null hypothesis). For example, if p-value is lower than 10%, the null hypothesis can be rejected, and the factor beta is statistically different from zero at 10% significance level
151 Assessed based on the p-value of the F-statistics. An F-test assesses how well the set of independent variables, as a group, explains the variation in the dependent variable. If p-value of F-statistics is lower than the significance level, there is sufficient evidence that the independent variables are jointly statistically different from zero.

- For the February 2020 cut-off date, MRP and ROE are in all cases individually statistically different from zero at the 1% significance level. The factors are jointly significant and add to the explanatory power of the model.
- The investment factor loadings are not statistically different from zero, suggesting that for regulated utility stocks, the investment style is neutral, i.e. neither aggressive nor passive.
- Whilst the investment factor is not significant for regulated utility stocks, the q-factor model specifies that all four factors in combination are required to explain observed stock returns across the market as a whole. As such, removing the investment factor from the model would introduce omitted variable bias and be inconsistent with the original design of the model.
- The above suggests that additional factors included in the q-factor model are useful in explaining observed returns, individually and jointly. The market beta in the CAPM is also significant in all cases, as expected.
- The results of the q-factor model represent a significant improvement in the statistical significance of the factor loadings compared to previous regulatory analysis on the FF3F where there was, at best, weak statistical evidence on SMB and HML factors in terms of factor loadings[152].

The additional factors in MFMs are systematic in nature i.e., they are not associated with any particular sector but rather the market as a whole. As such, there is general *economic intuition* for why the additional factors are significant in explaining observed returns across the market but this intuition is challenging to apply within the context of any particular sector. Consequently, the significance levels of the (additional) factor betas for a particular sector cannot be explained intuitively, and are purely a result of an empirical exercise. Similarly, in the single factor CAPM, it can be challenging to fully explain intuitively the exact size of the market beta for any particular sector.

The tables below illustrate that the returns for utilities under the q-factor model are consistently driven by the market and ROE factors, offset somewhat by the size factor. Critically, the positive loading on the ROE factor does not stem from the historical profit levels of the two water companies and as discussed above it is difficult to apply general *economic intuition* as to why utilities specifically load positively on this factor. As discussed in Section 4.2 the additional factors in MFMs are proxies for (directly unobservable) systematic risk exposures. An *interpretation* of what a positive loading on the ROE factor means is that a company is exposed to additional systematic risk (relative to the CAPM) which is best proxied by the ROE measure. This could indicate for example that a certain level of profitability is expected for investors to commit capital into a stock *given* the systematic risk exposure.

Table 14: The CAPM and q-factor raw factor loadings and associated statistical significance for the value-weighted SVT/UUW portfolio (cut-off date March 31, 2022)

Comparator	Estimation window	CAPM	q-factor				
		MRP	MRP	Size	Investment	ROE	Jointly significant at 1%?
Water portfolio	10-year	0.590***	0.595***	-0.302***	0.018	0.314***	✓
	p-value	(0.000)	(0.000)	(0.000)	(0.685)	(0.000)	(0.000)
	5-year	0.558***	0.569***	-0.265***	0.022	0.248***	✓
	p-value	(0.000)	(0.000)	(0.000)	(0.751)	(0.000)	(0.000)
	2-year	0.482***	0.528***	-0.275***	-0.058	0.270***	✓
	p-value	(0.000)	(0.000)	(0.000)	0.495	(0.000)	(0.000)

Source: KPMG analysis
Note: * is p-value <5%, ** is p-value <1%, *** is p-value <0.1%

152 PR09 Cost of capital and financeability (Nov 2009), Europe Economics; CAA's price control reference for Heathrow and Gatwick airports, 2008-2013, Supporting paper II (March 2007); Report on the Cost of Capital provided to Ofgem (September 2006), Smithers & Co

Table 15: The CAPM and q-factor raw factor loadings and associated statistical significance for the value-weighted SVT/UUW (cut-off date February 28, 2020)

Comparator	Estimation window	CAPM	q-factor				
		MRP	MRP	Size	Investment	ROE	Jointly significant at 1%?
Water portfolio	10-year	0.592***	0.633***	-0.147***	0.065	0.439***	✓
	p-value	(0.000)	(0.000)	(0.000)	(0.131)	(0.000)	(0.000)
	5-year	0.685***	0.749***	-0.045	0.049	0.456***	✓
	p-value	(0.000)	(0.000)	(0.431)	(0.435)	(0.000)	(0.000)
	2-year	0.640***	0.764***	0.186	0.005	0.468***	✓
	p-value	(0.000)	(0.000)	(0.078)	(0.971)	(0.000)	(0.000)

Source: KPMG analysis
Note: * is p-value <5%, ** is p-value <1%, *** is p-value <0.1%

6 Discussion and interpretation of the results for water companies

To assess the implications of the evidence from the q-factor model for the allowed CoE at PR24, this section compares CoE estimates derived using the CAPM and q-factor models as at each date for the SVT/UUW value weighted portfolio.

The differentials across all estimation windows and both cut-off dates are positive, i.e. returns implied by the q-factor model are higher than implied by the CAPM. The indicates that the proxies proposed by Ofwat for the UK water sector have a higher systematic risk exposure than that implied by the CAPM. All else equal, this implies that where the CAPM is used to set allowed CoE in a price control setting, the model is likely to *under-estimate* systematic risk exposure, and this should be taken into account in selecting a point estimate for the allowed CoE.

The differential between q-factor and CAPM CoE is 0.39 – 2.96%. This differential can be added to the mid-point of Ofwat's CAPM-derived CoE range to assess what the MFM cross-check implies for allowed returns at PR24.

The qualitative and quantitative evidence set out in this Report implies that (1) MFMs warrant inclusion in the suite of cross-checks for PR24 and (2) given that MFMs have stronger explanatory power than the CAPM (which does not apply to for example the MAR cross-check included in the as a cross-check in the PR24 DM), MFMs should be treated as a primary cross-check. This implies in cross-checking the CAPM that weight should be attached to evidence implied by the q-factor model over and above other cross-checks. In practice MFM evidence indicates that the point estimate for the allowed CoE for PR24 should be 0.39 – 2.96% *higher* than the mid-point of Ofwat's CAPM-derived CoE range.

This section sets out a discussion and interpretation of the results from the MFM analysis and potential implications for the allowed CoE at PR24. In doing so it considers both qualitative evidence and the results of the empirical analysis of the q-factor model for UK regulated water utilities.

6.1 Comparison of returns implied by the q-factor model and the CAPM

To assess the impact and implications of the evidence from the q-factor model for the allowed CoE at PR24, the Report compares CoE estimates derived using the CAPM and q-factor as at each date for the SVT/UUW portfolio. For both the CAPM and the q-factor, the CoE is calculated by combining the notional factor loadings, factor premia and an estimate of the risk-free rate as discussed in Section 5.4. The resulting differentials between the two models are set out in the table below.

A positive differential means that the CoE derived using the q-factor model exceeds that derived using the CAPM, whereas a negative differential means that CAPM-derived CoE exceeds that derived using the q-factor model. The results set out in Table 16 show that the differentials across all estimation windows and both cut-off dates are positive. . Overall, empirical analysis based on UK data finds that the CoE derived using the q-factor model is 0.39 – 2.96% higher than that derived using the CAPM. This points to a conclusion that the proxies proposed by Ofwat for the UK water sector, have a higher systematic risk exposure than that implied by the CAPM.

Table 16: CoE differentials between estimates derived using the CAPM and q-factor

Cut-off date	Estimation window	Water portfolio
March 31, 2022	10-year	0.47%
	5-year	0.39%
	2-year	0.52%
February 28, 2020	10-year	1.73%
	5-year	2.20%
	2-year	2.96%

Source: KPMG analysis
Note: The estimates presented for each estimation window represent the spot rate.

It is noted that there is material variance between the differentials across the two cut-off dates. This is primarily driven by the structural break associated with Covid, which has resulted in a marked 'flight to safety' effect. Excluding data from the Covid period the implied differential is 1.73% - 2.96%.

All else equal, this implies that the CAPM as the primary basis for estimation of allowed CoE in price control setting is likely to under-estimate systematic risk exposure, and this should be taken into account in setting the point estimate for CoE and cross-checking returns implied by CAPM.

The variance in returns implied by the two models can be viewed in the context of the extensive academic research which explored empirical shortcomings and contradictions of the CAPM, which has limited power to explain observed returns (which ultimately led to the genesis of MFMs, as set out in Section 4). The q-factor model has been shown to have stronger empirical performance than the CAPM based on UK data, and the variances set out in the table above should be considered in this context.

It is also important to note that the CAPM predicts that risky companies will earn returns above the market and less risky companies will earn returns below the market, but the average company will earn returns in line with the market. In a sense, it predicts a zero-sum game across the market. The same logic applies to MFMs. In the context of regulated utilities, the q-factor model predicts that investors perceive regulated utilities as riskier than implied by the CAPM and so require higher returns to hold them. As a result, inherently there will be sectors other than utilities in the q-factor model with *lower* implied returns than the CAPM.

6.2 Implications of the results for setting CoE at PR24

The following evidence set out in this Report is considered in coming to a view on MFMs as an alternative, robust cross-check on CAPM-derived CoE for PR24:

- CAPM is used by all UK regulators as the primary methodology for setting the allowed CoE for price controls, reflecting its simplicity, straightforward interpretation and ease of use. However, academic research has over time identified a number of empirical shortcomings in the CAPM to explain observed returns. These shortcomings are also acknowledged by sector regulators, the CMA and practitioners. As a result, MFMs have been used as the preferred asset pricing models in academia for over almost thirty years and are being increasingly relied upon by practitioners.
- UKRN and the CMA have also recognised the stronger power of MFMs compared to the CAPM. Whilst MFMs have been considered in the past[153] by UK regulators (Ofwat, CAA, Ofgem, Ofcom) as a tool which could be used to estimate regulatory CoE, regulatory analysis of MFMs was predominantly concentrated in the early 2000s and has not been substantively revisited thereafter as MFMs have developed.

153 For example, as part of PR04, PR09 in water, Q5 appeal in aviation, TPCR4 in energy.

- These previous analyses focussed on the FF3F which was established around thirty years ago (1993). MFMs have moved on significantly since then. In particular, they have undergone a process of development and refinement which has been informed by a long series of academic studies by several authors over the 1980s-2000s. The leading MFMs which are now favoured in academia include a broadly common set of factors and are significantly more empirically and theoretically robust than the FF3F.
- Empirical analysis based on UK data finds that the q-factor model – which is one of the leading MFMs in academic research – performs better than the CAPM and the FF5F on statistical tests. Both the factor premia and factor loadings (for all companies and UK regulated utilities) are statistically robust and are useful in explaining observed returns.
- These results are consistent with the findings from the US which show that the q-factor model (1) has stronger empirical performance than older MFMs and (2) outperforms the FF5F in head-to-head spanning tests.
  - Only the q-factor model passes the spanning test which means that the additional three factors in the q-factor model endow the model with greater explanatory power than the CAPM, while the FF3F and FF5F do not, for the UK market observed returns.
  - The GRS test shows that on average, the adjusted R^2 of the q-factor model is 64.5%, meaning that 64.5% variability of portfolio returns could be explained by the q-factor model, while for the CAPM, the adjusted R^2 is 45.2%.
  - Statistical analysis of the factor loadings for the water portfolio indicates that these factors are individually significant in most cases and jointly significant in all. This suggests that additional factors included in the q-factor model are useful in explaining observed returns, individually and jointly.
- MFM evidence *improves* the explanatory power of the CAPM based on a more granular and nuanced assessment of risk than the CAPM.
- The Report tests the performance of the q-factor model using an approach that is consistent with that applied in academic research where MFMs are evaluated as potential replacements for the CAPM as the primary methodology for estimation of returns. As a result, the bar applied in this Report to MFM evidence as a potential cross-check is significantly greater than for any cross-check being considered as part of the PR24 draft methodology. All else equal this suggests that MFM evidence should be considered to be a *primary* cross-check. This implies in cross-checking the CAPM that weight should be attached to evidence implied by the q-factor model over and above other cross-checks.
- If there is a risk that CAPM could over- or under-state returns, there is a requirement for a robust model to sense-check CAPM-derived returns and alternative models such as MFMs can help to select for example a point estimate within a range. It is possible (given that MFMs can improve on the explanatory power of CAPM) that MFM models could ultimately replace CAPM as the primary methodology for setting returns however this would represent a material change to the regulatory framework for returns estimation and this Report starts by considering what MFMs indicate in relation to CAPM results.
- The evidence set out above implies that in cross-checking the CAPM that weight should be attached to evidence implied by the q-factor model over and above other cross-checks. In practice MFM evidence indicates that the point estimate for the allowed CoE for PR24 should be 0.39 – 2.96% higher than the mid-point of Ofwat's CAPM-derived CoE range.

6.3 Comparison of MFMs to alternative cross-checks and potential role of MFMs

The cross-checks adopted for PR19 (broker forecasts and MAR) and proposed for PR24 (MAR only) have weaker explanatory power than the CAPM and cannot improve upon the CAPM-derived CoE. In contrast MFMs are empirically more robust based on US and UK data, are adopted by academics as a primary methodology for estimation of returns and are supported by economic theory.

Table 17 sets out a brief overview and limitations for each of the two cross-checks applied by Ofwat at PR19. In both cases, there is reliance on a number of assumptions and judgments which reduce robustness and render drawing clear conclusions on CoE difficult. Moreover these assumptions and judgments reduce the robustness of the MAR and broker forecast cross-checks *relative to* MFMs.

Table 17: Commentary on limitations of Ofwat's cross-checks applied to the PR19 CoE

Specification of cross-check	Key limitations of the cross-check and assumptions required
MAR-implied CoE CoE inferred from the MAR estimates on Affinity Water and Dee Valley Water transactions as well as from the trading premiums on United Utilities and Severn Trent. Assumptions for other variables that could impact a RCV premium, such as for growth and for outperformance, are made to infer CoE.	• Whilst both traded and transaction MARs have limitations, for example, reliance on an extensive set of assumptions which are individually and together estimated and/or assumed with significant uncertainty, the latter is exposed to additional challenges such as: • Transactions of non-listed companies are very infrequent and suffer from a significant selection bias as to when they occur, and which assets are traded. • These transactions do not reflect the true underlying market price that would be revealed from continuous, efficient, and liquid market, as opportunities arise on a sporadic bespoke basis at isolated points in time carefully chosen by the sellers, which results in a significant selection bias. • The ability of private transactions' MARs data to explain the fundamentals of a particular asset, including the regulatory price control parameters determining its cash flows, not to mention informing specifically a single input parameter, is therefore significantly compromised. • The transactions cover a small subset of the assets which is not necessarily representative of the rest of the sector. • Further complications arise from the private transactions process set-up (including the impact of winner's curse [154]), high transaction barriers and relatively few participants, private information and illiquid assets at a certain size of the equity cheque, management determination to succeed, the use subjective and optimistic assumptions related to future operational and financial performance, agency problems[155], the control premium effect[156].
Broker forecasts of CoE The investor CoE assumed by market analysts to value utilities can be used to infer the level of return required by investors.	• The broker forecasts focus on only the two listed water companies (United Utilities and Severn Trent) which may not be representative of the sector. • The investor CoE estimates used in broker valuations may be specifically tailored to particular investors or house views rather than representing the cost of capital demanded by the average or marginal investor in the sector. • There may be circularity in these estimates if analysts assume a CoE close to those set by the regulator rather than conducting their own assessments.

Source: KPMG analysis

The table below sets out an assessment of MFMs as a cross-check against robust criteria (the basis for each criterion is set out in Appendix 7). The same assessment is carried out for the MAR cross-check which Ofwat proposes to apply at PR24 (based on the DM) on a basis consistent with KPMG's *Use of market-to-asset ratios (MARs) as a cross-check in the context of regulatory price controls* Report (7 September).

154 The winner's curse is the tendency that the winning bidder pays an amount higher than the intrinsic value of an item
155 The agency problem is a conflict of interest between a company's management and the company's stockholders.
156 Control premium refers to an amount that a buyer is willing to pay in excess of the fair market value of shares to gain a controlling ownership interest.

Table 18: A high level assessment of potential cross-checks against criteria

Criterion	MFMs	MAR
Transparent		
Targeted		
Objective		
Incentive		
Consistent		

Source: KPMG analysis
Note: Green indicates that the cross-check meets the criterion well; Amber that it partially does so; and Red that it does not do so.

Overall, the assessment of the MFM asset pricing models against robust criteria indicates that estimates derived using the preferred MFM model offer robust estimates of returns and hence a good cross-check on the estimates of allowed returns. The MFM model:

- Is based on a methodology with stronger explanatory power than the primary methodology for estimation of returns (CAPM),
- Is based on a transparent methodology and can be grounded in established academic literature,
- Is a targeted cross-check on and unbiased estimator of CoE as the output of the MFM analysis is an estimate of CoE,
- Captures a more granular and nuanced view of risk for water companies, and
- Is the standard approach adopted by academics for the estimation of expected returns.

By comparison, MAR as the primary cross-check currently under consideration for PR24 does not represent a comparably robust cross-check. This is because:

- There are many unknowns in the determination of a company's value, and the calculated MAR cannot be solely attributed to a difference between investors' assumed return of equity and the allowed return. As a result, the MAR cross-check is not targeted.
- Academic literature and research are generally clear that MAR cannot be used to observe CoE without controlling for all other factors which influence companies' values. This is not possible in the case of MARs as the factors may not be quantifiable and controllable.
- The application of MARs by the regulator to revise the allowed CoE (whether up or down) could cause regulated companies to behave in ways which were not intended and not in the best interests of consumers.

In case of transaction MAR, there are additional issues with transparency and objectivity of the cross-check. The input information to derive traded MARs is generally not publicly available (especially for private companies) and MARs are also by nature biased towards the winning bidder's aims, assumptions, and strategic advantages.

The detailed assessment of MFMs and MAR against the criteria is set out in the tables below.

Table 19: Assessment of MFMs against cross-check criteria

Criterion	Assessment of MFMs	Suitability
Transparent	MFMs such as the q-factor model use publicly available observed data and prescriptive methodologies set out in seminal academic papers. They require limited judgment relative to other cross-checks and are not dependent on assumptions regarding the future (as they rely on past data only). The only judgement required is on how to tailor the factor construction approaches to the UK market (e.g., the frequency of accounting reporting) but this can be informed by existing academic papers and established methodologies as noted in Appendix 3.	
Targeted	MFMs are fully targeted as they assess the required returns directly based on the risks faced by water companies. MFMs are based on the same core underlying principle as CAPM and describe the return on an asset in terms of the risk of the asset with respect to a set of factors. MFMs ultimately represent extensions of CAPM i.e., they augment CAPM (which is based on the market factor only) with additional explanatory factors – i.e. represent a more granular analysis of systematic risk for water companies.	
Objective	This Report demonstrates that MFMs have stronger explanatory power than the CAPM, with better empirical performance than the CAPM across the two standard statistical tests. The genesis of MFMs was the consistent finding from academic research of a series of empirical shortcomings in the CAPM to explain observed returns. By design, MFMs are a more unbiased estimator of required returns than the CAPM.	
Incentive compatible	MFMs imply at least the same degree of compatibility with incentives and regulatory objectives as the CAPM and additionally provide statistically more robust evidence of the required returns.	
Consistent	MFMs are well-established and are widely adopted amongst both academics and practitioners to explain and estimate returns. In academia, MFMs have been used as the preferred asset pricing model for almost thirty years. MFMs are being used increasingly by practitioners to substitute and supplement the CAPM. For example, large assets managers, including those who have historically invested in regulated utilities, now use MFMs extensively to manage their portfolios. In the regulatory context, UKRN and CMA have recognised the stronger explanatory power of MFMs compared to the CAPM, however some regulators have previously dismissed them as too complex to apply and interpret.	

Source: KPMG analysis

Table 20: Assessment of MAR against cross-check criteria

Criterion	Assessment of MFMs	Suitability
Transparent	The information to calculate traded MARs—equity market value and net debt, which makes up the EV, and the RCV—often appear available. However, such figures do not necessarily reflect the MAR for the regulated business alone as listed companies undertake other activities beyond the regulated ones. In practice, it appears that the value decomposition cannot be effectively carried out. This is highlighted by the review of equity analysts' reports for the traded water companies, which shows that there is a wide range of estimates of the relative value of the regulated and unregulated businesses. This means that it is very difficult to reliably and consistently separate out the cost of capital as a driver of the underlying regulated MAR. Overall, therefore, MARs appear to only partially fulfil the criterion of transparency.	
Targeted	In order to be able to conclude that the 'true' rate of return assumed by an investor is different from the allowed return based on MAR, many long-term assumptions on the future cash flows, operational performance and others would need to be made with a high degree of certainty. This untargeted nature of MAR, particularly as it is difficult to isolate the cause of the premium from the 'noise', is inconsistent with the principles of better regulation introduced by the Better Regulation Taskforce. Overall, inferences from MARs do not meet the criterion for targeted regulation as they cannot reliably isolate the required information from other factors.	
Objective	Transaction MARs are by their nature biased towards the winning bidder's aims, assumptions, and strategic advantages. By their very nature these are not unbiased indicators of a company's value. Conversely, traded MARs (when taken in aggregate rather than one individual assessment) do reflect a much broader set of market expectations and assumptions for the company and companies in the sector and could thus be considered to having a wider, more informed view of the various valuation assumptions. Traded MARs also suffer from certain problems and limitations as efficient market signals but less so than private transactions. In this regard, it would suggest that MARs do partially meet the objective criterion but they also vary significantly over time and face the same limitations in linking observed or derived value to the allowed rate of return as a driver so they cannot be considered an unbiased indicator of the required return.	
Incentive compatible	The application of MARs by Ofwat to revise the allowed cost of equity (whether up or down) could cause regulated companies to behave in ways which were not intended and not in the best interests of consumers. Specifically, if it was feared that any outperformance – whether in totex, ODIs, or financing – which pushed the company's MAR above 1 (as explored in the stylised sensitivity analysis above) would subsequently result in a further reduction in the allowed cost of equity and the introduction of the 'ratchet effect'. This would further dampen incentives for improved performance as companies may fear that any rewards gained during a price control will be clawed back at the next price control through a reduced cost of equity. This effect would tend to suggest that MARs are not incentive compatible as a cross-check.	

Criterion	Assessment of MFMs	Suitability
Consistent	Academic literature and research has been generally clear that MAR cannot be used to observe cost of equity without controlling for other factors which influence companies' values. There is significant 'noise' in both traded markets and private transactions that cannot be directly controlled (behavioural biases, winners' curse, empire building mindsets, etc.) which directly influence the MAR. On this basis, while MARs have been occasionally used by regulators as a cross-check, such usage was only deemed reliable when part of wider set of indicators, and the academic dismissal of them as a basis for the cost of equity would suggest that they partially meet the consistency standard.	

Source: KPMG analysis

7 Appendix 1: Scope of work

Water UK has asked KPMG to develop a report on Multi-factor Models as a cross-check on allowed returns at PR24, to assist Water UK in its considerations regarding the PR24 DM and, in particular, a robust approach of cross-checking the CAPM-derived CoE that is best supported by the evidence provided by relevant financial literature, regulatory principles, and empirical analysis.

The evaluation of MFMs as a potential cross-check involves several steps to select the right model(s) with the best theoretical underpinnings and to inform the analytical approach to derive unbiased estimates of their impact on CoE. The Report assesses whether MFMs warrant inclusion in the cross-checks for PR24 and what weight they might merit in three steps:

- First, it considers (1) MFMs as an asset pricing tool, their development and evolution, (2) the rationale for these asset pricing models as cross-checks on the CAPM-derived CoE and (3) two of the leading MFMs in academic research as potential candidates for inclusion in cross-checks for PR24.
- Second, it develops an approach and methodology for empirical analysis based on two of the leading MFMs using UK data, covering data collection, model calibration and statistical robustness testing based on the approaches followed in academic research; and
- Third, it estimates returns for PR24 based on MFMs and consider the implications of the results from the MFM analysis and potential implications for the allowed cost of equity at PR24.

8 Appendix 2: Derivation of the FF5F

The analysis starts by considering a basic Dividend Discount Model of Valuation, where $\boldsymbol{M_0}$ is market value of the firm today and $\boldsymbol{D_t}$ is the expected dividend for year *t*:

$$\boldsymbol{M_0} = \sum_{t=1}^{\infty} \frac{\boldsymbol{D_t}}{(\boldsymbol{1}+\boldsymbol{r})^t}$$

As $\boldsymbol{D_t}$ is equivalent to "clean surplus" earnings[157], $\boldsymbol{Y_t}$, less the change in book value, $\boldsymbol{dB_t}$, the above equation could be rewritten as follows:

$$\boldsymbol{M_0} = \sum_{t=1}^{\infty} \frac{(\boldsymbol{Y_t} - \boldsymbol{dB_t})}{(\boldsymbol{1}+\boldsymbol{r})^t}$$

This simple relationship implies two things. First, for any given market value and investment, a higher earnings number, Yt implies a higher expected return, r. Given standard asset pricing theory, where required returns are influenced only by systematic risk factors, this would imply that the higher profitability is required to compensate for some (directly unobserved) systematic risk factor.

The second implication is that for any given market value and profit, a higher investment, dB_t, is associated with a lower ex higher expected return, *r*.

Expressing this in terms of the market to book ratios yields, $\frac{\boldsymbol{M_0}}{\boldsymbol{B_0}}$, we have the following derivation:

$$\frac{\boldsymbol{M_0}}{\boldsymbol{B_0}} = \sum_{t=1}^{\infty} \frac{(\boldsymbol{Y_t} - \boldsymbol{dB_t})/(\boldsymbol{1}+\boldsymbol{r})^t}{\boldsymbol{B_0}}$$

The equation presented above, as Fama and French (2015) note, implies the following relationships:

- When we hold everything constant except $\boldsymbol{M_0}$ and $\boldsymbol{r}$: a lower $\boldsymbol{M_0}$ (i.e., a higher $\frac{\boldsymbol{B_0}}{\boldsymbol{M_0}}$ ratio) implies a *higher* expected return $\boldsymbol{r}$
- When we hold everything constant except $\boldsymbol{Y_t}$ and $\boldsymbol{r}$: a higher $\boldsymbol{Y_t}$ implies a *higher* expected return $\boldsymbol{r}$
- When we hold everything constant except $\boldsymbol{dB_t}$ and $\boldsymbol{r}$: a higher $\boldsymbol{dB_t}$ implies a *lower* expected return $\boldsymbol{r}$

These relationships hold regardless of whether pricing is rational or irrational, as the implication is merely that if two firms have different share prices and the same expected dividends (which derive from earnings and investment), the one with the lower price must have higher risk reflected in the required return $\boldsymbol{r}$.

For Fama and French, the relationships suggest that there are three "natural" components to risk pricing, which includes B/M ratio, profitability, and investment. When considering these components with the standard CAPM market factor ($R_{Mt} - R_{Ft}$) alongside a size factor, which reflects the empirically observed result of small stocks having higher returns than large stocks[158], we arrive at the model of returns as follows:

Rit – RFt = ai + bi(RMt – RFt) + siSMBt + hiHMLt + riRMWt + ciCMAt + eit.

157 "Clean surplus" means that all valuation changes (e.g., depreciation and revaluation) in book value must flow through the P&L account.
158 Explanations for this differ, but liquidity is a reasonably compelling one.

where:

- *SMBt* is the difference between the returns on diversified portfolios of small stocks (S) and big stocks (B)
- *HMLt* is the difference between the returns on diversified portfolios of high book to market (H) and low book to market (L) stocks
- *RMWt* is the difference between the returns on diversified portfolios of stocks with robust (R) and weak (W) profitability
- *CMAt* is the difference between the returns on diversified portfolios of the stocks of low (conservative, C) and high (aggressive, A) investment firms

9 Appendix 3: Comparison of sample formation approaches

This appendix sets out a detailed comparison of sample formation approaches between Fama-French, Hou et al and this Report. Overall, the approaches established by these papers have been followed to the extent possible given the differences between the US and UK markets. Where it was not possible to align the approach exactly due to these differences, the Report has preserved the principles of the US approach whilst adjusting for UK specific circumstances. The appendix only covers sample formation approaches as in all other respects (such as factor formation), the Report has maintained full consistency with Fama-French and Hou et al.

Table 21: Relevance and applicability of the past criticisms of MFMs[159]

	Hou et al	Fama-French	This Report	Commentary on relevance and applicability
Breakpoints for the data	Breakpoints established using the 30th and 70th percentiles of *NYSE stocks* for the book-to-market, profitability and investment sorts, and the median of these stocks for size.		Breakpoints established for the "investable universe" which is the largest 250 firms each year by market capitalisation.	Whilst the main market listing on LSE has a minimum market capitalisation requirement, this is much lower than the threshold applied at NYSE[160]. As a result, LSE includes companies that would not be included in the population of stocks considered by Hou et al and Fama-French. Focusing on the largest 250 listed companies in the UK is a means of proxying the cut-off implied in Fama-French and Hou et al's focus on NYSE stocks. This is similar to the approach taken in Gregory et al (2013).[161]

159 CMA (Aug 2020), NATS (En Route) Plc/CAA Regulatory Appeal Final report - Appendix D:
Technical note on betas and gearing.CAA (Oct 2013), Estimating the cost of capital: a technical appendix to the CAA's Final Proposal for economic regulation of Heathrow and Gatwick after April 2014.
Europe Economics (Nov 2009,. Cost of Capital and Financeability at PR09 Updated Report by Europe Economics.
Europe Economics (Mar 2007), CAA's price control reference for Heathrow and Gatwick airports, 2008-2013, Supporting paper II Cost of capital – analysis of responses to CAA's initial proposals.
Competition Commission (Nov 2007), BAA Ltd - A report on the economic regulation of the London airports companies, Appendix F (cost of capital).
Smithers & Co (Sept 2006), Report on the Cost of Capital provided to Ofgem by Stephen Wright, Robin Mason, Steve Satchell, Kenjiro Hori and Meltem Baskaya.
Ofwat (Aug 2004), Future water and sewerage charges 2005 -10 Final determinations, Appendix 5.
Smithers & Co (Feb 2003), A Study into certain aspects of the cost of capital for regulated utilities in the UK - A report prepared for OFT, CAA, OFWAT, Ofgem, Oftel, ORR and OFREG.
Competition Commission (Feb 2003), The Competition Commission's report on the charges made by mobile operators for terminating calls - 18 February 2003.

160 Listing on the main market of the London Stock Exchange - An overview (withersworldwide.com)

161 With the difference being that the 2014 study used the top 350 firms. However, the available data from the combination of resources considered in this Report yielded fewer than 350 companies in some years, indicating that a focus on the largest 250 companies was appropriate throughout the horizon covered by the analysis.

	Hou et al	Fama-French	This Report	Commentary on relevance and applicability
Exclusion of AIM listed stocks	N/A as an equivalent market does not exist for NYSE		Excluded	Excluded as AIM stocks have not historically been viewed as investible by many fund managers (due to e.g., high failure rates based on findings from Gregory et al (2010)[162] and poorer standards of reporting) and the premise behind the factor analysis is to build factors to price stocks in the investible universe. This follows Gregory et al (2013), who note: *"we only include Main Market stocks and exclude financials, foreign companies and AIM stocks following Nagel (2001) and Dimson, Nagel and Quigley (2003)".*
Other exclusions	"Financials" and stocks with negative book to market ratios			
Earnings data frequency	Quarterly	Annual	Annual	Quarterly reporting is not available in the UK and neither is semi-annual data reliably available. The use of annual data is consistent with Fama and French.
Financial year ends (FYE) for accounting data	31 December		31 March	The Report preserves the principles of the US approach whilst adjusting for UK specific circumstances. In an approach long established by FF, six months is allowed to elapse between the market capitalisation values used, and the FYE data used. This means that in all cases, market capitalisation data is observed at end June year *t*, and matched with FYE data available at December year *t*-1. However, the most common FYE in the UK is 31st March, which leads Gregory et al (2013) to adopt a different approach, namely matching *September* market capitalisation data for year *t* with FYE data

162 Gregory, Alan and Guermat, Cherif and Al-Shawawrah, Fawaz, UK IPOs: Long Run Returns, Behavioural Timing and Pseudo Timing (2009-11). Journal of Business Finance & Accounting, Vol. 37, Issue 5-6, pp. 612-647, June/July 2010

	Hou et al	Fama-French	This Report	Commentary on relevance and applicability
				available at *March* year *t*. This is uncontentious in the UK (as this was the approach used in the University of Exeter factor data), but it does lead to a potentially confusing notation situation[163].
Investment measure	Change in total assets / opening total assets			
Profitability measure	Net income before extraordinary items divided by opening book equity	Operating profit (defined as revenue – total operating expense – net interest expense) divided by book equity[164]	Both measures replicated	Where extraordinary items are missing in Datastream, it is assumed to be zero. For some early years operating *profit* was missing but operating *margin* was present. In such cases it was assumed (reasonably) that operating profit was the product of operating margin and revenue.
Size measure	Market value			
Value measure	N/A	Book to market		

Source: KPMG analysis

163 Specifically, for the FF definition of profitability, the profit number is ROE^{FF}_{t-1} = Profit FF_{t-1}/TA_{t-1} whereas in terms of this Report, March belongs to year t, so this becomes: ROE^{FF}_{t}= Profit $^{FF}_{t}/TA_{t}$, whilst for the Hou et al definition the definition is ROE^{q}_{t-1} = Profit $^{q}_{t-1}/TA_{t-2}$, which for this Report (notationally) becomes: ROE^{q}_{t} = Profit $^{q}_{t}/TA_{t-1}$. The same logic applies to other factors.

164 Note that this definition implies no lag, i.e. it does not appear to be defined on opening equity.

10 Appendix 4: Differences in factor construction

This appendix sets out the key differences in factor construction for the additional factors employed in the FF5F and the q-factor model.

The factor construction for the additional factors in both models based on UK data requires three steps: (1) the measurement of the factors; (2) the formation of the factor portfolios; and (3) the calculation of the factor premia. In principle, the factor premium for any given factor should be consistent across both models where the specification of all three steps is consistent.

The table below provides an overview of the key differences in factor construction for the additional factors in both models, with reference to the three steps above.

Table 22: Differences in factor construction for additional factors in FF5F and q-factor model

Factors	Measurement of the factors*	Formation of the factor portfolios	Calculation of the factor premia
Size	No differences In both cases, size is measured by: Market value as of September $year_t$	Differences The factor portfolios for each FF5F factor are formed based on the intersection of two factors (size with each of investment, profitability and value). By contrast, the factor portfolios for each q-factor are formed based on the intersection of three factors (size, investment and profitability).	No differences For each factor, the premium is calculated by (1) averaging the daily (or monthly) returns of portfolios that belong to the characteristic associated with higher risk (e.g. small size) and the characteristic associated with lower risk (e.g., big size); then (2) deducting the latter from the former to construct the factor premium.
Investment	No differences In both cases, investment is measured by: (Total assets as of March $year_t$ – total assets as of March $year_{t-1}$) / total assets as of March $year_{t-1}$		
Profitability	Differences Profitability is measured: In FF5F by, Operating profit** as of March $year_t$ / book value of equity as of March $year_t$ In q-factor by, Net income before extraordinary income as of March $year_t$ / book value of equity as of March $year_{t-1}$		
Value	Differences Included in FF5F but not in q-factor.		

Source: KPMG analysis
* In all cases, the measurement is based on the largest 250 stocks
** Operating profit = (revenue – total operating expense – net interest expense)

In both models, the size and investment factors are measured in the same way however the profitability factor is measured differently. The measurement of the profitability factor differs in two ways: (1) in FF5F, profit (numerator) is close to a pre-tax operating profit definition whereas in q-factor it is a post-tax pre-extraordinary items definition; and (2) in FF5F, book value of equity (denominator) is taken from year t whereas in q-factor it is taken from year t-1. The value factor is only incorporated in FF5F and therefore no comparison across the models can be drawn.

The factor portfolios for each additional q-factor and FF5F factor are formed differently. At the same time, the factor premia for the additional factors in both models are calculated in the same way. The formation of factor portfolios and calculation of factor premia for each model is discussed in greater detail in Section 5.2.

In conclusion, for every additional factor, the specification of at least one of the three steps differs across the models and thus none of the factor premia across the models match.

11 Appendix 5: Estimating the factor risk premia for the q-factor model

This section sets out the factor risk premia for the q-factor model and the statistical significance of these premia.

The first full year of observation starts in October 1981, giving 40 complete years of data plus one-half year (to March 2022). In order to assess the statistical significance of these data, the monthly returns and their standard errors are calculated. The table below sets out monthly statistics in relation to the statistical significance of q-factor model factor risk premia.

Table 23: Assessment of statistical significance of q-factor model premia (based on monthly data)

Variable	No. observations	Mean	Standard Error	Significance
MRP	486	0.556%	0.195%	**
Size	486	0.172%	0.136%	ns
Investment	486	0.207%	0.113%	*
ROE	486	0.238%	0.125%	*

Source: KPMG analysis

The table above shows that in the UK, MRP is significant at the 5% level, whilst the Investment and ROE factors are significant at the 10% level. However, size is not statistically significant.

In principle, factor prices should be established on the annualised compound averages of monthly returns in each year, using the longest run of data available. Ideally, factor premia would be estimated over the same period as the market risk premium is estimated, i.e. since 1900. However, such data is not available because of the lack of accounting data going back that far and this study is limited to data since 1981. In the US, q-factors are available since 1961, and the Report uses those as a cross-check. The table below sets out a comparison of factor premia based on UK and US data.

Table 24: Comparison of UK and US factor premia

Country	Size	Investment	ROE
UK	2.05%	2.81%	3.00%
US, since 1967	3.57%	4.51%	7.00%
US, since 1981	1.63%	4.12%	6.60%

Source: Averages based on KPMG analysis for UK data and calculated averages based upon the US data available at Factors (global-q.org)

This data suggest that the investment factor is plausibly close to the US post 1981 data. The profitability factor premia are roughly half that of the US. The factor premia (size excepted) do not vary substantively by extending data back from 1981 to 1967 based on the US data.

12 Appendix 6: Relevance of previous regulatory commentary on MFMs

This appendix provides an overview of key criticisms of MFMs based on past regulatory precedent and considers the extent to which they could remain relevant and applicable to the analysis set out in this Report. The criticisms have been collated and aggregated into themes based on publicly available regulatory documents which comment on MFMs. These criticisms, their sources and comments regarding their relevance and applicability to analysis in this Report are set out in Table 25.

It is important to note that previous analysis of MFMs in a regulatory context has largely been in early 2000s and as a replacement for the CAPM. As a result, these analyses did not consider the models which are now favoured in academia as the most empirically and theoretically robust and did not consider their merits as cross-checks to the CoE implied by CAPM. Overall, the assessment finds that previous criticisms are of limited relevance to the analysis set out in the Report, primarily due to the robust selection and testing of the MFMs considered as potential cross-checks.

Table 25: Relevance and applicability of the past criticisms of MFMs[165]

Criticism	Source				Commentary on relevance and applicability
	Sector	Price control	Year	Author	
Lack of theoretical foundation	Water	PR04	Aug 2004	Ofwat	This criticism is of limited relevance to the analysis set out in the Report as the investment and profitability factors included in the q-factor model have strong theoretical foundations: • The economic foundation for the q-factor model is effectively the NPV rule of Corporate Finance. • The inclusion of investment and profitability factors is further supported by vast accounting and valuation literature such as Peasnell (1982).
	Telecoms	Mobile termination charge	Feb 2003	Competition Commission	

165 CMA (Aug 2020), NATS (En Route) Plc/CAA Regulatory Appeal Final report - Appendix D: Technical note on betas and gearing.
CAA (Oct 2013), Estimating the cost of capital: a technical appendix to the CAA's Final Proposal for economic regulation of Heathrow and Gatwick after April 2014.
Europe Economics (Nov 2009,. Cost of Capital and Financeability at PR09 Updated Report by Europe Economics.
Europe Economics (Mar 2007), CAA's price control reference for Heathrow and Gatwick airports, 2008-2013, Supporting paper II Cost of capital – analysis of responses to CAA's initial proposals.
Competition Commission (Nov 2007), BAA Ltd - A report on the economic regulation of the London airports companies, Appendix F (cost of capital).
Smithers & Co (Sept 2006), Report on the Cost of Capital provided to Ofgem by Stephen Wright, Robin Mason, Steve Satchell, Kenjiro Hori and Meltem Baskaya.
Ofwat (Aug 2004), Future water and sewerage charges 2005 -10 Final determinations, Appendix 5.
Smithers & Co (Feb 2003), A Study into certain aspects of the cost of capital for regulated utilities in the UK - A report prepared for OFT, CAA, OFWAT, Ofgem, Oftel, ORR and OFREG.
Competition Commission (Feb 2003), The Competition Commission's report on the charges made by mobile operators for terminating calls - 18 February 2003.

Criticism	Source				Commentary on relevance and applicability
	Sector	Price control	Year	Author	
					• Fama and French agree the investment and profitability factors are theoretically justified and incorporate them in the FF5F. Hou et al incorporate the size factor in the q-factor model based on empirical evidence that it helps to explain observed returns, but they note its incremental effect in capturing return anomalies is small. As such, they consider the size factor only plays a secondary role in the model.
Lack of explanatory power (across factor loadings, premia and overall)	Water	PR09	Nov 2009	Europe Economics	This criticism is of limited relevance to the analysis set out in the Report. This is because: • The Report considers two of the leading MFMs in academic literature which have stronger empirical performance than the CAPM based on US data. • The Report tests the performance of the models for (1) all companies used to calibrate the model and (2) for regulated utility stocks. • In the case of (1) The Report undertakes a two-stage statistical testing to evaluate the empirical performance of q-factor model relative to the CAPM based on UK data. The analysis, using tests widely recognised by the academic literature, finds that the q-factor model performs better in tests and has higher explanatory power than CAPM. Both the factor premia and factor loadings (for all companies) are for almost all factors statistically significant and are useful in explaining observed returns • On average, the adjusted R^2 of the q-factor model is 64.5%, meaning that 64.5% variability of portfolio returns could be explained by the q-factor model, while for the CAPM, the adjusted R^2 is 45.2%. • In the case of (2) statistical analysis of the factor loadings for UK regulated utilities indicates that the factors in the q-factor model are individually significant in most cases and jointly significant in all. This suggests that additional factors included in the q-factor model are useful in explaining observed returns, individually and jointly.
	Airports	Q5	Nov 2007	Competition Commission	
	Airports	Q5	Mar 2007	Europe Economics	
	Electricity Transmission	TPCR4	Sept 2006	Smithers & Co	
	Telecom	Mobile termination charge	Feb 2003	Competition Commission	

Criticism	Source				Commentary on relevance and applicability
	Sector	Price control	Year	Author	
Exposed to over-fitting	Airports	Q6	Oct 2013	CAA	This criticism is of limited relevance to the analysis set out in the Report: • "Over fitting" refers to a situation where too many independent variables fit into a limited amount of data. • In the case of the q-factor model, there are only 4 independent variables, while there are more than 40 years of daily market wide UK stocks return data. In contrast, previous analysis has used monthly data which implies fewer observations and greater exposure to risk of overfitting.
Exposed to time period dependency	Airports	Q6	Oct 2013	CAA	This criticism is of limited relevance to the analysis set out in the Report as the analysis has been undertaken for two cut off dates and several estimation windows.
Requires more judgment than the CAPM (which factors to include and how to measure the risk-free rate)	Airports	Q5	Mar 2007	Europe Economics	This criticism is of limited relevance to the analysis set out in the Report: • The models explored in the Report are largely aligned in terms of included factors. • In terms of estimation of the risk-free rate, the regression analysis to estimate factor loadings uses a short-term rate consistent with academic convention. However, the derivation of CoE for comparison between the CAPM and q-factor maintains consistency with Ofwat's approach to set the risk-free rate.
Difficult to populate	Air Traffic Control	RP3	Aug 2020	CMA	The analyses set out in this Report indicates that it is possible to populate MFMs based on UK data and generate statistically robust results. The data used to develop the MFM analysis in this report could in due course be made public and updated annually.

Source: KPMG analysis

In addition to the above, previous analyses relied upon monthly data to estimate the factors and identified the use of daily data as a potential future area of improvement. The Report utilises models calibrated based on daily data.

13 Appendix 7: Criteria for evidence to inform setting of allowed rate of return

One of the common objectives of public utility regulation is to ensure that investors are sufficiently remunerated, but not more than necessary, for the funds they provide to a regulated firm.[166] This ensures that investors have adequate incentives to invest in regulated public utilities, but consumers are not paying more than required for the services they receive.

Cross-checks can be an effective tool in an economic regulator's toolkit to assist in the calibration of regulatory determinations. However, not all crosschecks are effective, accurate, or unbiased. Therefore, there is a need to develop a set of criteria that can distinguish and judge the reliability and effectiveness of different cross-checks.

The cross-check criteria can be developed to check if the evidence can be relied upon to support many estimates, including the cost of equity. The principles identified aim to be robust, defensible and in line with established economic regulatory framework. The process for developing the criteria used in this report was as follows:

First, we had regard to existing regulatory principles, particularly best practice guidance that is legislated in the current water economic regulation framework. The principles of good regulation, that are prescribed in legislation include the following: **proportionality**, **accountability**, **consistency**, **transparency** and **targeting**. [167] The essence of three of these principles is relevant for consideration of robustness and effectiveness of cost of capital cross-checks.

Second, we considered the principles that UKRN have regard to in determining the cost of capital, which are designed to ensure that regulators take an effective and efficient approach in setting the cost of capital in all their respective sectors. The cost of capital principles include **consistency** between regulators, being **risk reflective**, whereby the reward reflect the allocation of risk, relying on **evidence**, facilitating **investment** in the interest of consumers, being clear and transparent in **communication** and **review** the principles and approaches regularly so that they are in line with **good practice.**[168]

Third, recent regulatory determinations were reviewed, such as decisions by Ofwat, Ofgem and the CMA, to infer how cross-checks were used, the intent and the context surrounding their use. The review provided guidance on the current issues, including contention surrounding their use and the resulting outcomes. For example, in the CMA's RIIO-2 final decision on the cost of equity, concerns around cross-checks related to lack of transparency, their interdependence on the variable they are intended to assess or their inconsistency with regulatory precedent.[169] Relying on recent regulatory determinations means that the criteria is contemporaneous and is in line with regulatory best practice.

Fourth, a cross-check is a regulatory tool that would support the mechanism that discharges regulators' financing and consumer duties. The focus was to consolidate and identify the principles that would characterise a regulatory tool rather than what would describe the manner with which a regulator regulates (i.e., the principles of proportionality and accountability were deemed to be more pertinent to the regulator's approach to economic regulation rather than a regulatory tool criterion).

166 This is in reference to the financing duty of Ofwat. Water Industry Act 1991 Section 2.
167 Better Regulation Task Force, Principles of Good Regulation available here. Ofwat must have regard to these principles according to the Water Industry Act 1991 Section (2)
168 UKRN, Cost of capital and price controls. See here.
169 For example, the use of nominal gilts as cross-check to estimate and the use of SONIA swap as a cross-check for the RFR. See CMA, Final determination: Volume 2A: Joined Grounds: Cost of equity, October 2021

Based on the four-step process above, the developed set of criteria below is objective and can be used to evaluate cross-checks in any upcoming price control review. The criteria set out above are used to evaluate cross-checks in Section 6 of this Report:

- ***Transparent*** *in that it would use information that can be widely observed and verified, and the results can be replicated consistently, including its approach and calculation methodology. For a cross-check to be transparent, any of the underlying calculations and mechanics need to be traceable, verifiable and the variation in outcomes (if any) explainable by reference to plausible and defensible assumptions. Transparency is an important element of the criteria as it ensures that all parties understand how any given assumptions influence a cross-check, which is in line with clear and transparent regulatory determinations.*
- ***Targeted****, which is the extent to which the indicator can isolate the effect in question from other factors and can therefore give accurate results. It would mean that there is limited to no doubt that the cross-check is assessing what is supposed to be assessed. Significant associated noise, surrounding the assessed variable, would render it difficult to distinguish between what is being measured and other non-relevant factors and would not be robust justification and evidence. In addition, it would provide an indication of accuracy, or how close the observed value is to the true value.*
- ***Objective*** *and can be relied on as an unbiased indicator. Objectivity ensures that the cross-check's underlying assumptions do not skew towards a specific pre-determined outcome and that there is a degree of independence between the cross-check and the variable that it is intended to assess.*
- ***Incentive compatible****, particularly whether it is compatible with regulatory objectives, one that facilitates investment in consumers' interests. Incentive based regulation is based on a series of carefully calibrated incentive mechanisms applied to costs and service levels. All incentives are designed to provide an economic motivation for companies to apply additional effort where the cost of that effort is less than the benefit to consumers. This aligns incentives between consumers and network companies to promote overall welfare gains for both. An indicator that is likely to distort incentives, or one that is associated with a ratchet effect,* [170] *or discourage investments in infrastructure that consumers want is unlikely to be incentive compatible.*
- ***Consistent*** *with established regulatory precedent and academic research. Consistency as a principle is largely based on the principles of good regulation, cost of capital principles and is an important consideration for investors, regulated entities, and regulators alike. It ensures that cross-checks would be used in a predictable way that provides certainty and stability to those who are regulated. It would also provide confidence that regulatory approaches are best practice, or in the case of new regulation that it considers international precedent if any.*

170 Ratchet effect is a term that is used to describe the approach of using current performance as a partial basis for setting future targets, which creates a dynamic incentive problem for the enterprise. Martin L. Weitzman, The "ratchet principle" and performance incentives, the Bell Journal of Economics.

www.kpmg.com/uk