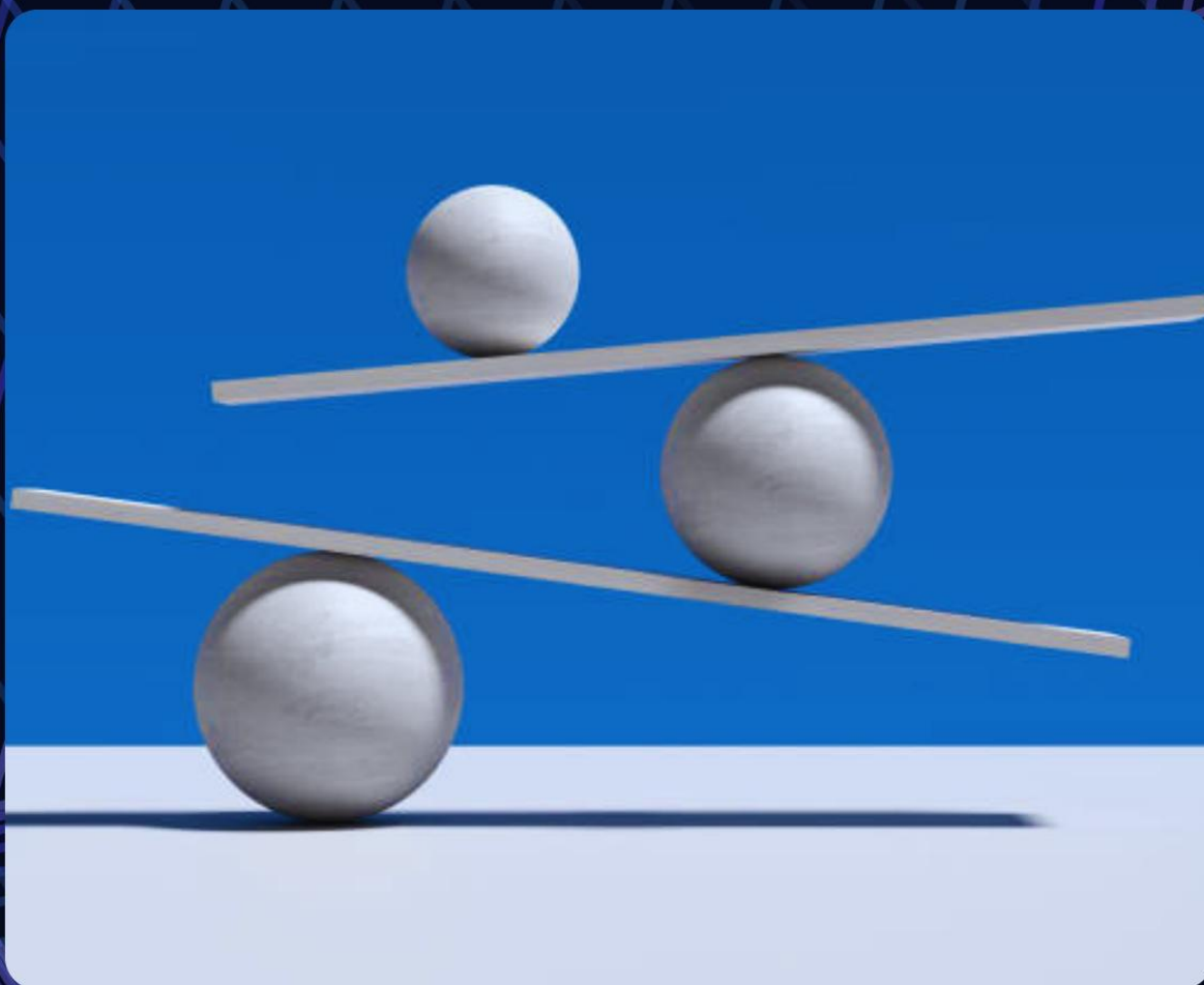


Stable Total Market Return or Stable Equity Risk Premium?

A Framework for Exercising Regulatory Judgement Under Uncertainty. A discussion paper on the estimation of the cost of equity in economic regulation



Stable Total Market Return or Stable Equity Risk Premium?

A Framework for Exercising Regulatory Judgement Under Uncertainty

A discussion paper on the estimation of the cost of equity in economic regulation

- Economic regulation is decision-making under uncertainty, and the cost of equity is a clear example: no technique can pin it down with precision, yet small differences matter a great deal to consumers and investors.
- Much of the UK debate over ‘Stable TMR’ versus ‘Stable ERP’ asks whether the historical record favours one formulation over the other – but historical data cannot discriminate between them with statistical confidence, however carefully realised returns are analysed.
- This paper sets out a framework for choosing between them, by triangulating three complementary perspectives: the historical record, contemporary market evidence, and theoretical coherence.
- Within the market evidence, valuation data says little about how returns should move as interest rates change. Corporate bond yield data may bear more directly on that question and merit closer attention, though noise makes it less helpful in evaluating overall model outputs.
- The technical and academic evidence leans towards the Stable ERP approach as the more defensible formulation. This perspective is important – particularly considering investors’ incentives at a critical time for incremental infrastructure investment.
- That said, the UK’s recent regulatory practice of Stable TMR has virtues of its own – including continuity of approach and a ‘through-the-cycle’ rationale, which have also been welcomed by past investors, and provide for more stable consumer outcomes. This paper does not evaluate or consider the trade-off with those practical considerations – focusing instead on its narrower technical conclusion.
- This is published as an independent contribution to the UK Regulators Network current consultation on the topic. It has not been prepared for, or at the request of, any party to that consultation.

Introduction

Economic regulation involves decision-making under uncertainty. Regulators routinely set prices, financing boundaries and investment incentives on evidence that is informative but rarely conclusive. The cost of equity illustrates the point sharply: differences of a few basis points matter materially to consumers and investors, yet no technique can determine the appropriate allowance with precision. The question is not *whether* judgement should be exercised, but *how* it should be exercised when there is reasonable disagreement about what the evidence implies.

A current UK regulatory debate centres on a choice between two formulations of the Capital Asset Pricing Model (CAPM: Sharpe, 1964): the ‘Stable Total Market Return’ (Stable TMR) approach and the ‘Stable Equity Risk Premium’ (Stable ERP) approach – terminology used in the UK Regulators Network’s current [consultation](#). Much of the material presented by stakeholders within that debate considers whether the historical record favours one formulation over the other. That framing risks overlooking a more fundamental question. This paper does not ask:

Can the historical record establish that one formulation of the CAPM is objectively correct?



but rather, asks:

How should regulators choose between competing formulations when no single source of evidence is capable of resolving the question conclusively?

Two distinct questions are at stake here, and this paper is concerned mainly with the first: which formulation is better supported by the historical record, market evidence and theoretical coherence, taken together? A second, separate question – whether the practical benefits of regulatory continuity and stability should outweigh that technical conclusion – is a legitimate question in its own right, and one this paper returns to only briefly, towards its close.

Confidence in questions of this kind rarely rests on one source of evidence; it grows when independent sources converge. This paper develops a framework built on three complementary perspectives:

- The Historical Record – what can we infer from long-run empirical evidence?
- Contemporary Market Evidence – what are investor behaviours telling us today?
- Theoretical Coherence – what is the most coherent way to interpret both?

This paper is published as an independent discussion paper. It has not been prepared for, or at the request of, any party to the UKRN consultation, and its conclusions reflect Flint Global's own methodological assessment. It addresses the Stable TMR/Stable ERP choice and briefly considers the value of two cross-checks – the 'Market-to-Asset Ratio' (MAR) and the comparison of returns on debt and equity. It does not address other aspects of the consultation, such as its other cross-checks, its framework for assessing them, aiming up, or the inflation leverage effect on the cost of debt; these merit separate treatment.

A Framework for Evaluating Alternative CAPM Formulations

Where sources of evidence are each informative but imperfect, judgement should not depend predominantly on any one of them. This reflects a broader principle of empirical enquiry: confidence rarely arises because a single dataset or methodology proves a proposition beyond doubt, but because independent approaches – each with recognised strengths and limitations – converge on a common conclusion. Independence matters, because genuinely different sources are less likely to share the same biases. Where sources diverge, that divergence is informative in itself: rather than identifying one perspective as authoritative, it clarifies where judgement must play the larger role.

The question is therefore not simply what the evidence says, but what conclusion is supported when the available evidence is considered in the round.

Three complementary perspectives

The framework evaluates competing CAPM formulations through three questions: what can we infer from the historical record; what are markets telling us today; and what is the most coherent way to interpret both. Each reflects a distinct source of evidence, each contributes something the others cannot, and each has limitations that prevent it resolving the problem alone.



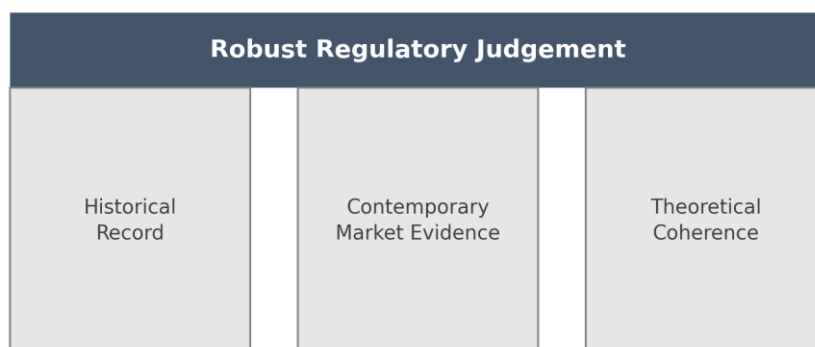


Figure 1: Three complementary perspectives – the historical record, contemporary market evidence and theoretical coherence – converging on robust regulatory judgement.

Where they converge, confidence increases; where they diverge, regulators gain a clearer view of the uncertainty that remains. The remainder of the paper applies this to the CAPM debate, beginning with the historical record – the longest and most comprehensive source of evidence on realised equity returns, and the empirical foundation on which any cost-of-equity estimate must be built.

It is worth being clear about what this framework can and cannot settle. Triangulating historical, market and theoretical evidence can identify which formulation is the more technically and academically defensible one – a focal point for this paper. However, it cannot, on its own, weigh that technical conclusion against the separate, practical benefits of regulatory continuity, predictability, and the avoidance of unnecessary change. That second, more overtly policy-facing judgement is an important consideration and is one this paper turns to briefly, having set out the technical case in full.

The Historical Record

What can we infer from the historical record?

Long-run evidence on realised equity returns provides the empirical foundation for virtually all cost-of-equity estimates (e.g. Dimson, Marsh and Staunton, 2002, 2025). No other source captures such a broad range of economic conditions, monetary regimes and financial crises. But its importance should not be mistaken for its ability to resolve every aspect of the estimation problem: the current CAPM debate appears to implicitly assume that sufficiently long datasets can establish which formulation of the CAPM, and the resulting estimates, are more accurate. That assumption deserves scrutiny.

Estimating parameters and selecting models

Marshalling the evidence from the historical record involves two related but different statistical exercises . Parameter estimation: given a long sequence of observations, what is the best estimate of the long-run average return; and model selection: which economic framework best interprets those returns, and provides the more reliable basis for estimating future expected returns?

Success in the first does not imply success in the second. While historical data clearly become more informative about long-run averages as more accumulates, the same data may still remain comparatively weak at discriminating between two closely related conceptual frameworks.

The signal and the noise

Expected returns cannot be observed directly and evolve only gradually. Realised returns reflect expected returns plus the cumulative effect of unexpected shocks, and so fluctuate far more. This is a signal-to-



noise problem (Merton, 1980): the greater the noise relative to the signal, the more data is needed to estimate the underlying parameter with confidence.

Equity markets display exactly this feature – substantial annual volatility against expected returns that may vary comparatively little. Even very long datasets therefore contain less statistical information about expected returns than their duration suggests. Fama and French (2002) found their dividend- and earnings-based estimates of the 1951–2000 US equity premium diverged sharply from the premium implied by average realised returns – consistent with Mehra and Prescott's (1985) finding that realised returns are hard to reconcile with plausible underlying risk aversion. Dimson, Marsh and Staunton's own decomposition (2002, 2025) attributes part of the twentieth century's realised premium to a valuation-multiple expansion they regard as unlikely to repeat – which, if excluded, implies a lower prospective premium than the raw historical average suggests. All illustrate that reasonable, different ways of reading the same historical data can yield materially different views of the premium investors should expect going forward for bearing equity risk.

A question of scale

Annual equity volatility has historically been of the order of 20% (Dimson, Marsh and Staunton, 2002, 2025), while regulatory debates on the cost of equity typically concern differences of tens of basis points – fundamentally different orders of magnitude. Even more than a century of annual observations leaves confidence intervals measured in percentage points, not basis points, as Table 1 illustrates (figures are illustrative rather than definitive; what matters is the order of magnitude).

Annual volatility	Years of observations	Approx. standard error of mean observation	Approx. 95% confidence interval
20%	25	4.0%	±8.0%
20%	50	2.8%	±5.6%
20%	100	2.0%	±4.0%
20%	150	1.6%	±3.2%

Table 1: Illustrative confidence intervals for annual equity market returns

The volatility data should be seen alongside the observed mean annual returns of 6.6% and 8.4% in the UK and US equity markets respectively. To interpret the table data crudely:

If we assume that the past is an indicator of the future, based on around 120 years of evidence, we can be 95% confident that the expected return on equity in the UK lies between 3.0% and 10.2%.

These confidence levels are not overwhelming. Any conclusions drawn from the historical dataset should therefore be viewed with appropriate caution and due recognition of its statistical limitations.

It is uncontroversial to state that any comparison of different *conceptual models* for explaining variation in returns over the period is beset by similar statistical uncertainty. A more direct test targets the parameter that specifically distinguishes Stable TMR from Stable ERP: the slope of TMR on the risk-free rate, which is zero under Stable TMR and one under Stable ERP. The standard error of that slope, estimated from realised annual returns, is of the order of 0.6 to 0.9. A slope of zero and a slope of one are therefore separated by little more than one to two standard errors: it is statistically difficult to distinguish from a century of historical data, in a more direct sense than the simpler illustration above.

Annex A sets out a fuller version of the evidence behind this, targeting directly the comparison of the Stable TMR and Stable ERP CAPM formulations.



Even further caution may sometimes be needed. Analytical conclusions – either on model form or parameter values – that involve subdivision of the already statistically weak realised returns dataset may be even less reliable than suggested by the statistics above. Efforts designed to support further precision have sometimes been based on such an approach and should be interpreted with great care. A specific example of this, linked to the ERP-RFR correlation, is covered below.

The role of historical evidence

None of this diminishes historical analysis – it instead clarifies its role. Historical evidence anchors judgement in long-run market experience, narrows the range of plausible estimates and constrains purely theoretical reasoning. What it is less well-placed to do is identify, in isolation, a uniquely correct CAPM formulation – not through any weakness in the record itself, but because of the statistical properties of the data. It follows that, where one important source of evidence carries unavoidable statistical uncertainty, confidence should depend increasingly on whether independent perspectives converge. The next section turns to the second pillar: contemporary market evidence, which addresses the return investors require today rather than what they have earned historically.

Contemporary Market Evidence

What are investors telling us today?

Regulators are not simply seeking to explain realised returns; their actual task is prospective – estimating the return needed to attract capital for future investment, at no more cost to customers than necessary. Regulators typically treat contemporary market evidence as a mere cross-check on historical analysis. However, that underplays its importance: while historical evidence asks what investors have earned over long periods, market evidence offers information to help inform us of what they currently require – different questions, answered with different information. Where the two converge, confidence increases because two genuinely independent perspectives agree, not because the same evidence has been examined twice.

Market prices, like historical returns, do not interpret themselves. They reflect investors' collective expectations about future cash flows, financing conditions, operational performance, macroeconomic developments, regulatory risk and future investment opportunities – so extracting a cost-of-equity signal requires an economic framework capable of separating these multiple influences.

Two examples help illustrate this point:

Enterprise value and regulated capital value

One frequently discussed indicator, linked to this, is the ratio of enterprise value (EV) to regulated capital value (RCV) – prominent in the CMA's Ofwat PR19 and PR24 redeterminations (CMA, 2021, 2026) and termed the 'Market-to-Asset Ratio' (MAR) in the UKRN consultation.

Enterprise value should not be read simply as a valuation of today's asset base: conceptually it comprises the present value of cash flows from existing regulated assets, the net present value of future regulated investment, and the anticipated net present value of activities outside the regulated business. An EV materially above RCV therefore needs careful interpretation – it need not imply that future investment is expected to earn above investors' true cost of capital – existing assets often embody historical decisions or operational advantages that new projects do not share, so investors need not expect identical returns from both. But the market-to-asset ratio can provide useful information, nonetheless.

Of course, in practice this pillar is not always available – for entities with no listed comparators or recent transactions, differences between the cohort subject to the regulatory decision and the entities for which such information is available will require careful assessment. In such cases, judgement must rest more heavily on the other two perspectives.



Debt vs equity returns

The analysis of debt premia bears more helpfully on the question of model form, since it speaks more directly to how required returns move as rates change – the point on which Stable TMR and Stable ERP disagree. Broadly stable credit spreads as interest rates move would be consistent with a Stable ERP view; narrowing spreads that hold the total cost of debt roughly constant would point the other way. Unhelpfully, credit spreads reflect other complex influences – default risk, liquidity and the credit cycle – so the evidence should be read as informative rather than conclusive. In keeping with this paper's wider theme: no single source, examined alone, settles the question.

As with the market-to-asset ratio, this evidence is not universally available in a clean form. Liquid, duration-matched public debt is not issued by every regulated entity, so the comparison often relies on proxies from comparable issuers – raising a version of the same comparability challenge already noted for MAR. Where such data exists, it offers a genuinely different lens on model choice than MAR provides; where it does not, judgement must again rest more heavily on the other two perspectives.

What market evidence can and cannot show

The real contribution of market evidence is an independent, forward-looking perspective on investor expectations that complements the historical record without duplicating it.

A static market-to-asset ratio *may* reflect whether the currently allowed return is adequate, given whatever specification produced it – but it says nothing about how that return should move as interest rates change, which is precisely what distinguishes the Stable TMR from the Stable ERP formulation of the CAPM. Static MAR evidence addresses the level of the allowed return, not its dynamic form, and cannot discriminate between the two approaches.

An exception applies to changes in MAR values over time, rather than their static level. For price controls that are not indexed to the risk-free rate, movements in MARs against gilt yields – controlling for other value drivers – can speak directly to the dynamic question that their static level cannot. For example, recent evidence points to a marked negative correlation between MAR levels and gilt yields, which seems more consistent with a Stable ERP view than a Stable TMR one, though it remains a narrower and more provisional point than the debt/equity comparison discussed below.

Evidence from the debt markets may help us judge between models – but its inherent noise may make it less effective as an input for the assessment of overall model outcomes. The ongoing debate on the relationship between debt and equity returns speaks clearly to the challenges of direct inference from debt markets to the cost of equity and the limitations of the data in this regard.

The difference in explanatory power and relevance to the different facets of cost of capital estimation means that it is important to keep the analysis of the different sources of evidence distinct. Historical returns, market-implied relationships and theoretical reasoning should remain separately identifiable throughout the reasoning process, rather than combined into intermediate blended statistics. Premature conflation of different types of evidence can obscure rather than clarify: it becomes harder to see which evidence is doing the work, and correspondingly harder for others to scrutinise the reasoning. For example, averaging an ERP estimate derived from historical returns with a superficially similar estimate derived from an assessment of market prices via a dividend growth model risks disguising two fundamentally different statistics as one, and any resulting average is likely less informative than either component viewed on its own terms, with potentially spurious weighting applied to each source. Where possible, evidence of different types should inform judgement side by side, rather than being blended into it.

Taken together with the historical record and theoretical coherence, however, this comparison is a genuine and useful third strand of market evidence, and merits close empirical attention. The remaining task is to interpret the historical record within a coherent conceptual framework – the role of financial theory.



Theoretical Coherence

What is the most coherent way to interpret the evidence?

Neither historical nor market evidence determines its own interpretation. That is the role of financial theory: not a third empirical dataset, but the intellectual structure within which independent evidence is interpreted consistently. This role matters most precisely where the empirical evidence itself is uncertain. The question is not whether theory can prove one formulation correct, but which formulation is the most coherent interpretation of the evidence within the broader framework of modern finance.

The architecture of the CAPM

Since Sharpe (1964), the CAPM has been organised consistently – across mainstream corporate finance textbooks, valuation guidance and regulatory practice – around two conceptually distinct quantities: the risk-free rate (the time value of money) and the equity risk premium (the market price of systematic risk), combined via beta (e.g. Brealey, Myers and Allen, 2020; Berk and DeMarzo, 2020). This is not merely notational: the risk-free rate captures the opportunity cost of risk-free capital, while the ERP compensates for undiversifiable risk. That separation is one of the CAPM's central strengths. Notably, it is not how the UKRN's current guidance operates in practice: Wright et al. (2018), and the UKRN's own methodology that draws on it, instead estimate the total market return directly and treat the CAPM identity in reverse, backing out an implied ERP as a residual (UKRN, 2023) – itself a clear illustration of the Stable TMR convention this paper questions.

Theory as a tiebreaker

This does not establish that any particular calibration is correct, nor that Stable TMR is theoretically invalid. But where competing models might each explain the evidence similarly well, preference naturally shifts towards the one that integrates more coherently with established knowledge – a principle used throughout the empirical sciences, not unique to finance. If either the historical record or market evidence pointed decisively to one formulation, theoretical considerations would matter little. The preceding two sections showed that neither does. Under those circumstances, conceptual coherence becomes a legitimate, and increasingly important, basis for choice.

Viewed this way, Stable ERP has a clear conceptual advantage: it aligns with how the CAPM is understood and applied throughout mainstream finance, via the separate estimation of the time value of money and the market price of risk. Stable TMR organises the same problem differently – not invalid, but less familiar within the wider literature. If overwhelming empirical evidence favoured Stable TMR, this difference would carry little weight; since no such evidence exists, theoretical coherence becomes an important third perspective.

Bringing the Evidence Together

This paper has not sought to show that one CAPM formulation is objectively correct and the other objectively wrong. Its question has been broader:

How should regulators exercise judgement when no individual source of evidence is capable of resolving the issue conclusively?

The answer: robust judgement should emerge from the disciplined integration of independent perspectives – each contributing different information, each subject to different limitations. Historical evidence anchors judgement in long-run market experience. Market evidence provides support for an independent, forward-looking view of the returns investors currently require – and may offer clues on model formulation. Theory provides the interpretative structure through which both become meaningful. None is sufficient alone;



confidence should derive from the consistency of conclusions across genuinely independent perspectives rather than the apparent precision of any one of them.

Implications for the CAPM

The historical record remains indispensable but, given its statistical properties, does not discriminate confidently between Stable TMR and Stable ERP. Market evidence speaks to a different question – whether the currently allowed return is priced as adequate, not how it should move as rates change – so market-to-asset ratios cannot resolve the choice, and the more conceptually relevant debt/equity risk premia comparison faces a similar signal-to-noise problem. Theoretical coherence provides a further criterion where empirical evidence remains open to more than one interpretation, and here Stable ERP aligns more naturally with the CAPM's conceptual architecture.

Taken individually, none of these observations is decisive. Taken together, they point towards a consistent conclusion: not that any one perspective proves Stable ERP superior, but that it is more strongly supported when the evidence is considered as a whole. That distinction matters. Indeed, the UKRN consultation itself acknowledges that the historical record is inconclusive on this choice – yet its proposed response is to seek further empirical triangulation (a hybrid mechanistic weighting of Stable TMR and Stable ERP evidence, subject to further stakeholder submissions) rather than to ask whether additional scrutiny of the same underlying data could ever resolve the question. No amount of additional care in interpreting realised returns can manufacture statistical precision that the data does not contain; the limitation is one of statistical power, not analytical effort. Recognising this points towards giving greater relative weight to theoretical coherence – and, on that basis, to Stable ERP – rather than towards further empirical refinement of an inherently uncertain exercise.

This conclusion should be read as a technical and academic one, rather than a complete account of the practical case for regulatory policy. Stable TMR has real virtues of its own within the UK regulatory setting, of a different kind to the evidence considered above. Proponents can reasonably point to a 'through-the-cycle' rationale: that the total return demanded by equity investors reflects a broadly stable long-run relationship between risk and reward, largely insulated from short-term interest rate movements, and that anchoring the allowed return to this stable relationship better serves long-term financeability than allowing it to move with each interest rate cycle. There is also a more pragmatic argument from consistency: Stable TMR has been the basis for UK regulatory decisions for some years, and continuity of approach across successive price controls has its own value, in reducing regulatory uncertainty and the scope for dispute at each review. These are legitimate considerations, and they do not depend on resolving the technical question addressed in this paper. The conclusion reached here is narrower: that where the choice between Stable TMR and Stable ERP is assessed on technical and academic merits alone, the evidence leans towards Stable ERP being the more defensible formulation. Whether that technical case should displace an established practice that has other things to recommend it is a separate, and properly regulatory, question that this paper does not attempt to answer.

A note on the ERP-RFR correlation

Some support for the through-the-cycle case draws on evidence that risk-free rates and equity risk premia have historically moved inversely, keeping the total market return comparatively stable (see, for example, Oxera, 2018). It is worth separating two distinct objects immediately: the ex-post correlation between realised equity and bond returns in a given historical sample, and the ex-ante relationship between the risk-free rate and the equity risk premium that the Stable TMR/Stable ERP choice is actually about. The two are related, but evidence about the former does not automatically settle the latter. Other recent regulatory analysis has observed that the realised correlation appears to have weakened in recent decades, and on some measures has turned negative – which, taken at face value, might seem to undermine the empirical case for Stable TMR. But this observation invites the same two questions this paper has already put to the historical record. First, a correlation estimated from a rolling window carries a wide margin of error of its own: using a Fisher z-transformation, the 95% confidence interval around a correlation estimated from a 30-year window is of the order of ± 0.35 to ± 0.40 – wide enough that much of



the apparent swing from positive to negative territory may reflect sampling noise rather than a genuine change in the underlying relationship. Second, and more fundamentally, the realised correlation between equity and bond returns is known to depend on the composition of the shocks that happen to hit markets in a given period: inflation shocks tend to move stock and bond prices together, while growth or ‘flight-to-quality’ shocks tend to move them apart (Yang, Zhou and Wang, 2009). A period dominated by growth and risk shocks will therefore show a weaker or more negative realised correlation, without this implying anything about the ex-ante relationship that Stable TMR and Stable ERP actually compete to describe. This is a further instance of the ex-post/ex-ante conflation just noted – and a reminder that recent data deserve the same caution as the long-run averages considered above.

Beyond the CAPM

The same logic recurs elsewhere within the CAPM – perhaps most notably in beta estimation. A comparable challenge arose in estimating Heathrow Airport’s beta after COVID-19, where historical comparator data was itself distorted; the approach developed for the CAA’s H7 price control decision (Flint Global, 2022), on which the author acted as the CAA’s expert witness, combined historical evidence with judgement about its temporary distortions, and was upheld in full by the Competition and Markets Authority on appeal (CMA, 2023). More broadly, decisions within regulatory economics require a similar synthesis of different perspectives; on efficient expenditure, financeability, incentive design, risk allocation and long-term investment. They all require evidence to be weighed under significant uncertainty. Robust regulatory judgement is unlikely to emerge from identifying one decisive source; it emerges from evaluating whether independent perspectives point towards a coherent overall conclusion.

Conclusion

The estimation of the cost of equity will always involve uncertainty. No historical dataset eliminates statistical ambiguity; no market valuation can be interpreted without economic judgement; no theoretical framework determines empirical outcomes independently of evidence. The objective of good regulation is not to remove judgement from the process, but to exercise it transparently, systematically and on the strongest available evidence. The CAPM debate illustrates that principle, but its relevance extends further: wherever regulators face important decisions on incomplete evidence, the quality of those decisions depends not on finding certainty where none exists, but on integrating diverse evidence within a coherent framework for judgement. The same approach could extend to the other elements of the UKRN consultation not addressed here – the remaining cross-checks, the framework for assessing them, and aiming up – though that lies beyond the present scope.

The central contribution of this paper is therefore not simply a preference for one CAPM formulation over another. It is a framework for exercising regulatory judgement under uncertainty. By bringing together historical evidence, contemporary market evidence and theoretical coherence, the paper argues that robust decisions emerge not from reliance on any single methodology, but from the disciplined synthesis of independent perspectives.

Author

This discussion paper was written by Flint Specialist Partner Craig Lonie. Craig draws on four decades of experience in the utility infrastructure, communications and media sectors, advises clients on cost-of-capital and regulatory finance matters, and is a member of the CMA’s corporate finance expert panel.



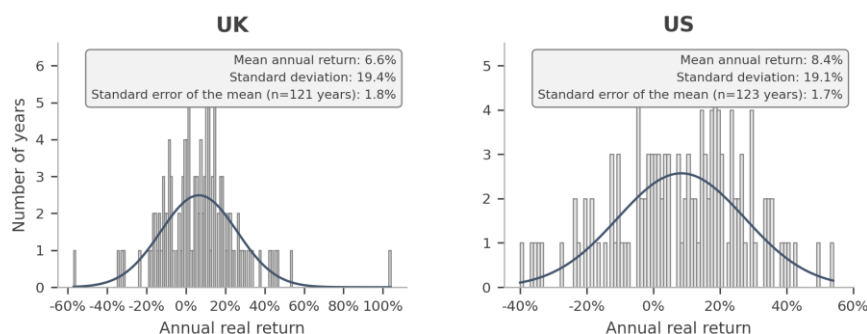
References

- Berk, J. and DeMarzo, P. (2020) *Corporate Finance*, 5th edn. Harlow: Pearson.
- Brealey, R.A., Myers, S.C. and Allen, F. (2020) *Principles of Corporate Finance*, 13th edn. New York: McGraw-Hill.
- Competition and Markets Authority (CMA) (2021) *Anglian Water Services Limited, Bristol Water plc, Northumbrian Water Limited and Yorkshire Water Services Limited Price Determinations: Final Report*. London: CMA.
- Competition and Markets Authority (CMA) (2023) *H7 Heathrow Airport Licence Modification Appeals: Final Determination*. London: CMA.
- Competition and Markets Authority (CMA) (2026) *Anglian Water Services Limited, Northumbrian Water Limited, South East Water Limited, Southern Water Services Limited and Wessex Water Services Limited Price Determinations: Final Report*. London: CMA.
- Dimson, E., Marsh, P. and Staunton, M. (2002) *Triumph of the Optimists: 101 Years of Global Investment Returns*. Princeton, NJ: Princeton University Press.
- Dimson, E., Marsh, P. and Staunton, M. (2025) *Global Investment Returns Yearbook 2025*. Zurich: UBS.
- Fama, E.F. and French, K.R. (2002) 'The Equity Premium', *The Journal of Finance*, 57(2), pp. 637–659.
- Flint Global (2022) *H7 Updated Beta Assessment*. Report for the UK Civil Aviation Authority (CAP2366B).
- Mehra, R. and Prescott, E.C. (1985) 'The Equity Premium: A Puzzle', *Journal of Monetary Economics*, 15(2), pp. 145–161.
- Merton, R.C. (1980) 'On Estimating the Expected Return on the Market: An Exploratory Investigation', *Journal of Financial Economics*, 8(4), pp. 323–361.
- Oxera (2018) *The Cost of Equity for RIIO-2: A Review of the Evidence*. Report for the Energy Networks Association, 28 February.
- Sharpe, W.F. (1964) 'Capital Asset Prices: A Theory of Market Equilibrium under Conditions of Risk', *The Journal of Finance*, 19(3), pp. 425–442.
- UK Regulators Network (UKRN) (2023) *Guidance for Regulators on the Methodology for Setting the Cost of Capital*. London: UKRN.
- Wright, S., Burns, P., Mason, R. and Pickford, D. (2018) *Estimating the Cost of Capital for Implementation of Price Controls by UK Regulators*. Report for the UK Regulators Network.
- Yang, J., Zhou, Y. and Wang, Z. (2009) 'The Stock-Bond Correlation and Macroeconomic Conditions: One and a Half Centuries of Evidence', *Journal of Banking & Finance*, 33(4), pp. 670–680.



Annex A. Understanding the Statistical Limits of Historical Evidence

As explained in the section on the historical record, estimating expected returns from historical observations is a signal-to-noise problem (Merton, 1980): annual realised returns are dominated by short-term variation that swamps the small, gradual movements in the expected return itself.



Source: author's calculations from Robert Shiller's US stock market dataset (S&P Composite, 1900–2022, n=123 years) and the Jordà-Schularick-Taylor Macrohistory Database (UK equities, 1900–2020, n=121 years).

Figure A1: Distribution of annual real equity returns, UK (1900–2020) and US (1900–2022). Source: author's calculations from the Jordà-Schularick-Taylor Macrohistory Database (UK) and Robert Shiller's US equity dataset (US).

Table A1 illustrates the scale of the challenge, assuming annual equity volatility of approximately 20% – consistent with Dimson, Marsh and Staunton (2002, 2025) – and reports the approximate precision of long-run average return estimates at different sample lengths. The figures are illustrative, not definitive; their significance lies in their order of magnitude. Even exceptionally long datasets leave uncertainty measured in percentage points, where many regulatory debates concern basis points – one of the principal motivations for the framework developed in this paper: recognising the statistical limits of historical evidence increases, rather than diminishes, the value of considering it alongside other independent sources.

Annual volatility	Years of observations	Approx. standard error	Approx. 95% confidence interval
20%	25	4.0%	±8.0%
20%	50	2.8%	±5.6%
20%	100	2.0%	±4.0%
20%	150	1.6%	±3.2%

Table A1: Illustrative precision of long-run average equity return estimates

The volatility data should be seen alongside the observed mean annual returns of 6.6% and 8.4% in the UK and US equity markets respectively. To interpret the data crudely, if we were to assume that the past is an indicator of the future, based on around 120 years of evidence, we can be 95% confident only that the expected return on equity in the UK lies between 3.0% and 10.2%.

A sharper test: the slope of TMR on the risk-free rate

The illustration above speaks to the imprecision of the historical average return itself. A more direct test targets the parameter that actually distinguishes Stable TMR from Stable ERP: the slope of TMR on the



risk-free rate, which is zero under Stable TMR and one under Stable ERP. The standard error of that slope, estimated from realised annual returns, is given by:

$$SE(\hat{\beta}) \approx \frac{\sigma_{\epsilon}}{\sigma_{RFR} \sqrt{T}}$$

where σ_{ϵ} is the residual volatility of TMR, σ_{RFR} is the standard deviation of the risk-free rate, and T is the sample length in years.

Using illustrative but reasonable assumptions – residual volatility of around 20%, a risk-free rate standard deviation of two to three percentage points, and approximately 120 years of data – this standard error is of the order of 0.6 to 0.9. A slope of zero and a slope of one are therefore separated by little more than one to two standard errors: statistically difficult to distinguish from a century of historical data, in a more direct sense than the illustration above.

