Estimating the market risk premium for valuations: arithmetic or geometric mean or something in between?

Christoph Kaserer[1]

Abstract

It is well-known that using arithmetic averages of yearly return observations leads to downward biased discount rates estimations. Well-known corrections, however, lead to upward biased results under the presence of negative serial correlation. Using a simulation analysis, we first show that a specific variant of the Cooper estimator, labelled as C4 in this paper, leads to robust estimations even under the presence of both serial correlation and heteroscedasticity. We also show that among the simple estimators, i.e. the arithmetic (AM) or geometric mean (GM) or the mean of both (MoM), the first one tends to perform best unless there is a high degree of negative serial correlation. In that case using the so-called mean of means rule would be better. Secondly, using data from Jordà et al. (Q J Econ 134(3):1225–1298, 2019) we find negative serial correlation and heteroscedasticity in market risk premia to be a widely spread phenomenon. Finally, we use this data to derive presumably least biased market risk premia estimations based on the C4 estimator. For the majority of the countries we find that these estimations are somewhere between the arithmetic and geometric average. When comparing these simple estimators among each other based on the empirical data, we find the arithmetic mean and mean of means to perform almost equally well, while the geometric mean clearly underperforms. Moreover, we found some evidence that the MoM is slightly outperforming the AM under a local CAPM perspective, while the opposite tends to be true under a global CAPM perspective. This leads us the cautious conclusion that the mean of means rule used by practitioners has some empirical rationale when there is evidence for substantial negative serial correlation.

✉ Christoph Kaserer
Christoph.Kaserer@tum.de
https://www.fm.wi.tum.de

1 Department of Financial Management and Capital Markets, TUM School of Management, Technische Universität München, Munich, Germany

1 Introduction

When estimating the cost of capital based on the CAPM the market risk premium is a necessary and pivotal input parameter. In the so-called historical method a long record of historical observations, in developed countries in some cases more than 100 years, is used. Typically, the market risk premium is then extracted as the arithmetic average of these observations.

However, starting with Blume (1974) there is an ongoing debate on how an unbiased estimation for the market risk premium has to be derived. Interestingly, he showed that even under ideal conditions, i.e. iid-returns, the unbiased estimator for a long-run compounding rate is an appropriately weighted average of the arithmetic and geometric mean return derived from historical return observations. Hereby, the weight on the arithmetic mean is the larger the longer the observation period relative to the compounding period is. A derivation and detailed analysis of an unbiased compounding estimator under the iid lognormal assumption can be found in Jacquier et al. (2003) and Jacquier et al. (2005). One compelling feature of the Blume estimator consists in the fact that the weights are always between 0 and 100%. Hence, by applying this estimator to the market risk premium the estimation used in the compounding problem is never larger than its arithmetic mean and never smaller than its geometric mean.

Unfortunately, this result does not hold anymore if applied to discounting problems, as it is the case in any valuation problem. Actually, Cooper (1996) shows that under the iid assumption the arithmetic mean is downward biased, i.e. valuations are upward biased.[1] Again, he shows that the unbiased estimator can be represented as a weighted average of the arithmetic and the geometric mean. However, the weight on the arithmetic mean is larger than 100%. Given that the geometric mean is smaller than the arithmetic mean, the unbiased estimation for the market risk premium must always be larger than the arithmetic mean of the historical observations, except for the one period discount rate estimation. And this distance is increasing in the length of the discounting period.

Since then a small number of papers dealing with these biases in discount rate estimations have been published. Elsner and Krumholz (2013) give a comprehensive overview and derive a specific version of the Cooper estimator under the assumption of lognormally distributed returns. Breuer et al. (2014) and Breuer et al. (2017) analyze the performance of a specific version of the Cooper estimator relative to their own "truncated" arithmetic mean estimator based on bootstrapped empirical data. While they find biases in present values to be significant overall, only for low growth firms the Cooper estimator does a reasonable job. When looking at high growth firms, using the arithmetic average but ignoring cash flows beyond year 30

---

[1] In an earlier paper Butler and Schachter (1989) already show the downward bias in discounting rates and how it can be overcome under specific distributional assumptions.

or 50, i.e. "truncating" the estimation, leads to the least biased results.[2] Moreover, it seems that the estimator of Elsner/Krumholz performs relatively well, if the cost of equity is estimated on the basis of a total return approach (cf. Breuer et al. 2017, p. 738).

While this work is closely related to our paper, there is also an important difference. Breuer et al. (2014, 2017) are interested in the valuation bias created at the single firm level. This is, of course, the ultimate question in the context of a valuation problem. Therefore, the bias finally not only depends on a flawed market risk premium estimation, but it is also determined by the level of the risk-free rate, the systematic risk of the company as well as the time pattern of a company's cash flows. In this respect their work is very insightful, as they show that the bias is most important for firms with high growth rates.

Nevertheless, in this paper we want to focus on the bias at the level of the market risk premium estimation. Therefore, we ignore any additional effects stemming from firm and cash flow characteristics as well as the risk-free term structure. Nevertheless, our results might give some guidance on how to mitigate estimation problems at the level of the market risk premium. This, however, does not rule out that estimation problems might be amplified when applied to specific valuation problems.

Practitioners rarely apply the Cooper estimator when deriving the discount rates or the underlying market risk premium. From a German perspective, one reason might be that there is no official statement by competent institutions, like the FAUB, recommending the usage of the Cooper-estimator. Moreover, as Breuer et al. (2014, p. 592), point out, there are several unknowns when it comes to the practical application of this estimator. In Pratt and Grabowski (2014, p. 156), a well-known handbook for practitioners, it is argued that the bias caused by using the arithmetic average in practical valuation problems tends to be small as most of the value comes from cash flows in the first ten years. Koller et al. (2020), another well-known source for practical valuation, do even not mention the Cooper estimator at all, while they briefly explain the Blume estimator. However, there is no recommendation as to whether, and if so, under which circumstances, this estimator should be used.

While, to the best of our knowledge, there is no survey giving a picture of how practitioners are handling this estimation problem, it seems that they are using simple estimators as the arithmetic or geometric mean. Moreover, it might well be that some of them are using a simplified version of the Blume estimator, as it is the case in Germany with the so-called Mean of the Means, i.e. the simple average of the arithmetic and geometric average; cf. Wagner et al. (2006), or Stehle (2016). From an international perspective, recommendations for using such a rule of thumb are harder to find. However, a cautious indication in this direction can be found in Welch (2017, p. 199). Hence, one might conclude that most valuations are actually upward biased, as the discount rates tend to be downward biased.

This conclusion, however, might be premature. The estimators discussed so far have been derived under the condition of iid returns. In reality, we have to take into account that stock returns and, therefore, also market risk premia, over the long-run might display mean reverting behavior, making them negatively serially correlated;

---

[2] The bias created for young, high growth firms is more deeply analyzed in Breuer and Mark (2013).

cf. among others Shiller (2014), Spierdijk et al. (2012), Pástor and Stambaugh (2012), Campbell (2003, p. 803), Campbell and Shiller (1988), Poterba and Summers (1988), Shiller (1981). Intuitively it is not so hard to see that negative serial correlation tends to inflate the market risk premium estimation as long as it is based on the arithmetic mean; cf. Wenger (2005), Cooper (1996), and Indro and Lee (1997). Therefore, Cooper (1996) derived an estimator under the assumption of lognormally distributed returns displaying negative serial correlation. This estimator, which in this paper will be labelled C4, leads to smaller discount rates the larger the degree of serial correlation is.[3]

It should be noted in this regard that when estimating market risk premia two opposite effects have to be taken into account. At the one side, extracting an unbiased discount rate estimation for an N-year discounting period from T one-year historical return observations leads to a correction which increases in N and yields estimations higher than the arithmetic average. However, taking into account negative serial correlation at the other side, causes the unbiased discount rate to decrease in the degree of serial correlation. Therefore, under realistic conditions the net effect on the unbiased estimation is unclear and in general we cannot say whether the unbiased estimation finally is larger or smaller than the arithmetic mean.

This is where this paper wants to make contribution. First, by running a simulation analysis we are comparing different existing estimators, the arithmetic and the geometric mean, the mean of means, the Blume estimator, the Elsner/Krumholz estimator as well as four different variants of the Cooper estimator, including the above mentioned C4 estimator. Our results clearly indicate that under the presence of serial correlation and heteroscedasticity in the market risk premium, the C4 estimator performs best. Interestingly, in this horse race the mean of means estimator arrives second in those cases where there is considerable negative serial correlation. In all other cases, however, it is clearly inferior to many of the other estimators. When taking an agnostic perspective with respect to the empirical properties of market return behavior, i.e. by comparing the rankings of all estimators in 24 different baseline parameter combinations, we again find the C4 estimator to perform best. It is followed by the C2 and C3 estimator, the Blume estimator and the arithmetic mean.

Second, we then apply an empirical analysis to a broad set of countries and regions from a global CAPM perspective. For that purpose, we measure historical market risk premia realizations in three different base currencies. The results indicate that negative serial correlation and heteroscedasticity is a wide-spread phenomenon, even though in some cases we also find positive serial correlation. This is especially true, if we include pre-World War periods and use the Euro/Mark as the base currency.

---

[3] It should be noted that the same problem arises when estimating compounding rates, i.e. the terminal wealth is upward biased in the presence of negative serial correlation. Similar to the approach used in this paper, Indro and Lee (1997) run a simulation analysis in order to study this problem. They show that the Blume estimator even under the presence of negative serial correlation does a fairly good job, at least compared to the pure arithmetic or geometric average.

Third, we use the C4 estimator in order to derive the presumably least biased market risk premia estimations for the countries or regions in our data set. Interestingly, we find that in the majority of the cases this presumably least biased market risk premia estimation is somewhere in between the geometric and the arithmetic average. More precisely, when comparing the arithmetic mean and the mean of means on the basis of their estimation errors, we find that they perform on an almost equal footing, with the mean of means being better in situations of pronounced negative serial correlation. This leads us the cautious conclusion that the so-called mean of means rule proposed among others by Wagner et al. (2006) and Stehle (2016) and used by practitioners might have some empirical rationale.

The rest of paper proceeds as follows. Section 2 is laying down the problem of why arithmetic averages generate biased estimations in the context of a simple example. Section 3 reviews the literature and explains the different estimators that have been developed in the literature. In Sect. 4 the results of the simulation analysis are presented, while Sect. 5 describes the data used for the empirical analysis and discusses the results. Finally, Sect. 6 concludes.

2 Sketching the problem

In order to illustrate the problem, we develop the following simple example. Assume a perfect capital market with one representative stock. The returns of this stock are binomially distributed with the upward return factor being $u = 1.2$ and the downward return factor $d = 0.9$. The risk-free rate is assumed to be zero. Hence, it follows that the risk-neutral probability for an upward movement is $q_n = 1/3$. Assuming the physical probability[4] of an upward movement being $q = 1/2$, the expected return of the stock is 5%, which is equal to the arithmetic average of the returns.

Now, in this economy the value of a new project has to be determined. The expected cashflows evolve according to the binomial distribution given above starting from a value of 50 for the last expired period. The distribution of these cash flows is resumed in Fig. 1. According to the fundamental asset pricing theorem the value of this project is equal to the expected present value of future cash flows where the expectation is calculated under the risk-neutral measure and the discount rate is equal to the risk-free rate.[5] Hence, the value is $[(1/3 \times 60 + 2/3 \times 45)/1 + (1/9 \times 72 + 4/9 \times 54 + 4/9 \times 40.5)/1^2] = 100$.

Alternatively, the value of this project could be determined as the expected present value using the physical probability distribution and discounting with the expected return, i.e. the arithmetic average. At this point, because all distributional parameters are perfectly known, this leads to the same value. Noting that the expected cash flow in $t = 1$ is $CF_1 = 52.5$ and in $t = 2$ is $CF_2 = 55.125$, it follows: $52.5/1.05 + 55.125/1.05^2 = 100$. So, in this world without any estimation problems, valuation using physical probabilities for calculating expected cash flows and

---

[4] Physical probabilities sometimes are also labelled as real or objective probabilities.

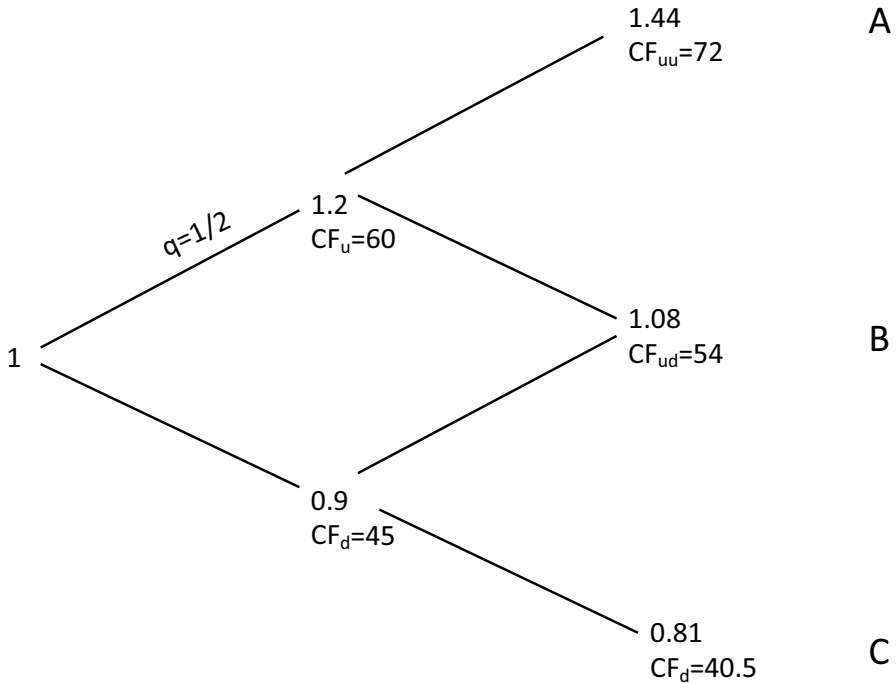[5] Cf. e.g. Cochrane (2005, p. 51 n).

Fig. 1 Binomially distributed stock returns and project cash flows

arithmetic averages for determining the risk-adjusted discount rate would work well, i.e. would lead to unbiased results.

However, things change once we introduce estimation problems. Assume that the true distributional parameters are not known, and we can only observe random paths of two period return realizations. Hence, in this example we would either observe two upward steps (path A) with an arithmetic average of realized returns of 20%. For the case of one up- and one downward step (path B) the arithmetic average would be 5% and for two downward steps (path C) –10%. The probability that path A or C is observed is 25% each, while for path B the probability is 50%.

Now, if by chance we would have observed path A the project would have been valued at $52.5/1.2 + 55.125/1.2^2 = 82.03$, while if we would have observed path C the value would have been $52.5/0.9 + 55.125/0.9^2 = 126.39$. Only in case of path B being realized we would have inferred the correct value of 100. Hence, the expected project value using the arithmetic average of realized returns as discount rates would be $82.03 \times 0.25 + 100 \times 0.5 + 126.39 \times 0.25 = 102.11$. As one can easily see, using the arithmetic average leads to an upward biased project value or, which is the same, to a downward biased discount rate estimation. And if we would have used the geometric average, which in this case would have been 3.92%, overvaluation would have even been worse.

The underlying reason for this result is the fact that present values are a convex function in the discount factors making the absolute valuation error of an

underestimation of the true discount rate not being symmetric to the absolute error caused by an overestimation of the same amount. In more formal terms this could be stated as follows. Assume $\widehat{D}_N$ is the unbiased discount factor over N periods and the expected annualized N-period return estimation is $\mu_N = E\left[1 + \widetilde{r}_N\right]$, e.g. the arithmetic mean. Then according to Jensen's inequality Butler and Schachter (1989, p. 15), show that the following holds:

$$\widehat{D}_N = E\left[\left(\frac{1}{1 + \widetilde{r}_N}\right)^N\right] \geq \left[\frac{1}{E\left[1 + \widetilde{r}_N\right]}\right]^N = \left[\frac{1}{1 + \mu_N}\right]^N \tag{1}$$

Hence, using the arithmetic average of observed yearly return realizations and plugging this into an N-period valuation problem causes an upward bias in the resulting value.

3 Unbiased estimators

3.1 The case of iid returns

The observation illustrated by the preceding example was the starting point of the paper of Blume (1974), who developed an unbiased estimator for the interest rate to be used in a compounding problem. He showed that under iid returns the unbiased compounding rate estimation is an appropriately weighted average of the arithmetic and the geometric mean. By extending this analysis to valuation problems, i.e. by analyzing how unbiased discount rates can be estimated, Butler and Schachter (1989) and Cooper (1996) were the first to come up with explicit solutions. Most importantly, under the assumption of iid returns Cooper (1996) derived the following unbiased estimator for the N-period discount factor:

$$\widehat{D}_{C1,N} = bA^{-N} + (1 - b)G^{-N} \text{ with } b = \frac{N + T}{T - 1} \tag{2}$$

Here, N is the length of the discounting period expressed in years, while $T \geq N$ is the number of yearly return observations on which the discount rate estimation is based upon.[6] Assuming return observations are expressed as return factors, i.e. 1 + yearly return, A stands for the arithmetic mean and G for the geometric mean of these return factors. Expressing the yearly return observation as $\widetilde{r}_t$ and having a total of T observations, these two means are defined as follows:

$$A = \frac{1}{T} \sum_{t=1}^{T} \left(1 + \widetilde{r}_t\right) \tag{3}$$

---

[6] It should be noted that for large N this estimator might become negative. This happened in a very few cases in our simulations causing non defined results.

$$G = \left[\prod_{t=1}^{T} \left(1 + \tilde{r}_t \right) \right]^{1/T} \tag{4}$$

For expositional convenience the unbiased discount factor can be transformed into an unbiased yearly discount rate in the following way:

$$\widehat{R}_{C1,N} = \left[bA^{-N} + (1-b)G^{-N} \right]^{-\frac{1}{N}} - 1 \tag{5}$$

Two remarks have to be made here. First, as the weighting factor b is always larger than 1, it must hold that the unbiased yearly discount rate according to (5) is always larger than the arithmetic mean, i.e. $\widehat{R}_{C1,N} > A - 1$, holds. However, in the special case of $N=1$ and T being relatively large, the discount rate is almost equal to the arithmetic average. Second, the unbiased discount rate is a function of the length of the discounting period. This creates the somewhat counterintuitive result that even though returns are drawn from a stationary distribution, yearly discount rates used in valuation problems are not constant.

By adding the assumption that returns are lognormally distributed, i.e. $\ln\left(1 + \tilde{r}\right) \sim \phi(\mu, \sigma)$, Cooper (1996) is able to derive the following more specific unbiased estimators[7]:

$$\widehat{D}_{C2,N} = G^{-N} e^{-(T+N)\frac{N\sigma^2}{2T}} \tag{6}$$

$$\widehat{R}_{C2,N} = G e^{(T+N)\frac{\sigma^2}{2T}} - 1 \tag{7}$$

Finally, under the assumption of lognormal returns the relationship $A = G e^{\frac{1}{2}\sigma^2}$ holds and this estimator can alternatively be written as:

$$\widehat{D}_{C3,N} = A^{-N} e^{-\frac{N^2\sigma^2}{2T}} \tag{8}$$

$$\widehat{R}_{C3,N} = A e^{\frac{N\sigma^2}{2T}} - 1 \tag{9}$$

This last equation again clearly points out that using the arithmetic average leads to an upward bias in valuations or to a downward bias in the discount rate.

Finally, it should be mentioned that (Elsner and Krumholz 2013), besides providing a comprehensive overview on different estimators, also analytically develop an unbiased estimator under conditions of iid distributed returns. Their approach is slightly different as they are interested in an unbiased estimation of the cost of capital. However, by assuming that all their assumptions apply to the market risk premium as well, their estimator, which we call EK, can be written as follows:

---

[7] Similar derivations can also be found in Jacquier et al. (2005) and Elsner and Krumholz (2013).

$$\widehat{D}_{EK,N} = \left\{ e^{\left[e^{\left[0.5(ln(1+A))^2 \frac{T}{\sigma^2} \right]} \sqrt{2\pi \frac{T}{\sigma^2}} \left(1 - \Phi\left[0; -ln(1+A)\frac{T}{\sigma^2}; \sqrt{\frac{T}{\sigma^2}} \right] \right) \right]^{-1}} \right\}^{-N} \tag{10}$$

$$\widehat{R}_{EK,N} = e^{\left[e^{\left[0.5(ln(1+A))^2 \frac{T}{\sigma^2} \right]} \sqrt{2\pi \frac{T}{\sigma^2}} \left(1 - \Phi\left[0; -ln(1+A)\frac{T}{\sigma^2}; \sqrt{\frac{T}{\sigma^2}} \right] \right) \right]^{-1}} - 1 \tag{11}$$

Here, Φ is the standard normal cumulative distribution function. It should be noted that (Elsner and Krumholz 2013) have developed an unbiased cost of capital estimator for the terminal value calculation. Hence, their estimator gives a discount rate for a perpetual cash flow stream. In our simulations we are interested in estimating discount rates for one-time cash flows accruing after N years. Therefore, comparing this estimator with the others presented above to some extend is unfair. We have decided to include the EK estimator for informational reasons nevertheless. Due to construction, we expect it to work better for longer discounting periods. As a drawback when it comes to simulations it should be mentioned that non existing values may arise for this estimator.

3.2 The case of non-iid returns

While the derivation of unbiased discount rate estimations in the case of iid-returns is clearly understood, it is still an open debate how to adjust these estimations, when it comes to take empirical distributions into account. Most importantly, two stylized facts have to be addressed in this regard. First, starting with Shiller (1981), Poterba and Summers (1988) and Campbell and Shiller (1988) there is plenty of literature pointing out that stock market returns display mean-reverting behavior, especially in the long-run.[8] Moreover, in a more recent strand of literature it is shown that this empirical behavior of stock returns can be seen as an equilibrium outcome which goes along with time-varying market risk-premia (cf. Campbell and Cochrane 1999; Campbell 2003; Wachter 2006; Santos and Veronesi 2006; Lettau and Wachter 2011; Berkman et al. 2011).

This literature raises fundamental questions on how market risk premia can be inferred from empirical data. Most importantly, if risk premia are time-varying a method based on historical observations might be conceptually flawed. This is why a new strand of literature has evolved trying to infer market risk premia either from analyst expectations (cf. Claus and Thomas 2001; Gebhardt et al. 2001; Fama and French 2002; Easton 2004; Ohlson and Juettner-Nauroth 2005),[9] bond prices (cf. Campello et al. 2008) or derivatives (cf. Berg and Kaserer 2013; van Binsbergen et al. 2013). However, this literature on the so-called implied market risk premium is

---

[8] An overview on this literature can be found in Campbell (2003, p. 830 n); see also the literature quote in Sect. 1.

[9] For an overview on this method and an empirical application cf. Jäckel et al. (2013).

beyond the scope of this paper. Therefore, it will not be considered here as an alternative to the historical method.

Coming back to the historical method, it is interesting to note that Cooper (1996) already analyzed the implications of serial correlation on the estimation of discount factors. And in fact, by simply assuming that returns still follow a lognormal distribution, even though with time-varying distributional parameters, i.e. $\ln(1 + \tilde{r}_N) \sim \phi(\mu_N, \sigma_N)$, he derived the following unbiased estimators[10]:

$$\widehat{D}_{C4,N} = G^{-N} e^{-(T+N)\frac{N\sigma_N^2}{2T}} \tag{12}$$

$$\widehat{R}_{C4,N} = G e^{(T+N)\frac{\sigma_N^2}{2T}} - 1 \tag{13}$$

Note that these estimators are equivalent to the one in Eqs. (6) and (7) just with the adjustment that the annualized N-period variance σ_N^2 is used instead of the constant annualized variance σ^2.[11] As this estimator will play an important role in this paper, we will label it as the C4 estimator.

As a second empirical phenomenon, heteroscedasticity, including fat-tailed distributions, has to be taken into account. However, in the literature so far no unbiased estimators controlling for heteroscedasticity have been developed. It is, therefore, an interesting question to see how sizeable the biases created by heteroscedasticity are.

4 Simulation analysis

4.1 Methods

In this section we run a simulation analysis in order to compare the biases caused by different estimators assuming specific characteristics in the stocks' return generating process. It should be noted here that we focus on the bias created by the market risk premium estimation only. We ignore any additional problems stemming from the level or the curvature of the risk-free term structure.

Following Campbell et al. (1997, p. 483 n), the return generating process for the market risk premium, defined as $r = \ln(1 + \tilde{r}_m) - \ln(1 + r_f)$, where r_m is the market return and r_f is the risk-free rate, is defined as follows:

$$r_t = \mu + \gamma(\mu - r_{t-1}) + \sigma_t \epsilon_t \tag{14}$$

---

[10] It should be noted that this is a very general distributional assumption used by Cooper (1996). As we will make clear in Sect. 4, the simulation is based on a GARCH(1,1) return process. There, the unconditional moments are constant. Nevertheless, there is serial correlation in the simulated return process implying variance ratios not being equal to one.

[11] In the simulation we use independent N-period returns to derive annualized variance: $\sigma_N^2 = Var[\tilde{r}_N]/N$, where $\tilde{r}_N$ is the return over N periods.

$$\sigma_t^2 = \omega^2 + \alpha\sigma_{t-1}^2 + (\alpha - \beta)\sigma_{t-1}^2(\varepsilon_t^2 - 1) \tag{15}$$

with $-1 \leq \gamma < 1$, $0 \leq \alpha < 1$, $0 \leq \beta < \alpha$, and $\epsilon, \varepsilon \sim N(0, 1)$

Note, this process captures all of the above-mentioned return characteristics. First, if $\gamma = \alpha = \beta = 0$ we have an iid-distributed market risk premium with an expected continuously compounded return equal to μ and standard deviation ω. Return innovations in this case are only temporary. Second, by setting $0 < \gamma < 1$ we get a weakly stationary mean-reverting process with a negative autocorrelation coefficient, i.e. $\rho_{t,t-1} = -\gamma$.[12] This captures the stylized empirical fact of variance ratios below 1, as will be shown in Sect. 5.3. In this case return innovations have a persistent component. Third, by setting $0 \leq \beta < \alpha < 1$ we allow for heteroscedasticity leading to market risk premia which display excess kurtosis and fat tails. In this case, also variance innovations have a persistent component. As a special case we will also run some simulations using $-1 < \gamma < 0$. In this case returns display positive serial correlation.

In every single run, i.e. for every combination of the parameters mentioned in Eqs. (14) and (15) we randomly draw a yearly return path over $T = 100$ years. This number is chosen for two reasons. First, it resembles typical historical records available for the estimation of market risk premia. Of course, for some countries, like the US, even longer periods are available, while for others, like emerging markets, periods are much shorter. However, for most industrial countries historical periods in this range are available. Second, for the unbiased as well as the C4 estimator we need return observations over non-overlapping periods. We choose these periods as integer fractions of the simulated historical period $T = 100$, i.e. we use all periods N for which T/N is an integer. Hence, we simulate discount rate estimations for periods N of 1, 2, 4, 5, 10, 20 and 25 years.[13] The fact that 100 years could be split-up in 7 such integer fractions is an additional reason for choosing this specific simulation period.

In each of these single runs we generate $T = 100$ random return drawings based on the return generating process defined in (14) and (15). Then, based on these simulated returns, we calculate the arithmetic (AM) and the geometric mean (GM) according to Eqs. (3) and (4). Also, in each single run we calculate the N-period discount factors based on the estimators introduced in Sect. 3, i.e. the estimator according to Blume (1974), the mean of geometric and arithmetic average (MoM), the Cooper estimator according to Eq. (2) (C1), the Cooper estimator according to Eq. (6) (C2), the Cooper estimator according to Eq. (8) (C3), the Cooper estimator according to Eq. (12) (C4), and the EK estimator according to Eq. (10). Note that for the C4 estimator we estimate the annualized N-period variance σ_N^2 by using in each single run non-overlapping intervals equal N.

---

[12] This type of return generating process is called an AR(1) model. For a more detailed analysis cf. e.g. Tsay (2010, Sect. 2.4).

[13] We could also have used 50 years. However, in this case, for one simulation run we would only have two independent observations, making the estimation outcome extremely noisy. Therefore, we ignore the case with N = 50.

We then repeat each single run 200,000 times ending up with a total of 200,000 estimates.[14] Then, after having completed all runs, the average over all runs for each of these estimators for the N-period discount factors is calculated. And finally, for expositional reasons, each N-period discount factor estimate is then transformed into a yearly discount rate.

We then compare the average outcome of these yearly discount rate estimates over all simulations with the unbiased estimation also resulting from our simulations. This unbiased estimation is calculated in the same way as the "simple estimator" in Blume (1974). This approach has the advantage that we do not need any assumptions about the independence of the one-period returns. In each single run we compute the N-period compounding factor for each discounting period N. Each of these compounding factors is then averaged over all 200,000 runs and, finally, transformed into a yearly discount rate.

Finally, the different estimators have to be compared. Dittmann and Maug (2008) have shown that when it comes to compare valuation methods it makes a difference how error measurement is defined. As we are interested in the first place in the impact of the estimation error on the valuation error, we define the error as the following present value difference:

$$\Delta = ln\left[\frac{(1+u)^{-N}}{(1+x)^{-N}}\right] = N[\ln(1+x) - \ln(1+u)] \tag{16}$$

Here, u stands for the unbiased discount rate estimation, i.e. the average of the unbiased discount rate over all simulation runs, and x stands for the discount rate given by one specific estimator, again averaged over all simulation runs. Hence, assuming that the market risk premium could be directly used as a discount rate, Eq. (16) gives the difference of the logarithm of the present value of a one-time one Euro cash flow accruing in N years calculated with one specific estimator to the unbiased present value. It should be noted that a positive sign indicates that the estimated market risk premium is higher than the unbiased one, leading to an underestimation of the present value.

Now, given that we are looking at estimates for different time horizons N, there might be a bias using this error definition. As we finally look at errors over different discounting periods, errors in the discount rate used for longer periods are amplified because of discounting. Hence, in some sense they are implicitly assigned a higher weight. Therefore, we use also a second error definition, where we basically normalize the error per discounting year. Starting from Eq. (16), this error is defined as:

$$\Delta = \ln(1+x) - \ln(1+u) \tag{17}$$

We will rank the estimators according to both error measures. For most estimators, there are no striking differences in the ranking depending on which of the two error measures is used.

---

[14] We use Python for doing these simulations.

Table 1 Simulated estimation errors produced by different estimators in a low volatility environment

Vol = 15%						α = 0								
	N	Unbiased	Kurtosis	Vol_emp	Correl.	Δ_AM	Δ_GM	Δ_Blume	Δ_MoM	Δ_C1	Δ_C2	Δ_C3	Δ_C4	Δ_EK
γ = 0	1	5.13%	−0.06	15.00%	−0.98%	−0.02%	−1.14%	−0.02%	−0.58%	0.00%	0.00%	−0.01%	0.00%	0.39%
	2	5.13%	−0.06	15.00%	−1.00%	−0.03%	−1.15%	−0.05%	−0.59%	0.00%	0.00%	−0.01%	0.00%	0.38%
	4	5.12%	−0.06	14.99%	−1.01%	−0.06%	−1.17%	−0.09%	−0.61%	0.00%	0.00%	−0.01%	0.00%	0.36%
	5	5.13%	−0.06	14.99%	−1.03%	−0.07%	−1.18%	−0.11%	−0.62%	0.00%	0.00%	−0.01%	0.00%	0.35%
	10	5.12%	−0.06	15.00%	−1.03%	−0.13%	−1.24%	−0.22%	−0.68%	0.01%	0.00%	−0.01%	−0.02%	0.30%
	20	5.13%	−0.06	15.00%	−0.98%	−0.24%	−1.35%	−0.43%	−0.79%	0.03%	0.00%	−0.01%	−0.08%	0.21%
	25	5.12%	−0.06	14.99%	−1.03%	−0.29%	−1.40%	−0.54%	−0.85%	0.06%	0.00%	−0.01%	−0.14%	0.16%
	Mean	5.12%	−0.06	15.00%	−1.01%	−0.12%	−1.23%	−0.21%	−0.67%	0.01%	0.00%	−0.01%	−0.04%	0.31%
	Mean abs.					0.12%	1.23%	0.21%	0.67%	0.01%	0.00%	0.01%	0.04%	0.31%
γ = 0.2	1	5.18%	−0.06	15.34%	−20.43%	−0.02%	−1.18%	−0.02%	−0.60%	0.01%	0.01%	0.00%	0.01%	0.40%
	2	4.93%	−0.07	13.72%	−20.42%	0.21%	−0.95%	0.20%	−0.37%	0.25%	0.25%	0.23%	0.01%	0.63%
	4	4.85%	−0.06	13.15%	−20.42%	0.27%	−0.89%	0.24%	−0.31%	0.33%	0.33%	0.32%	0.01%	0.69%
	5	4.83%	−0.06	13.03%	−20.45%	0.28%	−0.88%	0.23%	−0.30%	0.35%	0.35%	0.34%	0.00%	0.70%
	10	4.80%	−0.06	12.78%	−20.40%	0.27%	−0.89%	0.17%	−0.31%	0.41%	0.40%	0.39%	0.00%	0.70%
	20	4.78%	−0.06	12.66%	−20.40%	0.21%	−0.95%	0.00%	−0.37%	0.49%	0.45%	0.44%	−0.04%	0.64%
	25	4.78%	−0.06	12.63%	−20.39%	0.17%	−0.99%	−0.08%	−0.41%	0.54%	0.47%	0.46%	−0.07%	0.61%
	Mean	4.88%	−0.06	13.33%	−20.41%	0.20%	−0.96%	0.11%	−0.38%	0.34%	0.32%	0.31%	−0.01%	0.63%
	Mean abs.					0.20%	0.96%	0.13%	0.38%	0.34%	0.32%	0.31%	0.02%	0.63%
γ = 0.5	1	5.52%	−0.09	17.35%	−49.49%	−0.01%	−1.50%	−0.01%	−0.75%	0.02%	0.02%	0.00%	0.02%	0.48%
	2	4.73%	−0.09	12.28%	−49.48%	0.73%	−0.76%	0.72%	−0.01%	0.78%	0.78%	0.76%	0.01%	1.22%
	4	4.63%	−0.09	11.49%	−49.49%	0.81%	−0.67%	0.77%	0.07%	0.89%	0.89%	0.87%	0.01%	1.30%
	5	4.61%	−0.09	11.33%	−49.45%	0.83%	−0.66%	0.77%	0.09%	0.92%	0.92%	0.90%	0.01%	1.31%
	10	4.54%	−0.09	10.70%	−49.49%	0.87%	−0.62%	0.74%	0.13%	1.05%	1.04%	1.02%	0.01%	1.36%
	20	4.51%	−0.09	10.36%	−49.50%	0.85%	−0.63%	0.60%	0.12%	1.23%	1.17%	1.15%	−0.01%	1.35%
	25	4.50%	−0.09	10.31%	−49.48%	0.83%	−0.65%	0.52%	0.09%	1.33%	1.22%	1.20%	−0.02%	1.33%

Table 1 (continued)

	N	Unbiased	Kurtosis	Vol_emp	Correl.	Δ_AM	Δ_GM	Δ_Blume	Δ_MoM	Δ_C1	Δ_C2	Δ_C3	Δ_C4	Δ_EK
Vol = 15%						**α = 0**								
	Mean	4.72%	−0.09	11.98%	−49.49%	0.70%	−0.78%	0.59%	**−0.04%**	0.89%	0.86%	0.84%	**0.01%**	1.19%
	Mean abs.					0.70%	0.78%	0.59%	**0.18%**	0.89%	0.86%	0.84%	**0.01%**	1.19%
	1	5.17%	−0.06	15.27%	18.37%	−0.04%	−1.19%	−0.04%	−0.61%	−0.01%	−0.01%	−0.02%	−0.01%	0.41%
	2	5.41%	−0.06	16.72%	18.36%	−0.29%	−1.44%	−0.30%	−0.86%	−0.25%	−0.25%	−0.26%	−0.01%	0.16%
	4	5.60%	−0.07	17.71%	18.38%	−0.49%	−1.65%	−0.53%	−1.07%	−0.43%	−0.43%	−0.45%	−0.02%	−0.04%
γ = −0.2	5	5.64%	−0.06	17.91%	18.37%	−0.54%	−1.70%	−0.59%	−1.12%	−0.47%	−0.47%	−0.49%	−0.02%	−0.09%
	10	5.72%	−0.06	18.32%	18.40%	−0.71%	−1.86%	−0.81%	−1.28%	−0.57%	−0.58%	−0.59%	−0.05%	−0.24%
	20	5.75%	−0.06	18.49%	18.36%	−0.92%	−2.07%	−1.12%	−1.49%	−0.64%	−0.68%	−0.69%	−0.21%	−0.43%
	25	5.76%	−0.06	18.55%	18.36%	−1.01%	−2.16%	−1.26%	−1.59%	−0.65%	−0.71%	−0.72%	−0.32%	−0.51%
	Mean	5.58%	−0.06	17.57%	18.37%	−0.57%	−1.72%	**−0.66%**	−1.15%	**−0.43%**	**−0.45%**	**−0.46%**	**−0.09%**	**−0.11%**
	Mean abs.					0.57%	1.72%	**0.66%**	1.15%	**0.43%**	**0.45%**	**0.46%**	**0.09%**	**0.27%**
Vol = 15%						**α = 0.6**								
	1	6.89%	0.65	23.55%	−1.02%	−0.06%	−2.82%	−0.06%	−1.43%	0.00%	−0.02%	−0.03%	−0.02%	0.68%
	2	6.90%	0.64	23.53%	−1.02%	−0.10%	−2.85%	−0.13%	−1.46%	−0.01%	−0.03%	−0.04%	−0.03%	0.65%
	4	6.91%	0.64	23.54%	−0.96%	−0.18%	−2.92%	−0.26%	−1.54%	−0.03%	−0.05%	−0.07%	−0.06%	0.58%
γ = 0	5	6.94%	0.64	23.54%	−0.99%	−0.22%	−2.97%	−0.32%	−1.58%	−0.04%	−0.06%	−0.08%	−0.09%	0.54%
	10	6.97%	0.64	23.54%	−0.99%	−0.41%	−3.14%	−0.63%	−1.76%	−0.06%	−0.11%	−0.14%	−0.21%	0.37%
	20	6.97%	0.64	23.54%	−1.01%	−0.70%	−3.42%	−1.11%	−2.04%	0.13%	−0.13%	−0.16%	−0.56%	0.13%
	25	6.94%	0.64	23.52%	−1.01%	−0.81%	−3.52%	−1.31%	−2.15%	0.36%	−0.11%	−0.14%	−0.80%	0.05%
	Mean	6.93%	0.64	23.54%	−1.00%	**−0.35%**	−3.09%	−0.55%	−1.71%	**0.05%**	**−0.07%**	**−0.09%**	**−0.25%**	**0.43%**
	Mean abs.					**0.35%**	3.09%	0.55%	1.71%	**0.09%**	**0.07%**	**0.09%**	**0.25%**	**0.43%**
	1	7.00%	0.61	24.05%	−20.29%	−0.04%	−2.92%	−0.04%	−1.47%	0.01%	0.00%	−0.02%	0.00%	0.70%

Table 1 (continued)

| Vol = 15% | | | | | α = 0.6 | | | | | | | | |
N	Unbiased	Kurtosis	Vol_emp	Correl.	Δ_AM	Δ_GM	Δ_Blume	Δ_MoM	Δ_C1	Δ_C2	Δ_C3	Δ_C4	Δ_EK
γ = 0.2													
2	6.40%	0.61	21.53%	−20.32%	0.51%	−2.37%	0.48%	−0.92%	0.60%	0.58%	0.56%	−0.01%	1.25%
4	6.22%	0.61	20.66%	−20.35%	0.64%	−2.23%	0.56%	−0.79%	0.79%	0.77%	0.75%	−0.03%	1.39%
5	6.17%	0.62	20.47%	−20.29%	0.66%	−2.21%	0.55%	−0.76%	0.85%	0.82%	0.80%	−0.03%	1.41%
10	6.09%	0.61	20.07%	−20.28%	0.63%	−2.23%	0.40%	−0.78%	1.00%	0.94%	0.92%	−0.07%	1.40%
20	6.04%	0.61	19.86%	−20.29%	0.45%	−2.38%	0.03%	−0.95%	1.34%	1.04%	1.02%	−0.27%	1.25%
25	6.04%	0.62	19.85%	−20.31%	0.35%	−2.48%	−0.16%	−1.04%	1.62%	1.08%	1.06%	−0.42%	1.17%
Mean	6.28%	0.61	20.93%	−20.31%	**0.46%**	−2.40%	**0.26%**	−0.96%	0.89%	0.75%	0.73%	**−0.12%**	1.23%
Mean abs.					**0.47%**	2.40%	**0.32%**	0.96%	0.89%	0.75%	0.73%	**0.12%**	1.23%
γ = 0.5													
1	7.87%	0.44	27.22%	−49.22%	−0.03%	−3.71%	−0.03%	−1.86%	0.04%	0.02%	0.00%	0.02%	0.79%
2	5.90%	0.44	19.29%	−49.22%	1.79%	−1.88%	1.76%	−0.03%	1.91%	1.89%	1.86%	0.01%	2.62%
4	5.66%	0.45	18.05%	−49.24%	1.99%	−1.68%	1.89%	0.18%	2.19%	2.16%	2.14%	0.01%	2.82%
5	5.60%	0.44	17.81%	−49.25%	2.02%	−1.64%	1.88%	0.21%	2.26%	2.22%	2.20%	0.01%	2.85%
10	5.43%	0.44	16.80%	−49.23%	2.11%	−1.53%	1.83%	0.32%	2.61%	2.50%	2.47%	−0.01%	2.95%
20	5.34%	0.44	16.29%	−49.23%	2.03%	−1.58%	1.52%	0.26%	3.30%	2.77%	2.74%	−0.11%	2.88%
25	5.32%	0.44	16.18%	−49.26%	1.97%	−1.61%	1.37%	0.22%	3.89%	2.88%	2.85%	−0.17%	2.83%
Mean	5.88%	0.44	18.80%	−49.23%	1.70%	−1.95%	1.46%	**−0.10%**	2.31%	2.06%	2.04%	**−0.03%**	2.53%
Mean abs.					1.71%	1.95%	1.47%	**0.44%**	2.31%	2.06%	2.04%	**0.05%**	2.53%
γ = −0.2													
1	7.01%	0.61	23.96%	18.23%	−0.09%	−2.94%	−0.09%	−1.51%	−0.03%	−0.05%	−0.06%	−0.05%	0.70%
2	7.65%	0.60	26.24%	18.26%	−0.74%	−3.59%	−0.77%	−2.15%	−0.65%	−0.66%	−0.68%	−0.09%	0.06%
4	8.16%	0.60	27.75%	18.27%	−1.30%	−4.14%	−1.38%	−2.71%	−1.15%	−1.16%	−1.18%	−0.17%	−0.49%
5	8.26%	0.61	28.08%	18.28%	−1.44%	−4.28%	−1.55%	−2.85%	−1.25%	−1.28%	−1.30%	−0.20%	−0.62%
10	8.50%	0.61	28.72%	18.27%	−1.88%	−4.71%	−2.10%	−3.28%	−1.51%	−1.57%	−1.60%	−0.41%	−1.02%
20	8.68%	0.61	29.05%	18.27%	−2.50%	−5.31%	−2.92%	−3.88%	−1.62%	−1.90%	−1.94%	−1.23%	−1.54%

Table 1 (continued)

					α = 0.6									
	Vol = 15%													
N	Unbiased	Kurtosis	Vol_emp	Correl.	Δ_AM	Δ_GM	Δ_Blume	Δ_MoM	Δ_C1	Δ_C2	Δ_C3	Δ_C4	Δ_EK	
25	8.67%	0.61	29.10%	18.28%	−2.73%	−5.54%	−3.24%	−4.11%	−1.49%	−2.00%	−2.03%	−1.76%	−1.71%	
Mean	8.13%	0.61	27.56%	18.27%	−1.53%	−4.36%	−1.72%	−2.93%	−1.10%	−1.23%	−1.26%	−0.56%	−0.66%	
Mean abs.					1.53%	4.36%	1.72%	2.93%	1.10%	1.23%	1.26%	0.56%	0.88%	

This table is based on a simulation of 200,000 random yearly return realizations according to Eqs. (14) and (15) as described in Sect. 4. Parameters γ and α are varied according to the values given in the table. The long-term return volatility ω is set at 15%, for the long-term return we define $\mu = 0.05{-}0.5\omega^2$. Moreover, we set $\beta = 0.5\alpha$. In each panel the unbiased estimator (Unbiased), the excess kurtosis and the N-period volatility (Vol_emp) of realized returns as well as their one-period correlation coefficient (Correl.) is given for different discounting periods N. Moreover, the estimation error produced by the following estimators is tabulated as well: Δ_AM: arithmetic mean; Δ_GM: geometric mean; Δ_Blume: estimator according to Blume (1974); Δ_MoM: mean of geometric and arithmetic average; Δ_C1: Cooper estimator according to Eq. (2); Δ_C2: Cooper estimator according to Eq. (6); Δ_C3: Cooper estimator according to Eq. (8); Δ_C4: Cooper estimator according to Eq. (12); Δ_EK: EK estimator according to Eq. (10). The estimation error is defined according to Eq. (17); a positive sign indicates that the estimated market risk premium is higher than the unbiased one. All estimators leading to a mean estimation error of less than 0.5% are in bold. Mean abs. is the average of the unsigned estimation error

Table 2 Simulated estimation errors produced by different estimators in a high volatility environment

	N	Unbiased	Kurtosis	Vol_emp	Correl.	Δ_AM	Δ_GM	Δ_Blume	Δ_MoM	Δ_C1	Δ_C2	Δ_C3	Δ_C4	Δ_EK
						Vol = 20%								
						α = 0								
γ = 0	1	5.13%	−0.06	20.00%	−1.00%	−0.04%	−2.02%	−0.04%	−1.03%	0.00%	0.00%	−0.02%	0.00%	0.66%
	2	5.12%	−0.06	20.00%	−1.03%	−0.06%	−2.04%	−0.08%	−1.05%	0.00%	0.00%	−0.02%	0.00%	0.64%
	4	5.12%	−0.06	20.00%	−1.02%	−0.10%	−2.08%	−0.16%	−1.08%	0.00%	0.00%	−0.02%	−0.01%	0.61%
	5	5.13%	−0.06	20.01%	−0.95%	−0.12%	−2.10%	−0.20%	−1.11%	0.00%	0.00%	−0.02%	−0.01%	0.59%
	10	5.12%	−0.06	20.00%	−1.02%	−0.22%	−2.20%	−0.39%	−1.21%	0.02%	0.00%	−0.02%	−0.05%	0.51%
	20	5.13%	−0.06	19.99%	−1.01%	−0.43%	−2.40%	−0.75%	−1.41%	0.11%	−0.01%	−0.03%	−0.25%	0.35%
	25	5.12%	−0.06	19.96%	−1.01%	−0.52%	−2.49%	−0.91%	−1.50%	0.21%	0.00%	−0.02%	−0.40%	0.28%
	Mean	5.12%	−0.06	19.99%	−1.00%	**−0.21%**	−2.19%	**−0.36%**	−1.20%	**0.05%**	**0.00%**	**−0.02%**	**−0.10%**	0.52%
	Mean abs.					**0.21%**	2.19%	**0.36%**	1.20%	**0.05%**	**0.00%**	**0.02%**	**0.10%**	0.52%
γ = 0.2	1	5.21%	−0.06	20.44%	−20.40%	−0.03%	−2.10%	−0.03%	−1.06%	0.01%	0.01%	−0.01%	0.01%	0.67%
	2	4.78%	−0.06	18.29%	−20.40%	0.37%	−1.69%	0.35%	−0.66%	0.44%	0.44%	0.41%	0.01%	1.07%
	4	4.64%	−0.06	17.55%	−20.38%	0.48%	−1.59%	0.42%	−0.55%	0.59%	0.58%	0.56%	0.01%	1.18%
	5	4.61%	−0.06	17.37%	−20.42%	0.49%	−1.57%	0.41%	−0.54%	0.62%	0.62%	0.60%	0.00%	1.20%
	10	4.55%	−0.06	17.06%	−20.42%	0.47%	−1.59%	0.30%	−0.55%	0.73%	0.70%	0.68%	−0.02%	1.19%
	20	4.51%	−0.07	16.87%	−20.41%	0.36%	−1.70%	0.03%	−0.66%	0.93%	0.80%	0.78%	−0.12%	1.11%
	25	4.51%	−0.06	16.82%	−20.42%	0.29%	−1.76%	−0.11%	−0.72%	1.07%	0.83%	0.81%	−0.21%	1.05%
	Mean	4.69%	−0.06	17.77%	−20.40%	**0.35%**	−1.71%	**0.20%**	−0.68%	0.63%	0.57%	0.55%	**−0.04%**	1.07%
	Mean abs.					**0.36%**	1.71%	**0.24%**	0.68%	0.63%	0.57%	0.55%	**0.05%**	1.07%
γ = 0.5	1	5.82%	−0.09	23.14%	−49.45%	−0.02%	−2.67%	−0.02%	−1.33%	0.03%	0.03%	0.01%	0.03%	0.75%
	2	4.42%	−0.09	16.37%	−49.47%	1.30%	−1.35%	1.27%	−0.02%	1.38%	1.38%	1.35%	0.02%	2.07%
	4	4.25%	−0.09	15.34%	−49.45%	1.44%	−1.20%	1.37%	0.13%	1.58%	1.58%	1.55%	0.02%	2.22%
	5	4.22%	−0.09	15.12%	−49.45%	1.47%	−1.17%	1.37%	0.16%	1.64%	1.63%	1.60%	0.02%	2.25%
	10	4.08%	−0.09	14.26%	−49.50%	1.54%	−1.09%	1.33%	0.24%	1.88%	1.84%	1.81%	0.01%	2.33%
	20	4.02%	−0.09	13.84%	−49.48%	1.50%	−1.12%	1.10%	0.20%	2.29%	2.05%	2.03%	−0.03%	2.30%
	25	4.01%	−0.09	13.74%	−49.50%	1.46%	−1.16%	0.97%	0.17%	2.58%	2.14%	2.12%	−0.08%	2.27%

Table 2 (continued)

Vol = 20%

					α = 0								
N	Unbiased	Kurtosis	Vol_emp	Correl.	Δ_AM	Δ_GM	Δ_Blume	Δ_MoM	Δ_C1	Δ_C2	Δ_C3	Δ_C4	Δ_EK
Mean	4.40%	−0.09	15.97%	−49.47%	1.24%	−1.40%	1.06%	**−0.07%**	1.63%	1.52%	1.49%	**0.00%**	2.03%
Mean abs.					1.25%	1.40%	1.06%	**0.32%**	1.63%	1.52%	1.49%	**0.03%**	2.03%
1	5.23%	−0.07	20.36%	18.43%	−0.06%	−2.11%	−0.06%	−1.08%	−0.02%	−0.02%	−0.04%	−0.02%	0.68%
2	5.66%	−0.06	22.30%	18.36%	−0.51%	−2.56%	−0.53%	−1.53%	−0.45%	−0.45%	−0.47%	−0.03%	0.24%
4	5.97%	−0.06	23.61%	18.41%	−0.88%	−2.93%	−0.94%	−1.89%	−0.77%	−0.77%	−0.79%	−0.04%	−0.11%
γ = −0.2　5	6.04%	−0.06	23.89%	18.38%	−0.97%	−3.02%	−1.05%	−1.99%	−0.84%	−0.85%	−0.87%	−0.05%	−0.20%
10	6.17%	−0.06	24.41%	18.38%	−1.26%	−3.30%	−1.43%	−2.27%	−1.00%	−1.03%	−1.05%	−0.14%	−0.45%
20	6.25%	−0.06	24.68%	18.40%	−1.64%	−3.68%	−1.97%	−2.65%	−1.07%	−1.21%	−1.23%	−0.56%	−0.76%
25	6.28%	−0.06	24.74%	18.39%	−1.83%	−3.87%	−2.24%	−2.84%	−1.06%	−1.30%	−1.32%	−0.91%	−0.90%
Mean	5.94%	−0.06	23.43%	18.39%	−1.02%	−3.07%	−1.17%	−2.04%	−0.75%	−0.80%	−0.82%	**−0.25%**	**−0.22%**
Mean abs.					1.02%	3.07%	1.17%	2.04%	0.75%	0.80%	0.82%	**0.25%**	**0.48%**

Vol = 20%

					α = 0.6								
N	Unbiased	Kurtosis	Vol_emp	Correl.	Δ_AM	Δ_GM	Δ_Blume	Δ_MoM	Δ_C1	Δ_C2	Δ_C3	Δ_C4	Δ_EK
1	8.28%	0.64	31.37%	−1.02%	−0.12%	−5.02%	−0.12%	−2.54%	−0.02%	−0.06%	−0.07%	−0.06%	0.97%
2	8.37%	0.64	31.39%	−1.00%	−0.23%	−5.13%	−0.28%	−2.65%	−0.07%	−0.12%	−0.13%	−0.12%	0.87%
4	8.39%	0.64	31.38%	−1.01%	−0.37%	−5.25%	−0.50%	−2.77%	−0.09%	−0.15%	−0.17%	−0.19%	0.75%
γ = 0　5	8.41%	0.65	31.37%	−1.03%	−0.44%	−5.32%	−0.61%	−2.84%	−0.10%	−0.17%	−0.20%	−0.24%	0.68%
10	8.50%	0.65	31.38%	−1.01%	−0.80%	−5.64%	−1.15%	−3.17%	−0.08%	−0.28%	−0.31%	−0.56%	0.38%
20	8.66%	0.64	31.35%	−1.01%	−1.50%	−6.29%	−2.11%	−3.84%	0.59%	−0.50%	−0.55%	−1.68%	n.d.
25	8.47%	0.64	31.38%	−1.04%	−1.59%	−6.36%	−2.31%	−3.91%	2.11%	−0.36%	−0.41%	−2.13%	−0.24%
Mean	8.44%	0.64	31.38%	−1.02%	−0.72%	−5.57%	−1.01%	−3.10%	**0.33%**	**−0.23%**	**−0.26%**	−0.71%	**0.57%**
Mean abs.					0.72%	5.57%	1.01%	3.10%	**0.44%**	**0.23%**	**0.26%**	0.71%	0.65%
1	8.53%	0.61	32.07%	−20.29%	−0.09%	−5.21%	−0.09%	−2.61%	0.02%	−0.03%	−0.04%	−0.03%	0.99%

Table 2 (continued)

					α = 0.6								
Vol = 20% N	Unbiased	Kurtosis	Vol_emp	Correl.	Δ_AM	Δ_GM	Δ_Blume	Δ_MoM	Δ_C1	Δ_C2	Δ_C3	Δ_C4	Δ_EK
γ = 0.2													
2	7.43%	0.61	28.71%	−20.28%	0.88%	−4.23%	0.84%	−1.64%	1.05%	1.00%	0.99%	−0.04%	1.96%
4	7.11%	0.61	27.54%	−20.31%	1.09%	−4.01%	0.95%	−1.42%	1.38%	1.32%	1.29%	−0.10%	2.18%
5	7.03%	0.61	27.28%	−20.28%	1.13%	−3.96%	0.95%	−1.37%	1.49%	1.41%	1.39%	−0.10%	n.d.
10	6.88%	0.61	26.74%	−20.31%	1.07%	−3.99%	0.71%	−1.41%	1.83%	1.60%	1.57%	−0.24%	2.20%
20	6.83%	0.62	26.46%	−20.34%	0.70%	−4.30%	0.07%	−1.74%	2.97%	1.73%	1.69%	−0.82%	1.90%
25	6.82%	0.61	26.43%	−20.32%	0.53%	−4.45%	−0.21%	−1.90%	4.72%	1.80%	1.75%	−1.18%	1.76%
Mean	7.23%	0.61	27.89%	−20.30%	0.76%	−4.31%	**0.46%**	−1.73%	1.92%	1.26%	1.23%	**−0.36%**	1.83%
Mean abs.					0.79%	4.31%	0.55%	1.73%	1.92%	1.27%	1.24%	**0.36%**	1.83%
γ = 0.5													
1	10.08%	0.44	36.27%	−49.22%	−0.08%	−6.63%	−0.08%	−3.29%	0.06%	0.00%	−0.01%	0.00%	1.07%
2	6.52%	0.45	25.71%	−49.26%	3.18%	−3.37%	3.12%	−0.03%	3.39%	3.33%	3.31%	0.00%	4.33%
4	6.09%	0.44	24.08%	−49.25%	3.51%	−3.00%	3.34%	0.32%	3.89%	3.79%	3.77%	−0.01%	4.66%
5	6.01%	0.45	23.75%	−49.22%	3.55%	−2.95%	3.33%	0.38%	4.02%	3.90%	3.88%	−0.02%	4.71%
10	5.68%	0.45	22.39%	−49.22%	3.69%	−2.75%	3.26%	0.56%	4.74%	4.36%	4.32%	−0.08%	4.86%
20	5.52%	0.45	21.70%	−49.24%	3.51%	−2.81%	2.80%	0.46%	7.39%	4.79%	4.74%	−0.33%	4.71%
25	5.49%	0.45	21.56%	−49.23%	3.37%	−2.90%	2.55%	0.36%	24.59%	4.93%	4.87%	−0.54%	4.59%
Mean	6.48%	0.44	25.07%	−49.24%	2.96%	−3.49%	2.62%	**−0.18%**	6.87%	3.59%	3.55%	**−0.14%**	4.13%
Mean abs.					2.98%	3.49%	2.64%	0.77%	6.87%	3.59%	3.56%	**0.14%**	4.13%
γ = −0.2													
1	8.52%	0.60	31.95%	18.33%	−0.18%	−5.26%	−0.18%	−2.68%	−0.07%	−0.11%	−0.13%	−0.11%	0.98%
2	9.77%	0.61	34.98%	18.26%	−1.39%	−6.46%	−1.43%	−3.89%	−1.22%	−1.27%	−1.28%	−0.24%	−0.22%
4	10.82%	0.61	37.03%	18.23%	−2.51%	−7.57%	−2.65%	−5.00%	−2.23%	−2.29%	−2.31%	−0.56%	−1.31%
5	11.13%	0.60	37.47%	18.31%	−2.90%	−7.95%	−3.08%	−5.38%	−2.55%	−2.62%	−2.65%	−0.75%	−1.68%
10	11.63%	0.61	38.28%	18.31%	−3.75%	−8.76%	−4.11%	−6.21%	−3.00%	−3.21%	−3.25%	−1.41%	−2.44%
20	11.88%	0.61	38.74%	18.29%	−4.80%	−9.76%	−5.43%	−7.22%	−2.57%	−3.75%	−3.82%	−3.37%	−3.28%

Table 2 (continued)

	Unbiased	Kurtosis	Vol_emp	Correl.	Δ_AM	Δ_GM	Δ_Blume	Δ_MoM	Δ_C1	Δ_C2	Δ_C3	Δ_C4	Δ_EK
Vol = 20%					**α = 0.6**								
N													
25	11.87%	0.61	38.83%	18.30%	−5.18%	−10.13%	−5.92%	−7.59%	−1.08%	−3.89%	−3.97%	−4.47%	−3.54%
Mean	10.80%	0.61	36.75%	18.29%	−2.96%	−7.99%	−3.26%	−5.42%	−1.82%	−2.45%	−2.49%	−1.56%	−1.64%
Mean abs.					2.96%	7.99%	3.26%	5.42%	1.82%	2.45%	2.49%	1.56%	1.92%

This table is based on a simulation of 200,000 random yearly return realizations according to Eqs. (14) and (15) as described in Sect. 4. Parameters γ and α are varied according to the values given in the table. The long-–term return volatility ω is set at 20%, for the long-term return we define μ = 0.05–0.5ω[2]. Moreover, we set β = 0.5α. In each panel the unbiased estimator (Unbiased), the excess kurtosis and the N-period volatility (Vol_emp) of realized returns as well as their one-period correlation coefficient (Correl.) is given for different discounting periods N. Moreover, the estimation error produced by the following estimators is tabulated as well: Δ_AM: arithmetic mean; Δ_GM: geometric mean; Δ_Blume: estimator according to (Blume, 1974); Δ_MoM: mean of geometric and arithmetic average; Δ_C1: Cooperestimator according to Eq. (2); Δ_C2: Cooper estimator according to Eq. (6); Δ_C3: Cooper estimator according to Eq. (8); Δ_C4: Cooper estimator according to Eq. (12); Δ_EK: EK estimator according to Eq. (10). The estimation error is defined according to Eq. (17); a positive sign indicates that the estimated market risk premium is higher than the unbiased one. All estimators leading to a mean estimation error of less than 0.5% are in bold. Mean abs. is the average of the unsigned estimation error

4.2 Results

Baseline simulation results are presented in Tables 1 and 2 of this paper. The following parameters were chosen. First, the arithmetic average of the market risk premium is set to 5% in continuous compounding. Empirical numbers calculated in Table 5 in Sect. 5.3 display a median of 5.67% over all cases considered. But of course, there is substantial cross-country variation. Hence, we take a somehow intermediate number. For the structural results of our simulations this should be of minor importance.

Second, long-term return volatility ω is set at 15 or 20%, respectively. In this way, a low and high volatility environment is captured. As can be seen from Table 5 in Sect. 5.3, volatilities calculated in USD or GBP are rarely above 20%, while those calculated in Euro/Mark in almost all cases are above this number. The median over all cases considered in Table 5 is 16%, with 90% of all observations being in the range of 16−25%. By taking into account that realized volatilities in our simulations are in the range of 10−40%, we think that we cover a large spectrum of observed volatility ranges.

Third, the mean reverting coefficient γ is set to be equal −0.2, 0, 0.2, and 0.5. By using these values in our simulation, we get variance ratios, i.e. ratios of one year to N-period annualized volatility, in the range of 35−148% with smaller discounting periods being associated with a smaller range. This is exactly what we would have expected based on Eq. (22) in Sect. 5.3, where the structural relationship between the mean reverting coefficient, the length of the discounting period and the variance ratio is more deeply analyzed.

When comparing this simulation-based range according to Eq. (22) with the empirical range of variance ratios, it can be seen that almost all country/discounting period observations fall into this range. The specific ratios are reported in Table 6. Therefore, we think that with the simulation parameters described above we cover most of the empirically relevant range of mean reverting behavior.

Fourth, the heteroscedasticity parameter α is set to be equal 0 or 0.6 in combination with β = 0.5α. In the latter case this leads to an excess kurtosis in the range of 0.44−0.65. This is in line with the excess kurtosis found for Germany for one specific case in the after-World War II period in Table 5 in Sect. 5.3. However, compared to other countries this number seems to be relatively low. Therefore, we also run a robustness simulation producing a quite higher kurtosis of 0.83−1.50. It should be noted that the median kurtosis in Table 5 is 2.2; however, when only considering the after-World War II periods, the median is 1.5.

Structurally, the following results can be observed. First, in the case of iid-distributed returns all estimators except the geometric average and the mean of means work very well. Of course, by construction the Cooper estimators 1 to 3 generate the best results. This is true independently of whether there is a high or low volatility environment. For the EK estimator we can see that errors are small as well and, due to construction, decreasing in the length of the discounting period. The arithmetic mean and the Blume estimator behave oppositely. They work well for short discounting periods, but errors increase with the length of the discounting period. For the arithmetic mean this is the well-known bias already discussed in the introduction. As the Blume estimator is based on a weighted average of the arithmetic and

the geometric mean with weights on the second increasing in the length of the discounting period, it follows from the behavior of the arithmetic mean that this estimator deteriorates with longer discounting periods. For a similar reason also the MoM estimator generates errors increasing in the discounting period.

Second, once we add negative serial correlation estimation errors increase significantly for all estimators except the C4 estimator and, to some extent, also the Blume and MoM estimator. This, by itself, is not surprising given the construction of these estimators. However, it might be interesting to see how stable the C4 estimator performs over different degrees of negative serial correlation and simulated volatility ranges. This, in principle, is also true for positive serial correlation, even though the error produced by the C4 estimator in the high volatility environment becomes somewhat larger.

Third, by adding heteroscedasticity to independently distributed returns, we see that all estimators belonging to the family of Cooper estimators continue to produce small errors as long as simulated volatility is not too high. While this statement also applies to the arithmetic average, it does not apply to the GM as well as MoM estimator.

Fourth, if we combine negative serial correlation with heteroscedasticity errors in most of the cases increase. Only the C4 estimator continues to produce relatively low errors as long as volatility is not too high. Moreover, while the AM estimator deteriorates with an increasing degree of negative serial correlation, the GM estimator improves. As a consequence, also the MoM estimator is positively affected by an increasing degree of negative serial correlation. In fact, for a high degree of negative serial correlation the MoM estimator performs very well and comes close to the C4 estimator. Compared to that the Blume estimator achieves its best results for a medium degree of negative serial correlation and outperforms the MoM estimator in these cases.[15]

Fifth, if we combine positive serial correlation with heteroscedasticity errors become very large for all estimators. But again, in this case the C4 estimators clearly outperforms all the others; and also the EK estimator is doing remarkably well in these cases.

In order to get a better structural picture of the different estimators a ranking based on the error measurement according to Eq. (17), i.e. the error in present value terms over one year discounting periods, is summarized in Table 3. In this ranking we simply order the estimators form the best to the worst result and then look at their ranks. Moreover, when we look at the ranking of estimators over different parameter combinations, we calculate their rank sum and rank them accordingly. Corroborating to what we have said before, it can be seen that under an iid environment the C2 estimator generates the best results independently of the volatility level. However, once we introduce serial correlation without any heteroscedasticity, the C4 estimator under all circumstances produces the best results. Interestingly, this also holds true if we allow for heteroscedasticity. In fact, when looking at the rank

---

[15] The sensitivity of the Blume estimator with respect to return volatility has also been emphasized by Antonczyk and Mark (2010).

Table 3 Ranking of different estimators

	AM	GM	Blume	MoM	C1	C2	C3	C4	EK	AM	GM	Blume	MoM	C1	C2	C3	C4	EK
Vol= 20%			α= 0									α= 0.6						
γ= 0	5	9	6	8	3	1	2	4	7	6	9	7	8	3	1	2	5	4
γ= 0.2	3	9	2	7	6	5	4	1	8	3	9	2	6	8	5	4	1	7
γ= 0.5	4	5	3	2	8	7	6	1	9	4	5	3	2	9	7	6	1	8
γ= −0.2	6	9	7	8	3	4	5	1	2	6	9	7	8	2	4	5	1	3
Vol= 15%			α= 0									α= 0.6						
γ= 0	5	9	6	8	3	1	2	4	7	5	9	7	8	2	1	3	4	6
γ= 0.2	3	9	2	7	6	5	4	1	8	3	9	2	7	6	5	4	1	8
γ= 0.5	4	5	3	2	8	7	6	1	9	4	5	3	2	8	7	6	1	9
γ= −0.2	6	9	7	8	3	4	5	1	2	6	9	7	8	3	4	5	1	2
All volatilities			α= 0									α= 0.6						
γ= 0	5	9	6	8	3	1	2	4	7	6	9	7	8	2	1	2	4	5
γ= 0.2	3	9	2	7	6	5	4	1	8	3	9	2	6	7	5	4	1	8
γ= 0.5	4	5	3	2	8	7	6	1	9	4	5	3	2	8	7	6	1	8
γ= −0.2	6	9	7	8	3	4	5	1	2	6	9	7	8	2	4	5	1	2
All volatilities and γ's			α= 0									α= 0.6						
	4	9	4	7	6	2	2	1	8	4	9	5	8	6	2	3	1	7
All volatilities and α's - based on equation (16)																		
γ= 0	5	9	7	8	3	1	2	4	6	6	9	8	8	3	1	2	5	4
γ= 0.2	3	9	2	7	6	5	4	1	8	3	9	1	4	7	6	5	1	7
γ= 0.5	4	5	3	2	8	7	6	1	9	5	4	3	2	8	7	6	1	8
γ= −0.2	6	9	7	8	3	4	5	1	2	6	9	7	8	3	4	5	1	2
All volatilities, γ's and α's - based on equation (16)																		
	4	9	5	7	6	2	3	1	7	5	9	4	8	6	2	2	1	7

This table gives the ranks of the different estimators based on the mean absolute error measured as defined in Eq. (17); the grey shaded area gives the aggregated results based on Eq. (16). In case where more results are analyzed simultaneously the estimators are ranked according to their rank sum. The estimators are: AM: arithmetic mean; GM: geometric mean; Blume: estimator according to (Blume, 1974); MoM: mean of geometric and arithmetic average; C1: Cooper estimator according to Eq. (2); C2: Cooper estimator according to Eq. (6); C3: Cooper estimator according to Eq. (8); C4: Cooper estimator according to Eq. (12); Δ_EK: EK estimator according to Eq. (10)

Fig. 2 Biases of different estimators under high or low volatility, negative serial correlation and heteroscedasticity. This figure is based on a simulation of 200,000 random yearly return realizations according to Eqs. (14) and (15) as described in Sect. 4. Parameters σ, γ and α are varied according to the values given below. The long-term return volatility ω is set at 15 or 20%, for the long-term return we define μ = 0.05–0.5ω². Moreover, we set β = 0.5α. The errors according to Eq. (17) for the N-period discount rate estimations are show. The estimators are: Δ_AM: arithmetic mean; Δ_GM: geometric mean; Δ_MoM: mean of geometric and arithmetic average; Δ_C4: Cooper estimator according to Eq. (12). **A** Low risk, high negative serial correlation, no heteroscedasticity (σ = 0.15, γ = 0.5, α = 0). **B** Low risk, no negative serial correlation, high heteroscedasticity (σ = 0.15, γ = 0, α = 0.6). **C** Low risk, high negative serial correlation, high heteroscedasticity (σ = 0.15, γ = 0.5, α = 0.6). **D** High risk, high negative serial correlation, high heteroscedasticity (σ = 0.20, γ = 0.5, α = 0.6). **E** High risk, high negative serial correlation, no heteroscedasticity (σ = 0.20, γ = 0.5, α = 0). **F** High risk, no negative serial correlation, high heteroscedasticity (σ = 0.20, γ = 0, α = 0.6)

sum over all 16 parameter combinations we have analyzed here, the C4 estimator is clearly the winner.

Moreover, as an additional remark it should be pointed out that in all 16 different parameter combinations used for compiling Table 3, one estimator belonging to the Cooper family occupied the first place. This corroborates the statement that this family of estimators performs very well.

According to this overall ranking the GM estimator produces the worst estimates. While the GM estimator produces bad results under most parameter combinations,

the Blume estimator works relatively well as long as there is negative serial correlation. However, for none or positive serial correlation it performs badly.

For the MoM estimator it can be said that it produces good results in the case of high negative serial correlation independently of the heteroscedasticity and variance level. However, for all other cases, i.e. low negative or positive serial correlation, the estimator generates flawed results.

Interestingly, these rankings do not change much, if instead of an error term normalized by the length of the discounting period we look at the error in present value terms. In fact, the grey shaded area in Table 3 gives the aggregated rankings based on Eq. (16). There are no striking differences.

Finally, a graphical resume of these results, especially with respect to the C4 and MoM estimator, is given in Fig. 2. Again, it can be seen there that the C4 estimator works remarkably well under all circumstances, while the MoM estimator only does a somehow acceptable job in the case of high negative serial correlation. The AM estimator comes close to the C4 estimator in case of no negative serial correlation.

From a practitioners' perspective, these results can be summarized as follows. Under conditions of ambiguity, i.e. it is unclear to what extent market returns are subject to serial correlation and heteroscedasticity, it would be best to work with the C4 estimator, as it outperforms the other estimators over a large range of different parameter combinations. Based on the analysis in the next section, where we show that negative serial correlation is a somewhat pervasive phenomenon, using the C4 estimator is clearly supported. If serial correlation could be ruled out, using the C2 estimator would be recommendable, while if heteroscedasticity could be ruled, the C4 estimator should again be preferred.

However, it has been argued that practitioners prefer to use simple estimators, like the AM, GM or MoM. Of course, this analysis has shown that none of these estimators constitutes a perfect solution, as each of them comes with a risk. In fact, the GM estimator performs badly unless there is a very high degree of negative serial correlation. In the 16 baseline parameter combinations analyzed here, it was always outperformed either by the AM or the MoM estimator.

Results are clearly better for the AM estimator, as it outperforms the GM and MoM estimator in all those circumstances where negative serial correlation is not too high. Hence, using the mean of means——as some practitioners seem to be doing——can only be justified on the basis of empirical evidence pointing to a substantial degree of negative serial correlation. It will be shown in the next section that negative serial correlation is a widely spread, even though not ubiquitous, phenomenon. However, it seems that the empirically measured degree of negative serial correlation is not sufficiently high in order to give the MoM estimator a clear preference over the AM estimator.

Finally, as an additional robustness test we have re-run the simulation by setting the parameters in a way that much higher heteroscedasticity is produced, i.e. we use $\alpha = 0.4$, $\beta = 0$ and $\alpha = 0.5$, $\beta = 0$. In fact, kurtosis in these simulations is in a range of 0.81 to 1.50. Results are reported in Table 4. It can be seen that they are qualitatively unchanged. Again, the C4 estimator overall performs better than any other estimator

Table 4 Ranking of different estimators in high heteroskedasticity cases

| | Vol= 20% | | | | | | | | Vol= 15% | | | | | | | |
| | α= 0.4 | | β= 0 | | | | | | α= 0.5 | | β= 0 | | | | | |
	AM	GM	Blume	MoM	C1	C2	C3	C4	AM	GM	Blume	MoM	C1	C2	C3	C4
γ= 0	5	9	7	8	1	2	3	4	5	8	6	7	1	2	3	4
γ= 0.2	3	9	2	7	6	5	4	1	2	8	3	7	6	5	4	1
γ= 0.5	4	5	3	2	8	7	6	1	4	6	3	2	8	7	5	1
γ= -0.2	6	9	7	8	3	4	5	1	5	8	6	7	2	3	4	1
All γ's (α= 0.4 / α= 0.5, β= 0)	2	8	6	7	2	2	2	1	2	8	6	7	4	4	2	1
All γ's and α's	2	8	6	7	4	4	2	1	2	8	3	7	6	3	3	1

(All γ's and α's – based on equation (16))

This table gives the ranks of the different estimators based on the mean absolute error measured as defined in Eq. (17); the grey shaded area gives the aggregated results based on Eq. (16). In case where more results are analyzed simultaneously the estimators are ranked according to their rank sum. The estimators are: AM: arithmetic mean; GM: geometric mean; Blume: estimator according to (Blume, 1974); MoM: mean of geometric and arithmetic average; C1: Cooper estimator according to Eq. (2); C2: Cooper estimator according to Eq. (6); C3: Cooper estimator according to Eq. (8); C4: Cooper estimator according to Eq. (10). For cases with α=0.4 and β=0 the average (max) kurtosis is 0.83 (0.96) and the average (max) simulated volatility is 25 (32)%, for cases with α=0.5 and β=0 it is 1.51 (1.70) and 20 (26)%, respectively

under consideration.[16] Moreover, it might be worth to mention that the performance of the Blume as well as the AM estimator slightly improve with respect to the base case were kurtosis was clearly lower.

As a final disclaimer it should be mentioned that we have not done an extensive analysis of all possible return generating processes. It might well be that processes can be found where also the C4 estimator leads to significantly biased results. Answering this question must be left to future research.

5 International evidence

5.1 Background

In the preceding section it has been shown that under many circumstances the C4 estimator works best. If we are restricted to use a simple estimator, the choice depends most importantly on whether there exists significant negative serial correlation. In that case the MoM estimator outperforms the AM and GM, while in most other cases the AM estimator seems to work best among these three alternatives. Hence, it is an empirical question which of these estimators are more suitable for a practical valuation problem.

Therefore, in this section we focus on the simultaneous impact of correcting the discount rate estimation for biases stemming from observational problems and serial correlation. The question we are interested in is to empirically infer the relationship between the unbiased market risk premium estimation in the context of valuation problems with its geometric and arithmetic average as well as the mean of means. Of course, in an empirical application we do not know the true unbiased estimation. However, according to our results presented in Sect. 4 we assume that the C4 estimator produces estimations close to the unobserved unbiased estimation.

For our analysis we use an international sample of long historical records on stock- and bond-market returns as provided by Jordà et al. (2019). Because of the international nature of these data we use the global CAPM as an analytical starting point for the analysis. Following the work of Solnik (1974), Grauer et al. (1976), Stulz (1981), Solnik (1982), Adler and Dumas (1983), Dumas and Solnik (1995), and others, we write the securities market line equation in an international context as follows:

$$r_i^c = r_f^c + \beta_{i,w}^c \left(r_w^c - r_f^c \right) = r_f^c + \beta_{i,w}^c mrp_w^c \tag{18}$$

Here, c is the currency indicator making clear that all components of the securities market line equation have to be measured in one single currency, i.e. the base currency. This has an important implication: as there is only one risk-free asset in each currency, the risk-free rate r_f^c refers to the domestic government bond issued in

---

16 It should be noted that we had to leave out the EK estimator in this robustness test. The reason is that because of high heteroscedasticity the estimator in several cases generates numerically invalid results.

the country with the currency c as the legal tender. Hence, depending on the home country of the respective investor, the market risk premium mrp_w^c is measured as the difference of the return of a worldwide equity portfolio expressed in the domestic currency c minus the domestic risk-free rate. The realized market risk premium from the perspective of a German investor reflects the premium he was able to earn by holding the world stock market portfolio above the German government bond. As a consequence, from the perspective of a German investor the historically realized world market risk premium is different than from the perspective of a US-investor or any other investor with a different home currency.

It should be noted that the global CAPM according to Eq. (18) is a simplified version of the international CAPM as developed by the authors given above. In principle, these models show that the covariance risk of exchange rates with asset returns are priced adding an additional factor to the pricing equation given above. However, under specific conditions this risk factor can be ignored, for instance if exchange rates reflect purchasing power parity (cf. Stulz 1981; Dumas and Solnik 1995). Empirically it seems, however, that exchange rate risk might be a priced factor (cf. Dumas and Solnik 1995). We nevertheless stick to the simplified version of a global CAPM given in Eq. (18) as even under a truly international CAPM the world market risk premium is a pricing factor (cf. Adler and Dumas 1983), making its empirical estimation a relevant issue.

Finally, the question whether a global or local CAPM is more appropriate depends on the degree of market integration on an international level. While some markets might be highly integrated, others probably are not. And independently of market integration behavioral restrictions, such as the well-known home bias, might prevent investors from fully exploiting international diversification opportunities. While in reality this might cause the appropriate asset pricing model to some kind of mixture between a global and local CAPM,[17] we use a simpler approach by paying attention either to the global or to the local CAPM perspective. Therefore, we also measure the market risk premium of single countries or regions.

5.2 Data

As our base case we use the data provided by Jordà et al. (2019),[18] which we will refer to as the JST data. This is one of the few databases with a very long historical record of stock- and bond-market returns, exchange rates and also real GDP. In fact, this data set covers the 146 year period from 1870 to 2015. Over this period, we have stock and bond market returns for 16 countries with some years with missing entries.[19] However, for the sake of clarity we conduct our analysis from the perspective of a

[17] A more detailed analysis of such a possible relationship can be found in Lau et al. (2010).

[18] For more details on this database see https://www.macrohistory.net/database/. License terms can also be found there. Data has been downloaded on May 7, 2021.

[19] Countries are: Australia, Belgium, Denmark, Finland, France, Germany, Italy, Japan, Netherlands, Norway, Portugal, Spain, Sweden, Switzerland, UK, and US. It should be noted that the JST data also includes Canada and Ireland. However, because of the large number of missing observations these two countries were dropped from our dataset. We group the European countries in a separate European region (EUR).

US, UK, or German investor. Hence, we use the US-Dollar, the British Pound or a combination of Euro/Deutsche Mark/Reichsmark as base currencies. Moreover, we calculated average real GDP weighted returns in order to come up with estimates for the domestic as well as a European and a World market risk premium.[20] For all countries and regions we calculate up to 145 yearly market risk premium realizations in three different base currencies according to the following equation:

$$mrp_{it}^c = \ln\left(\frac{S_{i,t}^c}{S_{i,t-1}^c} \right) - \ln\left(\frac{B_{c,t}}{B_{c,t-1}} \right) \tag{19}$$

Here, c stands for one of the following three currencies: US-Dollar, British Pound or a combination of Euro, Deutsche Mark and its predecessors; i stands for one of the five regions/countries mentioned above. $S_{i,t}^c$ gives the value of a broad stock market portfolio in country/region i in year t expressed in currency c. $B_{c,t}$ gives the value of a portfolio of long-term government bonds at the end of year t issued in the country where c is the domestic currency. Hence, it represents the return of an investment in a portfolio of risk-free bonds issued by the country where the currency c is the legal tender. Because of the global or local CAPM perspective, we do not use all possible currency/country combinations, as this would lead, for instance, to calculate the US market risk premium from the perspective of a German investor. Hence, we combine each currency with the world stock market as well as with those domestic markets, where this currency is the legal tender. The British Pound and the Euro/Mark currency are also combined with the European stock market portfolio.

Because of the special situation during the two World War periods——and also because of a missing data issue——we do all the calculations also for the post-war period from 1957 to 2015 separately. For Germany we also use the online data provided by Richard Stehle.[21] The data, which will be called the FTS-data, provides yearly total returns of a broadly diversified German stock market portfolio and covers the period 1953 to 2013. We use the data over the period 1957 to 2015 by extending the data for the last remaining three years with the returns of the CDAX. The return on the long-term German government portfolio is proxied by the REXP index.[22] However, this index is only available starting from the year 1968. For the missing years we approximate this return by using the yield on German government bonds reported by the Bundesbank and assuming that this represents the yield of a 10-year German government bond and the bond is quoted at par.[23] As this data is only available starting from 1956, we have to restrict the whole FTS data to start from that year.

---

[20] We label all the countries mentioned in the preceding footnote as representing the world (WD) capital market. The European capital market is represented by the subset of European (EUROPE) countries mentioned there.

[21] We have downloaded the data from https://www.wiwi.hu-berlin.de/de/professuren/bwl/bb/daten/data-library. For further description cf. Stehle and Schmidt (2015).

[22] This is a large portfolio of German government bonds with an average maturity of 5.49 years (cf. Deutsche Bundesbank Kapitalmarktstatistik, January 2020, p. 81).

[23] Specifically, we use the time series BBK01.WU004 reported by the Bundesbank.

5.3 Results

The first set of results can be found in Table 5. There we report the statistical distribution of the continuously compounded market risk premium according to Eq. (19) for different combinations of currencies and countries/regions. Not surprisingly, the average returns are clearly impacted by the currency in which they are defined. It should be noted that this might be due to long-term trends in exchange rates as well as persistent differences in the level of the risk-free interest rates.

In general, it could be said that the differences in average market risk premia are not too pronounced. In fact, leaving two cases aside, namely the World and European market risk premium calculated based on the EUR/Mark, the (arithmetic) average after-World War II risk premia are in a range of 4.63–5.73%.[24] However, the World or European market risk premium in EUR/Mar is significantly lower at 3.82%. This might be caused by the huge appreciation of the Deutsche Mark.

When calculating long-term realized market risk premia dispersion is a bit higher, with UK being at the one (4.63%) and Germany (6.86%) at the other end of the range. It should be noted here that for Germany we do not have data for the hyperinflation period 1992–1923 as well as for the World War II collapse period from 1944 to 1948. Moreover, while in most of the cases the volatility of the market risk premium is in the range of 15–20%, it gets somewhat higher (about 22%) for Germany.

As a side remark, it should be pointed out that a direct comparison of the results presented in Table 5 with other sources presenting market risk premia estimations has to be done carefully. As explained in the preceding section, we apply a global CAPM, which means that stock market returns are converted into one base currency (i.e. USD, EUR or GBP); from this return the return on the government bond in that base currency is deducted.[25] Only for the market risk premia of Germany, US and UK a direct comparison with other studies is possible, as for these countries stock returns measured in local currency are compared with local risk-free rates.

What is more interesting for our purposes is the fact that regardless of the country, region or base currency we have excess kurtosis, which seems to be more

---

[24] As our base observations are continuously compounded market risk premia according to Eq. (19), arithmetic average in our context is expressed in continuous compounding as well; if AM is the arithmetic average of the yearly market risk premia observations calculated as ($e^{mrp} - 1$), the arithmetic average given in Table 5 is expressed as $\ln(1 + AM)$. The geometric average in the table refers to the arithmetic average of the continuously compounded market risk premia according to Eq. (19).

[25] The reader interested in the German market might nevertheless wonder why the German market risk premium based on FTS data and measured in Euro in Table 5 is only 3.12%, given that Stehle and Schmidt (2015, p. 39), mention a geometric average for the market risk premium of 6.08%. The stock market data is the same in both cases. However, there is a twofold reason for this deviation. First, in the Stehle and Schmidt (2015) paper the risk premium is measured against a money market rate, while we use the long-term Government bond return as measured by the REXP data. As this bond data is only available since 1956, we couldn't start our analysis earlier. And this is the second reason for the difference mentioned above. As the year 1954 generated a stock market return of about 85%, it makes a difference whether this year is included in the analysis, as in the Stehle and Schmidt (2015) paper, or excluded, as in our analysis.

Table 5 Distribution of the market risk premium relative to different base currencies

Panel A: EUR/Mark

Region	DE	DE (short)	DE (FTS)	WD	WD (short)	EUROPE
Geom. Mean	4.32%	2.74%	3.12%	4.98%	2.44%	4.24%
Arith. Mean	6.86%	5.17%	5.52%	10.05%	3.82%	8.47%
Mean of Means	5.59%	3.95%	4.32%	7.52%	3.13%	6.36%
St.dev.	22.23%	22.79%	22.48%	26.07%	17.18%	24.52%
Max	81.90%	51.08%	54.28%	204.33%	37.09%	189.14%
Min	−65.14%	−65.14%	−59.68%	−88.45%	−52.47%	−67.74%
Excess kurtosis	2.8	1.1	0.6	27.4	1.8	24.5
Skew	0.4	−0.7	−0.5	3.4	−0.9	3.1
+2sigma fat tail	1.35%	−0.58%	−0.58%	−0.10%	−0.58%	−0.83%
−2sigma fat tail	0.62%	2.81%	2.81%	−0.83%	2.81%	−0.83%
N	138	59	59	138	59	138

Panel B: USD

Region	US	US (short)	WD	WD (short)
Geom. Mean	4.13%	3.23%	4.70%	3.89%
Arith. Mean	5.75%	5.03%	6.05%	5.73%
Mean of Means	4.94%	4.13%	5.37%	4.81%
St.dev.	18.38%	19.61%	16.67%	19.92%
Max	42.70%	42.04%	60.51%	40.56%
Min	−71.09%	−71.09%	−70.48%	−70.48%
Excess kurtosis	1.4	2.5	3.0	2.1
Skew	−0.6	−0.9	−0.5	−0.9
+2sigma fat tail	−0.89%	−2.28%	0.48%	−2.28%
−2sigma fat tail	0.50%	1.11%	0.48%	−0.58%
N	144	59	145	59

Panel C: GBP

Region	GB	GB (short)	WD	WD (short)	EUROPE	EUROPE (short)
Geom. Mean	3.56%	3.64%	5.28%	3.49%	4.45%	3.06%
Arith. Mean	4.63%	5.60%	6.70%	4.98%	6.58%	5.02%
Mean of Means	4.09%	4.62%	5.99%	4.23%	5.51%	4.04%
St.dev.	14.70%	20.07%	16.80%	17.71%	19.82%	20.39%
Max	60.63%	60.63%	73.35%	37.42%	99.75%	38.15%
Min	−52.55%	−52.55%	−42.95%	−42.95%	−46.72%	−46.72%
Excess kurtosis	3.5	1.5	2.3	0.2	5.7	−0.3
Skew	−0.3	−0.2	0.2	−0.7	1.0	−0.6
+2sigma fat tail	0.48%	−0.58%	−0.21%	−2.28%	−0.90%	−2.28%
−2sigma fat tail	−0.21%	1.11%	2.55%	1.11%	0.48%	1.11%
N	145	59	145	59	145	59

The table reports descriptive statistics of the continuously compounded market risk premium according to Eq. (19) with respect to one base currency according to Sect. 5 for the US, Germany (DE), United

Table 5 (continued)

Kingdom (GB), the World (WD) and EUROPE, from the perspective of a German (Panel A), US (Panel B) and British (Panel C) investor. Each Panel uses a different base currency. All means are based on continuous compounding; for the definition of the arithmetic average see footnote 24; Std.dev. is annualized. $\pm 2\sigma$ fat tail is the percentage of return realizations above (below) $+(-)$ two standard deviations minus the probability of such realizations under the normal distribution assumption. Raw data is from Jordà et al. (2019) and covers the period 1871 to 2015. For all results in EUR/Mark, estimations for the years 1922, 1923, and 1944 to 1948 are either not available or have been excluded. Columns labeled (short) use return data from 1957 to 2015. The column D (FTS) is based on the FTS series as described in Sect. 5.2 covering the period 1957 to 2015. The return of the long-term government portfolio is based on the REXP-data, which has been extended to the period 1956 using our own estimations

pronounced for pre-war period returns. We also find in 12 out of 15 cases negative skew. This is in line with downside fat tails.

Coming back to the focus of the paper we next investigate the existence of negative serial correlation. For that purpose we calculate the variance ratios of the market risk premium following the approach of Campbell et al. (1997, p. 48 n). It should be noted that under the null hypothesis, i.e. iid returns, variance ratios are not different from 1. In fact, it should not make a difference whether the variance is calculated on the basis of yearly or, let's say, bi-annual returns. The bi-annual variance should simply be twice the annual variance. However, under the presence of serial correlation this is not true anymore. Therefore, in the presence of negative (positive) serial correlation the variance ratio VR(q), intuitively expressed as the ratio of the annualized variance of q-yearly return observations to the variance of yearly return observations, should be the lower (higher) the higher the degree of negative (positive) serial correlation.

In general, the variance ratio can be approximated by a linear combination of the first $q-1$ autocorrelation coefficients with arithmetically declining weights (cf. Campbell et al. 1997, p. 54), i.e. for $q \geq 2$.

$$VR(q) \approx 1 + 2 \sum_{k=1}^{q-1} \left(1 - \frac{k}{q}\right) \rho(k) \, with \, \rho(k) = \frac{Cov(r_t, r_{t+k})}{Var(r_t)} \tag{20}$$

In case $q = 2$ this simplifies to:

$$VR(2) = 1 + \rho(1) \tag{21}$$

Hence, having in mind that VR(2) is the variance ratio of the annualized bi-annual return variance to the yearly return variance, a variance ratio of 0.8 would imply an empirical coefficient of autocorrelation close to -0.2.

Moreover, under an AR(1) process with a one-period autocorrelation coefficient γ, the variance ratio can be written as (cf. Campbell et al. 1997, p. 49):

$$VR(q) = 1 + \frac{2}{1-\gamma} \left[\gamma - \frac{\gamma^q}{q} - \frac{\gamma - \gamma^q}{q(1-\gamma)} \right] \tag{22}$$

We have already discussed in Sect. 4.2 that for our simulations this implies a range of variance ratios from 35 to 148%. Now we can compare this with the

Table 6 Variance ratios of the global and local market risk premium

Panel A: EUR/Mark

q \ Region	DE (short)	DE (FTS)	WD (short)	EUROPE (short)	Mean ratio
2	102%	100%	106%	100%	109%
3	90%	86%	99%	92%	102%
4	78%	72%	95%	91%	96%
5	61%	58%	89%	82%	87%
6	45%	45%	85%	72%	77%
7	40%	40%	83%	70%	75%
8	35%	33%	78%	68%	71%
9	36%	35%	82%	71%	78%
10	38%	35%	81%	70%	82%

Panel B: USD

q \ Region	US	US (short)	WD	WD (short)	Mean ratio	Mean ratio (short)
2	97%	82%	103%	85%	103%	90%
3	85%	65%	96%	67%	96%	73%
4	85%	64%	94%	62%	95%	72%
5	83%	64%	88%	55%	92%	71%
6	75%	57%	81%	40%	85%	60%
7	70%	53%	78%	36%	82%	57%
8	69%	49%	77%	33%	82%	54%
9	69%	50%	77%	33%	83%	58%
10	70%	48%	78%	32%	85%	58%
12	76%		88%		97%	
15	74%		93%		104%	
20	67%		92%		107%	

Panel C: GBP

q \ Region	GB	GB (short)	WD	WD (short)	EUROPE	EUROPE (short)	Mean ratio	Mean ratio (short)
2	93%	86%	94%	87%	92%	96%	96%	96%
3	77%	66%	90%	73%	86%	81%	88%	81%
4	73%	56%	93%	68%	84%	76%	88%	77%
5	73%	54%	96%	59%	80%	63%	89%	70%
6	70%	47%	99%	52%	78%	50%	90%	62%
7	67%	42%	102%	52%	77%	46%	90%	60%
8	64%	37%	104%	48%	77%	40%	92%	56%
9	67%	39%	107%	48%	75%	35%	94%	57%
10	71%	41%	110%	47%	72%	29%	97%	56%
12	80%		121%		77%		110%	
15	91%		137%		84%		129%	

Table 6 (continued)

Panel C: GBP

q \ Region	GB	GB (short)	WD	WD (short)	EUROPE	EUROPE (short)	Mean ratio	Mean ratio (short)
20	93%		148%		92%		149%	

The table reports the variance ratios calculated according to Campbell et al. (1997, p. 48 n), for different time lengths q of the continuously compounded market risk premium for Germany (DE), EUROPE, Great Britain (GB), the US, and the World (WD), from the perspective of a German (Panel A), US (Panel B) and British (Panel C) investor. Raw data is from Jordà et al. (2019) and covers the period 1871 to 2015; the period labelled (short) covers 1957 to 2016. For more details cf. Sect. 5. Moreover, the column DE (FTS) is based on the FTS series as described in Sect. 5.2 covering the period 1957 to 2016. Each Panel uses a different base currency. For returns in EUR/Mark only the short period could be used because of missing data. In the before last column, the ratio of those columns is given, which have a variance ratio lower than 100%. The last column the ratio of the mean observed variance ratio to the expected variance ratio according to Eq. (23) is given

empirical results, which are given in Table 6. Note that the variance ratio depends on the length of the time period over which the return is measured. The length of this period measured in years is labelled as q. It should be noted that when using Euro or the Deutsche Mark (and its predecessors) as a base currency, we are restricted to post-World War II periods as we have missing observations in the pre-World War II period.

Overall, it could be said that the picture in Table 6 is quite consistent. In about 7% of all region/frequency-observations we derive variance ratios above one. Almost all of these cases arise when using the British Pound as a base currency and calculating world-wide market risk premia. Maybe this could be related to the World War experience. In that case it might well be that we would observe similar results, if we were able to calculate Euro/Mark-based variance ratios for periods including the two World Wars. As a side remark, however, it should be noted that having serial correlation in the long run might give way to some serious questions about the rationality of the market equilibrium.[26]

As a final remark it should be said that we do not report any statistical test for the variance ratios. As has been shown by Campbell et al. (1997, p. 57 n.), there are serious difficulties in making long-horizon inferences based on these variance ratios. If the ratio q/T is not close to zero these tests will have little power. And, moreover, even under the null hypothesis variance ratios tend to be below one. In fact, Richardson and Stock (1989) and Campbell et al. (1997, p. 58), show that expected variance ratios under the null hypothesis converge to

---

[26] In fact, in such cases stock prices might become predictable over the long run. An extensive discussion of the implications caused by long-run negative serial correlation can be found in Shiller (2014, p. 1496 n).

$$E\big[VR(q,T)\big] = \left(1 - \frac{q}{T}\right)^2 \qquad (23)$$

Therefore, having q/T-ratios of up to 1/6 the fact that most of the empirical variance ratios are below one should not be overinterpreted. In order to cope with this issue an additional statistic is presented in Table 6. It is called the Mean Ratio and it gives the ratio of the average empirical variance ratio over the respective regions to the expected variance ratio according to Eq. (23). In case of negative serial correlation this ratio should be lower than 1. It can be seen that only in a few cases, i.e. 8 out of 51, this does not hold and ratios are above one. It is interesting to note that most of these cases happen for very long frequencies, i.e. for frequencies beyond 10 years.

To sum up, despite the poor statistical robustness the evidence presented here is broadly in line with the notion that negative serial correlation tends to be an ubiquitous phenomenon.

5.4 Practical implications

The evidence presented so far should not be ignored when it comes to the estimation of discount rates. Starting from this observation we now ask the question whether there is any practical guidance that can be drawn from our results. In this respect, two remarks emerge from our analysis.

First, because serial correlation seems to be a pervasive phenomenon ideally the C4 estimator according to Eq. (12) should be used for estimating discount rates in valuation problems. This is true even in the presence of heteroscedasticity, as we have shown. However, by using this estimator the valuation problem is enlarged by new degrees of freedom, as one needs to have an information about the extent of serial correlation over many different lags. This is not easy to implement, especially also because the analysis presented in this section showed that due to the relatively small number of return observations results are not very robust.

Therefore, as a second remark we address the question whether our analysis would allow for any kind of heuristic rule in order to simplify the problem. For that purpose, we run an additional analysis, where we estimate the presumably least biased market risk premium according to the C4 estimator in Eq. (12). The period specific variance σ_N^2 is estimated using the variance ratio results presented in Table 6. We then compare these presumably least biased estimates with the well-known AM, GM, and MoM estimators.

Some first graphical results are presented in Figs. 3 and 4. In the majority of the cases the least biased discount rate estimation is between the geometric and the arithmetic average. This corroborates the intuition that because of negative serial correlation the unbiased discount rate estimators have to be corrected downward. Interestingly, based on the empirical distributions found here, this correction in many cases leads to unbiased estimators below the arithmetic average. At the same time, it can be seen that the unbiased estimator is rarely below the MoM. Hence, form graphical inspection it is unclear, whether the AM or MoM estimator perform
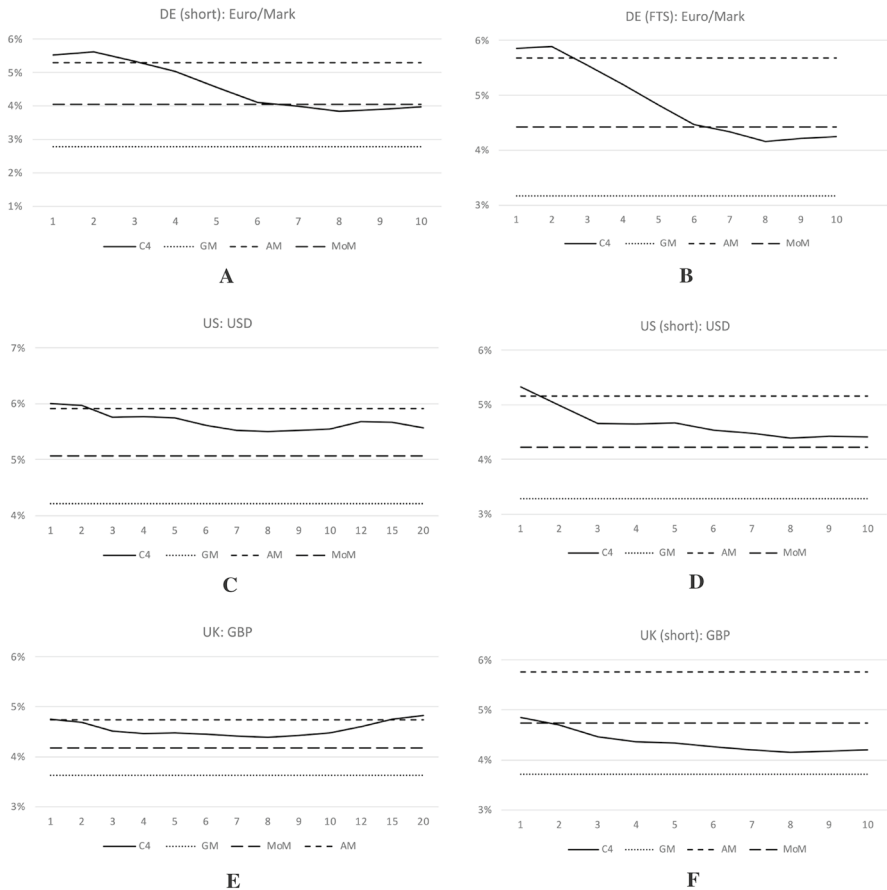
Fig. 3 Unbiased local market risk premia based on the C4-estimator for Germany, US, and UK. This figure draws the discretely compounded unbiased market risk premia estimated on the basis of the C4 estimator according to Eq. (12); for more details cf. Sect. 3.2. Raw data is from JST or FTS as described in Sect. 5.2 covering the period 1871 to 2015 or 1957 to 2015. GM, AM, and MoM indicate the discretely compounded geometric or arithmetic mean or the mean of both. **A** Germany (short JST data) with Deutsche Mark (and its predecessors) and Euro as base currency. **B** Germany (FTS data) with Deutsche Mark and Euro as base currency. **C** US (JST data) with US-Dollar as base currency. **D** US (short JST data)) with US-Dollar as base currency. **E** UK (JST data) with British Pound as base currency. **F** UK (short JST data) with British Pound as base currency

better. Of course, both are clearly biased, but which of the two tends to be less biased cannot simply be inferred from graphical inspection.

Therefore, we perform a more precise analysis by computing the estimation biases based on the measures introduced in Eqs. (16) and (17) in Sect. 4.1. A similar approach has already been used when evaluating the simulation results in Sect. 4.2. By using the currency/region combinations presented in Table 6 and assuming that the C4 estimator according to Eq. (9) produces the least biased results, we can calculate for every discounting period q a weighted or unweighted estimation error based
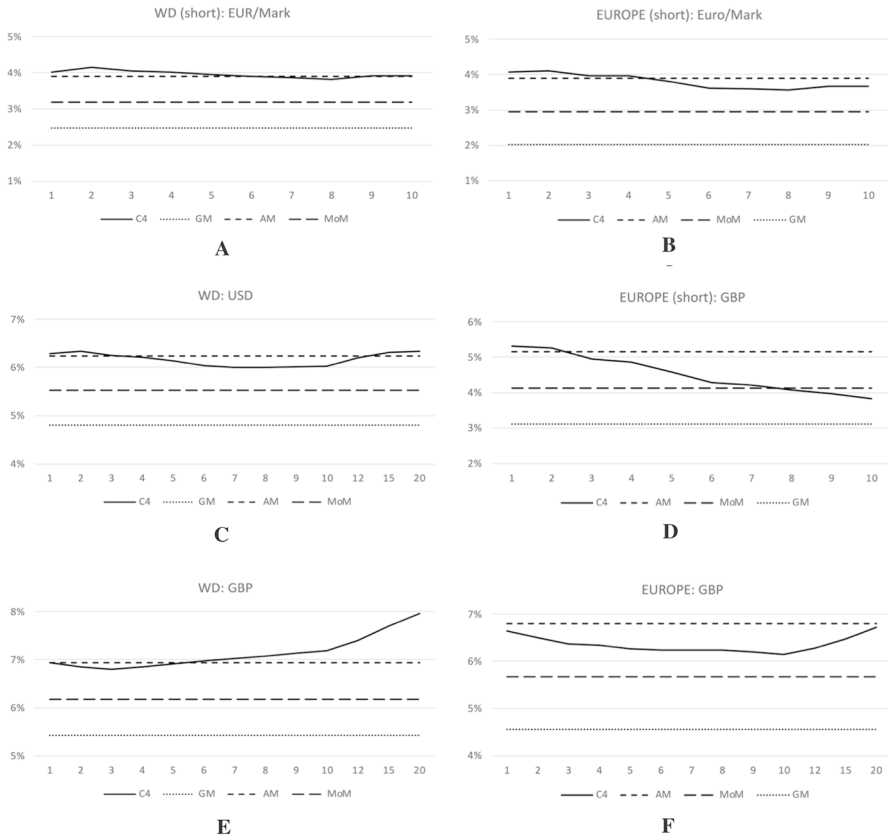
Fig. 4 Unbiased global market risk premia based on the C4-estimator. This figure draws the discretely compounded unbiased market risk premia estimated on the basis of the C4 estimator according to Eq. (12). Raw data is from JST as described in Sect. 5.2 covering the period 1871 to 2015 or 1957 to 2015. GM, AM, and MoM indicate the discretely compounded geometric or arithmetic mean or the mean of both. **A** World (short) with Deutsche Mark and Euro as base currency, **B** Europe (short) with Deutsche Mark and Euro as base currency. **C** World with US-Dollar as base currency, **D** Europe (short) with British Pound as base currency. **E** World with British Pound as base currency, **F** Europe with British Pound as base currency

on Eqs. (16) and (17). Further distributional data needed for this exercise is taken from Table 5. Results are reported in Table 7.

Several results emerge from this exercise. First, it can be seen that in practically all cases the GM produces the most biased results. Secondly, AM and MoM generate significantly smaller biases. Third, the MoM is more likely to outperform the AM, if there is more pronounced negative serial correlation. In fact, those cases were MoM is ranked on the first place are the cases where the variance ratios are strongly declining over time. However, as long as variance ratios stay relatively high, i.e. they are not clearly going below 70–80%, the AM performs better. It is important to emphasize that this result to some extent might also be driven by the length of the time series. In fact, in Panel B and C the MoM estimator outperforms

Table 7 Estimation errors and rankings of simple estimators

Panel A: EUR/Mark

		Region	DE (short)	DE (FTS)	WD (short)	EUROPE (short)	Avg. Rank (local)	Avg. Rank (global)
GM	Eq. (16)	MAD	8%	7%	8%	9%		
	Eq. (16)	Rank	3	3	3	3	3.0	3.0
	Eq. (17)	MAD	2%	2%	1%	2%		
	Eq. (17)	Rank	3	3	3	3	3.0	3.0
AM	Eq. (16)	MAD	6%	6%	0%	1%		
	Eq. (16)	Rank	2	2	1	1	2.0	1.0
	Eq. (17)	MAD	1%	1%	0%	0%		
	Eq. (17)	Rank	2	2	1	1	2.0	1.0
MoM	Eq. (16)	MAD	1%	1%	4%	4%		
	Eq. (16)	Rank	1	1	2	2	1.0	2.0
	Eq. (17)	MAD	1%	0%	1%	1%		
	Eq. (17)	Rank	1	1	2	2	1.0	2.0
GM		avg. Rank	3	3	3	3	3.0	3.0
AM		avg. Rank	2	2	1	1	2.0	1.0
MoM		avg. Rank	1	1	2	2	1.0	2.0

Panel B: USD

		Region	US	US (short)	WD	WD (short)	Avg. Rank (local)	Avg. Rank (global)
GM	Eq. (16)	MAD	10%	7%	10%	5%		
	Eq. (16)	Rank	3	3	3	3	3.0	3.0
	Eq. (17)	MAD	1%	1%	1%	1%		
	Eq. (17)	Rank	3	3	3	3	3.0	3.0
AM	Eq. (16)	MAD	2%	3%	0%	5%		
	Eq. (16)	Rank	1	2	1	2	1.5	1.5
	Eq. (17)	MAD	0%	0%	0%	1%		
	Eq. (17)	Rank	1	2	1	2	1.5	1.5
MoM	Eq. (16)	MAD	4%	2%	5%	0%		
	Eq. (16)	Rank	2	1	2	1	1.5	1.5
	Eq. (17)	MAD	1%	0%	1%	0%		
	Eq. (17)	Rank	2	1	2	1	1.5	1.5
GM		avg. Rank	3	3	3	3	3.0	3.0
AM		avg. Rank	1	2	1	2	1.5	1.5
MoM		avg. Rank	2	1	2	1	1.5	1.5

Table 7 (continued)

Panel C: GBP

		Region	GB	GB (short)	WD	WD (short)	EUROPE	EUROPE (short)	Avg. Rank (local)	Avg. Rank (global)
GM	Eq. (16)	MAD	7%	3%	14%	5%	14%	14%		
	Eq. (16)	Rank	3	2	3	3	3	3	3.0	3.0
	Eq. (17)	MAD	1%	1%	2%	1%	2%	6%		
	Eq. (17)	Rank	3	2	3	3	3	3	3.0	3.0
AM	Eq. (16)	MAD	1%	8%	3%	3%	3%	3%		
	Eq. (16)	Rank	1	3	1	2	1	1	2.0	1.0
	Eq. (17)	MAD	0%	1%	0%	0%	0%	5%		
	Eq. (17)	Rank	1	3	1	2	1	2	2.0	1.0
MoM	Eq. (16)	MAD	3%	2%	9%	1%	5%	5%		
	Eq. (16)	Rank	2	1	2	1	2	2	1.0	2.0
	Eq. (17)	MAD	0%	0%	1%	0%	1%	1%		
	Eq. (17)	Rank	2	1	2	1	2	1	1.0	2.0
GM		avg. Rank	3	2	3	3	3	3	3.0	3.0
AM		avg. Rank	1	3	1	2	1	2	2.0	1.0
MoM		avg. Rank	2	1	2	1	2	2	1.0	2.0

The table reports average estimation errors of the simple estimators geometric mean (GM), arithmetic mean (AM), and the mean of both (MoM) relative to the C4 estimator according to Eq. (9) and their relative ranking. Errors are calculated based on (16) and (17) and averaged over all discount periods q given in Table 6. The mean absolute difference (MAD) is the absolute value of the sum of the errors over all discounting periods divided by number of observations. Avg. Rank (local) (global) indicates the rank for those cases where the local (global) CAPM has been applied, For calculating the C4 estimator variance ratios reported in Table 6 as well as the AM and GM reported in Table 5 are used. Raw data is from Jordà et al. (2019) and covers the period 1870 to 2015; the period labelled (short) covers 1957 to 2015. For more details cf. Sect. 5. Moreover, the column DE (FTS) is based on the FTS series as described in Sect. 5.2 covering the period 1957 to 2015. Panel A uses the Euro/Mark as a base currency, while Panel B the US-Dolar and Panel C the British Pound

the AM estimator in (almost) all those cases where short time series are used, while in the cases with the long time series the opposite is true.

Fourth, from graphical inspection one gets the impression that the MoM tends to perform better than the AM in those cases where the local CAPM is applied, i.e. Figure 3, while in the global CAPM cases, i.e. Fig. 4, the opposite tends to be true. This might be driven by the different degrees of negative serial correlation we have found in local vs. international markets. For this reason, we have calculated average ranks in Table 7 for the local and global CAPM cases separately. Actually, it can be said that in the local CAPM cases the MoM ranks slightly better than the AM. The average rank over all three Panels in Table 7 is 1.2 as compared to 1.8 for the AM. However, the opposite happens, if the global CAPM is applied. Here the average rank of the MoM is 1.8 as compared to 1.2 for the AM. This result also implies that

when doing a ranking over all cases considered in Table 7 the average rank of both estimators is equal. And the question whether the ranking is based on Eq. (16) or (17) is almost irrelevant.

Fifth, as a final remark it should also be said that at least in some cases shown in Figs. 3 and 4 it seems that the MoM gets the better the longer the discounting period is. This is not totally surprising as we have seen that variance ratios tend to decline in the length of the return measurement period. However, it cannot be taken as a general rule. Moreover, when it comes to the practical implementation of these estimators it might not be viable to switch among MoM and AM estimators depending on the length of the discounting period.

To summarize, based on the reasoning presented in this section we should be careful in making any unconditional statement regarding the relative performance of the two estimators. The analysis presented here, however, corroborates the perception that in those cases where we have a noteworthy degree of negative serial correlation the ranking would be tilted towards the MoM. And if variance ratios are decreasing in the length of the time periods, the usage of the MoM estimator would be even more important in cases a company with substantially growing cash flows has to be valued.

6 Conclusion

In this paper we have done a twofold analysis. First, we started from the well-known result of Cooper (1996) that estimating unbiased discount rates in a perfect capital market setting implies to use discount rates that display an increasing premium above the arithmetic average the longer the discounting period is. At the same time, however, it has been shown that a downward correction of these discount rates is warranted in the presence of negative serial correlation. The appropriate estimator has also been developed by Cooper (1996) and was labelled as the C4 estimator in this paper.

Taking into account that besides negative serial correlation also heteroscedasticity is a well-documented phenomenon in capital markets, we were interested in detecting the potential bias of the C4 estimator in the presence of both phenomena. Based on a simulation analysis we have presented some evidence that this estimator seems to work pretty well even under both serially correlated and heteroscedastic stock market returns. In any case, it clearly outperforms all other estimators analyzed in this paper under a wide range of different parameter combinations for the return generating process.

In the second part of the paper we then presented evidence of serial correlation and heteroscedasticity in a worldwide sample of realized market risk premia. By applying the global and local CAPM we provided evidence that negative serial correlation is a wide-spread phenomenon, even though we also found a few currency/region combinations displaying positive serial correlation. However, this only happened when taking the pre-World War II period into account. We have also shown that heteroscedasticity is a wide-spread phenomenon.

Finally, we applied these results to the so called C4 estimator, i.e. an estimator accounting for serial correlation, and found that in the majority of the cases the presumably least biased market risk premia delivered by this estimator is somewhere in between the MoM and the AM, with the AM performing the better the less prevalent negative serial correlation is, the shorter the discounting periods are, and the longer the historical return series are. Hence, using the MoM tends to produce better results for cases with a noteworthy degree of negative serial correlation and for longer discounting periods. However, by comparing the two estimators based on the empirical data collected for this paper we cannot give a clear indication which of the two likely produces less biased results. If ever, we found some evidence that the MoM is slightly outperforming the AM under a local CAPM perspective, while the opposite tends to be true under a global CAPM perspective. Overall, even though there is some rationale for the so-called mean of means rule used by practitioners any recommendation strongly depends on the distributional parameters of the market risk premium.

Of course, the results presented in this paper need further investigation. For instance, in our simulation analysis we restricted ourselves to a relatively small range of return generating processes. Also, we did not account for any firm specific characteristics or different time patterns of cash flows. In order to better understand the robustness of the C4 estimator it should be challenged by a much broader range of different processes and firm specific characteristics. Also, bringing the empirical analysis to a wider range of datasets with higher frequency should be helpful.

Declarations

Conflict of interest The author has no relevant financial or non-financial interests to disclose.

References

Adler M, Dumas B (1983) International portfolio choice and corporate finance: a synthesis. J Financ 38(3):925–984. https://doi.org/10.2307/2328091

Antonczyk RC and Mark K (2010) Reduktion von Schätzproblemen bei langfristigen Endvermögensprognosen im Rahmen von Kapitalanlageentscheidungen. Corporate Finance Biz (Heft 7), 477–488

Berg T, Kaserer C (2013) Extracting the expected equity premium from credit extracting the expected equity premium from credit spreads. J Deriv 21(1):8–26

Berkman H, Jacobsen B, Lee JB (2011) Time-varying rare disaster risk and stock returns. J Financ Econ 101(2):313–332. https://doi.org/10.1016/j.jfineco.2011.02.019

Blume ME (1974) Unbiased estimators of long-run expected rates of return. J Am Stat Assoc 69:634–638. https://doi.org/10.1080/01621459.1974.10480180

Breuer W, Mark K (2013) Valuation of young growth firms and firms in emerging economies. Rethink Valuat Pricing Models. https://doi.org/10.1016/B978-0-12-415875-7.00012-9

Breuer W, Fuchs D, Mark K (2014) Estimating cost of capital in firm valuations with arithmetic or geometric mean – or better use the Cooper estimator? Eur J Finance 20(6):568–594. https://doi.org/10.1080/1351847X.2012.733717

Breuer W, Kohn K, Mark K (2017) A note on corporate valuation using imprecise cost of capital. J Bus Econ 87(6):709–747. https://doi.org/10.1007/s11573-016-0832-6

Butler JS, Schachter B (1989) The investment decision: estimation risk and risk adjusted discount rates. Financ Manage 18(4):13–22. https://doi.org/10.2307/3665793

Campbell JY (2003) Consumption-based asset pricing. In: Constantinides GM, Harris M, Stulz RM (eds) Handbook of the economics of finance. Elsevier, Amsterdam, pp 801–885

Campbell JY, Cochrane JH (1999) By force of habit: a consumption-based explanation of aggregate stock market behavior. J Polit Econ 107(2):205–251. https://doi.org/10.1086/250059

Campbell JY, Shiller RJ (1988) Stock prices, earning, and expected dividends. J Finance 43(3):661–676. https://doi.org/10.1111/j.1540-6261.1988.tb04598.x

Campbell JY, Lo AW, MacKinlay AC (1997) The econometrics of financial markets. Princeton University Press, Princeton

Campello M, Chen L, Zhang L (2008) Expected returns, yield spreads, and asset pricing tests. Rev Financ Stud 21(3):1297–1338. https://doi.org/10.1093/rfs/hhn011

Claus J, Thomas J (2001) Equity premia as low as three percent? Evidence from analysts' earnings forecasts for domestic and international stock markets. J Finance 56(5):1629–1666. https://doi.org/10.1111/0022-1082.00384

Cochrane JH (2005) Asset pricing (Revised edition). Princeton University Press, Princeton

Cooper I (1996) Arithemtic versus geometric mean estimators: setting discount rates for capital budgeting. Eur Financ Manag 2(2):157–167. https://doi.org/10.1111/j.1468-036X.1996.tb00036.x

Dittmann I, Maug EG (2008) Biases and error measures: how to compare valuation methods (SSRN Library No. No. 2006-07). Available from: http://ssrn.com/abstract=947436

Dumas B, Solnik B (1995) The world price of foreign exchange risk. J Finance 50(2):445–479. https://doi.org/10.1111/1468-036X.00028

Easton P (2004) PE Ratios, PEG Ratios, and estimating the implied expected rate of return on equity capital. Account Rev 79(1):73–95. https://doi.org/10.2308/accr.2004.79.1.73

Elsner S, Krumholz HC (2013) Corporate valuation using imprecise cost of capital. J Bus Econ 83(9):985–1014. https://doi.org/10.1007/s11573-013-0687-z

Fama EF, French KR (2002) The equity premium. J Finance 57(2):637–659. https://doi.org/10.1111/1540-6261.00437

Gebhardt WR, Lee CMC, Swaminathan B (2001) Toward an implied cost of capital. J Account Res 39(1):135–176. https://doi.org/10.1111/1475-679X.00007

Grauer FLA, Litzenberger RH, Stehle RE (1976) Sharing rules and equilibrium in an international capital market under uncertainty. J Financ Econ 3(3):233–256. https://doi.org/10.1016/0304-405X(76)90005-2

Indro DC, Lee WY (1997) Biases in arithmetic and geometric averages as estimates of long-run expected returns and risk premia. Financ Manage 26(4):81–90

Jäckel C, Kaserer C, Mühlhäuser K (2013) Analystenschätzungen und zeitvariable Marktrisikoprämien - Eine Betrachtung der europäischen Kapitalmärkte. Die Wirtschaftsprüfung 8:365–383

Jacquier E, Kane A, Marcus AJ (2003) Geometric or arithmetic mean: a reconsideration. Financ Anal J 59(6):46–53. https://doi.org/10.2469/faj.v59.n6.2574

Jacquier E, Kane A, Marcus AJ (2005) Optimal estimation of the risk premium for the long run and asset allocation: a case of compounded estimation risk. J Financ Economet 3(1):37–55. https://doi.org/10.1093/jjfinec/nbi001

Jordà Ò, Knoll K, Kuvshinov D, Schularick M, Taylor AM (2019) The rate of return on everything, 1870–2015. Q J Econ 134(3):1225–1298. https://doi.org/10.1093/qje/qjz012

Koller T, Goedhart M, Wessels D (2020) Valuation: measuring and managing the value of companies. (McKinsey & Company Inc., Ed.) (7th ed.). Wiley, Hoboken

Lau ST, Ng L, Zhang B (2010) The world price of home bias. J Financ Econ 97(2):191–217. https://doi.org/10.1016/j.jfineco.2010.04.002

Lettau M, Wachter JA (2011) The term structures of equity and interest rates. J Financ Econ 101(1):90–113. https://doi.org/10.1016/j.jfineco.2011.02.014

Ohlson JA, Juettner-Nauroth BE (2005) Expected EPS and EPS Growth as Determinants of value. Rev Account Stud 10(2–3):349–365. https://doi.org/10.1007/s11142-005-1535-3

Pástor Ľ, Stambaugh RF (2012) Are stocks really less volatile in the long run? J Finance 67(2):431–478. https://doi.org/10.1111/j.1540-6261.2012.01722.x

Poterba JM, Summers LH (1988) Mean reversion in stock prices: evidence and implications. J Financ Econ 22(1):27–59. https://doi.org/10.1016/0304-405X(88)90021-9

Pratt SP, Grabowski RJ (2014) Cost of Capital: Applications and Examples: Fifth Edition. Cost of Capital: Applications and Examples, 5th edn. Wiley, Hoboken

Richardson M, Stock JH (1989) Drawing inferences from statistics based on multiyear asset returns. J Financ Econ 25:323–357

Santos T, Veronesi P (2006) Labor Income and Predictable Stock Returns. Rev Financ Stud 19(1):1–44. https://doi.org/10.1093/rfs/hhj006

Shiller RJ (1981) Do stock prices move too much to be justified by subsequent changes in dividends? Am Econ Rev 71(3):421–436

Shiller RJ (2014) Speculative asset prices. Am Econ Rev 104(6):1486–1517. https://doi.org/10.1257/aer.104.6.1486

Solnik B (1974) An equilibrium model of the international capital market. J Econ Theory 8(4):500–524. https://doi.org/10.1016/0022-0531(74)90024-6

Solnik B (1982) International arbitrage pricing theory. J Finance 38:449–457. https://doi.org/10.1111/j.1540-6261.1983.tb02251.x

Spierdijk L, Bikker JA, van den Hoek P (2012) Mean reversion in international stock markets: an empirical analysis of the 20th Century. J Int Money Financ 31(2):228–249

Stehle R, Schmidt MH (2015) Returns on German Stocks 1954 to 2013. Kredit Kapital 48(3):427–476

Stehle R (2016)Wissenschaftliches Gutachten zur Schätzung der Marktrisikoprämie (Equity risk premium) im Rahmen der Entgeltregulierung im Auftrag der BNetzA. Berlin

Stulz R (1981) A model of international asset pricing. J Financ Econ 9(4):383–406. https://doi.org/10.1016/0304-405X(81)90005-2

Tsay RS (2010) Analysis of financial time series. Analysis of financial time series: Third Edition, 3rd edn. Wiley, Hoboken

van Binsbergen J, Hueskes W, Koijen R, Vrugt E (2013) Equity yields. J Financ Econ 110(3):503–519

Wachter JA (2006) A consumption-based model of the term structure of interest rates. J Financ Econ 79(2):365–399. https://doi.org/10.1016/j.jfineco.2005.02.004

Wagner W, Jonas M, Ballwieser W, Tschöpel A (2006) Unternehmensbewertung in der Praxis - Empfehlungen und Hinweise zur Anwendung von IDW S1. Die Wirtschaftsprüfung 59(16):1–24

Welch I (2017) Corporate Finance (4th). https://book.ivo-welch.info/home/

Wenger E (2005) Verzinsungsparameter in der Unternehmensbewertung – Betrachtungen aus theoretischer und empirischer Sicht. *Die Aktiengesellschaft*, *Sonderheft*, 9–22