Boston QWAFAFEW

Dec. 17th 2013

Eric Jacquier
Boston University

Estimation of Long-Term Expected Returns and Asset Allocation

Compounded estimation errors in long-term expected returns

More bad news for the equity premium

OUTLINE

1. Motivation
2. Biases of arithmetic and geometric averages, Unbiased forecast
3. Efficient classical forecast (Minimum mean squared error)
4. Robustness to distributional assumptions
5. Optimal long-term allocation with estimation error: Utility Based Forecast

Some initial results in Jacquier, Kane, Marcus (FAJ 2001, J. Financial Econometrics 2005)

MOTIVATION

Estimates of **expected future long-term returns** are crucial inputs in empirical asset pricing:

- Wealth a portfolio is expected to generate over the long term:

  Pension funds, Social security retirement policy

- Wealth needed at a future horizon:

  Use expected future long-term return to back-out investment needed today.

- Input to asset allocation decision: optimal mix of risky and risk-free assets for the long run.

Current portfolio theory

- Very sophisticated models of time varying opportunity:

  Doubts on the applicability of complex intertemporal models (Garlappi & Uppal)

- Effect of parameter uncertainty on forecast and optimal decision not discussed as often.

  Barberis (2000)

Fact (i.i.d returns)

If the arithmetic mean one-period return of a portfolio is: $1 + E(R)$
The long-term H-period expected return is $[1 + E(R)]^H$

- But we use *estimates* of the mean return!

What is the best *estimate* of the H-period expected returns?

Recall: we want to estimate: $E[(1+R)^H]$

Compound the arithmetic average, or the geometric average, or something else?

Arithmetic average:

Estimate $1+E(R)$ by (for example) the sample average of $1+R_t$

Then compound $1 today is expected to grow to $(1+\bar{R})^H = 1.085^{60}$ = **$ 154**

Appeal:

Sample average is the Best estimator of one-period E(R) under i.i.d assumption

Maximum likelihood justification for compounding at $\bar{R}$.

Geometric average:

The rate of return per period G so that $(1+G)^T = P_T/P_1$

$\text{Log}(1+G) = 1/T \log(P_T/P_1)$ the sample mean of log-returns

$(1+G)^H = 1.07^{60}$ = **$ 58**

Appeal: Powerful intuition for compounding, it's what we actually earned in the sample!

- For log-normal returns $\log(1+R) \sim N(\mu, \sigma)$

$$(1+\bar{R}) \quad \text{estimates } \exp(\mu + 0.5\,\sigma^2)$$

$$\text{Log }(1+G) \text{ estimates } \mu \quad \Rightarrow \quad 1+G \text{ estimates } \exp(\mu)$$

$$(1+\bar{R}) = (1+G)\exp(0.5\,\sigma^2)$$

$\sigma = 20\% \rightarrow 1.02$
$\sigma = 30\% \rightarrow 1.046$

- Arithmetic average always larger than the geometric no matter the distribution.

So…. what do People do?

- Ibbotson SBBI yearbook: Simulates future values with arithmetic returns
- Academics favor arithmetic average even recently
- Practitioners lean toward geometric average.

Other Related Problem:

- Recent market downturns (2001, 2008) brings back the question of the equity premium:

  Fama and French (2002)

  - The one-period equity risk premium, a.k.a. the mean return in excess of the risk-free rate, is less than implied by the post-1926 average returns.
  - Estimation error in mean returns large even for long sample sizes
  - They talk about " What is the best $\bar{R}$? "
  - … not how to use the best $\bar{R}$ to compute long-term expected returns.

- What we do here ($1+\bar{R}$) applies beyond the standard sample average

  It applies to any quantity that estimates the one-period return with error.

- What we do **not** discuss:

  Predicting short or medium term returns - Market or Sector timing.
  What the best estimate of one-period average return is.

Biases of Arithmetic and Geometric Methods, Unbiased Forecast

1. Estimating a compound return:

$r_t = \log(1 + R_t) \sim N(\mu, \sigma^2)$, i.i.d.

$$V_H = \$1 \times \exp(\mu H + \sigma \sum_{i=1}^{H} \varepsilon_{t+i}), \qquad \varepsilon \sim \text{i.i.d. } N(0,1) \qquad (1)$$

$$E(V_H) = e^{(\mu + \frac{1}{2}\sigma^2)H} = [1 + E(R)]^H \qquad (2)$$

Ibbotson etc… uses (2) with an *estimate of E(R)*

- We ignore estimation error in σ for now.
  - μ: Only the calendar span can increase precision of estimation
  - σ: Sampling frequency increases precision of estimation

  High frequency data available, estimation of σ is a second order effect. Merton (1980)

- Jensen's Inequality: $E(\widehat{V_H}) = E((1+\bar{R})^H) > [1 + E(\bar{R})]^H = [1 + E(R)]^H = E(V_H)$. (3)=>

Arithmetic method biased upward: $\widehat{V_H}$ biased upward if $\bar{R}$ is unbiased

2 Our version of the arithmetic estimator based on log-normality

First approach: estimate E(R) substitute in $[1 + E(R)]^H$ $A_1 = (1 + \overline{R})^H$

Second approach: estimate μ substitute in $e^{(\mu + \frac{1}{2}\sigma^2)H}$ $A_2 = e^{(\hat{\mu} + \frac{1}{2}\sigma^2)H}$

Both are MLE estimators of $E(V_H)$, not exactly equal in small sample.

We will use A_2 for analytics:

3 Bias of Arithmetic and Geometric Estimators

- $$\hat{\mu} = \frac{1}{T} \sum_{i=1}^{T} \ln(1 + R_{-i}) = \frac{1}{T}\left(\mu T + \sigma \sum_{i=1}^{T} \varepsilon_{-i} \right).$$

  Sample mean is unbiased, with standard error $\sigma/\sqrt{T}$.

  $$\hat{\mu} = \mu + \omega\, \sigma/\sqrt{T}\,, \qquad \omega \sim N(0,1) \qquad (4)$$

- Arithmetic estimator:

  $$A = e^{(\hat{\mu} + \frac{1}{2}\sigma^2)H} = e^{(\mu + \omega\sigma/\sqrt{T} + \frac{1}{2}\sigma^2)H} = e^{(\mu + \frac{1}{2}\sigma^2)H}\, e^{(\omega\ \sigma/\sqrt{T})H}$$

  $$E(A) = e^{(\mu + \frac{1}{2}\sigma^2)H}\, E[e^{\omega\sigma H/\sqrt{T}}] = E(V_H) \qquad e^{\frac{1}{2}\sigma^2 H^2/T} \qquad (5)$$

  Bias

- Geometric estimator

$$E(G) = E(e^{\hat{\mu}H}) = E[e^{(\mu + \omega\sigma/\sqrt{T})H}] = e^{\mu H + \frac{1}{2}\sigma^2 H^2/T}$$

$$= E(V_H)\ e^{\frac{1}{2}\sigma^2(H^2/T - H)} \quad (6)$$

Bias

Biased: upward if H> T

downward if H < T

- Does it matter?

Table 1:
Bias induced by forecasting final portfolio value using arithmetic average return of portfolio over a sample period. Ratio of forecast to true expected value of cumulative return. [σ: annual standard deviation.

Sample period = 75 years

σ	Horizon (years) 10	20	30	40
0.15	1.015	1.062	1.145	1.271
0.2	1.027	1.113	1.271	1.532
0.25	1.043	1.181	1.455	1.948
0.3	1.062	1.271	1.716	2.612

Sample period = 30 years

σ	Horizon (years) 10	20	30	40
0.15	1.038	1.162	1.401	1.822
0.2	1.069	1.306	1.822	2.906
0.25	1.110	1.517	2.554	5.294
0.3	1.162	1.822	3.857	11.023

4 Unbiased estimator

- Simple inspection of (5) or (6): Compounding at $\hat{\mu} + ½\,\sigma^2(1 - H/T)$ removes bias.

- Formally: for sample size T, horizon H, Construct an unbiased estimator in the family

$$C = e^{(\hat{\mu} + ½\, k\, \sigma^2)H} \qquad \textbf{(7)}$$

Family nests G ($k = 0$) and A ($k = 1$).

- Unbiased estimator U: solve for the value k_U that so that E(C) = $E(V_H)$.

$$k_U = 1 - H/T \qquad (8)$$

- Bias vanishes iff H<<T (T/H $\rightarrow \infty$)

Estimates of compounding rates and future portfolio values.

A, G, U: Annual compounding rates with arithmetic and geometric average G, and unbiased estimator (H = 40 years).

V(A), V(G), V(U): Forecasts of future portfolio values, initial $1 invested for H = 40 years

Country/Index	T	Begin	End	Sample estimates $\hat{\mu}$	Sample estimates $\hat{\sigma}$	Annual growth rates A	Annual growth rates G	Annual growth rates U	Future portfolio value V(A)	Future portfolio value V(G)	Future portfolio value V(U)
Canada/TSE	52	1950	2001	6.6	14.9	8.0	6.8	7.1	21.8	14.0	15.5
France/SBF250	52	1950	2001	8.7	22.2	11.8	9.1	9.7	87.0	32.5	40.8
Germany/DAX	52	1950	2001	8.0	22.8	11.2	8.3	9.0	69.4	24.5	31.2
UK/FTAS	52	1950	2001	6.4	24.7	9.9	6.6	8.3	43.8	12.9	24.2
Japan/Nikkei	52	1950	2001	8.8	24.1	12.4	9.2	9.9	107.9	33.8	44.2
Hong Kong[b]	28	1973	2001	10.7	30.7	16.7	11.3	9.0	475.8	72.2	31.0
MSCI/$Emg Mkt	14	1988	2001	8.2	24.2	11.8	8.5	2.5	85.7	26.6	2.7

- Example: $1.08 = \exp(\mathbf{0.066} + ½ * 0.149^2)$ for 40 years : $ 21.8

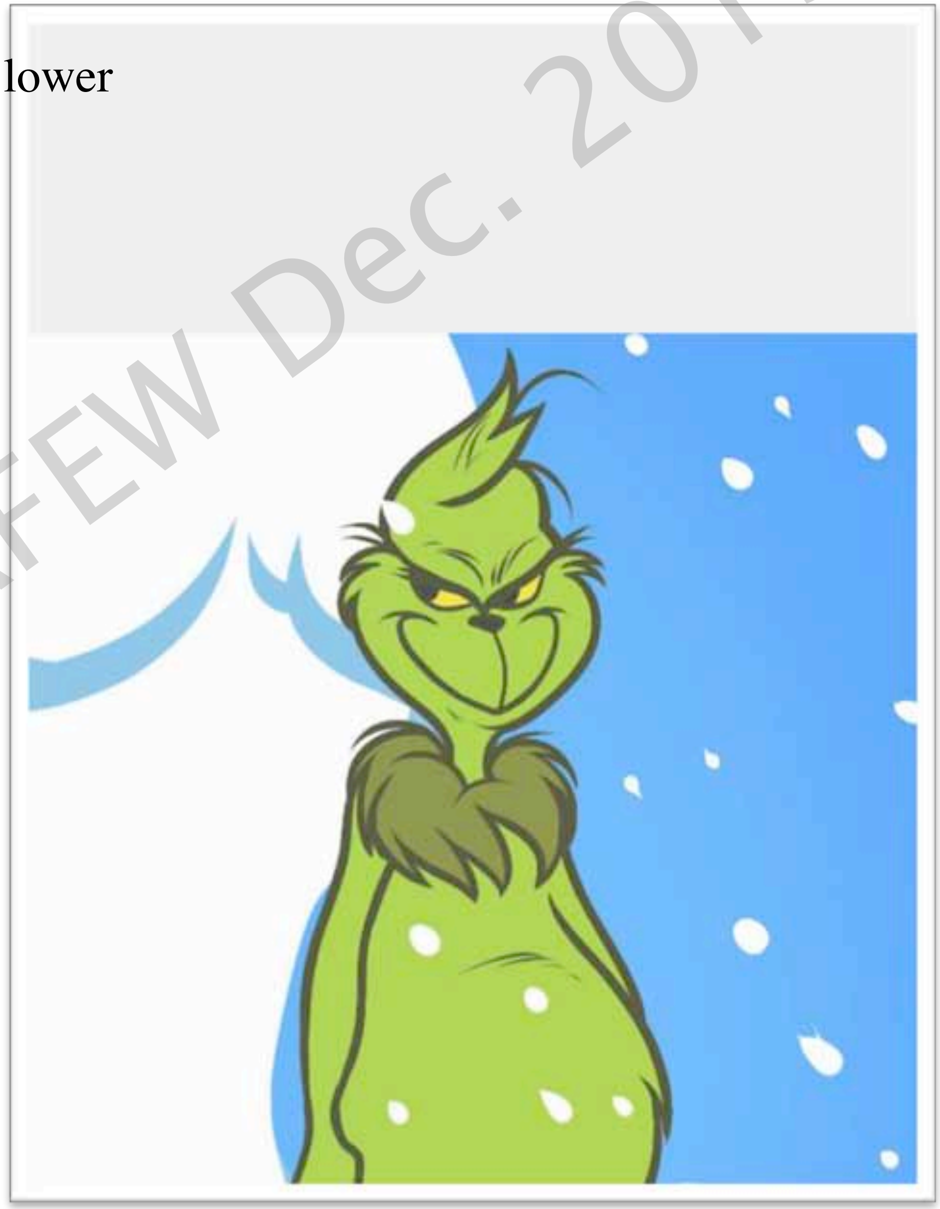
- Remedy: If you use a sample average, get more data !

Estimates of compounding rates and future portfolio values

Country/Index	T	Begin	End	Sample estimates $\hat{\mu}$	$\hat{\sigma}$	Annual growth rates A	G	U	Future portfolio value V(A)	V(G)	V(U)
Canada/TSE	78	1914	2001	4.8	16.7	6.4	4.9	5.6	11.9	6.8	9.0
Canada/TSE	52	1950	2001	6.6	14.9	8.0	6.8	7.1	21.8	14.0	15.5
France/SBF250	145	1857	2001	5.1	19.7	7.3	5.2	6.7	16.7	7.7	13.5
France/SBF250	82	1920	2001	8.5	24.7	12.2	8.9	10.6	101.5	30.0	56.0
France/SBF250	52	1950	2001	8.7	22.2	11.8	9.1	9.7	87.0	32.5	40.8
Germany/DAX	145	1857	2001	1.9	32.2	7.3	1.9	5.8	17.0	2.1	9.6
Germany/DAX	82	1920	2001	5.5	37.0	13.1	5.7	9.4	139.5	9.0	36.7
Germany/DAX	52	1950	2001	8.0	22.8	11.2	8.3	9.0	69.4	24.5	31.2
UK/FTAS	201	1801	2001	2.4	15.6	3.7	2.4	3.4	4.2	2.6	3.9
UK/FTAS	82	1920	2001	5.5	20.0	7.8	5.7	6.7	20.1	9.0	13.6
UK/FTAS	52	1950	2001	6.4	24.7	9.9	6.6	8.3	43.8	12.9	24.2

It does not get better:

Longer calendar sample means are uniformly lower



EFFICIENT ESTIMATION

- Unbiasedness is not a goal per se.
  - Used by statisticians to reduce possibly unmanageable estimation problems
  - Can lead to inferior estimators, (Sample mean vs. Shrinkage)

- Better to minimize a loss function – measure of average distance to the true parameter

Mean squared error: $E[(\beta^* - \beta)^2] = E([\beta^* - E(\beta^*)]^2) + [E(\beta^*) - \beta]^2$

Variance Squared Bias

Can be seen as a generalization of the Maximum Likelihood for small sample

What is the minimum MSE estimator for $E(V_H)$

- $\boldsymbol{C = e^{(\hat{\mu} + ½ k \sigma^2)H}}$
- $MSE(C) = E[C - E(V_H)]^2$

$$= E(e^{\hat{\mu}H + ½k\sigma^2 H} - e^{\mu H + ½\sigma^2 H})^2$$

$$= E(e^{2\hat{\mu}H + k\sigma^2 H} - e^{2\mu H + \sigma^2 H} - 2e^{\hat{\mu}H + ½k\sigma^2 H + \mu H + ½\sigma^2 H})$$

Now Substitute $\hat{\mu} = \mu + \omega\,\sigma/\sqrt{T}$, evaluate, ..

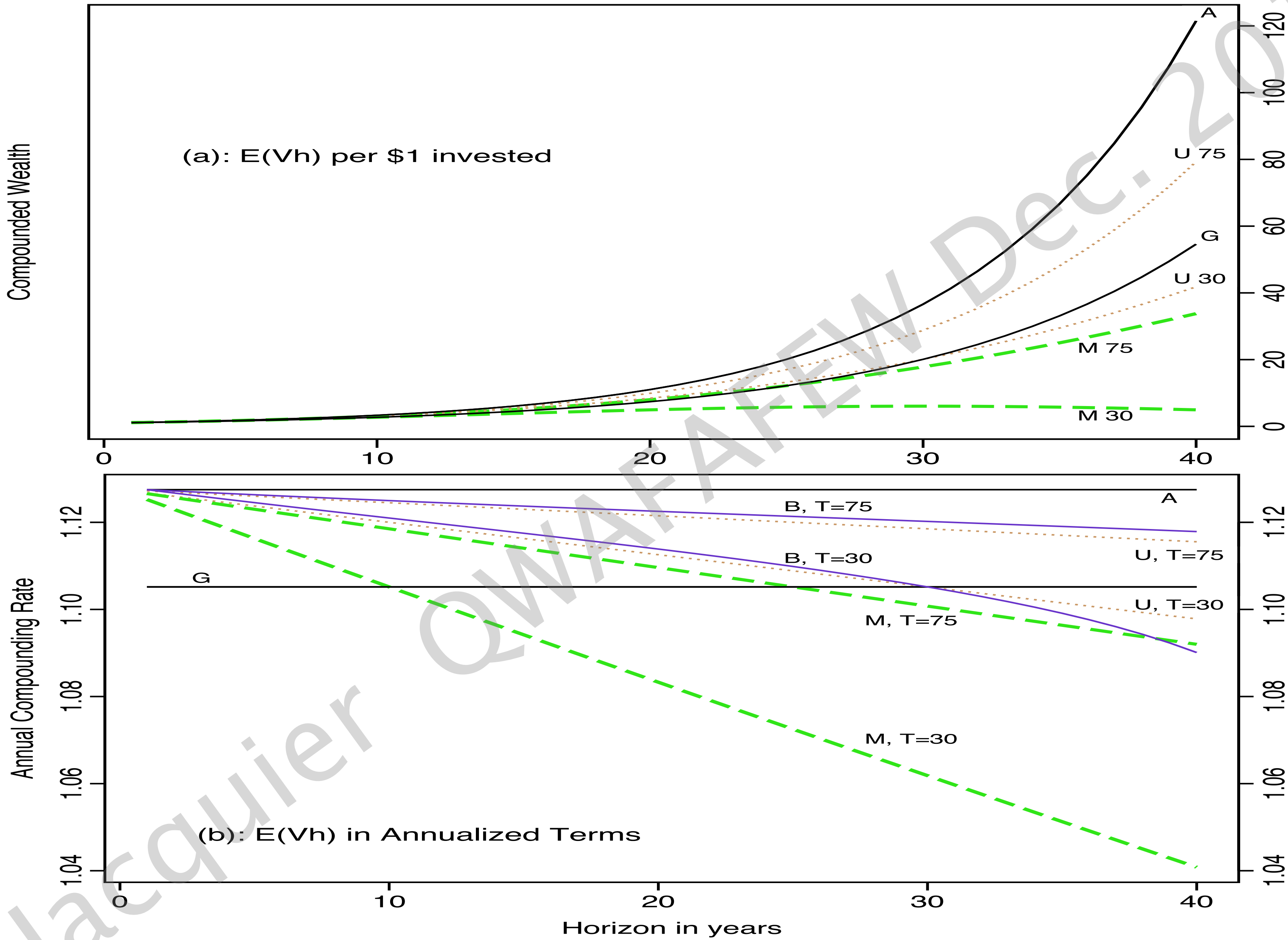
$$MSE(C) = e^{2\mu H + 2\sigma^2 H^2/T + k\sigma^2 H} + e^{2\mu H + \sigma^2 H} - 2\, e^{2\mu H + ½\sigma^2 H^2/T + ½k\sigma^2 H + ½\sigma^2 H}$$

- Find k that minimizes C :

$$\boldsymbol{k_M = 1 - 3\, H/T} \qquad (10)$$

- Stronger downward penalty than the unbiased estimator !

Figure 1: Long-Term wealth forecasts and annual compound rates for *A, G, U, M*. $\hat{\mu} = 0.1$, $\sigma = 0.2$.



Robustness to Distributional Assumptions

- Negative long term autocorrelation in returns

  Summers (1986), Poterba and Summers (1988), and Fama and French (1988) etc..

- Effect on the analysis:

  Autocorrelation enters through the sum of the H future returns and the sumt of the T past returns used to estimate μ.

- Easily corrected

Correlation matrix C: TxT for past returns, HxH for future returns
Vector of ones of lengths T and H: i
Variance of a sum of past and future returns: $\sigma^2 i'Ci$ instead of $T\sigma^2$ or $H\sigma^2$

- The future: $E(V_H)$ in (2), exponential term becomes $H(\mu + \frac{1}{2}\ i'C_H i\ \sigma^2/H)$.
- The past: $\hat{\mu}$ in (14): $\mu + \omega\ \sigma \sqrt{i'C_T i/T}$.
- Compute $F_T = i'C_T i/T$ and $F_H = i'C_H i/H$,

Then

$$k_U = 1 - \frac{H}{T} \times \frac{F_T}{F_H} \tag{21}$$

$$k_M = 1 - \frac{3H}{T} \times \frac{F_T}{F_H} \tag{22}$$

Other non-i.i.d specifications easily extend the ratio F_T / F_H.

Figure 3: estimated MA(4) on annual SP500 log-returns from 26-01.

- Forecasts are barely affected by the correction.

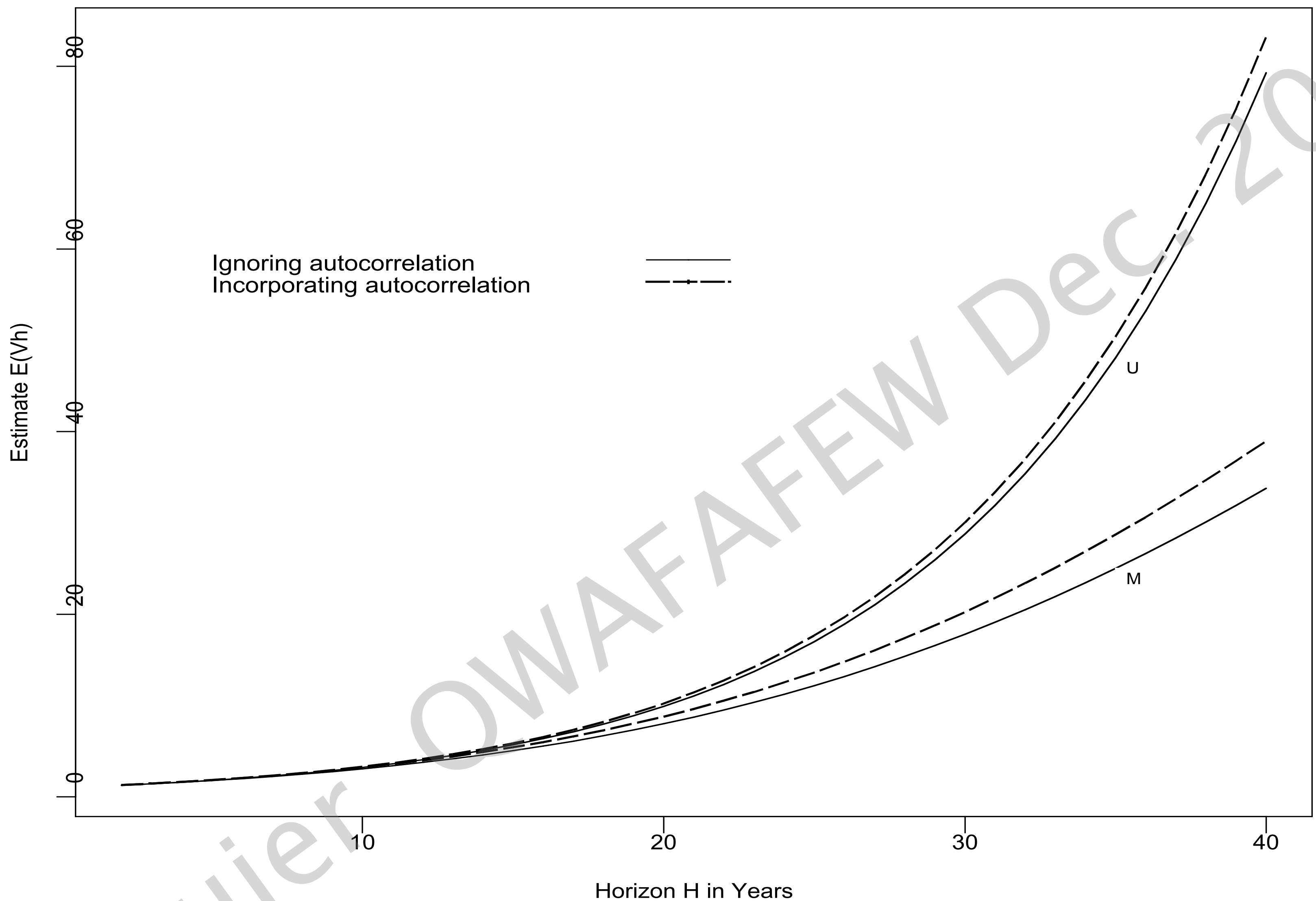


Figure 3: Effect of autocorrelation on estimators *U* and *M*, $\hat{\mu}$ = 0.1, σ = 0.2, *T* = 75, MA(4) on annual S&P returns: θ = (-0.16, -0.02, -0.16, -0.08) estimated on 1926-2001.

- **<u>Heteroskedasticity?</u>**

Little problem for large H: Volatility reverts to unconditional variance above a year

Increased Kurtosis increases the estimation uncertainty of σ.

- **<u>Alternative estimation of μ</u>**

Fama-French (2002): the dividend discount model helps reduce the variance of the estimator of E(R).

Just convert the variance reduction into an equivalent increase in T

- **<u>σ is estimated as well</u>**

Induces non-normality in the predictive distribution of log-returns?
Induces a variance inflation in the predictive distribution $\nu/(\nu-2)$.

Very large degrees of freedom if higher frequency of returns is used

Geometric estimator is robust, it doesn't use an estimate of σ!

Basic Merton framework: No estimation error

$\alpha = \mu + ½\, \sigma^2$

r_0 the risk-free return.

Power utility function and relative risk aversion γ,

Investor maximizes the expectation of her utility of final wealth:

$$U(V_H) = \frac{{V_H}^{1-\gamma}}{1-\gamma} = \frac{1}{1-\gamma} \exp[(1-\gamma) \ln(V_H)] \qquad (11)$$

w: risky portfolio and $(1 - w)$: risk free asset

- Portfolio value is then log-normal with parameters (assume continuous rebalancing):

$$\ln(V_H) \sim \mathrm{N}(\mu_H, \sigma_H^2) \equiv \mathrm{N}[(r_0 + w(\alpha - r_0) - ½\, w^2\sigma^2 H, H w^2 \sigma^2] \qquad (12)$$

- Expected utility is then:

$$E[U(V_H)] = \frac{1}{1-\gamma}\ \exp\{(1-\gamma)\ H\ [r_0 + w(\alpha - r_0) - \tfrac{1}{2}\ w^2\sigma^2 + \tfrac{1}{2}(1-\gamma)w^2\sigma^2]\} \qquad (13)$$

Maximize (13) with respect to w => **Merton optimal allocation** $w^* = \frac{\alpha - r_0}{\gamma\sigma^2}$.

- Independent of the horizon for i.i.d. returns, .. well known

But…..

  - Conventional advice is to **increase allocation with the horizon**.
  - Largely motivated in the literature via predictability in expected returns, e.g., Garcia et al. (2000), Wachter (2000) and others.
  - Conclusions most always assume knowledge of the parameters of the return distribution.

2 Long-term allocation with estimation error

- Optimal asset allocation is affected by estimation uncertainty in α, more so as H/T grows.

  Bawa, Brown, and Klein (1979) and others model it in one-period framework.

  Not very dramatic for a short horizon.

We need to incorporate the uncertainty in α in the above asset allocation:

- Practice of substituting a point estimate in the optimal allocation in place of the unknown α is incorrect.

- The investor, has a distribution for α that represents its uncertainty. a sampling distribution or, for a Bayesian, a posterior distribution.

  => $E[U(V_H)]$ in (13) is random, as a non-linear function of a random variable α.

- Basic decision theory:

  The correct expected utility to maximize, follows from first integrating α out of equation (13).

  In Bayesian jargon, this integration produces the expected utility of wealth given the *data,*

$$E[U(V_H \mid D)]$$

  … to be then optimized by the investor.

- Specifically:

$$E[U(V_H) \mid D] = \int E[U(V_H)/\alpha] \; p(\alpha \mid D)\, d\alpha \qquad (14)$$

- With diffuse priors, the posterior distribution of α is $N(\hat{\alpha}, \sigma^2/T)$.

- Integrate (14):

$$E[U(V_H \mid D)] = \frac{1}{1-\gamma}\exp\{(1-\gamma)H[r_0 + w(\hat{\alpha} - r_0) - ½\, w^2\sigma^2 + ½(1-\gamma)w^2\sigma^2(1 + \frac{H}{T})]\} \qquad (15)$$

Note:

1) α in (13) is replaced with $\hat{\alpha}$,

2) New term in (15) reflects the variance inflation due to the estimation of α.

- Finally maximize (15) for the optimal asset allocation:

$$\boldsymbol{w^* = \frac{\hat{\alpha} - r_0}{\sigma^2[\gamma(1 + \frac{H}{T}) - \frac{H}{T}]}} \qquad \textbf{(16)}$$

- $H << T$: Back to Merton.

$$w^* = \frac{\hat{\alpha} - r_0}{\sigma^2[\gamma(1 + \frac{H}{T}) - \frac{H}{T}]}$$

- $\gamma > 1$: . risky asset allocation decreased relative to known α case

  . The more so the greater the ratio H/T.

  . Contrary to the common advice to invest more in stocks for longer horizons.

  . Happens even if returns are unpredictable.

- Log-utility, $\gamma = 1$.

  Linear in α: estimation uncertainty and the horizon H do not affect the location of the optimum.

- Less than Log-utility investors: $0 < \gamma < 1$

  Actually derive benefit from the estimation error!

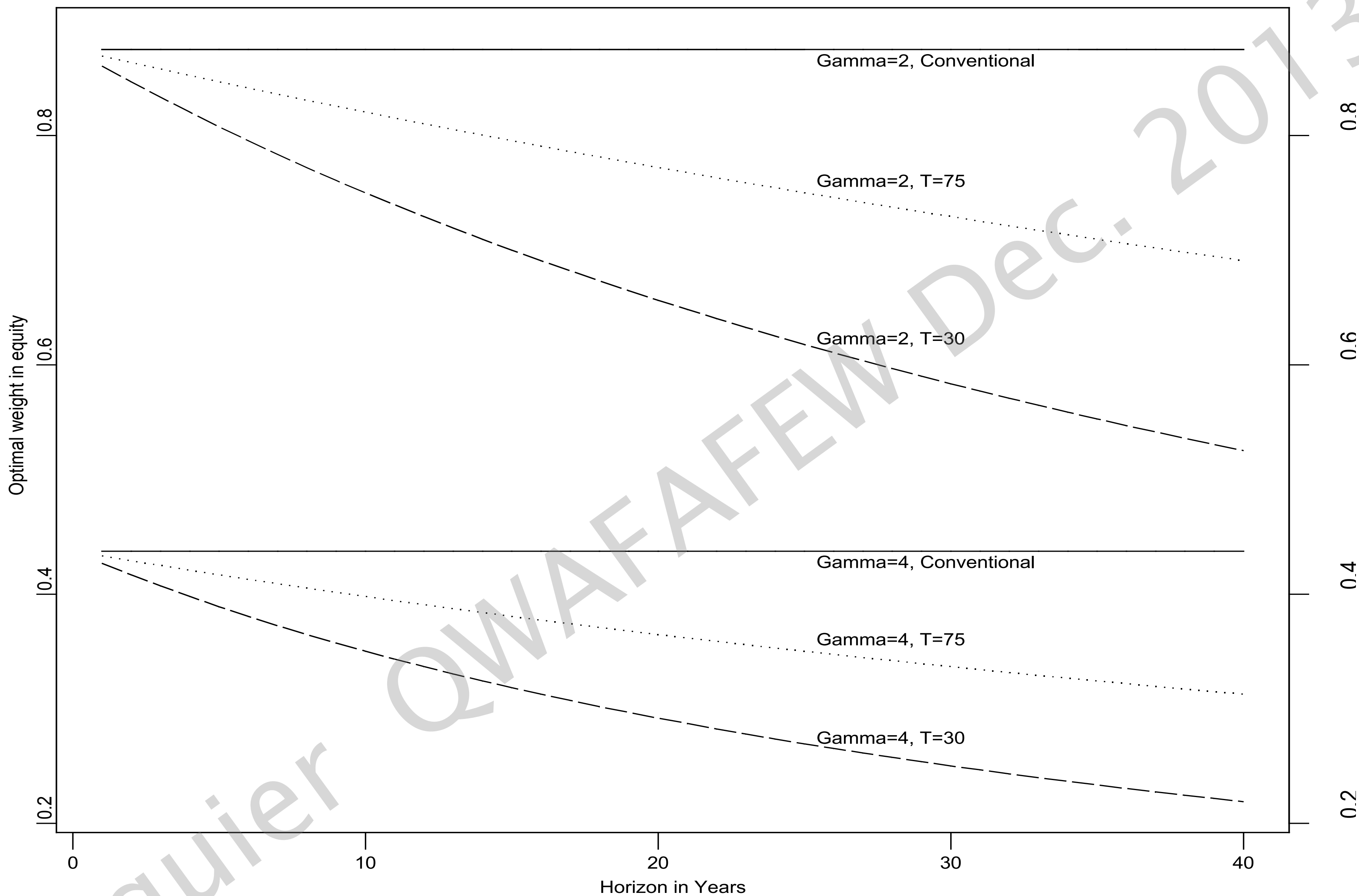


Figure 4: Joint effects of horizon and estimation error on optimal allocation, $\hat{\mu} = 0.1, \sigma = 0.2$

3 Estimator consistent with Optimal Asset Allocation

- Asset allocation with diffuse prior:

$$w^* = \frac{\widehat{\alpha} - r_0}{\sigma^2\gamma\left[1 + \frac{H}{T} - \frac{H}{T\gamma}\right]} \equiv \frac{\alpha^* - r_0}{\sigma^2\gamma},$$

  α^*: The estimate of annualized expected return *corrected for estimation risk*, by a change of measure consistent with the investor's risk aversion.

- Ignore r_0 without loss of generality: Then

$$\alpha^* = \frac{\widehat{\alpha}}{1 + \frac{H}{T}(1 - \frac{1}{\gamma})}$$

  We can rewrite $\alpha*$ as $\widehat{\alpha} - k\frac{\sigma^2}{2}\frac{H}{T}$ to compare with the classical estimators.

Unbiased	$\widehat{\alpha} - \frac{\sigma^2}{2}\frac{H}{T}$
Minimum MSE	$\widehat{\alpha} - 3\frac{\sigma^2}{2}\frac{H}{T}$
Utility	$\widehat{\alpha} - \left[\frac{\widehat{\alpha}}{\frac{\gamma}{\gamma-1} + \frac{H}{T}}\right]\frac{H}{T}$

- Loss function consistent with investor's utility.

Contrast with the “statistical estimator

- Prevents $\alpha^* < 0$ when $\widehat{\mu} > 0$.
  $\alpha^* \rightarrow 0^+$ when $\frac{H}{T} \rightarrow \infty$
- Magnitude of the penalty:

  For conventional μ, σ, γ, and reasonable H/T, the estimation risk penalty is closer to the MMSE than to the unbiased estimator.

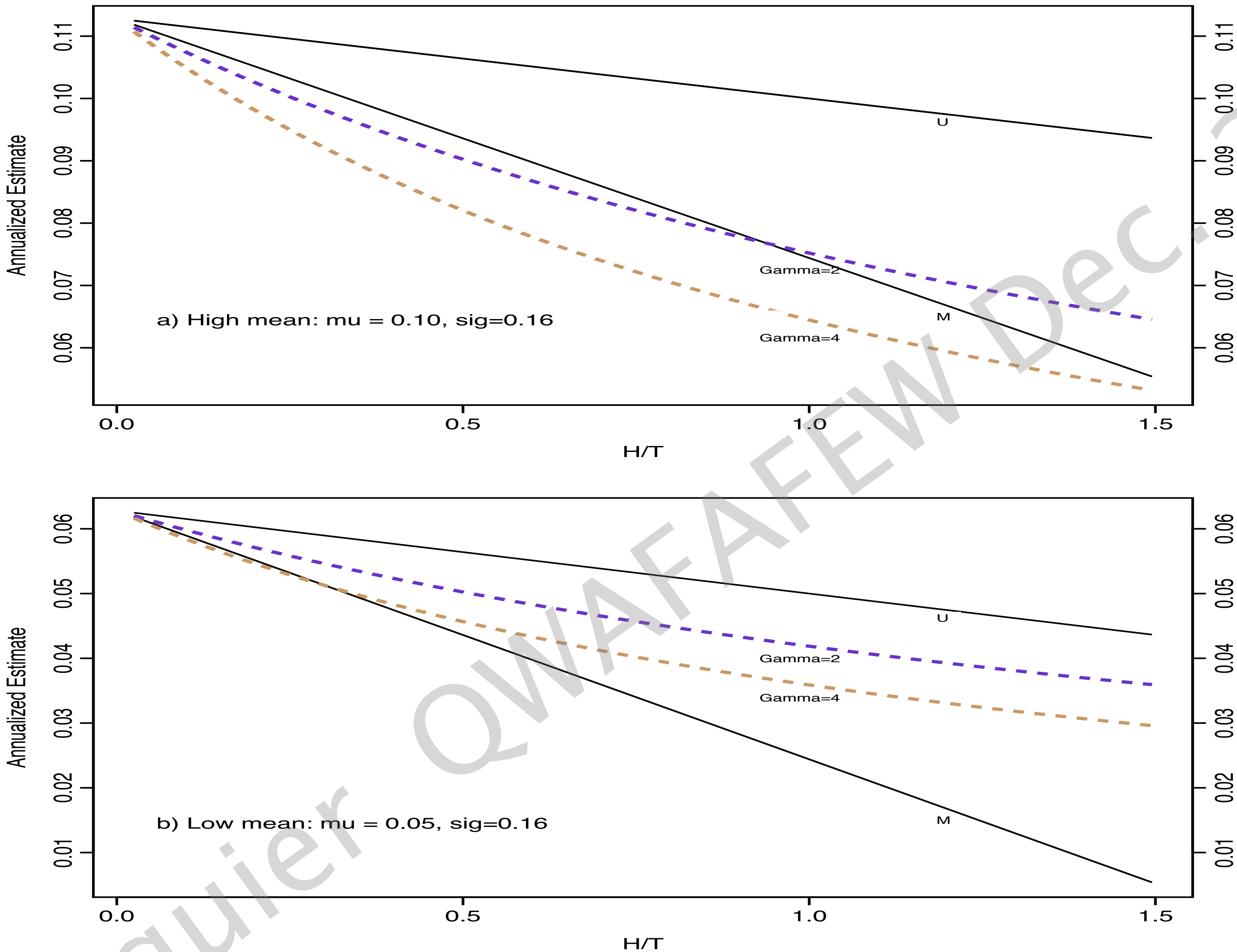


Figure 5: Annualized Optimal estimates consistent with Power Utility vs. H/T. γ=2,4.