

## Holding Period Return-Risk Modeling : Ambiguity in Estimation

Winfried G. Hallerbach

ERIM REPORT SERIES <i>RESEARCH IN MANAGEMENT</i>	
ERIM Report Series reference number	ERS-2003-063-F&A
Publication	2003
Number of pages	18
Email address corresponding author	hallerbach@few.eur.nl
Address	Erasmus Research Institute of Management (ERIM) Rotterdam School of Management / Faculteit Bedrijfskunde Rotterdam School of Economics / Faculteit Economische Wetenschappen Erasmus Universiteit Rotterdam P.O. Box 1738 3000 DR Rotterdam, The Netherlands Phone: + 31 10 408 1182 Fax: + 31 10 408 9640 Email: <a href="mailto:info@erim.eur.nl">info@erim.eur.nl</a> Internet: <a href="http://www.erim.eur.nl">www.erim.eur.nl</a>

Bibliographic data and classifications of all the ERIM reports are also available on the ERIM website:  
[www.erim.eur.nl](http://www.erim.eur.nl)

# ERASMUS RESEARCH INSTITUTE OF MANAGEMENT

## REPORT SERIES *RESEARCH IN MANAGEMENT*

BIBLIOGRAPHIC DATA AND CLASSIFICATIONS		
Abstract	<p>In this paper we explore the theoretical and empirical problems of estimating average (excess) return and risk of US equities over various holding periods and sample periods. Our findings are relevant for performance evaluation, for estimating the historical equity risk premium, and for investment simulation. Using a unique set of US equity data series, comprising monthly prices and dividends based on consistent definitions over the 132 year period 1871-2002, we investigate the complex effect of temporal return aggregation and sample estimation error. Our major finding is that holding period risk and return statistics show an extraordinary sensitivity to the choice of the starting point in calendar time. For example, over the period 1926-2002 there is a difference of almost 140 basis points between the average annual total return starting in January compared to starting in July, and a difference of almost 7 (!) percentage points in estimated annual volatility. This is yet another way in which stock price seasonality manifests itself, but this ambiguity in the underlying estimation process seems completely neglected in the current literature.</p>	
Library of Congress Classification (LCC)	5001-6182	Business
	4001-4280.7	Finance Management, Business Finance, Corporation Finance
	HG 4636	Stock prices
Journal of Economic Literature (JEL)	M	Business Administration and Business Economics
	G 3	Corporate Finance and Governance
	C13	Estimation
	C22	Time-series models
	C89	Data collection and data estimation methodology : other
European Business Schools Library Group (EBSLG)	G14	Information and market efficiency: event studies
	85 A	Business General
	220 A	Financial Management
	220 Q	Stock and bond markets
Gemeenschappelijke Onderwerpsontsluiting (GOO)		
Classification GOO	85.00	Bedrijfskunde, Organiseatiekunde: algemeen
	85.30	Financieel management, financiering
	85.33	Beleggingen
Keywords GOO	Bedrijfskunde / Bedrijfseconomie	
	Financieel management, bedrijfsfinanciering, beslissonde	
	Effectenhandel, koersen, risico's, Seizoensvloeden, tijdreeksen, 1871-2002	
Free keywords	Holding period return, equity risk premium, temporal aggregation, stock price seasonality	

# **Holding Period Return-Risk Modeling :**

## **Ambiguity in Estimation**

Winfried G. Hallerbach<sup>\*)</sup>

presented at the 32<sup>nd</sup> Meeting of the  
EURO Working Group on Financial Modeling,  
London, April 24-26, 2003

this version: September 15, 2003

---

<sup>\*)</sup> Dept of Finance, Erasmus University Rotterdam, POB 1738, NL-3000 DR Rotterdam, The Netherlands. Phone: +31.10.408-1290, facsimile: +31.10.408-9165, [hallerbach@few.eur.nl](mailto:hallerbach@few.eur.nl), <http://www.few.eur.nl/few/people/hallerbach/>. I'd like to thank Michiel de Pooter and Haikun Ning for excellent programming assistance, Jack Wilson for kindly providing me with the data sets, and Nico Van Der Sar for helpful discussions. Of course, the usual disclaimer fully applies.

# **Holding Period Return-Risk Modeling :**

## **Ambiguity in Estimation**

### **Abstract**

In this paper we explore the theoretical and empirical problems of estimating average (excess) return and risk of US equities over various holding periods and sample periods. Our findings are relevant for performance evaluation, for estimating the historical equity risk premium, and for investment simulation.

Using a unique set of US equity data series, comprising monthly prices and dividends based on consistent definitions over the 132 year period 1871-2002, we investigate the complex effect of temporal return aggregation and sample estimation error. Our major finding is that holding period risk and return statistics show an extraordinary sensitivity to the choice of the starting point in calendar time. For example, over the period 1926-2002 there is a difference of almost 140 basis points between the average annual total return starting in January compared to starting in July, and a difference of almost 7 (!) percentage points in estimated annual volatility. This is yet another way in which stock price seasonality manifests itself, but this ambiguity in the underlying estimation process seems completely neglected in the current literature.

*Keywords:* holding period return, equity risk premium, temporal aggregation, stock price seasonality

“So over the period 1926 – 2002, the total return on US equities was on average 11.6% p.a. with a volatility of 18.4%?”

- “Yes.”

“...Or was it 13.2% and 26.7%...?”

- “Yes!”

“?????”

## 1. Introduction and summary

In this paper we explore theoretical and empirical problems of estimating average (excess) return and risk of US equities over various holding periods and sample periods. Our findings are relevant for investment simulation (cf. Freeman [1992]), performance evaluation and for estimating the historical equity risk premium. We use a unique set of revised and corrected US equity data series, comprising monthly prices and dividends based on consistent definitions over the period 1871-2002 (132 years). These data are based on the S&P500 Index and Cowles’s extensions as described in Wilson & Jones [2002]. This long history enables us to avoid overlapping bias when estimating risk and return statistics for longer holding periods.

We investigate the complex effect of temporal return aggregation. This is relevant when estimating risk-return statistics (such as the expected equity return, or its standard deviation, or the equity risk premium) over longer holding periods, especially for one year. There are several methods for estimating the equity risk premium; see for example Welch [2000]. One popular approach is to gauge the average annual equity (excess) return over a long period; cf. Siegel [1992,1994]. Jones, Wilson & Lundstrum [2002] even document the full historical frequency distribution of equity returns over the period 1871-2001 to evaluate future return scenarios for different holding periods. The need for very long data series when estimating the mean – and even the quantiles – of the equity return distribution is apparent; cf. Merton [1980]. Aside from data problems this approach carries the burden of a questionable stationarity assumption; cf. Black [1993].

In addition we here pinpoint a stability problem: given a total sample period of  $T$  years, annual return statistics can be estimated by using January to January returns. This is common practice, cf. Ibbotson & Sinquefeld [1976a,b] and Dimson & Marsh [2001], for example. Alternatively, annual return statistics can be derived from yearly holding period returns starting in each of the other months. We do not expect to see much difference between considering January-January returns or even July-July returns, where the difference in months is greatest. However, our empirical findings show that holding period risk and return statistics show an extraordinary sensitivity to the choice of the starting point in calendar time. For example, over the period 1926-2002 there is a difference of almost 140 basis points between the average annual total return starting in January compared to starting in July, and a difference of almost 7 (!) percentage points in estimated annual volatility. This is yet another way in which stock price seasonality manifests itself, but this ambiguity in the underlying estimation process seems completely neglected in the current literature. The only example we are aware of is Ball, Kothari & Shanken [1995] who observe that the

returns on contrarian strategies are very sensitive to the choice of the starting point in calendar time. Our results suggest that the problem is much more general.

This paper is organized as follows. Section 2 introduces notation and summarizes various return definitions. Section 3 discusses and describes the data set. Section 4 explores the temporal aggregation issue.

## 2. Notation and definitions

### Discretely and continuously compounded returns

We introduce the following notation and definitions. We start from the price index  $PI$  which represents a equity price series adjusted for stock splits and stock dividends. The discretely compounded price return (capital appreciation or “price relative”)  $p_t$  over period  $t$  is defined by:

$$(1) \quad 1 + p_t = \frac{PI_t}{PI_{t-1}}$$

where  $PI_t$  and  $PI_{t-1}$  denote the price index at the end of period  $t$  and  $t-1$ , respectively. The discretely compounded dividend yield  $y_t$  over the period  $t$  is:

$$(2) \quad y_t = \frac{D_t}{PI_{t-1}}$$

where  $D_t$  is the cash dividend paid at the end of period  $t$ . Combining (1) and (2), the discretely compounded total return (or “value relative”) over period  $t$ ,  $r_t$ , is given by:

$$(3) \quad 1 + r_t = 1 + p_t + y_t = (1 + p_t)(1 + d_t)$$

where  $d_t$  denotes the dividend ratio:

$$(4) \quad d_t = \frac{y_t}{1 + p_t} = y_t \frac{PI_{t-1}}{PI_t} = \frac{D_t}{PI_t}$$

The dividend ratio relates the dividend to the stock price at the payment date, not to the previous price. The dividend ratio (and not the dividend yield) is relevant when considering continuous compounded returns.<sup>1</sup>

In addition to equity, we consider a riskfree investment opportunity (Treasury Bills, e.g.). Denoting the discretely compounded riskfree rate over the period  $t$  by  $b_t$ , the excess return on equities is defined as  $r_t - b_t$ . When the equities are representative

---

<sup>1</sup> Note that one plus the dividend yield is the *arithmetic* difference between the value and the price relatives, whereas one plus the dividend ratio is the *geometric* difference between the value and the price relatives.

for the stock market as a whole, the market risk premium is the expected excess return,  $E\{r_t - b_t\}$ .<sup>2</sup>

The continuously compounded return is obtained by taking the natural logarithm of one plus the discretely compounded return. Using (3), the continuously compounded total equity return over the period  $t$  is  $\ln(1 + r_t) = \ln(1 + p_t) + \ln(1 + d_t)$ . Considering the dividend ratio instead of the dividend yield allows us to express the log total return as the sum of the log price return and the log dividend ratio.<sup>3</sup>

### Arithmetic and geometric mean returns

When studying returns over a long horizon, the compounded average growth rate or geometric mean return becomes relevant. The geometric mean of the returns  $\{\tilde{r}_t\}_{t \in T}$  over  $T$  periods,  $G(\tilde{r}_t; T)$ , is defined as:

$$(5) \quad G(\tilde{r}_t; T) = \left[ \prod_{t=1}^T (1 + \tilde{r}_t) \right]^{1/T} - 1$$

When returns are intertemporally independent and identically distributed, then (according to the strong law of large numbers) the geometric mean converges almost surely to the constant  $\exp[E\{\ln(1 + \tilde{r}_t)\}] - 1$  as  $T \rightarrow \infty$ . This implies that the distribution of the geometric mean degenerates and converges to a point distribution.

In practice, this limit is not reached. However, when  $T$  is sufficiently large and when the stationarity and independence assumption is satisfied, the distribution of the logarithm of one plus the geometric mean is approximately normal with mean  $\mu$  and variance  $s^2/T$ , where  $\mu$  and  $s^2$  are the mean and variance of the log returns  $\ln(1 + \tilde{r}_t)$ . Hence, the distribution of the geometric mean is approximately lognormal with mean  $\exp[\mu + \frac{1}{2}s^2/T] - 1$  and variance  $\exp[2\mu + s^2/T] \{ \exp[s^2/T] - 1 \}$ ; see Michaud [1981], e.g. Note that a confidence interval of the log of one plus the geometric mean will be symmetric, whereas confidence intervals of the geometric mean and the end-of-horizon value will be asymmetric. For large  $T$ , this lognormal distribution has properties similar to that of a normal distribution; hence the expected value approaches the median.

The geometric mean is the rate of return that compounds initial value  $V_0$  to  $T$ -period terminal value  $V_T$ :  $V_T = V_0 [1 + G(\tilde{r}_t; T)]^T$ . Given the previous results, it follows that the asymptotic distribution of  $T$ -period terminal value is lognormal. The

---

<sup>2</sup> Since the risk premium is the return on a self-financing portfolio, it does not make sense to distinguish between a nominal and a real risk premium. After all, expected inflation is contained in the equity return as well as in the riskfree rate. Another way to see this is to consider a portfolio of  $x$  in equities and  $1-x$  in riskfree assets. The nominal portfolio return is  $r_{p,t} = x(r_t - b_t) + b_t$ , where expected inflation is contained in the riskfree rate.

<sup>3</sup> This approach is not to be confused with the dividend ratio model developed by Campbell & Shiller [1988a,b]; see also Campbell, Lo & MacKinlay [1997]. Since they want to model log dividend growth, they *approximate* the log of the sum of price and dividend with a weighted average of log price and log dividend.

$\alpha$ -quantiles of the distributions of final value and geometric mean are related by  $V_T^{(\alpha)} = V_0 \left[ 1 + G^{(\alpha)}(\tilde{r}_i; T) \right]^T$ . Since for large  $T$  the expected geometric mean approaches the median value, the expected geometric mean relates to the median terminal value,  $V_T^{(0.5)} = V_0 \left[ 1 + E \{ G(\tilde{r}_i; T) \} \right]^T$ . The arithmetic mean  $m$  of the discretely compounded returns, in contrast, relates to the expected terminal value,  $E \{ V_T \} = V_0 [1 + m]^T$ , from which we recognize the familiar valuation maxim. It also follows that the arithmetic mean is an estimator of the risk premium (and not the geometric mean). Since  $E \{ G(\tilde{r}_i; T) \} < m$  for  $s^2 > 0$ , the median of the final value distribution is lower than the mean, indeed implying a right-skewed distribution.

When the number of observations  $T$  used to estimate the geometric mean is not equal to the horizon  $T'$  over which the geometric mean is compounded in order to compute the expected terminal value, the geometric mean is a biased estimator. Blume [1974] derived an approximately unbiased estimator, which is a weighted average of the arithmetic mean  $m$  and the geometric mean:

$$(6) \quad \hat{E} \left( \frac{V_{T'}}{V_0} \right) \approx \frac{T - T'}{T - 1} (1 + m)^{T'} + \frac{T' - 1}{T - 1} (1 + G(\tilde{r}_i; T))^{T'}$$

Given the mean  $m$  and variance  $s^2$  of the discretely compounded returns, a very accurate approximation to the geometric mean can be obtained through:<sup>4</sup>

$$(7) \quad G(\tilde{r}_i; T) \approx (1 + m) \exp \left[ - \frac{1/2 s^2}{(1 + m)^2} \right] - 1$$

This approximation (actually, all of the derived approximations) clearly reveals the “variance slippage”: the negative relationship between the geometric mean and the variance of returns.

### 3. Data and descriptive statistics

The data set runs from December 31, 1871 through December 31, 2002.<sup>5</sup> We use a unique set of revised and corrected US equity data series, comprising monthly prices and dividends based on consistent definitions over the period 1871 through 2002 (132 years). These data are based on the S&P500 Index and Cowles’s extensions as described in Wilson & Jones [1987, 2002]. All prices are measured ultimo month except for the sub-period 1871:01 through 1885:02, for which only mid-month prices are available. However, important is that prices are not averaged over each month. Compared to other available data sets this is a distinguishing feature; it is well known

<sup>4</sup> See Michaud [1981] and Jean & Helms [1983] for a comparison of various approximations and further references to the literature.

<sup>5</sup> I thank Jack Wilson (College of Management at North Carolina State University, Raleigh NC 27695, [wilson@gw.fis.NCSU.EDU](mailto:wilson@gw.fis.NCSU.EDU)) for generously providing me with the equity and T-Bill data sets.



that the use of within-month averaged prices generates various statistical biases in the return series.<sup>6</sup> From the monthly prices, a price index is constructed. Monthly dividends were estimated from trailing quarterly dividends by Wilson & Jones [2002] and used to construct a cumulative total returns index with monthly reinvestment.

As a proxy for the riskfree rate we use the monthly total return on US Treasury Bills. Since T-Bills were only introduced in 1929, the riskfree rate series consists from 1870:12 through 1912:12 of 75% of the commercial paper yield, and from then on until 1928:12 of the yield on short-term government bonds.

We have divided the total sample period in various sub-periods. Since 1926 is the base year of the S&P Indexes (i.e. the S&P90 and from 1957:03 on the familiar S&P500) we consider 1871-1925 and 1926-2002. The period 1963-2002 is consistent with an evaluation horizon of 40 years. To allow putting recent developments in a broader historical context, we finally set a breakpoint at 1983.

Table 1 presents some descriptive statistics over various sub-periods, obtained from monthly return data. We have annualized the means and medians by simply multiplying monthly figures with 12. In this way, the average price return and average dividend yield sum to the average total return. All return series exhibit excess kurtosis and to some lesser degree skewness. A Jarque-Bera test rejects normality for all series and all (sub-) periods at  $p=0.0000$ , except for the riskfree return over the most recent sub-period 1983-2002 ( $p=0.085$ ). Comparing means and medians we see that the distributions of price returns, total returns and excess returns are skewed to the left, except for the period 1963-1982.<sup>7</sup> The distributions of the dividend yield and the dividend ratio (and to a lesser extent the T-Bill return), in contrast, are skewed to the right. This can be explained by the fact that these return figures are restricted to non-negative values.

Over the total sample period the total return on stocks was on average almost 10% p.a. with an annualized volatility of 16.7% (monthly volatility times  $\sqrt{12}$ ). The period after 1926 shows both a higher average return and a higher standard deviation. However, although the mean returns over 1926-1962 and 1963-2002 are almost the same, the volatility is substantially higher in the first sub-period.

The average excess return is an estimate of the annualized historical *monthly* equity risk premium, since the riskfree rate is measured over the same interval as the stock returns (we will consider the *annual* risk premium in section 4). Over the full 132 years it equals about 6% p.a. Over the most recent 40-year period it is about 5% p.a. where the risk premium of about 7.5% over the most recent 20 years sharply contrasts with the 2.5% over the period 1963-1982. In the latter period the average total return on stocks is 80 basis points below the overall period average whereas at the same time the average riskfree rate reached its historical high.

Comparing their statistics, the dividend yield and the dividend ratio are almost identical. The level of the average dividend yield / ratio has declined steadily over time. This seems consistent with Fama & French [2001] who argue that the propensity

---

<sup>6</sup> See for example Schwert [1990], Wilson, Jones & Lundstrum [2001] and Hallerbach [2003a].

<sup>7</sup> This is confirmed by visual inspection of the empirical frequency distributions. All skewness statistics, however, are positive. This is caused by some extreme observations in the right tails. Hence the positive skewness suggested by the positive third moment is not real but apparent. Indeed, a zero third order moment is a necessary and not a sufficient condition for distributional symmetry and knowledge of the third moment gives almost no clue about the shape of the distribution; see Mood, Graybill & Boes [1974, pp.75-76].

to pay cash dividends has declined over time. However, at the same time the level of the average price return has increased, most markedly over the last 20 years. Since the dividend yield is a function of both dividends and stock prices, dividend yields can also decrease because of increasing prices. Indeed, the level of S&P500 dividends has increased steadily over time, until September 2000 when stock prices started plunging and the dividend level stabilized.

The standard deviation of the dividend yield / ratio is very low, comparable to the volatility of the riskfree rate over time. Although the dividend yield does contribute its share to the total stock return, it does not contribute to the volatility: the standard deviations of price and total stock returns are virtually the same. Even when average total return remains the same, decreasing dividend yield (and hence increasing average price return) implies that a larger portion of the total return is subject to risk. The importance of dividends is further explored in Hallerbach [2003b].

Table 3 displays the annualized arithmetic mean and standard deviation of the price and value relatives. In contrast to Table 1, the means are now compounded to *per annum* figures.<sup>8</sup> In addition, the actual geometric mean return according to eq.(5), its approximation eq.(7) and its 95% confidence interval is provided. The effect of variance slippage is pronounced for the price return and total return, and almost absent for the dividend ratio and the riskfree rate. The geometric mean approximation according to eq.(7) is outstanding. For the price and total returns, the 95% confidence interval is quite wide, even for the overall period of 132 years. One dollar invested in the stock market in January 1871, with dividends reinvested, has grown to  $(1.0879)^{132} = \$ 67,679$  in December 2002; this is the median horizon value. The expected horizon value was a staggering  $(1.1031)^{132} = \$ 419,765$  and the difference with the median value clearly indicates how skewed the distribution of horizon value is. The 95% confidence interval of horizon value is between a modest  $(1.0574)^{132} = \$ 1,578$  and an astounding  $(1.1193)^{132} = \$ 2,902,588$  ... For some periods, the confidence interval of the geometric mean price return extends to negative compound growth rates, but for the total return the confidence interval is strictly positive.

Striking is that the overall period geometric means of the price return and the riskfree rate are almost the same. This implies that the end value obtained by a roll-over strategy of one-month riskfree Bills from 1871:12 on was approximately equal to the cumulative price return obtained in the stock market. Stated otherwise: the equity risk premium was fully generated by the (reinvested) cash dividends. Figure 1 plots the total return Bill index and the stock price index over time. Many empirical studies start their sample in 1926, but there are fundamental differences between the pre and post 1926 periods. Comparing the geometric means in Table 2, we see that in the period 1871-1925 the largest part of the equity return was generated by (reinvested) dividends, whereas in the period 1926-2002 the importance of dividends has decreased and the contribution of the risky price return to the total return was higher.<sup>9</sup> Finally note that in the 1926-2002 period, the 95% confidence intervals of the total equity return and the riskfree rate do not overlap; at this confidence level, equities “dominate” T-Bills.

---

<sup>8</sup> Under the simple annualization used in Table 1, the artifact can arise that the arithmetic mean p.a. is smaller than the geometric mean p.a.

<sup>9</sup> For more details we refer to Hallerbach [2003b].

## 4. Temporal return aggregation

Long data sets are relevant for gauging long term average returns which in turn can be used for estimates of long term expected returns and equity risk premiums. These estimates are interesting in their own right, but they can also be used as inputs for investment simulations *per se* (an early example is Ibbotson & Sinquefeld [1976b]), or in an asset-liability management context. The returns underlying the estimates are not only expressed *per annum* (i.e. annualized), they are also measured on an annual basis. This means that January on January value relatives are computed and the means and standard deviations of these annual returns are estimated.

This raises the issue of temporal return aggregation: do annual returns have the same statistical characteristics as monthly returns, and have January on January value relatives the same characteristics as annual holding period returns starting in each of the other months? Common practice is to estimate annual return statistics from January to January returns; likewise quarterly returns are estimated using January quarters, not February or March quarters. We do not expect to see much difference between considering January-January returns or evaluating February-February returns; after all these annual returns share eleven months. Since the starting point and end point of the index series is given, we wouldn't even expect to see much difference between January-January statistics or July-July statistics (where the difference in months is greatest). However, this argument applies to the geometric mean return, not to the arithmetic mean that is (correctly) used as an estimator of the risk premium or the mean of the annual return distribution.

In order to investigate the effect of temporal return aggregation we considered various annual return specifications. Firstly, we computed 12 different sets of annual returns, each starting in a different month. The first of this set, comprising January-to-January returns, is commonly used in empirical studies.

Secondly, annual return statistics can be derived from higher frequency (in our case: monthly) data under the *iid* assumption. From the monthly arithmetic average return  $m$  the annual average return is obtained by compounding over 12 months:  $(1 + m)^{12}$ . Since the monthly mean is an estimate, we also applied Blume's [1974] adjustment eq.(6). The annual volatility is obtained by scaling the monthly volatility with  $\sqrt{12}$ . This annualization of statistics goes at the cost of making assumptions such as intertemporal independence and stationarity. Christoffersen, Diebold & Schuermann [1998] and Estrada [2000] already showed that simple volatility scaling fails when returns are not *iid*. (They considered holding periods up to one month, we here consider longer holding periods.)

For this reason we also performed bootstrap analyses. In bootstrapping, the sample is substituted for the unknown population, where after repetitively same-sized sub-samples are drawn with replacement from the original sample; cf. Efron [1979]. Bootstrapping fails when the resampling scheme does not match the structure of the actual sampling mechanism. Here this feature is turned into a virtue: sampling monthly returns with replacement provides sub-samples in which annual returns are formed by aggregating independent monthly returns. Hence, the annual return statistics obtained from the bootstrapped samples satisfy the *iid* assumption and serve as a benchmark for the other estimates. We generated 1,000 sub-samples. Within each sample period we used the same set of sub-samples in order to guarantee comparability.

Table 3 presents estimates of annual(-ized) means and standard deviation of total returns. The arithmetic mean return, compounded from monthly to annual basis, corresponds very closely to the mean of the bootstrapped annual returns (in square brackets). Since the *iid* assumption underlies both approaches, this is indeed exactly what we expect. The differences with the actual January-January mean returns are somewhat larger, up to 40 basis points. The Blume correction proves not material for scaling monthly means to an annual basis, but perhaps we could have expected this. Turning to the annual(-ized) standard deviations the picture is different. Each period shows quite large differences between the  $\sqrt{12}$ -scaled monthly volatilities and the annual estimates from bootstrapping. Given that both estimators are based on the *iid* premise, this is surprising. In addition we observe small to substantial differences with the estimates based on actual January annual returns.

These results spurred us to consider estimates derived from actual annual returns, with varying starting months. The results are in Table 4. Below the estimates of the mean and standard deviation are the quantiles derived from the frequency distribution of bootstrapped estimates. For example, for the period 1926-2002, the mean of the actual January returns is 11.79% and the quantile indicates that 49.6% of the bootstrapped samples showed a lower mean. Turning to the July annual returns, the mean is 13.18%, which is surpassed in only 28.2% of the bootstrap samples. There is no clear pattern in the annual means. In the post 1926 periods the means of January returns correspond quite closely to those of the bootstrapped samples. However, the mean estimates for different starting months show variations between the 30% and 65% quantiles of the *iid* distribution from bootstrapping.

Would these results lead one to believe that the *iid* assumption is not too bad, the situation is radically different for the volatility estimates. The discrepancies between the different estimates are very large with quantiles ranging from 6.6% to 99.9%.... Given the relatively limited number of observations, the estimates from the periods 1963-1982 and 1983-2002 will suffer from sample error, but even the longer 1963-2002 and 1926-2002 periods clearly show the large differences that can arise between the point estimates. For the popular 1926-2002 period both the mean and standard deviation are much larger for the July returns than for the January returns. There is a difference of almost 7 percentage points in estimated annual volatility and a difference of almost 140 basis points between the average annual total returns. Of course, we are reasoning within the bounds of sampling error. After all, the standard deviation of the annual returns generates the standard error of the estimate of the mean return. We would therefore expect some variation in the estimates, which could be captured by constructing confidence intervals around the point estimate. However, we do not know the distribution of, say, the July annual returns and the standard error of their mean and their volatility is not easily obtained.<sup>10</sup> But although the reported quantiles relate to the *iid* benchmark, they suggest that the differences between the different point estimates are substantial. In addition, it is commonplace in practice to work with point estimates (and not with confidence intervals) and the results show that it is at least advisable to take account of the discrepancies in point estimates that arise from different starting points in calendar time.

Our results seem to uncover in an indirect way the complex effects of stock price instability and seasonality. We have tried to correct for the January effect by adjusting monthly returns in such a way that the means of the different monthly

---

<sup>10</sup> Confidence intervals for each of the 12 different annual mean returns could be obtained from bootstrapping separately from each of the corresponding annual return sets.

returns are identical (but preserving the overall arithmetic mean return). We then computed annual-ized return statistics from these adjusted data sets. However, this yielded comparable results.<sup>11</sup> Unfortunately, this ambiguity in the underlying estimation process seems unrecognized in the current literature. Uncovering the relationship between the return generating process and temporal aggregation seems to be an interesting area for further research.

---

<sup>11</sup> In addition, Gu [2003] shows that the January effect is declining since 1988. However, in our sub-period 1983-2002 we still observe the same phenomenon.

**Table 1 : Descriptive statistics.**

Mean, median and standard deviation of monthly discretely compounded price returns, dividend returns, dividend ratios, total returns and excess returns on stocks, and of T-Bill returns, expressed in percent per annum. Means / medians and standard deviations are annualized simply by multiplying monthly figures with 12 and  $\sqrt{12}$ , respectively. Normality is rejected (Jarque-Bera) for all periods at the  $p=0.0000$  level, except for the riskfree T-Bill return over the period 1983-2002 ( $p=0.085$ ).

in %	stocks					T-Bill
	price return	dividend yield	dividend ratio	total return	excess return	total return
1871-2002 : mean :	5.31	4.53	4.53	9.85	5.92	3.93
(132 yrs) median :	7.05	3.97	3.96	12.15	7.80	3.69
st.dev. :	16.73	0.77	0.78	16.71	16.72	0.82
1871-1925 : mean :	2.78	5.14	5.14	7.91	4.02	3.89
(55 yrs) median :	3.38	4.88	4.83	9.00	5.12	3.76
st.dev. :	13.03	0.71	0.71	13.02	13.00	0.51
1926-2002 : mean :	7.13	4.10	4.10	11.23	7.27	3.96
(77 yrs) median :	10.68	3.11	3.08	15.03	11.03	3.58
st.dev. :	18.93	0.80	0.81	18.91	18.93	0.99
1926-1962 : mean :	6.46	4.86	4.87	11.32	9.84	1.48
(37 yrs) median :	11.18	4.34	4.37	16.95	15.70	1.04
st.dev. :	22.38	0.88	0.90	22.35	22.35	0.43
1963-2002 : mean :	7.75	3.40	3.39	11.15	4.89	6.25
(40 yrs) median :	9.97	2.64	2.65	12.80	7.35	5.60
st.dev. :	15.06	0.64	0.64	15.07	15.08	0.89
1963-1982 : mean :	5.06	3.99	3.99	9.05	2.36	6.69
(20 yrs) median :	3.58	2.94	2.94	6.85	2.67	5.60
st.dev. :	14.52	0.77	0.77	14.50	14.53	1.08
1983-2002 : mean :	10.43	2.81	2.78	13.24	7.42	5.82
(20 yrs) median :	13.08	2.52	2.52	16.21	11.35	5.60
st.dev. :	15.58	0.43	0.42	15.63	15.60	0.63

**Table 2 : Arithmetic and geometric mean returns .**

Annualized arithmetic mean  $m$  and standard deviation  $s$  of monthly discretely compounded returns. One plus the arithmetic mean is annualized by exponentiating to the power 12; the standard deviation is annualized by multiplying with  $\sqrt{12}$ .  $G(\cdot;T)$  is the actual geometric mean according to eq.(5).  $G(\cdot;T) \sim$  denotes the approximation by eq.(7) on the basis of monthly data. Both the geometric mean and the approximation are annualized by exponentiating to the power 12. “95% confid.” indicates the 95% confidence interval of the geometric mean.

%		price return		dividend ratio		total return		T-Bill return	
1871-2002 : (132 yrs)	$m$	5.45		4.63		10.31		4.00	
	$s$	16.73		0.78		16.71		0.82	
	$G(\cdot;T)$	3.98		4.62		8.79		4.00	
	$G(\cdot;T) \sim$	3.99		4.62		8.80		4.00	
	95% confid.	1.05	7.00	4.48	4.76	5.74	11.93	3.86	4.15
1871-1925 : (55 yrs)	$m$	2.81		5.26		8.21		3.96	
	$s$	13.03		0.71		13.02		0.51	
	$G(\cdot;T)$	1.94		5.26		7.30		3.96	
	$G(\cdot;T) \sim$	1.95		5.26		7.31		3.96	
	95% confid.	-1.53	5.53	5.06	5.45	3.66	11.06	3.82	4.10
1926-2002 : (77 yrs)	$m$	7.36		4.18		11.83		4.03	
	$s$	18.93		0.81		18.91		0.99	
	$G(\cdot;T)$	5.47		4.17		9.87		4.03	
	$G(\cdot;T) \sim$	5.48		4.17		9.88		4.03	
	95% confid.	1.09	10.03	3.99	4.36	5.34	14.60	3.80	4.26
1926-1962 : (37 yrs)	$m$	6.65		4.98		11.93		1.49	
	$s$	22.38		0.90		22.35		0.43	
	$G(\cdot;T)$	4.04		4.98		9.21		1.49	
	$G(\cdot;T) \sim$	4.04		4.97		9.22		1.49	
	95% confid.	-3.20	11.82	4.67	5.28	1.66	17.33	1.35	1.63
1963-2002 : (40 yrs)	$m$	8.03		3.44		11.74		6.44	
	$s$	15.06		0.64		15.07		0.89	
	$G(\cdot;T)$	6.81		3.44		10.48		6.43	
	$G(\cdot;T) \sim$	6.82		3.44		10.50		6.43	
	95% confid.	1.91	11.94	3.23	3.64	5.43	15.78	6.14	6.72
1963-1982 : (20 yrs)	$m$	5.18		4.06		9.44		6.90	
	$s$	14.52		0.77		14.50		1.08	
	$G(\cdot;T)$	4.09		4.06		8.32		6.89	
	$G(\cdot;T) \sim$	4.08		4.06		8.31		6.89	
	95% confid.	-2.30	10.90	3.71	4.41	1.70	15.37	6.39	7.39
1983-2002 : (20 yrs)	$m$	10.95		2.82		14.08		5.98	
	$s$	15.58		0.42		15.63		0.63	
	$G(\cdot;T)$	9.60		2.82		12.69		5.98	
	$G(\cdot;T) \sim$	9.63		2.82		12.72		5.98	
	95% confid.	2.28	17.45	2.63	3.01	5.15	20.77	5.68	6.27

**Table 3 : Annual and annualized means and standard deviations of total return.**

Actual annual returns are measured from January to January. Between the square brackets are the mean and standard deviation averaged over 1,000 bootstrap samples. One plus the monthly arithmetic / geometric mean returns are compounded annually by raising to the power 12. The volatility is annualized by scaling with  $\sqrt{12}$ .

in %	m e a n				standard deviation	
	actual annually Jan.-Jan.	annually compounded monthly means		Blume estimator eq.(9)	actual annually Jan.-Jan.	annualized (mo · $\sqrt{12}$ )
		arithmetic	geometric			
<b>1871-2002 :</b> (1584 mo)	10.45 [10.34]	10.31	8.79	10.30	18.67 [18.40]	16.71
<b>1871-1925 :</b> (660 mo)	8.56 [8.16]	8.21	7.30	8.19	16.65 [13.98]	13.02
<b>1926-2002 :</b> (924 mo)	11.79 [11.84]	11.83	9.87	11.80	19.99 [21.06]	18.91
<b>1926-1962 :</b> (444 mo)	11.79 [12.19]	11.93	9.21	11.86	23.20 [24.72]	22.35
<b>1963-2002 :</b> (480 mo)	11.80 [11.70]	11.74	10.48	11.71	16.80 [16.72]	15.07
<b>1963-1982 :</b> (240 mo)	9.63 [9.64]	9.44	8.32	9.39	16.85 [15.48]	14.50
<b>1983-2002 :</b> (240 mo)	13.96 [13.92]	14.08	12.69	14.01	16.89 [17.32]	15.63

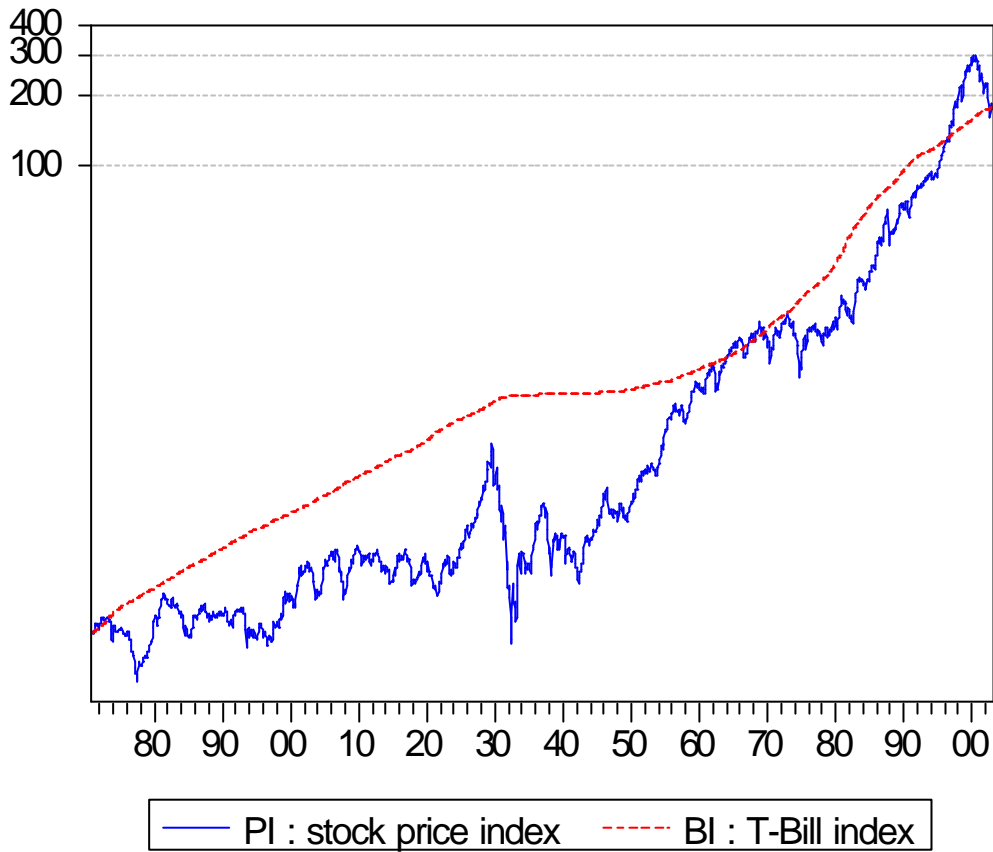


**Table 4 : Annual means and standard deviations of total return.** Actual annual returns are measured from the month in the column heading. The quantiles are derived from 1,000 bootstrapped samples (assuming independence of monthly returns; same sub-samples for each sub-period).

	in %	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
<b>1871-2002 :</b>	mean <i>m</i>	10.45	10.62	10.74	11.02	10.69	11.03	11.19	10.68	10.44	10.43	10.36	10.46
	quantile	52.80	57.00	59.90	66.40	58.70	66.50	70.40	58.40	52.50	52.30	50.50	52.90
	st.dev. <i>s</i>	18.67	18.31	19.68	20.82	18.89	21.39	23.34	19.95	18.76	18.95	18.04	18.58
	quantile	61.30	49.20	82.60	94.90	66.60	97.90	99.90	85.90	63.70	68.50	41.90	58.00
	median	10.23	10.48	10.62	10.13	9.53	9.52	10.04	8.96	10.02	9.88	12.29	11.91
<b>1871-1925 :</b>	mean <i>m</i>	8.56	8.04	8.13	7.97	7.81	7.99	8.24	8.31	8.16	8.26	8.65	8.48
	quantile	59.50	48.30	49.90	46.20	43.00	46.90	51.90	53.20	50.20	52.30	61.60	57.40
	st.dev. <i>s</i>	16.65	15.32	16.22	16.31	15.45	16.23	17.64	17.34	16.47	16.73	17.74	16.97
	quantile	97.20	83.70	94.30	95.80	86.00	94.60	99.70	99.30	96.30	97.40	99.70	98.70
	median	6.50	7.01	4.29	6.07	4.91	4.81	5.04	6.88	6.53	6.41	10.65	8.53
<b>1926-2002 :</b>	mean <i>m</i>	11.79	12.26	12.50	13.08	12.61	13.14	13.18	12.24	11.93	11.88	11.61	11.86
	quantile	49.60	57.90	61.80	70.10	63.40	71.10	71.80	57.70	52.10	51.00	46.70	50.70
	st.dev. <i>s</i>	19.99	20.13	21.82	23.48	20.95	24.41	26.71	21.67	20.27	20.46	18.38	19.75
	quantile	33.80	36.40	66.90	86.30	52.90	92.80	98.40	64.00	38.90	42.50	10.60	28.90
	median	13.80	14.63	12.69	14.08	10.85	13.35	13.39	12.84	12.15	11.63	13.15	15.46
<b>1926-1962 :</b>	mean <i>m</i>	11.79	12.62	13.36	14.19	13.30	14.06	13.86	12.20	11.53	11.28	11.11	11.75
	quantile	48.00	56.00	62.40	69.40	61.70	68.50	66.40	51.80	45.80	43.10	41.10	47.70
	st.dev. <i>s</i>	23.20	25.20	28.40	29.97	26.17	31.61	34.45	25.31	22.71	21.72	20.82	23.01
	quantile	36.40	57.60	83.90	91.60	67.60	95.60	98.60	58.90	31.60	23.00	15.60	34.10
	median	11.95	13.54	13.79	16.08	19.13	15.69	10.64	13.76	9.34	9.37	10.27	10.42
<b>1963-2002 :</b>	mean <i>m</i>	11.80	12.25	12.14	12.42	12.02	12.03	12.09	12.00	11.93	11.93	11.47	11.71
	quantile	50.80	57.90	55.70	59.90	53.70	54.10	54.90	53.30	52.30	52.30	46.70	49.80
	st.dev. <i>s</i>	16.80	14.50	13.71	15.93	15.32	15.85	17.45	18.26	18.17	19.51	15.90	16.74
	quantile	51.00	13.90	6.60	35.20	24.90	33.00	64.70	78.50	77.00	90.90	34.30	50.40
	median	15.39	14.87	12.16	14.39	10.34	11.78	13.40	12.73	12.48	12.62	14.92	16.47
<b>1963-1982 :</b>	mean <i>m</i>	9.63	8.24	7.85	7.75	7.59	7.43	7.54	7.37	7.84	8.57	8.51	9.00
	quantile	52.10	36.00	31.00	30.10	28.60	26.80	27.90	26.20	31.00	40.20	38.70	45.80
	st.dev. <i>s</i>	16.85	14.44	12.70	14.13	12.90	13.65	14.58	13.77	14.36	18.02	15.39	17.37
	quantile	70.50	37.90	15.70	31.70	17.40	25.70	39.40	27.00	36.70	82.60	51.10	75.80
	median	13.38	11.85	7.82	6.82	6.90	8.09	10.73	7.35	9.85	10.19	13.72	11.91
<b>1983-2002 :</b>	mean <i>m</i>	13.96	15.44	15.05	15.41	14.50	14.50	14.06	14.13	14.33	13.58	13.55	13.69
	quantile	48.30	64.70	60.80	64.50	54.70	54.70	49.20	49.80	52.60	43.90	43.80	45.20
	st.dev. <i>s</i>	16.89	13.93	13.01	15.72	15.04	14.97	16.46	19.05	20.26	20.15	16.37	16.35
	quantile	46.30	11.50	7.10	31.90	22.50	22.00	40.80	72.90	82.60	82.10	39.60	39.20
	median	17.59	15.23	14.65	17.92	17.43	16.55	16.45	12.77	16.29	13.25	14.92	20.30

**Figure 1: Stock price index and T-Bill index.**

Stock price index *PI* is the S&P500 price index, cash dividends excluded. The T-Bill index *BI* is the total return index from a roll-over strategy in one-month T-Bills. The series start ultimo 1871:12 at 1.00 (logscale) and ends ultimo 2002:12.



## References

- Ball, R., S.P. Kothari & J. Shanken, 1995, "Problems in Measuring Portfolio Performance, An Application to Contrarian Investment Strategies", *Journal of Financial Economics* 38, pp.79-107
- Black, F., 1993, "Estimating Expected Return", *Financial Analysts Journal* Sept-Oct, pp.36-38
- Blume, M.E., 1974, "Unbiased Estimators of Long-Run Expected Rates of Return", *Journal of the American Statistical Association* 69/347, Sept, pp.634-638
- Campbell, J.Y. & R.J. Shiller, 1988a, "The Dividend-Price Ratio and Expectations of Future Dividends and Discount Factors", *The Review of Financial Studies* 1/3, pp.195-228
- Campbell, J.Y. & R.J. Shiller, 1988b, "Stock Prices, Earnings, and Expected Dividends", *The Journal of Finance* 43/3, pp.661-676
- Campbell, J.Y., A.W. Lo & A.C. MacKinlay, 1997, "*The Econometrics of Financial Markets*", Princeton University Press, Princeton
- Christoffersen, P.F., F.X. Diebold & T. Schuermann, 1998, "Horizon Problems and Extreme Events in Financial Risk Management", *FRBNY Economic Policy Review* Oct., pp.109-118
- Dimson, E. & P. Marsh, 2001, "U.K. Financial Market Returns, 1955-2000", *The Journal of Business* 74/1, pp.1-31
- Efron, B., 1979, "Bootstrap Methods: Another Look at the Jackknife", *The Annals of Statistics* 7, pp.1-26
- Estrada, J., 2000, "The Temporal Dimension of Risk", *The Quarterly Review of Economics and Finance* 40, pp.189-204
- Fama, E.F. & K.R. French, 2001, "Disappearing Dividends: Changing Firm Characteristics or Lower Propensity to Pay?", *Journal of Financial Economics* 60, pp.3-43
- Freeman, J.D., 1992, "Behind the Smoke and Mirrors: Gauging the Integrity of Investment Simulations", *Financial Analysts Journal* Nov-Dec, pp.26-31
- Gu, A.Y., 2003, "The Declining January Effect: Evidences From the U.S. Equity Markets", *The Quarterly Review of Economics and Finance* 43, pp.395-404
- Hallerbach, W.G., 2003a, "Cross- and Auto-Correlation Effects Arising From Averaging: the Case of US Interest Rates and Equity Duration", *Applied Financial Economics* 13/4, pp.287-294
- Hallerbach, W.G., 2003b, "Holding Period Return-Risk Modeling: The Importance of Dividends", ERIM Report ERS-2003-064-F&A, Erasmus University Rotterdam
- Ibbotson, R.G. & R.A. Sinquefeld, 1976a, "Stocks, Bonds, Bills, and Inflation: Year-by-Year Historical Returns (1926-1974)", *The Journal of Business* 49/1, pp.11-47
- Ibbotson, R.G. & R.A. Sinquefeld, 1976b, "Stocks, Bonds, Bills, and Inflation: Simulations of the Future (1976-2000)", *The Journal of Business* 49/3, pp.313-338
- Jean, W.H. & B.P. Helms, 1983, "Geometric Mean Approximations", *Journal of Financial and Quantitative Analysis* 18/3, Sept., pp.287-293
- Jones, C.P., J.W. Wilson & L.L. Lundstrum, 2002, "Estimating Stock Returns", *The Journal of Portfolio Management* Fall, pp.40-50
- Merton, R., 1980, "On Estimating the Expected Return on the Market: An Exploratory Investigation", *Journal of Financial Economics* 8, pp.323-361

- Michaud, R.O., 1981, "Risk Policy and Long-Term Investment", *Journal of Financial and Quantitative Analysis* 16/2, June, pp.147-167
- Mood, A.M., F.A. Graybill & D.C. Boes, 1974, "Introduction to the Theory of Statistics", McGraw-Hill, Auckland
- Schwert, G.W., 1990, "Indexes of U.S. Stock prices from 1802 to 1987", *Journal of Business* 63/3, pp.399-426
- Siegel, J.J., 1992, "The Equity Premium: Stock and Bond Returns Since 1802", *Financial Analysts Journal* Jan-Feb, pp.28-46
- Siegel, J.J., 1994, "Stocks for the Long Run", Irwin, Burr Ridge Ill.
- Welch, I., 2000, "Views of Financial Economists on the Equity Premium and on Professional Controversies", *Journal of Business* 73/4, pp.501-537
- Wilson, J.W. & C.P. Jones, 1987, "A Comparison of Annual Common Stock Returns: 1871-1925 with 1926-85", *Journal of Business* 60/2, pp.239-258
- Wilson, J.W. & C.P. Jones, 2002, "An Analysis of the S&P 500 Index and Cowles's Extensions: Price Indexes and Stock Returns, 1870-1999", *Journal of Business* 75/3, pp.505-533
- Wilson, J.W., C.P. Jones & L.L. Lundstrum, 2001, "Stochastic Properties of Time-Averaged Financial Data: Explanation and Empirical Demonstration Using Monthly Stock Prices", *The Financial Review* 38, pp.175-190

## Publications in the Report Series Research\* in Management

ERIM Research Program: "Finance and Accounting"

2003

*COMMENT, Risk Aversion and Skewness Preference*

Thierry Post and Pim van Vliet

ERS-2003-009-F&A

<http://hdl.handle.net/1765/319>

*International Portfolio Choice: A Spanning Approach*

Ben Tims, Ronald Mahieu

ERS-2003-011-F&A

<http://hdl.handle.net/1765/276>

*Portfolio Return Characteristics Of Different Industries*

Igor Pouchkarev, Jaap Spronk, Pim van Vliet

ERS-2003-014-F&A

<http://hdl.handle.net/1765/272>

*Asset prices and omitted moments*

*A stochastic dominance analysis of market efficiency*

Thierry Post

ERS-2003-017-F&A

<http://hdl.handle.net/1765/430>

*A Multidimensional Framework for Financial-Economic Decisions*

Winfried Hallerbach & Jaap Spronk

ERS-2003-021-F&A

<http://hdl.handle.net/1765/321>

*A Range-Based Multivariate Model for Exchange Rate Volatility*

Ben Tims, Ronald Mahieu

ERS-2003-022-F&A

<http://hdl.handle.net/1765/282>

*Macro factors and the Term Structure of Interest Rates*

Hans Dewachter and Marco Lyrio

ERS-2003-037-F&A

<http://hdl.handle.net/1765/324>

*The effects of decision flexibility in the hierarchical investment decision process*

Winfried Hallerbach, Haikun Ning, Jaap Spronk

ERS-2003-047-F&A

<http://hdl.handle.net/1765/426>

---

\* A complete overview of the ERIM Report Series Research in Management:  
<http://www.erim.eur.nl>

ERIM Research Programs:  
LIS Business Processes, Logistics and Information Systems  
ORG Organizing for Performance  
MKT Marketing  
F&A Finance and Accounting  
STR Strategy and Entrepreneurship

*Takeover defenses and IPO firm value in the Netherlands*  
Peter Roosenboom, Tjalling van der Goot  
ERS-2003-049-F&A  
<http://hdl.handle.net/1765/433>

*The Cost of Technical Trading Rules in the Forex Market: A Utility-based Evaluation*  
Hans Dewachter and Marco Lyrio  
ERS-2003-052-F&A  
<http://hdl.handle.net/1765/435>

*Stress testing with Student's  $t$  dependence*  
Erik Kole, Kees Koedijk and Marno Verbeek  
ERS-2003-056-F&A  
<http://hdl.handle.net/1765/923>

*Fat Tails in Power Prices*  
Ronald Huisman and Christian Hurman  
ERS-2003-059-F&A  
<http://hdl.handle.net/1765/924>