EQUITY RISK PREMIUM: EXPECTATIONS GREAT AND SMALL

Richard A. Derrig* and Elisha D. Orr†

ABSTRACT

The equity risk premium (ERP) is an essential building block of the market value of risk. In theory, the collective action of all investors results in an equilibrium expectation for the return on the market portfolio excess of the risk-free return, the ERP. The ability of the valuation actuary to choose a sensible value for the ERP, whether as a required input to capital asset pricing model valuation, or any of its descendants, is as important as choosing risk-free rates and risk relatives (betas) to the ERP for the asset at hand.

The historical realized ERP for the stock market appears to be at odds with pricing theory parameters for risk aversion. Since 1985, there has been a constant stream of research, each of which reviews theories of estimating market returns, examines historical data periods, or both. Those ERP value estimates vary *widely* from about −1% to about 9%, based on a geometric or arithmetic averaging, short or long horizons, short- or long-run expectations, unconditional or conditional distributions, domestic or international data, data periods, and real or nominal returns.

This paper examines the principal strains of the recent research on the ERP and catalogues the empirical values of the ERP implied by that research. In addition, the paper supplies several time series analyses of the standard Ibbotson Associates 1926–2002 ERP data using short Treasuries for the risk-free rate. Recommendations for ERP values to use in common actuarial valuation problems also are offered.

"What I actually think is that our prey, called the equity risk premium, is extremely elusive."
—*Stephen A. Ross (2002, p. 22)*

1. INTRODUCTION

The equity risk premium (ERP) is an essential building block of the market value of risk. In theory, the collective action of all investors results in an equilibrium expectation for the return on the market portfolio excess of the risk-free return, the ERP. The ability of the valuation actuary to choose a sensible value for the ERP—whether as a required input to capital asset pricing model (CAPM) valuation or any of its descendants[1]—is as important as choosing risk-free rates and risk relatives (betas) to the ERP for the asset at hand. Risky discount rates, asset allocation models, and project costs of capital are common actuarial uses of ERP as a benchmark rate.

The ERP should be of particular interest to actuaries. For pensions and annuities backed by bonds and stocks, the actuary needs to have an understanding of the ERP and its variability compared to fixed-horizon bonds. Variable products, including guaranteed minimum death benefits, require accurate projections of returns to ensure adequate future assets. With the latest research producing a relatively low ERP, the rationale for including equities in insurers' asset holdings is being tested.

In describing individual investment account guarantees, LaChance and Mitchell (2003) point out an underlying assumption of pension asset investing that, based only on the historical record, future equity returns will continue to outperform bonds; they clarify that those higher ex-

* Richard A. Derrig is Senior Vice President, Automobile Insurers Bureau of Massachusetts, 101 Arch Street, Boston, MA 02110, e-mail: richard@aib.org.

† Elisha D. Orr is a Research Assistant, Automobile Insurers Bureau of Massachusetts, 101 Arch Street, Boston, MA 02110, e-mail: eorr@aib.org.

[1] The multifactor arbitrage pricing theory of Ross (1976), the three-factor model of Fama and French (1992) and the recent Mamaysky (2002) five-factor model for stocks and bonds are all examples of enhanced CAPM models.

pected equity returns come with the additional higher risk of equity returns. Ralfe et al. (2003) support the risky equity view and discuss their pension experience with an all-bond portfolio. Recent projections in some literature of a zero or negative ERP challenge the assumptions underlying these views.

By reviewing some of the most recent and relevant work on the issue of the ERP, actuaries will have a better understanding of how these values were estimated, critical assumptions that allowed for such a low ERP, and the time period for the projection (see Appendix B). Actuaries can then make informed decisions for expected investment results going forward.

In 1985, Mehra and Prescott published their work on the *equity risk premium puzzle*: the fact that the historical realized ERP for the stock market from 1889–1978 appeared to be at odds with and, relative to Treasury bills, far in excess of asset pricing theory values based on investors with reasonable risk aversion parameters. Since then, there has been a constant stream of research, each of which reviews theories of estimating market returns, examines historical data periods, or both (for example, see Cochrane 1997, Cornell 1999, or Equity Risk Premium Forum 2002). Those ERP value estimates vary *widely,* from about −1% to about 9%, based on geometric or arithmetic averaging, short or long horizons, short- or long-run means, unconditional or conditional expectations, using domestic or international data, differing data periods, and real or nominal returns. Brealey and Myers (2000), in the sixth edition of their standard corporate finance textbook, believe a range of 6–8.5% for the U.S. ERP is reasonable for practical project valuation. Is that a fair estimate?

Current research on the ERP is plentiful. This paper covers a selection of mainstream articles and books that describe different approaches to estimating the *ex ante* ERP. We select examples of the research that cover the most important approaches to the ERP. We begin by describing the methodology of using historical returns to predict future estimates. We identify the many varieties of ERPs in order to alert the reader to the fact that numerical estimates of the ERP that appear different may instead be about the same under a common definition. We examine the well-known Ibbotson Associates 1926–2002 data series for stationarity, that is, time invariance of the mean ERP. We show by several statistical tests that stationarity cannot be rejected and the best estimate going forward, *ceteris paribus,* is the realized mean. This paper will examine the principal strains of the recent research on the ERP and catalogue the empirical values of the ERP implied by that research (see Appendix B).

We first discuss how the Social Security Administration derives estimates of the ERP. Then, we survey the puzzle research, that is, the literature written in response to the equity premium puzzle suggested by Mehra and Prescott (1985). We cover five major approaches from the literature. Next, we report from two surveys of "experts" on the ERP. Finally, after describing the main strains of research, we explore some of the implications for practicing actuaries.

We do not discuss the important companion problem of estimating the risk relationship of an individual company, line of insurance, or project with the overall market. Within a CAPM or Fama-French framework, the problem is estimating a market beta.[2] Actuaries should be aware, however, that simple 60-month regression betas are biased low where size or nonsynchronous trading is a substantial factor (Kaplan and Peterson 1998, Pratt 1998, p. 86). Adjustments are made to historical betas in order to remove the bias and derive more accurate estimates. Elton and Gruber (1995, p. 148) explain that by testing the relationship of beta estimates over time, empirical studies have shown that an adjustment toward the mean should be made to project future betas.

2. THE EQUITY RISK PREMIUM

Based on the definition in Brealey and Myers (2000), the ERP is the "expected additional return for making a risky investment rather than a safe one" (p. 1071). In other words, the ERP is the difference between the market return and a risk-free return. Market returns include both dividends and capital gains. Because both the historical ERP and the prospective ERP have been

[2] According to CAPM, investors are compensated only for nondiversifiable, or market, risk. The market beta becomes the measurement of the extent to which returns on an individual security co-vary with the market. The market beta times the ERP represents the nondiversifiable expected return from an individual security.

referred to simply as the ERP, the terms *ex post* and *ex ante* are used to differentiate between them but are often omitted. Table 1 shows the historical annual average returns from 1926 to 2002 for large company equities (S&P 500), Treasury bills and bonds, and their arithmetic differences using data from Ibbotson Associates (2003a,b); the entire series is shown in Appendix A.

In 1985, Mehra and Prescott introduced the idea of the ERP puzzle. The puzzling result is that the historical realized ERP for the stock market using 1889–1978 data appeared to be at odds with and, relative to Treasury bills, far in excess of asset pricing theory values based on normal parametrizations of risk aversion. When using standard frictionless return models and historical growth rates in consumption, the real risk-free rate, and the ERP, the resulting relative risk aversion parameter appears too high. By choosing a maximum relative risk aversion parameter to be 10 and using the growth in consumption, Mehra and Prescott's model produces an ERP much lower than the historical premium.[3]

Their result inspired a stream of finance literature that attempts to solve the puzzle. Two different research threads have emerged. One thread, including behavioral finance, attempts to explain the historical returns with new models and different assumptions about investors (see, e.g., Benartzi and Thaler 1995 and Mehra 2002). A second thread is from a group that provides estimates of the ERP that are derived from historical data and/or standard economic models. Some in this latter group argue that historical returns may have been higher than those that should be required in the future. In a curiously asymmetric way, there are no serious studies yet concluding that the historical results are too low to serve as *ex ante* estimates.

Although both groups have made substantial and provocative contributions, the behavioral models do not give any *ex ante* ERP estimates other than explaining and supporting the historical returns. We presume, until results show otherwise, that the behaviorists support the historical average as the *ex ante* unconditional long-run expectation. Therefore, we focus on the latter to catalogue ERP estimates other than those based on the historical approach,[4] but we will discuss both as important strains for puzzle research.

Table 1
U.S. Equity Risk Premia 1926–2002

Annual Equity Returns and Premia versus Treasury Bills, Intermediate, and Long-Term Bonds			
Horizon	Equity Returns	Risk-Free Return	ERP
Short	12.20%	3.83%	8.37%
Intermediate	12.20	4.81	7.40
Long	12.20	5.23	6.97

Source: Authors' calculations using Ibbotson Associates (2003a, p. 38–39, 177, 238–39, 246–47).

3. ERP TYPES

Many different types of ERP estimates can be given, even though they are labeled by the same general term. These estimates vary widely; currently the estimates range from about 9% to a small negative. When ERP estimates are given, one should determine the type before comparing to other estimates. Here are seven important types to look for when given an ERP estimate:

- Geometric versus arithmetic averaging.
- Short versus long investment horizon.
- Short- versus long-run expectation.
- Unconditional versus conditional on some related variable.
- Domestic United States versus international market data.
- Data sources and periods.
- Real versus nominal returns.

The average market return and ERP can be stated as a geometric or arithmetic mean return. An *arithmetic mean return* is a simple average of a series of returns. The *geometric mean return* is the compound rate of return; it is a measure of the actual average performance of a portfolio over a given time period. Arithmetic returns are the same or higher than geometric returns, so it is not appropriate to make a direct comparison between an arithmetic estimate and a geometric estimate. However, those two returns can be transformed one to the other. For example, arithmetic returns

[3] Campbell, Lo, and MacKinlay (1997, pp. 307–308) performed a similar analysis and found a risk-aversion coefficient of 19, larger than the reasonable level suggested in Mehra and Prescott's paper.

[4] See Appendix C.

can be approximated from geometric returns by the formula

$$AR = GR + \frac{\sigma^2}{2}, \sigma^2 \text{ the variance of the (arithmetic) return process}$$

(see Welch 2000, Dimson et al. 2002, and Ibbotson and Chen 2003). Arithmetic averages of periodic returns are to be preferred when estimating next period returns since they, not geometric averages, reproduce the proper probabilities and means of expected returns.[5] ERPs can be generated by arithmetic differences (Equity − Risk Free) or by geometric differences ([(1 + Equity)/(1 + Risk Free)] − 1). Usually, the arithmetic and geometric differences produce similar estimates.[6]

A second important difference in ERP estimate types is the horizon. The *horizon* indicates the total investment or planning period under consideration. For estimation purposes, the horizon relates to the term or maturity of the risk-free instrument that is used to determine the ERP. Ibbotson Associates (2003a, p. 177) provides definitions for three different horizons. The *short-horizon* expected ERP is defined as "the large company stock total returns minus U.S. Treasury bill *total* returns." Note, the income return and total return are the same for U.S. Treasury bills. The *intermediate-horizon* expected ERP is "the large company stock total returns minus intermediate-term government bond *income* returns." Finally, the *long-horizon* expected ERP is "the large company stock total returns minus long-term government bond *income* returns." (Table 1 displays the short-horizon ERP.)

For the Ibbotson data, Treasury bills have a maturity of approximately one month, intermediate-term government bonds have a maturity around five years, and long-term government bonds have a maturity of about 20 years. Although the Ibbotson definitions may not apply to other research, we will classify ERP estimates based on these guidelines to establish some consistency among the current research. The reader should note that Ibbotson Associates recommends the income return (or the yield) when using a bond as the risk-free rate rather than the total return.[7]

A third type is the length of time of the ERP forecast. We distinguish between short-run and long-run expectations. *Short-run* expectations refer to the current ERP or, for this paper, a prediction of up to 10 years. In contrast, the *long-run* expectation is a forecast over 10 years to as many as 75 years for social security purposes. Ten years appears an appropriate breaking point based on the current literature surveyed.

The next difference is whether the ERP estimate is *unconditional* or *conditioned* on one or more related variables. In defining this type, we refer to an admonition by Constantinides (2002) of the differences in these estimates:

> "First, I draw a sharp distinction between *conditional, short-term forecasts* of the mean equity return and premium and *estimates of the unconditional mean.* I argue that the currently low conditional short-term forecasts of the return and premium do not lessen the burden on economic theory to explain the large unconditional mean equity return and premium, as measured by their sample average over the past one hundred and thirty years" (p. 1568).

Many of the estimates we catalogue below will be conditional ones, conditional on dividend yield, expected earnings, capital gains, or other assumptions about the future.

ERP estimates can also exhibit a *U.S. versus international* market type depending on the data used for estimation purposes and the ERP being estimated. Dimson et al. (2002) notes that, at the start of 2000, the U.S. equity market, while dominant, was slightly less than one-half (46.1%) of the total international market for equities, capitalized at $52.7 trillion. Table 2 shows a comparison of historical ERP values for the United States and the world. Data from the non-U.S. equity markets are clearly different from those of U.S. markets and, hence, will produce different estimates for returns

[5] For a complete discussion of the arithmetic/geometric choice, see Ibbotson Associates (2003b, pp. 71–3). See also Dimson et al. (2002, p. 35), and Brennan and Schwartz (1985).

[6] The arithmetic difference is the geometric difference multiplied by 1 + Risk Free.

[7] The reason for this is two-fold. First, when issued, the yield is the expected market return for the entire horizon of the bond. No net capital gains are expected for the market return for the entire horizon of the bond. No capital gains are expected at the default-free maturity. Second, historical annual capital gains on long-term government bonds average near zero (0.4%) over the 1926–2002 period (Ibbotson Associates 2003a, tables 6–7).

and ERP.[8] Results for the entire world equity market will, of course, be a weighted average of the U.S. and non-U.S. estimates.

The next type is the *data source and period* used for the market and ERP estimates. Whether given an historical average of the ERP or an estimate from a model using various historical data, the ERP estimate will be influenced by the length, timing, and source of the underlying data used. The time series compilations are primarily annual or monthly returns. Occasionally, daily returns are analyzed, but not for the purpose of estimating an ERP. Some researchers use as much as 200 years of history; the Ibbotson data currently uses S&P 500 returns from 1926 to the present.[9]

As an example, Siegel (2002) examined a series of real U.S. returns beginning in 1802.[10] He used three sources to obtain the data. For the first period, 1802–1870, characterized by stocks of financial organizations involved in banking and insurance, he cites Schwert (1990). The second period, 1871–1925, incorporates Cowles stock indexes compiled in Shiller (1989). The last period, beginning in 1926, uses data from the Center for Research in Security Prices (CRSP), University of Chicago Graduate School of Business; these are the same data underlying Ibbotson Associates calculations.

Goetzmann et al. (2001) constructed an NYSE data series for 1815–1925 to add to the 1926–1999 Ibbotson series. They concluded that the pre-1926 and post-1926 data periods show differences in both risk and reward characteristics. They highlighted the fact that inclusion of pre-1926 data will generally produce lower estimates of ERPs than relying exclusively on the Ibbotson post-1926 data, similar to that shown in Appendix A. Several studies that rely on pre-1926 data, catalogued in Appendix B, show the magnitudes of these lower estimates.[11] Table 3 displays Siegel's ERPs for three subperiods. He notes that subperiod III, 1926–2001, shows a larger ERP (4.7%), or a smaller real risk-free mean (2.2%), than the prior subperiods.[12]

Table 2

Worldwide Equity Risk Premia, 1900–2000 Annual Equity Risk Premium Relative to Treasury Bills

Country	Geometric Mean	Arithmetic Mean
United States	5.8%	7.7%
World	4.9	6.2

Source: Dimson et al. (2002, pp. 166–67)

Smaller subperiods will show much larger variations in equity, bill, and ERP returns. Table 4 displays the Ibbotson returns and short-horizon risk premia for subperiods as small as five years. The scatter of results is indicative of the underlying large variation (20% std dev) in annual data.

In calculating an expected equity risk premium by averaging historical data, projecting historical data using growth models, or even conducting a survey, one must determine a proxy for the "market." Common proxies for the U.S. market include the S&P 500, the NYSE index, and the NYSE, AMEX, and NASDAQ index (Ibbotson Associates 2003b, p. 92). For the purpose of this paper, we use the S&P 500 and its antecedents as the market. However, in the various research surveyed, many different market proxies were assumed. We have already discussed using international versus ERP domestic data when describing different MRP types. With international data, different proxies for other country, region, or world markets are used. For example, Dimson (2002) and Claus and Thomas (2001) use international market data.

[8] One qualitative difference can arise from the collapse of equity markets during war time.

[9] For the Ibbotson analysis of the small stock premium, the NYSE/AMEX/NASDAQ combined data are used, with the S&P 500 data falling within deciles 1 and 2 (Ibbotson Associates 2003b, pp. 66 and Chapter 7.)

[10] A more recent alternative is Wilson and Jones (2002), as cited by Dimson et al. (2002, p. 39).

[11] Using Wilson and Jones' 1871–2002 data series, time series analyses show no significant ERP difference between the 1871–1925 period and the 1926–2002 period; one cannot distinguish the old from the new. The overall average is lower with the additional 1871–1925 data, but on a statistical basis, they are not significantly different. Assuming the equivalency of the two data series for 1871–1925 (Goetzmann et al. 2001 and Wilson and Jones 2002), the risk difference found by Goetzmann et al. must be determined by a significantly different ERP in the pre-1871 data. The 1871–1913 return that is prior to personal income tax and that appears to be about 35% lower than the 1926–2002 period average of 11.8%, might simply reflect a zero valuation for income taxes in the pre-1914 returns. Adjusting the pre-1914 data for taxes would most likely make the ERP for the entire period (1871–2002) approximately equal to 7.5%, the 1926–2002 average.

[12] The low risk-free return is indicative of the "risk-free rate puzzle," the twin of the ERP puzzle. For details see Weil (1989).

Table 3
Short-Horizon Equity Risk Premium by Subperiods

	Subperiod I 1802–1870	Subperiod II 1871–1925	Subperiod III 1926–2001
Real Geometric Stock Returns	7.0%	6.6%	6.9%
Real Geometric Long-Term Governments	4.8	3.7	2.2
Equity Risk Premium	2.2	2.9	4.7

Source: Siegel (2002, pp. 13 and 15).

For domestic data, different proxies have been used over time as stock market exchanges have expanded. (For a data series that is a mixture of the NYSE exchange, NYSE, AMEX, and NASDAQ stock exchange, and the Wilshire 5000, see Dimson 2002, p. 306.) Fortunately, as shown by Ibbotson Associates (2003b), the issue of a U.S. market proxy does not have a large effect on the ERP estimate because the various indices are highly correlated. For example, the S&P 500 and the NYSE have a correlation of 0.95, the S&P 500 and NYSE/AMEX/NASDAQ 0.97, and the NYSE and NYSE/AMEX/NASDAQ 0.90 (Ibbotson Associates 2003b, p. 93, using data from October 1997–September 2002). Therefore, the equity proxy selected is one reason for slight differences in the estimates of the market risk premium.

As a final note, stock returns and risk-free rates can be stated in *nominal* or *real* terms. Nominal includes inflation; real removes inflation. The ERP should not be affected by inflation because either the stock return and risk-free rate both

Table 4
Average Short-Horizon Risk Premium over Various Time Period

Year		Common stocks Total Annual Returns	U.S. Treasury Bills Total Annual Returns	Short-Horizon Risk Premium
All Data	1926–2002	12.20%	3.83%	8.37%
50-year	1953–2002	12.50	5.33	7.17
40-year	1963–2002	11.80	6.11	5.68
30-year	1943–1972	14.55	2.54	12.02
	1973–2002	12.21	6.61	5.60
15-year	1928–1942	5.84	0.95	4.89
	1943–1957	17.14	1.20	15.94
	1958–1972	11.96	3.87	8.09
	1973–1987	11.42	8.20	3.22
	1988–2002	13.00	5.03	7.97
10-year	1933–1942	12.88	0.15	12.73
	1943–1952	17.81	0.81	17.00
	1953–1962	15.29	2.19	13.11
	1963–1972	10.55	4.61	5.94
	1973–1982	8.67	8.50	0.17
	1983–1992	16.80	6.96	9.84
	1993–2002	11.17	4.38	6.79
5-year	1928–1932	−8.25	2.55	−10.80
	1933–1937	19.82	0.22	19.60
	1938–1942	5.94	0.07	5.87
	1943–1947	15.95	0.37	15.57
	1948–1952	19.68	1.25	18.43
	1953–1957	15.79	1.97	13.82
	1958–1962	14.79	2.40	12.39
	1963–1967	13.13	3.91	9.22
	1968–1972	7.97	5.31	2.66
	1973–1977	2.55	6.19	−3.64
	1978–1982	14.78	10.81	3.97
	1983–1987	16.93	7.60	9.33
	1988–1992	16.67	6.33	10.34
	1993–1997	21.03	4.57	16.46
	1998–2002	1.31	4.18	−2.88

Source: Authors' calculations using Ibbotson Associates (2003a, p. 38–39).

Table 5
ERP Using Same Historical Data (1926–2002)

RFR Description	ERP Description	ERP Historical Return
Short nominal	Arithmetic short-horizon	8.4%
Short nominal	Geometric short-horizon	6.4
Short real	Arithmetic short-horizon	8.4
Short real	Geometric short-horizon	6.4
Intermediate nominal	Arithmetic inter-horizon	7.4
Intermediate nominal	Geometric inter-horizon	5.4
Intermediate real	Arithmetic inter-horizon	7.4
Intermediate real	Geometric inter-horizon	5.4
Long nominal	Arithmetic long-horizon	7.0
Long nominal	Geometric long-horizon	5.0
Long real	Arithmetic long-horizon	7.0
Long real	Geometric long-horizon	5.0

Source: Authors' calculations using Ibbotson Associates (2003a, p. 112).

include the effects of inflation (both stated in nominal terms) or neither have inflation (both stated in real terms). If both returns are nominal, the difference in the returns is generally assumed to remove inflation. Otherwise, both terms are real, so inflation is removed prior to finding the ERP. While numerical differences in the real and nominal approaches may exist, their magnitudes are expected to be small.

4. Equity Risk Premia 1926–2002

As an example of the importance of knowing the types of ERP estimates under consideration, Table 5 displays ERP returns that each use the same historical data, but are based on arithmetic or geometric returns and the type of horizon. The ERP estimates are quite different.[13]

5. Historical Methods

The historical methodology uses averages of past returns to forecast future returns. Different time periods may be selected, but the two most common periods arise from data provided by either Ibbotson or Siegel. The Ibbotson series begins in 1926 and is updated each year. The Siegel series begins in 1802, with the most recent compilation using returns through 2001.

Appendix A provides ERP estimates using Ibbotson data for the 1926–2002 period that we use in this paper for most illustrations. We begin with a look at the ERP history through a time series analysis of the Ibbotson data.

6. Time Series Analysis

Much of the analysis addressing the ERP puzzle relies on the annual time series of market, risk-free and risk premium returns. Two opposite views can be taken of these data. One view would have the 1926–2002 Ibbotson data or the 1802–2001 Siegel data represent one data point; that is, we have observed one path for the ERP through time from the many possible 77- or 200-year paths. This view rests upon the existence or assumption of a stochastic process with (possibly) intertemporal correlations.

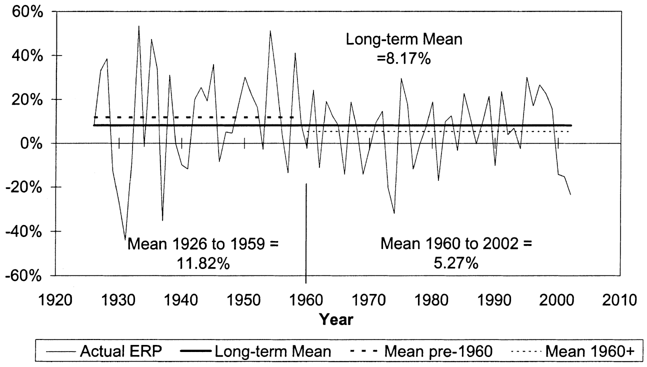
While mathematically sophisticated, this model is particularly unhelpful without some testable hint at the details of the generating stochastic process. The practical view is that the observed returns are random samples from annual distributions that are i.i.d. (independent and identically distributed) about the mean. The obvious advantage is that we have at hand 77 or 200 observations on the i.i.d. process to analyze. We adopt the latter view.

Some analyses adopt the assumption of stationarity of ERP; that is, the true mean does not change with time. Figure 1 displays the Ibbotson ERP data and highlights two subperiods, 1926–1959 and 1960–2002.[14] While the mean ERP for

[13] The nominal and real ERPs are identical in Table 5 because the ERPs are calculated as arithmetic differences, and the same value of inflation will reduce the market return and the risk-free return equally. Geometric differences would produce minimally different estimates for the same types.

[14] The ERP shown here are the geometric differences (calculated) rather than the simple arithmetic differences in Table 1; i.e., ERP = $[(1 + r_m)/(1 + r_f)] - 1$. The test results are qualitatively the same for the arithmetic differences.

Figure 1
Short-Horizon Equity Risk Premium



Source: Authors' calculations using Ibbotson Associates (2003a, p. 38–39), geometric differences.

the two subperiods appear quite different (11.82% versus 5.27%), the large variance of the process (20.24% std dev) should make them indistinguishable, statistically speaking.

7. T-TESTS

The standard t-test can be used for the null hypothesis H_o : mean 1960–2002 = 8.17%, the 77-year mean.[15] The outcome of the test is shown in Table 6; the null hypothesis cannot be rejected. Another t-test can be used to test whether the subperiod means are different in the presence of unequal variances.[16] The result is similar to Table 6 and the difference of subperiod means equal to zero cannot be rejected.[17]

8. TIME TRENDS

The supposition of stationarity of the ERP series can be supported by ANOVA regressions. The results of regressing the ERP series on time is shown in Table 7. There are no significant time trends in the Ibbotson ERP data.[18]

Table 6
T-Test under the Null Hypothesis That ERP (1960–2002) = ERP (1926–2002) = 8.17%

Sample mean 1960–2002	5.27%
Sample s.d. 1960–2002	15.83%
T-value (DF = 42)	−1.20
PR > |T|	0.2374
Confidence Interval 95%	(0.0040, 0.1014)
Confidence Interval 90%	(0.0121, 0.0933)

9. ARIMA MODEL

Time series analysis using the well-established Box-Jenkins approach can be used to predict future series values through the lag correlation structure (see Harvey 1990, p. 30). The SAS ARIMA procedure applied to the full 77 time series data shows:

1. No significant autocorrelation lags.
2. An identification of the series as white noise.
3. ARIMA projection of year 78 + ERP is 8.17%, the 77 year average.

All of the above single time series tests point to the reasonability of the stationarity assumption for (at least) the Ibbotson ERP 77-year series.[19]

10. SOCIAL SECURITY ADMINISTRATION

In the current debate on whether to allow private accounts that may invest in equities, the Office of the Chief Actuary (OCACT) of the Social Security Administration (SSA) has selected certain assumptions to assess various proposals (Goss 2001). The relevant selection is to use 7% as the real (geometric) annual rate of return for equities (compare Table 3, subperiod III). This assumption is based on the historical return of the 20th century. SSA received further support that showed the historical return for the last 200 years is consistent with this estimate, along with the Ibbotson series beginning in 1926.

For SSA, the calculation of the ERP uses a long-run real yield on Treasury bonds as the risk-free rate. From the assumptions in the 1995 Trustees Report, the long-run real yield on Treasury bonds that the Advisory Council proposals

[15] Standard statistical procedures in SAS 8.1 have been used for all tests.

[16] Equality of variances is rejected at the 1% level by an F test (F = 2.39, DF = 33,42).

[17] T-value 1.35, PR > |T| = 0.1850 (Cochran method).

[18] The result is confirmed by a separate Chow test on the two subperiods.

[19] The same tests applied to the Wilson and Jones 1871–2002 data series show similar results: Neither the 1871–1925 period nor the 1926–2002 period is different from the overall 1871–2002 period. The overall period and subperiods also show no trends over time.

Table 7
ERP ANOVA Regressions on Time

Period	Time Coefficient	P-Value
1926–1959	0.004	0.355
1960–2002	0.001	0.749
1926–2002	−0.001	0.443

use is 2.3%. Using a future Treasury securities real yield of 2.3% produces a geometric ERP of 4.7% over long-term Treasury securities. More recently, the Treasury securities assumption has increased to 3% (Social Security Trustees Report 1999), yielding a 4% geometric ERP over long-term Treasury securities.

At the request of the OCACT, John Campbell, Peter Diamond, and John Shoven were engaged to give their expert opinions on the assumptions Social Security made. Each economist begins with the Social Security assumptions and then explains any difference he or she feels would be more appropriate.

Campbell (2001) considered valuation ratios as a comparison to the returns from the historical approach. The current valuation ratios are at unusual levels, with a low dividend-price ratio and high price-earnings ratio. He reasoned that the prices are what have dramatically changed these ratios. Campbell presented two views as to the effect of valuation ratios in their current state. One is that valuations will remain at the current level, suggesting much lower expected returns. The second view is a correction to the ratios, resulting in less favorable returns until the ratios readjust. He decided to give some weight to both possibilities, so he lowered the geometric equity return estimate to 5–5.5% from 7%. For the risk-free rate, he used the yield on the long-term inflation-indexed bonds of 3.5% or the OCACT assumption of 3% (see discussion of current yields on Treasury Inflation Protection Securities (TIPS) in Section 16 below). Therefore, his geometric equity premium estimate was around 1.5–2.5%.

Diamond (1999, 2001) used the Gordon growth formula to calculate an estimate of the equity return. The classic Gordon dividend growth model (Brealey and Myers 2000, p. 67) follows.

$$K = (D_1/P_0) + g$$

K = Expected return or discount rate

P_0 = Price this period

D_1 = Expected dividend next period

g = Expected growth in dividends in perpetuity

Based on analysis, he felt that the equity return assumption of 7% for the next 75 years is not consistent with a reasonable level of stock value compared to GDP. Even when increasing the GDP growth assumption, he still did not feel that the equity return was plausible. By reasoning that the next decade of returns will be lower than normal, only then is the equity return beyond that time frame consistent with the historical return. By considering the next 75 years together, he would lower the overall projected equity return to 6–6.5%. He argued that the stock market is overvalued, and a correction is required before the long-run historical return is a reasonable projection for the future. By using the OCACT assumption of 3% for the long-term real yield on Treasury bonds, Diamond estimated a geometric ERP of about 3–3.5%.

Shoven (2001) began by explaining why the traditional Gordon growth model is not appropriate and suggested a modernized Gordon model that allows share repurchases to be included, instead of only using the dividend yield and growth rate. By assuming a long-term price-earnings ratio between its current and historical value, he came up with an estimate for the long-term real equity return of 6.125%. Using his general estimate of 6–6.5% for the equity return and the OCACT assumptions for the long-term bond yield, he projected a long-term ERP of approximately 3–3.5%.

All the SSA experts begin by accepting the long-run historical ERP analyses and then modifying that by changes in the risk-free rate or by decreases in the long-term ERP based on their own personal assessments. We now turn to the major strains in ERP puzzle research.

11. ERP PUZZLE RESEARCH

Campbell and Shiller (2001) began with the assumption of mean reversion of dividend/price and price/earnings ratios. Next, they explained the result of prior research (Campbell and Shiller 1988) that found that the dividend-price ratio predicts future prices, and historically, the price corrects the ratio when it diverts from the mean. Based on this result, they then used regressions of the dividend-price ratio and the price-smoothed-earnings ratio—"smoothed" by using 10-year averages—to predict future stock prices out 10 years. Both regressions predict large losses in stock prices for the 10-year horizon.

Although Campbell and Shiller (2001) did not rerun the regression on the dividend-price ratio to incorporate share repurchases, they pointed out that the dividend-price ratio should be upwardly adjusted, but the adjustment only moves the ratio to the lower range of the historical fluctuations (as opposed to the mean). They concluded that the valuation ratios indicate a bear market in the near future.[20] They predicted negative real stock returns for the next 10-year period. They also cautioned that, because valuation ratios have changed so much from their normal level, they may not completely revert to the historical mean, but this does not change their pessimism about the next decade of stock market returns.

Arnott and Ryan (2001) took the perspective of fiduciaries, such as pension fund managers, with an investment portfolio. They began by breaking down the historical stock returns (for the 74 years since December 1925) by analyzing dividend yields and real dividend growth. They pointed out that the historical dividend yield is much higher than the current dividend yield of about 1.2%. They argued that the changes from stock repurchases, reinvestment, and mergers and acquisitions, which affect the lower dividend yield, can be represented by a higher dividend growth rate. However, they capped real dividend or earnings growth at the level of real economic growth. They added the dividend yield and the growth in real dividends to come up with an estimate for the future equity return; the current dividend yield of 1.2% and the economic growth rate of 2% add to the 3.2% estimated real stock return. This method corresponds to the dividend growth model or earnings growth model and does not take into account changing valuation levels. They cite a TIPS yield of 4.1% for the real risk-free rate return (see Section 16). These two estimates yield a negative geometric long-horizon conditional ERP.

Arnott and Bernstein (2002) began by arguing that, in 1926, investors were not expecting the realized, historical compensation that they later received from stocks. They cited bonds' reaction to inflation, increasing valuations, survivorship bias (see Brown et al. 1992, 1995), and changes in regulation as positive events that helped investors during this period. They only used the dividend growth model to predict a future expected return for investors. They did not agree that the earnings growth model is better than the dividend growth model, both because earnings are reported using accounting methods and earnings data before 1870 are inaccurate. Even if the earnings growth model is chosen instead, they found that the earnings growth rate from 1870 only grows 0.3% faster than dividends, so their results would not change much. Because of the Modigliani-Miller theorem (Brealey and Myers 2000, p. 447; also see the discussion in Ibbotson and Chen 2003), a change in dividend policy should not change the value of the firm. Arnott and Bernstein concluded that managers benefited in the "era of 'robber baron' capitalism" (p. 66) instead of the conclusion reached by others that the dividend growth model underrepresents the value of the firm.

By holding valuations constant and using the dividend yield and real growth of dividends, Arnott and Bernstein (2002) calculated the equity return that an investor might have expected during the historical time period starting in 1802. They used an expected dividend yield of 5%, close to the historical average of 1810–2001. For the real growth of dividends, they chose the real per capita GDP growth less a reduction for entrepreneurial activity in the economy plus stock repurchases. They concluded that the net adjustment is negative, so the real GDP growth is reduced from 2.5–3% to only 1%. A fair expectation of the stock return for the historical period is close to 6.1% by adding 5% for the dividend yield and a net real GDP per capita growth of 1.1%. They used a TIPS yield of 3.7% for the real risk-free rate, which yields a geometric intermediate-horizon ERP of 2.4% as a fair expectation for investors in the past. They considered this a "normal" ERP estimate. They also opined that the current ERP is zero; that is, they expected stocks and (risk-free) bonds to return the same amounts.

[20] The stock market correction from year-end 1999 to year-end 2002 is a decrease of 37.6%, or 14.6% per year. Presumably, the "next 10 years" refers to 2000 to 2010.

Fama and French (2002) used both the dividend growth model and the earnings growth model to investigate three periods of historical returns: 1872–2000, 1872–1950, and 1951–2000. Their ultimate aim was to find an unconditional ERP. They cited that, by assuming the dividend-price ratio and the earnings-price ratio follow a mean reversion process, the result follows that the dividend growth model or earnings growth model produce approximations of the unconditional equity return. Fama and French's analysis of the earlier period of 1872–1950 shows that the historical average equity return and the estimate from the dividend growth model are about the same.

In contrast, they found that the 1951–2000 period has different estimates for returns when comparing the historical average and the growth models' estimates. The difference in the historical average and the model estimates for 1951–2000 was interpreted to be "unexpected capital gains" over this period. They found that the unadjusted growth model estimates of the ERP, 2.55% from the dividend model and 4.32% from the earnings model, fell short of the realized average excess return for 1951–2000.

Fama and French preferred estimates from growth models instead of the historical method because of the lower standard error using the dividend growth model. Fama and French provided 3.83% as the unconditional expected ERP return (referred to as the annual bias-adjusted ERP estimate) using the dividend growth model with underlying data from 1951–2000. They gave 4.78% as the unconditional expected ERP return, using the earnings growth model with data from 1951–2000. Note that using a one-month Treasury bill instead of commercial paper for the risk-free rate would increase the ERP by about 1% to nearly 6% for the 1951–2000 period.

Ibbotson and Chen (2003) examined the historical real geometric long-run market and long risk-free returns using their "building block" methodology.[21] They used the full 1926–2000 Ibbotson Associates data and considered as building blocks all of the fundamental variables of the prior researchers. Those blocks include (not all simultaneously):

- Inflation.
- Real risk-free rates (long).
- Real capital gains.
- Growth of real earnings per share.
- Growth of real dividends.
- Growth in payout ratio (dividend/earnings).
- Growth in book value.
- Growth in ROE.
- Growth in price/earnings ratio.
- Growth in real GDP/population.
- Growth in equities excess of GDP/POP.
- Reinvestment.

Their calculations show that a forecast real geometric long-run return of 9.4% is a reasonable extrapolation of the historical data underlying a realized 1926–2000 return of 10.7%, yielding a long-horizon arithmetic ERP of 6%, or a short-horizon arithmetic ERP of about 7.5%.

Ibbotson and Chen (2003) constructed six building-block methods; that is, they used combinations of historic estimates to produce an expected geometric equity return. They highlighted the importance of using both dividends and capital gains by invoking the Modigliani-Miller theorem. The methods, and their component building blocks are:

- Method 1: Inflation, real risk-free rate, realized ERP.
- Method 2: Inflation, income, capital gains and reinvestment.
- Method 3: Inflation, income, growth in price/earnings, growth in real earnings per share and reinvestment.
- Method 4: Inflation, growth rate of price/earnings, growth rate of real dividends, growth rate of payout ratio dividend yield and reinvestment.
- Method 5: Inflation, income growth rate of price/earnings, growth of real book value, ROE growth and reinvestment.
- Method 6: Inflation, income, growth in real GDP/POP, growth in equities excess GDP/POP and reinvestment.

All six methods reproduce the historical long-horizon geometric mean of 10.70% as shown in Appendix D. Since the source of most other researchers'

[21] See Appendix D for a summary of their estimates. Also see Pratt (1998) for a discussion of the building block, or build-up model, cost of capital estimation method.

lower ERP is the dividend yield, Ibbotson and Chen (2003) recast the historical results in terms of *ex ante* forecasts for the next 75 years. Their estimate of 9.37% using supply side methods 3 and 4 is approximately 130 basis points lower than the historical result. Within their methods, they also show how the substantially lower expectation of 5.44% for the long mean geometric return is calculated by omitting one or more relevant variables. Underlying these *ex ante* methods are the assumptions of stationarity of the mean ERP return and market efficiency, the absence of the assumption that the market has mispriced equities. All of their methods are aimed at producing an unconditioned estimate of the *ex ante* ERP.

As opposed to short-run, conditional estimates from Campbell and Shiller and others, Constantinides (2002) sought to estimate the unconditional ERP, more in line with the goal of Fama and French (2002) and Ibbotson and Chen (2003). He began with the premise that the unconditional ERP can be estimated from the historical average using the assumption that the ERP follows a stationary path. He suggested that most of the other research produces conditional estimates, conditioned upon beliefs about the future paths of fundamentals such as dividend growth, price-earnings ratio, and the like. While interesting in themselves, they add little to the estimation of the unconditional mean ERP.

Constantinides (2002) used the historical return and adjusted downward by the growth in the price-earnings ratio to calculate the unconditional ERP. He removed the growth in the price-earnings ratio because he was assuming no change in valuations in the unconditional state. He gave estimates using three periods. For 1872–2000, he used the historical ERP, which is 6.9%, and, after amortizing the growth in the price-dividend ratio or price-earnings ratio over a period as long as 129 years, the effect of the potential reduction was no change. Therefore, he found an unconditional arithmetic, short-horizon ERP of 6.9% using the 1872–2000 underlying data. For 1951–2000, he again started with the historical ERP, which is 8.7%, and lowered this estimate by the growth in the price-earnings ratio of 2.7% to find an unconditional arithmetic, short-horizon ERP of 6.0%. For 1926–2000, he used the historical ERP, which is 9.3%, and reduced this estimate by the growth in the price-earnings ratio of 1.3% to find an unconditional arithmetic, short-horizon ERP of 8.0%. He appealed to behavioral finance to offer explanations for such high unconditional ERP estimates.

From the perspective of giving practical investor advice, Malkiel (1999) discussed "the age of the millennium" to give some indication of what investors might expect for the future. He specifically estimated a reasonable expectation for the first few decades of the 21st century. He estimated the future bond returns by giving estimates if bonds are held to maturity with corporate bonds of 6.5–7%, long-term zero-coupon Treasury bonds of about 5.25%, and TIPS with a 3.75% return.

Depending on the desired level of risk, Malkiel indicated bondholders should be more favorably compensated in the future compared to the historical returns from 1926 to 1998. Malkiel used the earnings growth model to predict future equity returns. He used the then-current dividend yield of 1.5% and an earnings growth estimate of 6.5%, yielding an 8% equity return estimate, compared with an 11% historical return. Malkiel's estimated range of the ERP is from 1% to 4.25%, depending on the risk-free instrument selected. Although his ERP is lower than the historical return, his selection of a relatively high earnings growth rate is similar to Ibbotson and Chen's (2003) forecasted models. In contrast with Ibbotson and Chen, Malkiel allowed for a changing ERP and advised investors not to rely solely on the past "age of exuberance" as a guide for the future. Malkiel pointed out the impact of changes in valuation ratios but did not attempt to predict future valuation levels.

Finally, Mehra (2002) summarized the results of the research since the ERP puzzle was posed. The essence of the puzzle is the inconsistency of the ERPs produced by descriptive and prescriptive economic models of asset pricing, on the one hand, and the historical ERPs realized in the U.S. market, on the other. Mehra and Prescott (1985) speculated that the inconsistency could arise from the inadequacy of standard models to incorporate market imperfections and transaction costs. Failure of the models to reflect reality rather than failure of the market to follow the theory seems to be Mehra's conclusion as of 2002. Mehra points to two promising threads of model-modifying research. Campbell and Cochrane (1999) incorporated economic cycles and changing risk aversion while Constantinides et al. (2002) proposed a life cycle investing modification, replacing the representative agent by segmenting investors into young, middle-aged, and older cohorts. Mehra summed up as follows:

> "Before we dismiss the premium, we not only need to have an understanding of the observed phenomena but also why the future is likely to be different. In the absence of this, we can make the following claim based on what we know. Over the long horizon the equity premium is likely to be similar to what it has been in the past and the returns to investment in equity will continue to substantially dominate those in bonds for investors with a long planning horizon" (p. 146).

12. Financial Analyst Estimates

Claus and Thomas (2001) and Harris and Marston (2001) both provided equity premium estimates using financial analysts' forecasts. However, their results were rather different. Claus and Thomas used an abnormal earnings model with data from 1985 to 1998 to calculate an ERP, as opposed to using the more common dividend growth model. Financial analysts project five-year estimates of future earnings growth rates. When using this five-year growth rate for the dividend growth rate in perpetuity in the Gordon growth model, Claus and Thomas explained that there is a potential upward bias in estimates for the ERP. Therefore, they chose to use the abnormal earnings model, instead, and only let earnings grow at the level of inflation after five years. The abnormal earnings model replaced dividends with "abnormal earnings" and discounted each flow separately instead of using a perpetuity. The average estimate that they found was 3.39% for the ERP.

Although it is generally recognized that financial analysts' estimates have an upward bias, Claus and Thomas (2001) proposed that, in the current literature, financial analysts' forecasts have underestimated short-term earnings in order for management to achieve earnings estimates in the slower economy. Claus and Thomas concluded that their findings of the ERP using data from the past 15 years were not in line with historical values.

Harris and Marston (2001) used the dividend growth model with data from 1982 to 1998. They assumed that the dividend growth rate should correspond to investor expectations. By using financial analysts' longest estimates (five years) of earnings growth in the model, they attempted to estimate these expectations. They argued that, if investors are in accord with the optimism shown in analysts' estimates, even biased estimates do not pose a drawback because these market sentiments will be reflected in actual returns. Harris and Marston found an ERP estimate of 7.14%, with fluctuations in the ERP over time. Because their estimates were close to historical returns, they contended that investors would continue to require a high ERP.

13. Survey Methods

One method to estimate the *ex ante* ERP is to find the consensus of experts. Graham and Harvey (2002) surveyed chief financial officers to determine the average cost of capital used by firms. Welch (2000, 2001) surveyed financial economists to determine the ERP that academic experts in this area would estimate.

Graham and Harvey (2002) administered surveys from the second quarter of 2000 to the third quarter of 2002. For their survey format, they showed the current 10-year bond yield and then asked CFOs to provide their estimate of the S&P 500 return for the next year and over the next 10 years. CFOs are actively involved in setting a company's individual hurdle rate[22] and, therefore, are considered knowledgeable about investors' expectations. When comparing the survey responses of the one- and 10-year returns, the one-year returns have so much volatility that the authors, Graham and Harvey, concluded that the 10-year ERP is the more important and appropriate return of the two when making financial decisions such as estimating hurdle rates and cost of capital. The average 10-year ERP estimate varied from 3% to 4.7%.

In his most current survey, Welch (2001) compiled the responses of about 500 financial economists to determine their consensus ERP. He found the average arithmetic estimate for the 30-year ERP, relative to Treasury bills, to be 5.5% and the one-year arithmetic ERP consensus to be 3.4%. Welch deduced from the average 30-year geometric equity return estimate of 9.1% that the

[22] A "hurdle" rate is a benchmark cost of capital used to evaluate projects to accept (expected returns greater than hurdle rate) or reject (expected returns less than hurdle rate). Graham and Harvey (2002) claim three-fourths of the CFOs use CAPM to estimate hurdle rates.

Table 8
Differences in Forecasts across Expertise Level

Relative Expertise	Statistic	Stock Market 30-Year Geometric	Equity Premium 30-Year Arithmetic	Equity Premium 30-Year Geometric
188 Less Involved	Mean	8.5%	4.9%	4.4%
	Median	8	5	4
	IQ Range	6–10	3–6	2–5.5
235 Average	Mean	9.2	5.8	4.8
	Median	9	5	4
	IQ Range	7.5–10	3.5–7	3–6
72 Experts	Mean	10.1	6.2	5.4
	Median	9	5.4	5
	IQ Range	8–11	4–7.5	3.4–6

Source: Welch (2001, table 5).

arithmetic equity return forecast was approximately 10%.[23]

Welch's survey question allowed participants to self-select into different categories based on their knowledge of ERP. The results indicate that the responses of the less ERP-knowledgeable participants were more pessimistic than those of the self-reported experts. The experts gave 30-year estimates that are 30–150 basis points above the estimates of the nonexpert group.

Table 8 shows that there may be a "lemming" effect, especially among economists who are not directly involved in the ERP question. Stated differently, all the academic and popular press—together with the prior 1998 Welch survey (which had an ERP consensus of about 7%)—could have conditioned the nonexpert, or the "less involved," that the expected ERP was lower than historic levels.

14. THE BEHAVIORAL APPROACH

Benartzi and Thaler (1995) analyzed the ERP puzzle from the viewpoint of prospect theory (Kahneman and Tversky 1979). Prospect theory allows asymmetric "loss aversion"—the fact that individuals are more sensitive to potential loss than gain—as one of its central tenets (see Tversky and Kahneman 1991 and Barberis et al. 2001 for a current survey of the applications of prospect theory to finance). Once an asymmetry in risk aversion is introduced into the model of the rational representative investor or agent, the unusual risk aversion problem raised initially by Mehra and Prescott (1985) can be "explained" by parameters within this behavioral model of decision making under uncertainty.

Stated differently, given the historical ERP series, there exists a model of investor behavior that can produce those or similar results. Benartzi and Thaler (1995) combined loss aversion with "mental accounting"—the behavioral process people use to evaluate their status relative to gains and losses compared to expectations, utility, and wealth—to get "myopic loss aversion." In particular, mental accounting for a portfolio needs to take place infrequently in order to reduce the chances of observing loss versus gain. The authors concede that there is a puzzle with the standard expected utility-maximizing paradigm but that the myopic loss aversion view may resolve the puzzle. The authors' views are not free of controversy; any progress applying behavioral concepts to the ERP puzzle is sure to match the advance of behavioral economics as a whole.

The adoption of other behavioral aspects of investing also may provide support for the historical patterns of ERPs we see from 1802–2002. For example, as the true nature of risk and rewards has been uncovered by the virtual army of 20th century researchers, and as institutional investors held sway in the latter 50 years of the century, the demand for higher rewards seen in the later historical data may be a natural and rational response to the new and expanded information set. Dimson et al. (2002, figs. 4–6) displays increasing real U.S. equity returns of 6.7%, 7.4%, 8.2% and 10.2% for periods of 101, 75, 50 and 25

[23] For the Ibbotson 1926–2002 data, the arithmetic return is about 190 basis points higher than the geometric return, rather than the inferred 90 basis points. This suggests the participants' beliefs, in Welch's study, may not be internally consistent.

years, ending in 2001, consistent with this "risk-learning" view.

15. The Next 10 Years

The "next 10 years" is an issue that Campbell and Diamond discuss when reviewing Social Security's assumptions and Campbell and Shiller (2001) address, either explicitly or implicitly. Experts evaluating Social Security's proposals predicted that returns during the "next 10 years," indicating a period beginning around 2000, were likely to be below the historical return. However, a historical return was recommended as appropriate for the remaining 65 of the 75 years to be projected. The period Campbell and Shiller discussed is approximately 2000–2010. Based on the then-current state of valuation ratios, they predicted lower stock market returns over "the next 10 years."

These expert predictions, and other pessimistic low estimates, have already come to fruition as market results from 2000 through 2002.[24] The U.S. equities market has decreased 37.6% since 1999, or an annual decrease of 14.6%. Although these forecasts have proved to be accurate in the short term, for future long-run projections, the market is not at the same valuation today as it was when these conditional estimates were originally given. Therefore, actuaries should be wary of using the low long-run estimates made prior to the large market correction of 2000–2002.

16. Treasury Inflation Protection Securities

Several of the ERP researchers referred to TIPS when considering the real risk-free rates. Historically, they adjusted Treasury yields downward to a real rate by an estimate of inflation, presumably for the term of the Treasury security. The modern era data in Table 3 show a low real long-term, risk-free rate of return (2.2%). This contrasts with the initial TIPS issue yields of 3.375%.[25] Some researchers use those TIPS yields as (market) forecasts of real risk-free returns for intermediate and long-horizon, together with reduced (real) equity returns, to produce low estimates of *ex ante* ERPs. None consider the volatility of TIPS as indicative of the accuracy of their ERP estimate.

Table 9
Inflation-Indexed Treasury Securities

Maturity	Coupon Issue Rate	Yield to Maturity
1/2011	3.500	1.763
1/2012	3.375	1.831
7/2012	3.000	1.878
4/2028	3.625	2.498
4/2029	3.875	2.490
4/2032	3.375	2.408

Source: Wall Street Journal (2003)

Table 9 shows a 2003 market valuation of 10- and 30-year TIPS issued in 1998–2002. Note the large 90–180 basis point decrease in the current "real" yields from the issue yields even just a year later for some issues. While there can be several explanations for the change (revaluation of the inflation option, flight to Treasury quality, paucity of 30-year Treasuries), the use of these current "real" risk-free yields, with fixed expected returns, would raise ERPs by at least 1%.

17. Conclusion

This paper has sought to bring the essence of recent research on the ERP to practicing actuaries. The researchers covered here face the same ubiquitous problems that actuaries face daily: Do I rely on past data to forecast the future (costs, premiums, investments), or do I analyze the past and apply informed judgment as to future differences, if any, to arrive at actuarially fair forecasts? Most of the ERP estimates lower than the unconditional historical estimate have an undue reliance on recent lower dividend yields (without a recognition of capital gains[26]) and/or on data prior to 1926.

Despite a spate of research suggesting *ex ante* ERPs lower than recent realized ERPs, actuaries should be aware of the range of estimates covered here (Appendix B); be aware of the underlying

[24] The Social Security Advisory Board (2002) will revisit the 75-year rate of return assumption during 2003.

[25] TIPS were introduced by the Treasury in 1996 with the first issue in January 1997.

[26] Under the current U.S. tax code, capital gains are tax-advantaged relative to dividend income for the vast majority of equityholders (households and mutual funds are 55% of the total equityholders, according to the Federal Flow of Funds, 2002 Q3, Table L-213). Curiously, the reverse is true for property-liability insurers because of the 70% stock dividend exclusion afforded insurers.

assumptions, data, and terminology; and be aware that their independent analysis is required before adopting an estimate other than the historical average. We believe that the Ibbotson and Chen (2003) layout, reproduced here as Appendix D, offers the actuary both an understanding of the fundamental components of the historical ERP and the opportunity to change the estimates based on good judgment and supportable beliefs. We believe that reliance solely on "expert" survey averages, whether of financial analysts, academic economists, or CFOs, is fraught with risks of statistical bias in estimates of the *ex ante* ERP.

It is dangerous for actuaries to engage in simplistic analyses of historical ERPs to generate *ex ante* forecasts that differ from the realized mean.[27] The research we have catalogued in Appendix B, the common level ERPs estimated in Appendix C, and the building-block (historical) approach of Ibbotson and Chen (2003) in Appendix D all discuss important concepts related to both *ex post* and *ex ante* ERPs and cannot be ignored in reaching an informed estimate.

For example, Wendt (2002) concluded that a linear relationship with interest rates is a better predictor of future returns than is a "constant" ERP based on the average historical return. He arrived at this conclusion by estimating a regression equation relating long bond yields with 15-year geometric mean market returns starting monthly in 1960.[28] Wendt's findings are misleading. First, there was no significant relationship between short-, intermediate-, or long-term income returns over 1926–2002 (or 1960–2002) and annual ERPs, as evidenced by simple regressions using Ibbotson data.[29] Second, if the linear structural equation indeed held, there would be no need for an ERP since the (15-year) return could be predicted within small error bars. Third, there is always a negative bias introduced when geometric averages are used as dependent variables (Brennan and Schwartz 1985). Finally, the results are likely to be spurious due to the high autocorrelations of the target and independent variables; an autocorrelation correction would eliminate any significant relationship of long yields to the ERP.

Actuaries also should be aware of the variability of both the ERP and risk-free rate estimates discussed in this paper (see Tables 4 and 9). All too often, return estimates are made without noting the error bars, and that can lead to unexpected "surprises." As one example, recent research by Longstaff (2004) proposes that a 1991–2001 "flight to quality" has created a valuation premium (and lowered yields) in the entire yield curve of Treasuries. He finds a 10–16 basis point liquidity premium throughout the zero coupon Treasury yield curve. He translates that into a 10–15% pricing difference at the long end. This would imply a simple CAPM market estimate for the long horizon might be biased low.

Finally, actuaries should know that the research catalogued in Appendix B is not definitive. No simple model of ERP estimation has been universally accepted. Undoubtedly, there will be still more empirical and theoretical research into this data-rich financial topic. We await the potential advances in understanding the return process that the behavioral view may uncover.

18. POST SCRIPT: APPENDICES A–D

We provide four appendices that catalogue the ERP approaches and estimates discussed in the paper. Actuaries, in particular, should find the numerical values, and descriptions of assumptions underlying those values helpful for valuation work that adjusts for risk. Appendix A provides the annual data from 1926 through 2002 from Ibbotson Associates referred to throughout this paper. The equity risk premium shown is a simple difference of the arithmetic stock returns and the arithmetic U.S. Treasury bill total returns.

Appendix B is a compilation of articles and books related to the ERP. The puzzle research section contains the articles and books that were most related to addressing the ERP puzzle.[30] Appendix B

[27] ERPs are derived from historical or expected after-corporate-tax returns. Pre-tax returns depend uniquely on the tax schedule for the differing sources of income.

[28] Fifteen-year mean returns = 2.032 (Long Government Yield) − 0.0242, $R^2 = 0.882$.

[29] The *p*-values on the yield variables in an annual ERP/yield regression using 1926–2002 annual data are 0.1324, 0.2246, and 0.3604 for short-, intermediate-, and long-term yields, respectively, with adjusted R-square values virtually zero.

[30] Additional references are included, in the table, that were not previously discussed (see Cornell 1999, Dimson et al. 2002, Siegel 1999, Siegel 2002, and Grinold and Kroner 2002 (Barclays Global Investors).

gives each source, along with risk-free rate and ERP estimates and further details collected from each source. For example, we show the data period used, if applicable, and the projection period. We also list the general methodology used in the reference. Footnotes give additional details on the sources' intent.

Appendix C adjusts all the ERP estimates to a short-horizon, arithmetic, unconditional ERP estimate. We begin with the authors' estimates for a stock return (the risk-free rate plus the ERP estimate). Next, we make adjustments if the ERP "type" given by the author(s) is not provided in this format. For example, to adjust from a geometric to an arithmetic ERP estimate, we adjust upward by the 1926–2002 historical difference in the arithmetic large-company stocks' total return and the geometric large-company stocks' total return of 2%. Next, if the estimate is given in real instead of nominal terms, we adjust the stock return estimate upward by 3.1%, the 1926–2002 historical return for inflation.

We make an approximate adjustment to move the estimate from a conditional to unconditional estimate based on Fama and French (2002) where they make similar adjustments for the biases in a dividend or earnings growth model. For the 1951–2000 period, Fama and French use an adjustment of 1.28% for the dividend growth model and 0.46% for the earnings growth model (Table 4, p. 655). Using their adjustment method and the data provided in Fama and French's table 1, the 1872–2000 period would require a 0.82% adjustment and the 1872–1950 period would require a 0.54% adjustment using a dividend growth model. Therefore, we selected the lowest adjustment (0.46%) from the different time periods and models as a minimum adjustment from a conditional estimate to an unconditional estimate of market returns. Finally, we subtract the 1926–2002 historical U.S. Treasury bills' total return to arrive at an adjusted ERP.

These adjustments are only approximations because the various sources rely on different underlying data, but the changes in the ERP estimate should reflect the underlying concept that different "types" of ERPs cannot be directly compared and require some attempt to normalize the various estimates.

Appendix D reproduces a table from Ibbotson and Chen (2003) that breaks down historical returns using various methods discussed in their paper, including forward-looking estimates. Summarized formulas from Ibbotson and Chen's paper are also displayed.

ACKNOWLEDGMENTS

The authors acknowledge the helpful comments from participants in the 2003 Bowles Symposium, Louise Francis of Francis Analytics & Actuarial Data Mining Inc., and four anonymous referees. The authors would also like to thank Jack Wilson for supplying his data series from Wilson and Jones (2002).

Appendix A

Ibbotson Market Data 1926–2002

Year	Common stocks	U.S. Treasury Bills	Arithmetic Short-Horizon	Year	Common stocks	U.S. Treasury Bills	Arithmetic Short-Horizon
	Total Annual Returns	Total Annual Returns	Equity Risk Premia		Total Annual Returns	Total Annual Returns	Equity Risk Premia
1926	11.62%	3.27%	8.35%	1966	−10.06%	4.76%	−14.82%
1927	37.49	3.12	34.37	1967	23.98	4.21	19.77
1928	43.61	3.56	40.05	1968	11.06	5.21	5.85
1929	−8.42	4.75	−13.17	1969	−8.50	6.58	−15.08
1930	−24.90	2.41	−27.31	1970	4.01	6.52	−2.51
1931	−43.34	1.07	−44.41	1971	14.31	4.39	9.92
1932	−8.19	0.96	−9.15	1972	18.98	3.84	15.14
1933	53.99	0.30	53.69	1973	−14.66	6.93	−21.59
1934	−1.44	0.16	−1.60	1974	−26.47	8.00	−34.47
1935	47.67	0.17	47.50	1975	37.20	5.80	31.40
1936	33.92	0.18	33.74	1976	23.84	5.08	18.76
1937	−35.03	0.31	−35.34	1977	−7.18	5.12	−12.30
1938	31.12	−0.02	31.14	1978	6.56	7.18	−0.62
1939	−0.41	0.02	−0.43	1979	18.44	10.38	8.06
1940	−9.78	0.00	−9.78	1980	32.42	11.24	21.18
1941	−11.59	0.06	−11.65	1981	−4.91	14.71	−19.62
1942	20.34	0.27	20.07	1982	21.41	10.54	10.87
1943	25.90	0.35	25.55	1983	22.51	8.80	13.71
1944	19.75	0.33	19.42	1984	6.27	9.85	−3.58
1945	36.44	0.33	36.11	1985	32.16	7.72	24.44
1946	−8.07	0.35	−8.42	1986	18.47	6.16	12.31
1947	5.71	0.50	5.21	1987	5.23	5.47	−0.24
1948	5.50	0.81	4.69	1988	16.81	6.35	10.46
1949	18.79	1.10	17.69	1989	31.49	8.37	23.12
1950	31.71	1.20	30.51	1990	−3.17	7.81	−10.98
1951	24.02	1.49	22.53	1991	30.55	5.60	24.95
1952	18.37	1.66	16.71	1992	7.67	3.51	4.16
1953	−0.99	1.82	−2.81	1993	9.99	2.90	7.09
1954	52.62	0.86	51.76	1994	1.31	3.90	−2.59
1955	31.56	1.57	29.99	1995	37.43	5.60	31.83
1956	6.56	2.46	4.10	1996	23.07	5.21	17.86
1957	−10.78	3.14	−13.92	1997	33.36	5.26	28.10
1958	43.36	1.54	41.82	1998	28.58	4.86	23.72
1959	11.96	2.95	9.01	1999	21.04	4.68	16.36
1960	0.47	2.66	−2.19	2000	−9.11	5.89	−15.00
1961	26.89	2.13	24.76	2001	−11.88	3.83	−15.71
1962	−8.73	2.73	−11.46	2002	−22.10	1.65	−23.75
1963	22.80	3.12	19.68				
1964	16.48	3.54	12.94	Mean	**12.20**	**3.83**	**8.37**
1965	12.45	3.93	8.52	Std dev	**20.49**	**3.15**	**20.78**

Source: Authors' calculations using Ibbotson Associates (2003a, pp. 38–39).

Appendix B
Compilation of Equity Risk Premium Estimates

Source	Risk-free-rate	ERP Estimate	Real Risk-Free Rate	Nominal Risk-Free Rate	Geometric	Arithmetic	Long Horizon	Short Horizon	Short-Run Expectation	Short-Run Expectation	Conditional	Unconditional
Historical												
Ibbotson Associates	3.8%[7]	8.4%[31]		X		X		X		X		X
Social Security												
Office of the Chief Actuary[1]	2.3%, 3%[8]	4.7%, 4%[32]	X		X		X			X		X
John Campbell[2]	3–3.5%[9]	1.5–2.5%, 3–4%[33]	X		X	X	X	X		X	X	
Peter Diamond	2.2%[10]	<4.8%[34]	X		X		X			X	X	
Peter Diamond[3]	3%[11]	3–3.5%[35]	X		X		X			X	X	
John Shoven[4]	3%, 3.5%[12]	3–3.5%[36]	X		X		X			X	X	
Puzzle Research												
Robert Arnott and Peter Bernstein	3.7%[13]	2.4%[37]	X		X		X			X	X	
Robert Arnott and Ronald Ryan	4.1%[14]	−0.9%[38]	X		X		X			X	X	
John Campbell and Robert Shiller	N/A	Negative[39]	X		?		?		X		X	
James Claus and Jacob Thomas	7.64%[15]	3.39% or less[40]		X		X	X			X	X	
George Constantinides	2%[16]	6.9%[41]	X			X		X		X		X
Bradford Cornell	5.6%, 3.8%[17]	3.5–5.5%, 5–7%[42]		X		X	X	X		X	X	
Dimson, Marsh, & Staunton	1.0%[18]	5.4%[43]	X			X		X		X	X	
Eugene Fama and Kenneth French	3.24%[19]	3.83% & 4.78%[44]	X			X		X		X		X
Robert Harris and Felicia Marston	8.53%[20]	7.14%[45]		X		X	X		X		X	
Roger Ibbotson and Peng Chen	2.05%[21]	4% and 6%[46]	X		X	X	X			X		X
Jeremy Siegel	4%[22]	−0.9% to −0.3%[47]	X		X		X			X	X	
Jeremy Siegel	3.5%[23]	2–3%[48]	X		X		X			?	X	
Surveys												
John Graham and Campbell Harvey	Δ by survey[24]	3–4.7%[49]		X		?	X		X		X	
Ivo Welch	N/A[25]	7%[50]		X		X		X		X	X	
Ivo Welch[5]	5%[26]	5–5.5%[51]		X		X		X		X	X	
Misc.												
Barclays Global Investors	5%[27]	2.5%, 3.25%[52]		X	X		X		X		X	
Richard Brealey and Stewart Myers	N/A[28]	6–8.5%[53]		X		X		X		X		X
Burton Malkiel	5.25%[29]	2.75%[54]		X	X		X			X	X	
Richard Wendt[6]	5.5%[30]	3.3%[55]		X		X	X			X	X	

Notes: Long-run expectation considered to be a forecast of more than 10 years. Short-run expectation considered to be a forecast of 10 years or less.

Footnotes:
[1]Social Security Administration.
[2]Presented to the Social Security Advisory Board.
[3]Presented to the Social Security Advisory Board. Update of 1999 article.
[4]Presented to the Social Security Advisory Board.
[5]Update to Welch (2000).
[6]Newsletter of the Investment Section of the Society of Actuaries.
[7]Arithmetic mean of U.S. Treasury bills annual total returns from 1926–2002.
[8]2.3% long-run real yield on Treasury bonds, used for Advisory Council proposals; 3% long-term real yield on Treasury bonds; used in 1999 Social Security Trustees Report.
[9]Estimate for safe real-interest rates in the future based on yield of long-term inflation-indexed Treasury securities of 3.5% and short-term real-interest rates recently averaging about 3%.
[10]Real long-term bond yield using 75-year historical average.
[11]Real yield on long-term Treasuries (assumption by OCACT).
[12]3% is the OCACT assumption; 3.5% is the real return on long-run (30-year) inflation-indexed Treasury securities.
[13]Long-term expected real geometric bond return (10-year-horizon).
[14]The yield on U.S. government inflation-indexed bonds (starting bond real yield in January 2000).
[15]Average 10-year government T-bond yield between 1985 and 1998 (yield of 11.43% in 1985 to 5.64% in 1998). The mean 30-year risk-free rate for each year of the U.S. sample period is 31 basis points higher than the mean 10-year risk-free rate.
[16]Rolled-over real arithmetic return of three-month Treasury bills and certificates.
[17]Historical 20-year Treasury bond return of 5.6%. Yield on 20-year Treasury bonds in 1998 was approximately 6%. Historical one-month Treasury bill return of 3.8%. Yield on one-month Treasury bills in 1998 was approximately 4%.
[18]U.S. historical arithmetic real Treasury bill return over 1900–2000 period. 0.9% geometric Treasury bill return.

Appendix B (continued)

Source	Risk-free Rate	ERP Estimate	Data Period	Methodology
Historical				
Ibbotson Associates	3.8%[7]	8.4%[31]	1926–2002	Historical
Social Security				
Office of the Chief Actuary[1]	2.3%, 3%[8]	4.7%, 4%[32]	1900–1995, Projecting out 75 years	Historical
John Campbell[2]	3–3.5%[9]	1.5–2.5%, 3–4%[33]	Projecting out 75 years	Historical & ratios (div/price & earn gr)
Peter Diamond	2.2%[10]	<4.8%[34]	Last 200 yrs for eq/75 for bonds, Proj 75 yrs	Fundamentals: div yld, GDP gr
Peter Diamond[3]	3%[11]	3–3.5%[35]	Projecting out 75 years	Fundamentals: div/price
John Shoven[4]	3%, 3.5%[12]	3–3.5%[36]	Projecting out 75 years	Fundamentals: P/E, GDP gr
Puzzle Research				
Robert Arnott and Peter Bernstein	3.7%[13]	2.4%[37]	1802 to 2001, normal	Fundamentals: div yld & gr
Robert Arnott and Ronald Ryan	4.1%[14]	−0.9%[38]	Past 74 years, 74 year projection[56]	Fundamentals: div yld & gr
John Campbell and Robert Shiller	N/A	Negative[39]	1871 to 2000, 10-year projection	Ratios: P/E and div/price
James Claus and Jacob Thomas	7.64%[15]	3.39% or less[40]	1985–1998, long-term	Abnormal earnings model
George Constantinides	2%[16]	6.9%[41]	1872 to 2000, long-term	Hist. and Fund.: price/div & P/E
Bradford Cornell	5.6%, 3.8%[17]	3.5–5.5%, 5–7%[42]	1926–1997, long-run forward-looking	Weighing theoretical and empirical evidence
Dimson, Marsh, & Staunton	1%[18]	5.4%[43]	1900–2000, prospective	Adj hist ret, var of Gordon gr model
Eugene Fama and Kenneth French	3.24%[19]	3.83% & 4.78%[44]	Estimate for 1951–2000, long-term	Fundamentals: dividends and earnings
Robert Harris and Felicia Marston	8.53%[20]	7.14%[45]	1982–1998, expectational	Fin analysts' est, div gr model
Roger Ibbotson and Peng Chen	2.05%[21]	4% and 6%[46]	1926–2000, long-term	Historical and supply side approaches
Jeremy Siegel	4%[22]	−0.9% to −0.3%[47]	1871 to 1998, forward-looking	Fundamentals: P/E, div yld, div gr
Jeremy Siegel	3.5%[23]	2–3%[48]	1802–2001, forward-looking	Earnings yield
Surveys				
John Graham and Campbell Harvey	Δ by survey[24]	3–4.7%[49]	2Q 2000 through 3Q 2002, 1 & 10-year projections	Survey of CFO's
Ivo Welch	N/A[25]	7%[50]	30-year forecast, surveys in 97/98 & 99	Survey of financial economists
Ivo Welch[5]	5%[26]	5–5.5%[51]	30-year forecast, survey around August 2001	Survey of financial economists
Misc.				
Barclays Global Investors	5%[27]	2.5%, 3.25%[52]	Long-run (10-year) expected return	Fundamentals: inc, earn gr, & repricing
Richard Brealey and Stewart Myers	N/A[28]	6–8.5%[53]	1926–1997	Predominantly historical
Burton Malkiel	5.25%[29]	2.75%[54]	1926 to 1997, estimate millennium[57]	Fundamentals: div yld, earn gr
Richard Wendt[6]	5.5%[30]	3.3%[55]	1960–2000, estimate for 2001–2015 period	Linear regression model

Footnotes:

[19]Average real return on six-month commercial paper (proxy for risk-free interest rate). Substituting the one-month Treasury bill rate for the six-month commercial paper rate causes estimates of the annual equity premium for 1951–2000 to rise by about 1%.

[20]Average yield to maturity on long-term U.S. government bonds, 1982–1998.

[21]Real, geometric risk-free rate. Geometric risk-free rate with inflation (nominal) 5.13%. Nominal yield equivalent to historical geometric long-term government bond income return for 1926–2000.

[22]The 10- and 30-year TIPS bond yielded 4% in August 1999.

[23]Return on inflation-indexed securities.

[24]Current 10-year Treasury bond yield. Survey administered from June 6, 2000 to June 4, 2002. The rate on the 10-year Treasury bond changes in each survey. For example, in the Dec. 1, 2000 survey, the current annual yield on the 10-year Treasury bond was 5.5%. For the June 6, 2001 survey, the current annual yield on the 10-year Treasury bond was 5.3%.

[25]Arithmetic per-annum average return on rolled-over 30-day T-bills.

[26]Average forecast of arithmetic risk-free rate of about 5% by deducting ERP from market return.

[27]Current nominal 10-year bond yield.

[28]Return on Treasury bills. Treasury bills yield of about 5% in mid-1998. Average historical return on Treasury bills 3.8%.

[29]Good quality corporate bonds will earn approximately 6.5–7%. Long-term zero-coupon Treasury bonds will earn about 5.25%. Long-term TIPS will earn a real return of 3.75%.

[30]1/1/01 long T-bond yield; uses initial bond yields in predictive model.

[31]Arithmetic short-horizon expected ERP. Arithmetic intermediate-horizon expected ERP 7.4%. Arithmetic long-horizon expected ERP 7.0%. Geometric short-horizon expected ERP 6.4%.

[32]Geometric equity premium over long-term Treasury securities. OCACT assumes a constant geometric real 7% stock return.

Appendix C

Estimating a Short-Horizon Arithmetic Unconditional Equity Risk Premium

Source	Risk-free Rate	ERP Estimate	Stock Return Estimate	Geometric to Arithmetic	Real to Nominal	Conditional to Unconditional[60]	Fixed Short-Horizon RFR	Short-Horizon Arithmetic Unconditional ERP Estimate
	I	II	III	IV	V	VI	VII	VIII
Historical								
Ibbotson Associates	3.8%[7]	8.4%[31]	12.2%	0.0%	0.0%	0.00%	3.8%	8.4%
Social Security								
Office of the Chief Actuary[1]	2.3%, 3%[8]	4.7%, 4%[32]	7.0%	2%	3.1%	0.00%	3.8%	8.3%
John Campbell[2]	3–3.5%[9]	1.5–2.5%, 3–4%[33]	6.0%–7.5%	0.0%	3.1%	0.46%	3.8%	5.8%–7.3%
Peter Diamond	2.2%[10]	<4.8%[34]	<7.0%	2%	3.1%	0.46%	3.8%	<8.8%
Peter Diamond[3]	3%[11]	3–3.5%[35]	6.0–6.5%	2%	3.1%	0.46%	3.8%	7.8%–8.3%
John Shoven[4]	3%, 3.5%[12]	3–3.5%[36]	6.0–7.0%	2%	3.1%	0.46%	3.8%	7.8%–8.8%
Puzzle Research								
Robert Arnott and Peter Bernstein	3.7%[13]	2.4%[37]	6.1%	2%	3.1%	0.46%	3.8%	7.9%
Robert Arnott and Ronald Ryan	4.1%[14]	−0.9%[38]	3.2%	2%	3.1%	0.46%	3.8%	5%
John Campbell and Robert Shiller	N/A	Negative[39]	Negative	N/A	N/A	N/A	N/A	N/A
James Claus and Jacob Thomas	7.64%[15]	3.39% or less[40]	11.03%	0.0%	0.0%	0.46%	3.8%	7.69%
George Constantinides	2%[16]	6.9%[41]	8.9%	0.0%	3.1%	0.00%	3.8%	8.2%
Bradford Cornell	5.6%, 3.8%[17]	3.5–5.5%, 5–7%[42]	8.8–10.8%	0.0%	0.0%	0.46%	3.8%	5.5–7.5%
Dimson, Marsh, & Staunton	1.0%[18]	5.4%[43]	6.4%[58]	0.0%	3.1%	0.46%	3.8%	6.2%[61]
Eugene Fama and Kenneth French	3.24%[19]	3.83% & 4.78%[44]	7.07–8.02%	0.0%	3.1%	0.00%	3.8%	6.37–7.32%
Robert Harris and Felicia Marston	8.53%[20]	7.14%[45]	12.34%[59]	0.0%	0.0%	0.46%	3.8%	9.00%
Roger Ibbotson and Peng Chen	2.05%[21]	4% and 6%[46]	8.05%	0.0%	3.1%	0.00%	3.8%	7.35%
Jeremy Siegel	4%[22]	−0.9% to −0.3%[47]	3.1–3.7%	2%	3.1%	0.46%	3.8%	4.9–5.5%
Jeremy Siegel	3.5%[23]	2–3%[48]	5.5–6.5%	2%	3.1%	0.46%	3.8%	7.3–8.3%
Surveys								
John Graham and Campbell Harvey	Δ by survey[24]	3–4.7%[49]	8.3–10.2%	N/A	0.0%	0.46%	3.8%	5–6.9%
Ivo Welch	N/A[25]	7%[50]	N/A	0.0%	0.0%	0.46%	0.0%	7.5%
Ivo Welch[5]	5%[26]	5–5.5%[51]	10.0–10.5%	0.0%	0.0%	0.46%	3.8%	6.7–7.2%
Misc.								
Barclays Global Investors	5%[27]	2.5%, 3.25%[52]	7.5%, 8.25%	2%	0.0%	0.46%	3.8%	6.16–6.91%
Richard Brealey and Stewart Myers	N/A[28]	6–8.5%[53]	N/A	0.0%	0.0%	0.00%	0.0%	6.0–8.5%
Burton Malkiel	5.25%[29]	2.75%[54]	8.0%	2%	0.0%	0.46%	3.8%	6.7%
Richard Wendt[6]	5.5%[30]	3.3%[55]	8.8%	0.0%	0.0%	0.46%	3.8%	5.5%

Column formulas: III = I + II; VIII = III + IV + V + VI − VII

Source for adjustments: Ibbotson Associates (2003a, table 2-1 p. 33); Fama and French (2002)—see footnote 60.

Footnotes (continued from Appendix B):

[33]Long-run average equity premium of 1.5–2.5% in geometric terms and 3–4% in arithmetic terms.

[34]Lower return over the next decade, followed by a geometric, real 7% stock return for remaining 65 years or lower rate of return for entire 75-year period (obscures pattern of returns).

[35]Most likely poor return over the next decade followed by a return to historic yields. Working from OCACT stock return assumption, he gives a single rate of return on equities for projection purposes of 6–6.5% (geometric, real).

[36]Geometric real stock return over the geometric real return on long-term government bonds.

[37]Expected geometric return over long-term government bonds. Their current risk premium is approximately zero, and their recommended expectation for the future real return for both stocks and bonds is 2–4%. The "normal" level of the risk premium is modest (2.4% or quite possibly less).

[38]Geometric real returns on stocks are likely to be in the 3–4% range for the foreseeable future (10–20 years).

[39]Substantial declines in real stock prices, and real stock returns below zero, over the next 10 years (2001–2010).

[40]The equity premium for each year between 1985 and 1998 in the United States. Similar results for five other markets.

[41]Unconditional, arithmetic mean aggregate equity premium over the 1872–2000 period. Over the period 1951 to 2000, the adjusted estimate of the unconditional mean premium is 6%. The corresponding estimate over the 1926 to 2000 period is 8%. Sharp distinction between conditional, short-term forecasts of the mean equity return and premium and estimates of the unconditional mean.

[42]Long-run arithmetic future ERP of 3.5–5.5% over Treasury bonds and 5–7% over Treasury bills. Compares estimates to historical returns of 7.4% for bond premium and 9.2% for bill premium.

[43]5.4% U.S. arithmetic expected future ERP relative to bills; 4% World (16 countries) arithmetic expected future ERP relative to bills; 4.1% U.S. geometric expected future ERP relative to bills; 3% World (16 countries) geometric expected future ERP relative to bills.

[44]3.83% unconditional expected annual simple equity premium return (referred to as the annual-bias adjusted estimate of the annual equity premium) using dividend growth model; 4.78% unconditional expected annual simple equity premium return (referred to as the annual-bias adjusted estimate of the annual equity premium) using earnings growth model. Compares these results against historical real equity risk premium of 7.43% for 1951–2000.

[45]Average expectational risk premium. Because of the possible bias of analysts' optimism, the estimates are interpreted as "upper bounds" for the market premium. The average expectational risk premium is approximately equal to the arithmetic (7.5%) long-term differential between returns on stocks and long-term government bonds.

Appendix D

Historical and Forecasted Equity Returns, Ibbotson and Chen (2003) Models (%)

Method/ Model	Sum	Inflation	Real Risk - Free Rate	Equity Risk Premium	Real Capital Gain	g (Real EPS)	g (Real Div)	−g (Pay-out Ratio)	g (BV)	g (ROE)	g (P/E)	g (Real GDP/ POP)	g (FS-GDP/ POP)	Income Return	Re-investment + Interaction	Additional Growth	Forecast Earnings Growth
Column #	I	II	III	IV	V	VI	VII	VIII	IX	X	XI	XII	XIII	XIV	XV	XVI	XVII
Historical																	
Method 1	10.70	3.08	2.05	5.24											0.33		
Method 2	10.70	3.08			3.02									4.28	0.32		
Method 3	10.70	3.08				1.75					1.25			4.28	0.34		
Method 4	10.70	3.08					1.23	0.51			1.25			4.28	0.35		
Method 5	10.70	3.08							1.46	0.31	1.25			4.28	0.31		
Method 6	10.70	3.08										2.04	0.96	4.28	0.32		
Forecast with Historical Dividend Yield																	
Model 3F	9.37	3.08				1.75								4.28	0.26		
Model 3F (ERP)	9.37	3.08	2.05	3.97											0.27		
Forecast with Current Dividend Yield																	
Model 4F	5.44	3.08					1.23							1.10[a]	0.03		
Model 4F (ERP)	5.44	3.08	2.05	0.24											0.07		
Model $4F_2$	9.37	3.08					1.23	0.51						2.05[b]	0.21	2.28	
Model $4F_2$ (FG)	9.37	3.08												1.10[a]	0.21		4.98

Source: The data and format was made available by Ibbotson and Chen and is reprinted with permission from Ibbotson Associates. Corresponds to Ibbotson and Chen (2003, table 2); column numbers have been added.
[1]2000 dividend yield.
[2]Assuming the historical average dividend-payout ratio, the 2000 dividend yield is adjusted up 0.95 percentage points.

Footnotes: (continued from Appendix B and C)
[46]4% geometric (real) and 6% arithmetic (real). Forward-looking long-horizon sustainable ERP.
[47]Using the dividend discount model, the forward-looking real long-term geometric return on equity is 3.3%. Based on the earnings yield, the forward-looking real long-term geometric return on equity is between 3.1% and 3.7%.
[48]Future geometric equity premium. Future real return on equities of about 6%.
[49]The 10-year premium. The one-year risk premium averages between 0.4% and 5.2%, depending on the quarter surveyed.
[50]Arithmetic 30-year forecast relative to short-term bills; 10-year same estimate. Second survey 6.8% for 30- and 10-year estimate. One-year horizon between 0.5% and 1.5% lower. Geometric 30-year forecast around 5.2% (50% responded to this question).
[51]Arithmetic 30-year equity premium (relative to short-term T-bills). Geometric about 50 basis points below arithmetic. Arithmetic one-year equity premium 3–3.5%.
[52]2.5% current (conditional) geometric equity risk premium. 3.25% long-run, geometric normal or equilibrium equity risk premium.
[53]Extra arithmetic return versus Treasury bills. "Brealey and Myers have no official position on the exact market risk premium, but we believe a range of 6–8.5% is reasonable for the United States. We are most comfortable with figures towards the upper end of the range."
[54]The projected geometric (nominal) total return for the S&P 500 is 8% per year.
[55]Arithmetic mean 15-year horizon.
[56]74 years since December 1925 and 74 years starting January 2000.
[57]Estimate the early decades of the 21st century.
[58]World estimate of 5%.
[59]Long risk-free of 5.2% plus 7.14%.
[60]For the 1951–2000 period, Fama and French (2002) adjust the conditional dividend growth model estimate upwards by 1.28% for an unconditional estimate, and they make a 0.46% upwards adjustment to the earnings growth model. We select the smaller of the two as an approximate minimum adjustment. For the longer period of 1872–2000, a comparable adjustment would be 0.82% for the dividend growth model and 0.54% for the 1872–1950 period using a dividend growth model. Earnings growth rates are shown by Fama and French only for the 1951–2000 period.
[61]World estimate of 4.8%.

Appendix D (continued)

Explanation of Ibbotson/Chen Table 2 Exhibit Using Column Numbers to Represent Formula

	Formula*	Description of Method
Historical		
Method 1	I = (1 + II)*(1 + III)*(1 + IV) − 1	Building Blocks Method: inflation, real risk-free rate, and ERP.
Method 2	I = [(1 + II)*(1 + V) − 1] + XIV + XV	Capital Gain and Income Method: inflation, real capital gain, and income return.
Method 3	I = [(1 + II)*(1 + VI)*(1 + XI) − 1] + XIV + XV	Earnings Model: inflation, growth in earnings per share, growth in price to earnings ratio, and income return.
Method 4	I = [(1 + II)*(1 + XI)*(1 + VII)/(1 − VIII) − 1] + XIV + XV	Dividends Model: inflation, growth rate of price earnings ratio, growth rate of the dollar amount of dividends after inflation, growth rate of payout ratio, and dividend yield (income return).
Method 5	I = [(1 + II)*(1 + XI)*(1 + IX)*(1 + X) − 1] + XIV + XV	Return on Book Equity Model: inflation, growth rate of price earnings ratio, growth rate of book value, growth rate of ROE, and income return.
Method 6	I = [(1 + II)*(1 + XII)*(1 + XIII) − 1] + XIV + XV	GDP Per Capita Model: inflation, real growth rate of the overall economic productivity (GDP per capita), increase of the equity market relative to the overall economic productivity, and income return.
Forecast with Historical Dividend Yield		
Model 3F	I = [(1 + II)*(1 + VI) − 1] + XIV + XV	Forward-Looking Earnings Model: inflation, growth in real earnings per share, and income return.
Model 3F (ERP)	IV = (1 + I)/[(1 + II)*(1 + III)]− 1	Using Model 3F result to calculate ERP.
Forecast with Current Dividend Yield		
Model 4F	I = [(1 + II)*(1 + VII) − 1] + XIV + XV	Forward-Looking Dividends Model: inflation, growth in real dividend, and dividend yield (income return); also referred to as Gordon model.
Model 4F (ERP)	IV = (1 + I)/[(1 + II)*(1 + III)] − 1	Using Model 4F result to calculate ERP.
Model $4F_2$	I = [(1 + II)*(1 + VII)*(1 + VIII) − 1] + XIV + XV + XVI	Attempt to reconcile Model 4F and Model 3F.
Model $4F_2$ (FG)	XVII = [(1 + I)/(1 + II) − 1] − XIV − XV	Using Method $4F_2$ result to calculate forecasted earnings.

REFERENCES

ARNOTT, ROBERT D., AND PETER L. BERNSTEIN. 2002. "What Risk Premium Is 'Normal'?" *Financial Analysts Journal* 58(2): 64–85.

ARNOTT, ROBERT D., AND RONALD J. RYAN. 2001. "The Death of the Risk Premium: Consequences of the 1990s," *Journal of Portfolio Management* 27(3): 61–74.

BARBERIS, NICHOLAS, MING HUANG, AND TANO SANTOS. 2001. "Prospect Theory and Asset Prices," *Quarterly Journal of Economics* 116(1): 1–53.

BENARTZI, SHLOMO, AND RICHARD H. THALER. 1995. "Myopic Loss Aversion and the Equity Premium Puzzle," *Quarterly Journal of Economics* 110(1): 73–92.

BREALEY, RICHARD A., AND STEWART C. MYERS. 2000. *Principles of Corporate Finance.* 6th ed. Boston: Irwin McGraw-Hill.

BRENNAN, MICHAEL J., AND EDUARDO S. SCHWARTZ. 1985. "On the Geometric Mean Index: A Note," *Journal of Financial and Quantitative Analysis* 20(1): 119–22.

BROWN, STEPHEN J., WILLIAM N. GOETZMANN, ROGER G. IBBOTSON, AND STEPHEN A. ROSS. 1992. "Survivorship Bias in Performance Studies," *Review of Financial Studies* 5(4): 553–80.

BROWN, STEPHEN J., WILLIAM N. GOETZMANN, AND STEPHEN A. ROSS. 1995. "Survival," *Journal of Finance* 50(3): 853–73.

CAMPBELL, JOHN Y. 2001. "Forecasting U. S. Equity Returns in the 21st Century," in *Estimating the Real Rate of Return on Stocks over the Long Term,* presented to Social Security Advisory Board, August. Online at http://www.ssab.gov/estimated%20rate%20of%20return.pdf.

CAMPBELL, JOHN Y., AND JOHN H. COCHRANE. 1999. "By Force of Habit: A Consumption-Based Explanation of Aggregate Stock Market Behavior," *Journal of Political Economics* 107(April): 205–51.

CAMPBELL, JOHN Y., ANDREW W. LO, AND A. CRAIG MACKINLAY. 1997. *The Econometrics of Financial Markets*. Princeton, NJ: Princeton University Press.

CAMPBELL, JOHN Y., AND ROBERT J. SHILLER. 1988. "The Dividend-Price Ratio and Expectations of Future Dividends and Discount Factors," *Review of Financial Studies* 1(Fall): 195–228.

———. 2001. "Valuation Ratios and the Long-Run Stock Market Outlook: An Update." Working paper #8221, National Bureau of Economic Research. Forthcoming in *Advances in*

Behavioral Finance, Vol. II, edited by Nicholas Barberis and Richard Thaler, Russell Sage Foundation, 2003.

CLAUS, JAMES, AND JACOB THOMAS. 2001. "Equity Premia as Low as Three Percent?" *Journal of Finance* 56(5): 1629–66.

COCHRANE, JOHN H. 1997. "Where is the Market Going? Uncertain Facts and Novel Theories," *Economic Perspectives* 21(6): 3–37.

CONSTANTINIDES, GEORGE M., JOHN B. DONALDSON, AND RAJNISH MEHRA. 2002. "Junior Can't Borrow: A New Perspective on the Equity Premium Puzzle," *Quarterly Journal of Economics* 117(1): 269–296.

CONSTANTINIDES, GEORGE M. 2002. "Rational Asset Prices," *Journal of Finance* 57(4): 1567–91.

CORNELL, BRADFORD. 1999. *The Equity Risk Premium: The Long-Run Future of the Stock Market.* New York: John Wiley & Sons.

DIAMOND, PETER A. 1999. "What Stock Market Returns to Expect for the Future?" Center for Retirement at Boston College. No. 2, September.

———. 2001. "What Stock Market Returns to Expect for the Future: An Update," in *Estimating the Real Rate of Return on Stocks over the Long Term,* presented to the Social Security Advisory Board, August. Online at http://www.ssab.gov/estimated%20rate%20of%20return.pdf.

DIMSON, ELROY, PAUL MARSH, AND MIKE STAUNTON. 2002. *Triumph of the Optimists: 101 Years of Global Investment Returns.* Princeton, NJ: Princeton University Press.

ELTON, EDWIN J., AND MARTIN J. GRUBER. 1995. *Modern Portfolio Theory and Investment Analysis.* 5th ed. New York: John Wiley & Sons.

EQUITY RISK PREMIUM FORUM. 2002 AIMR Online Publication (June) available at http://www.aimrpubs.org/ap/issues/v2002n1/toc.html.

FAMA, EUGENE F., AND KENNETH R FRENCH. 1992. "The Cross-Section of Expected Stock Returns," *Journal of Finance* 47(2): 427–65.

———. 2002. "The Equity Premium," *Journal of Finance* 57(2): 637–59.

GOETZMANN, WILLIAM N., ROGER G. IBBOTSON, AND LIANG PENG. 2001. "A New Historical Database for the NYSE 1815 to 1925: Performance and Predictability," *Journal of Financial Markets* 4(1): 1–32.

GOSS, STEPHEN C. 2001. "Equity Yield Assumptions Used by the Office of the Chief Actuary, Social Security Administration to Develop Estimates for Proposals with Trust Fund and/or Individual Account Investments," in *Estimating the Real Rate of Return on Stocks Over the Long Term,* presented to the Social Security Advisory Board, August. Online at http://www.ssab.gov/estimated%20rate%20of%20return.pdf.

GRAHAM, JOHN R., AND CAMPBELL R. HARVEY. 2002. "Expectations of Equity Risk Premia, Volatility and Asymmetry." Working paper, Duke University.

GRINOLD, RICHARD, AND KENNETH KRONER. 2002. "The Equity Risk Premium," *Barclays Global Investors' Investment Insights* 5(3): 3–24.

HARRIS, ROBERT S., AND FELICIA C. MARSTON. 2001. "The Market Risk Premium: Expectational Estimates Using Analysts' Forecasts," *Journal of Applied Finance* 11(1): 6–16.

HARVEY, ANDREW. 1990. *The Econometric Analysis of Time Series.* 2nd ed. Cambridge, MA: MIT Press.

IBBOTSON ASSOCIATES. 2003a. *Stocks, Bonds, Bills, and Inflation Yearbook.* Chicago: Ibbotson Associates.

———. 2003b. *Stocks, Bonds, Bills, and Inflation Yearbook,* Valuation Edition. Chicago: Ibbotson Associates.

IBBOTSON, ROGER G., AND PENG CHEN. 2003. "Long-Run Stock Returns: Participating in the Real Economy," *Financial Analysts Journal* 59(1): 88–98.

KAHNEMAN, DANIEL, AND AMOS TVERSKY. 1979. "Prospect Theory: An Analysis of Decisions Under Risk," *Econometrics* 47(2): 263–91.

KAPLAN, PAUL D., AND JAMES D. PETERSON. 1998. "Full-Information Industry Betas," *Financial Management* 27(2): 85–93.

LACHANCE, MARIE-EVE, AND OLIVIA S. MITCHELL. 2003. "Understanding Individual Account Guarantees." Working paper, WP 2003-035, University of Michigan Retirement Research Center.

LONGSTAFF, FRANCIS A. 2004. "The Flight-to-Liquidity Premium in U.S. Treasury Bond Prices," *Journal of Business* 77(3).

MAMAYSKY, HARRY. 2002. "A Model for Pricing Stocks and Bonds with Default Risk." Working Paper No. 02-13, Yale International Center for Finance.

MALKIEL, BURTON G. 1999. *A Random Walk Down Wall Street.* New York: W.W. Norton & Company.

MEHRA, RAJNISH. 2002. "Equity Premium Puzzle," in *Financial Times Mastering Investment,* by James Pickford. London: Financial Times-Prentice Hall.

MEHRA, RAJNISH, AND EDWARD C. PRESCOTT. 1985. "The Equity Premium: A Puzzle," *Journal of Monetary Economics* 15(March): 145–61.

PRATT, SHANNON, P. 1998. *Cost of Capital Estimation and Applications.* New York: John Wiley & Sons..

RALFE, JOHN, CLIFF SPEED, AND JON PALIN. 2003. "Pensions and Capital Structure: Why Hold Equities in the Pension Fund?" Society of Actuaries Symposium, Vancouver, June.

ROSS, STEPHEN A. 1976. "The Arbitrage Theory of Capital Asset Pricing," *Journal of Economic Theory* 13(3): 341–60.

———. 2002. "Theoretical Foundations: Discussion," Equity Risk Premium Forum. AIMR Online Publication (June). Available at http://www.aimrpubs.org/ap/issues/v2002n1/toc.html.

SCHWERT, G. WILLIAM. 1990. "Indexes of United States Stock Prices from 1802 to 1987," *Journal of Business* 63(July): 399–426.

SHILLER, ROBERT J. 1989. *Market Volatility.* Cambridge, MA: MIT Press.

SHOVEN, JOHN B. 2001. "What Are Reasonable Long-Run Rates of Return to Expect on Equities?" *Estimating the Real Rate of Return on Stocks over the Long Term,* presented to the Social Security Advisory Board, August. Online at http://www.ssab.gov/estimated%20rate%20of%20return.pdf.

SIEGEL, JEREMY J. 1999. "The Shrinking Equity Premium," *Journal of Portfolio Management* 26(1): 10–17.

———. 2002. *Stocks for the Long Run: The Definitive Guide to*

Financial Market Returns and Long-Term Investment Strategies. 3rd ed. New York: McGraw-Hill.

SOCIAL SECURITY ADVISORY BOARD. 2002. "Fiscal Year Annual Report." Online at http://www.ssab.gov/annualreport2002.pdf.

SOCIAL SECURITY TRUSTEES REPORT. 1999. Annual Report of the Board of Trustees of the Federal Old-Age and Survivors Insurance and Disability Insurance Trust Funds.

TIPS. 2003. "Treasury Bonds, Notes, and Bills." *Wall Street Journal,* 24 February.

TVERSKY, AMOS, AND DANIEL KAHNEMAN. 1991. "Loss Aversion and Riskless Choice: A Reference Dependent Model," *Quarterly Journal of Economics* 106(4): 1039–61.

WEIL, PHILIPPE. 1989. "The Equity Premium Puzzle and the Risk-free Rate Puzzle," *Journal of Monetary Economics* 24(3): 401–21.

WELCH, IVO. 2000. "Views of Financial Economists on the Equity Premium and on Professional Controversies," *Journal of Business* 73(4): 501–37.

———. 2001. "The Equity Premium Consensus Forecast Revisited." Discussion paper No. 1325, Cowles Foundation.

WENDT, RICHARD. 2002. "Understanding Equity Risk Premium," *Risk and Rewards* 38(February): 1–10.

WILSON, JACK W., AND CHARLES 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): 505–33.

Discussions on this paper can be submitted until July 1, 2004. The authors reserve the right to reply to any discussion. Please see the Submission Guidelines for Authors on the inside back cover for instructions on the submission of discussions.

"Equity Risk Premium: Expectations Great and Small," Richard A. Derrig and Elisha D. Orr, January 2004

SHANE F. WHELAN*

Derrig and Orr provide a comprehensive overview of current estimates of the equity risk premium (ERP), carefully distinguishing between the many different definitions in common use that lead to much confusion. Their survey concentrated almost entirely on the U.S. markets, with the statistical analysis based primarily on the 77 annual returns over the period 1926–2002 (inclusive) given in Ibbotson Associates (2003), with occasional reference to a longer series of annual returns from 1871 to 2002 in Wilson and Jones (2002) or data over the years 1802–2001 in Siegel (2002). In this discussion I address two issues:

1. I draw attention to another strand of research in this area that demonstrates that returns from capital markets are not a stationary series. If returns are nonstationary, then this undermines the direct use of simple historical averages or estimating the future ERP based on projections from stationary models fitted to the data. As the approach outlined by the authors as well as many of those surveyed assumed returns are stationary, this is a particularly devastating critique.
2. I explore a little further the alternative way of viewing the historical market returns suggested by the authors (Section 6), when the U.S. experience is treated as just one realized path of the grand stochastic process that is the capital markets. The past performance of other national capital markets traces other paths, which, though perhaps neither independent nor equally likely, can be used to shed light on the process of asset price formation and the evolving market price of risk. To provide added contrast to the Derrig and Orr study, I treat the experience of the smallest national market with a history as long as the U.S. market: the Irish capital markets. The Irish experience reinforces the earlier remarks on the nonstationarity of the ERP.

* Shane F. Whelan, FSA, FFA, FSAI, PhD, is Head of the Department of Statistics & Actuarial Science, University College Dublin, Belfield, Dublin 4, Ireland, e-mail: shane.whelan@ucd.ie.

IS THE PATH TRACED BY THE U.S. EQUITY RISK PREMIUM WEAKLY STATIONARY?

Derrig and Orr do provide some tests for the ERP being weakly stationary (Sections 6–9) but fail, in my opinion, to interpret them correctly. First, they report that equality of the sample variances over two subperiods can be rejected at the 1% significance level under a standard F-test (footnote 16). This is evidence, insofar as the normality assumption under the F-test is tenable, that the annual ERP does not form a stationary series and, in particular, cannot adequately be modeled as independent and identically distributed as suggested (Sections 6 and 9). Second, the t-test they employ to test equality of means in Table 6 (or more strictly, that the mean of the subperiod 1960–2002 equals the mean of the total period 1926–2002) is questionable in light of the reported difference of variances. However, even if the variances were equal and best estimated with just the 1960–2002 data, the test they employ has such low power that it could not reject the null at the 5% critical level if the true ERP in the 1926–1959 fell anywhere in the range –5.7% to +16.3%.[1] As this range encompasses all reasonable values for the ERP, the failure to reject the null of constancy of the ERP is really saying more

[1] The t-test in Table 6 of the paper is that the mean of the second subperiod equals that of the total period. The test statistic can therefore be decomposed into the mean of the second subperiod weighed by the number of data points plus the mean of the first subperiod weighed by its number of data points, i.e., $\frac{5.27\% - (43 \times 5.27\% + 31\bar{x}_{1926-1959})/77}{15.83\%/\sqrt{43}}$, which, as applied, is assumed to follow a t-distribution with 42 degrees of freedom. Using critical values of the statistic, it is straightforward to solve for the range of values that $\bar{x}_{1926-1959}$ can take without rejecting the null. The result is somewhat larger than their reported confidence intervals.

about the paucity of data than about the structure of the data.

If we move from annual return to monthly return data, then the data set increases 12-fold, and statistical testing can be more discriminating between alternative hypotheses. I am unaware of studies testing properties of the monthly ERP, but a considerable literature exists on the properties of monthly returns from both bonds and equities. Loretan and Phillips (1994) is a particularly pertinent study as it demonstrates that U.S. monthly stock returns (from January 1834 to December 1987) are not weakly stationary (even when allowance is made for the well-documented seasonality in means, second moment dependencies, and failure of the fourth moment of the unconditional return distribution). This finding is especially general as it rules out many classes of models popularly used to characterize return data, such as the ARMA suite, ARCH and GARCH processes where the unconditional second moment is constant, and many types of regime-switching models (where the unconditional model found from integrating over all possible regimes is stationary). As the ERP is the difference between volatile stock returns and less volatile cash (or bond) returns, one would expect the ERP series to inherit noncovariance stationarity from the stock return series.

Loretan and Phillips's testing procedure reports that the failure of the ERP's being weakly stationary is due to the nonconstancy of the unconditional variance of the return series, so their finding does not preclude the constancy of the unconditional mean of the returns (and thereby the ERP) over the period. However, if the ERP is a premium for assuming equity risk, and equity risk is measured by the volatility of excess returns,[2] then, on economic grounds, one would predict a higher ERP in those times when the equity return series exhibit higher volatility. Hence, we can infer nonconstancy of the ERP from the nonconstancy of the unconditional variance of the stock returns. Since Loretan and Phillips (1994) a number of papers, using different approaches, have appeared that confirm their finding of the nonstationarity of returns from capital markets, although most such studies are based on daily or higher-frequency returns. See, for instance, Ibrahim (2003) for another direct testing procedure that reports failure of weak stationarity in daily returns of the S&P500, or, more indirectly, the very considerable empirical evidence presented in Plerou et al. (1999a, 1999b), and the supporting evidence in Pagan (1996), based on the monthly returns of the S&P Composite Price Index in the period 1928–87, that the fourth moment of the unconditional return distribution of U.S. stocks and stock indices fails—a finding inconsistent with a weakly stationary series where the fourth moment of the innovations exist.

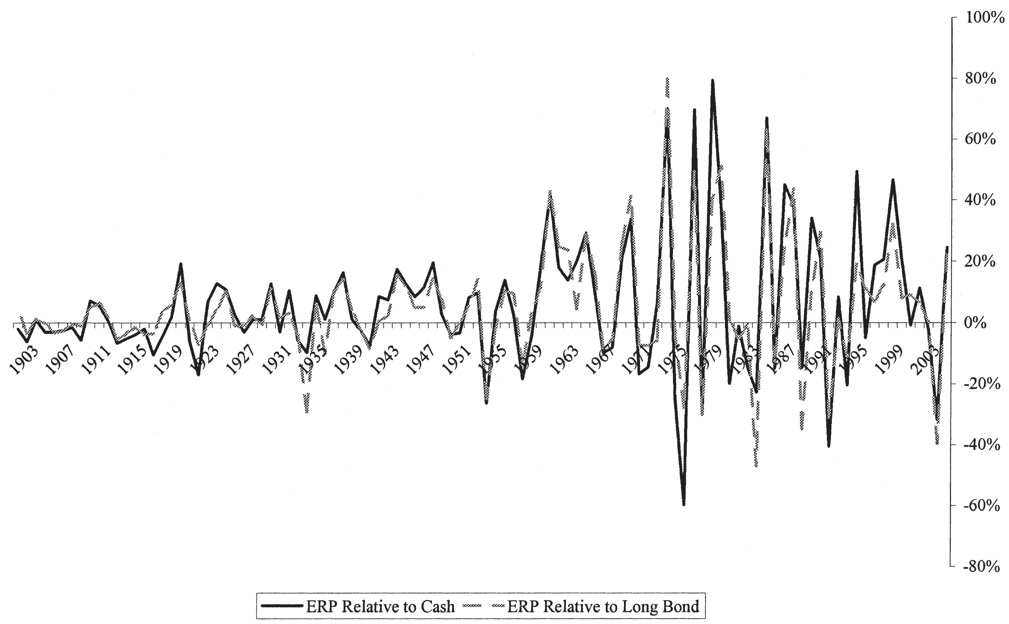
OTHER PATHS

The above considerations point to the conclusion that the path of the ERP in the United States forms a nonstationary series, casting doubt on many of the approaches used to forecast it that Derrig and Orr survey in their paper. Viewing the evolution of the ERP in the United States as just one realized path of a stochastic process as suggested in Section 6, knowledge of the ERP can be augmented by considering other market histories. Consider, for instance, the Irish capital markets, which, though small, have a history of continuous trading as long as that of the U.S. markets.[3] Especially relevant to this discussion is that the path of the ERP in the Irish market reinforces the above remarks on its nonstationarity, as illustrated in Figure 1. Inspection of the figure shows, without the need for formal statistical tests, that the series are obviously nonstationary. The variance of the returns in the latter half of the twentieth cen-

[2] Officer (1973) explored the relationship between the standard deviation and other measures of variability, comparing the rolling 12-month standard deviation of returns with the 12-month mean absolute deviation and the 12-month interpercentile range (from the 28th percentile to the 72nd percentile). He reports a reasonably stable relationship between the estimated variability on each of these measures using monthly data from U.S. market from February 1897 to June 1969. Accordingly, defining and estimating risk by other measures of the spread of the return distribution is unlikely to produce significantly different conclusions.

[3] The Irish equity market has a capitalization less than ½% of that of the U.S. market at the present time. The Dublin Stock Exchange was formally constituted in 1799, making it the sixth oldest surviving national stock market in the world according to Goetzmann and Jorion (1999).

Figure 1
Ex Post Equity Risk Premium on Irish Capital Markets Each Year, 1900–2003



Source: Calculated from the annual returns for each market in Whelan (2004), suitably updated.

tury is clearly significantly higher than that of the first half.

As noted by the authors, Dimson, Marsh, and Staunton (2002) provide the most complete synthesis to date of the twentieth-century experience of national capital markets, recording returns from the cash, bond, and equity markets in 16 countries that, in total, cover about 90% of the current world markets by capitalization. No doubt the path traced by the ERP in each of these markets will reinforce the above remarks. However, the domain of study perhaps can be cast even wider than just the paths traced by low-frequency returns of national markets over the long term. Investigations of the statistical properties of the return paths traced by equity markets have shown that many key properties are invariant with respect to a change in the timescale over which returns are measured (e.g., monthly returns exhibit the same patterns as daily or hourly returns), and markets as diverse as those for commodities, currencies, cash, bonds, and equities display remarkably similar properties. Cont (2001) provides an overview of key empirical regularities of the return paths of financial markets, pointing out, aside from their shared property of nonstationarity, that all returns over any timescale exhibit (a) a heavy-tailed distribution, where the variance exists but the kurtosis (fourth moment) does not, (b) a volatility that tends to cluster in time, and the decay from high bouts of volatility tends to follow a characteristic power law, (c) a negative correlation between the current return and future volatility, decaying to zero in a characteristic pattern as the time lag increases, (d) an asymmetry between large positive and negative movement, with the latter more frequent, and (e) a high correlation between volume traded and volatility. The invariance of these properties with respect to time scaling and between markets strongly suggests that the annual returns delivered by the U.S. markets over the long-term past are no different statistically from, say, hourly returns on the dollar-yen over the last few weeks. Modeling with the latter, however, reduces the problems associated with the paucity of data of the former. It is true that estimation of the ERP is based on the difference between two market returns (the risky and riskless), but parallels can be drawn between the ERP and the minimum enticement for market players (in whatever market) to increase their mismatch risks.

IMPLICATIONS OF NONSTATIONARITY

Nonstationarity of the ERP series and, more generally, returns series from capital markets, tell us that past performance is not a reliable guide to future performance. The riskiness of markets, if measured by the standard deviation of returns or other measures of spread, changes with time not just in temporary bouts (as captured in ARCH-type models), but structurally: the whole background volatility of the markets changes level with time. If the underlying risk is a function of time, then the risk premium must also be a function of time, implying, in turn, that simple averages of the historical ex post ERP must be too. To forecast the ERP by fitting a stationary model is therefore unstable in the sense that changing the time period used to calibrate the model will change the forecast ERP.

We must abandon the stationary assumption of asset returns. One obvious approach to forecasting returns from capital assets is to transform the original return data into a (near) stationary series based on estimates of the unconditional variance at each point in time, forecast the transformed series using standard stationary models, and then apply the inverse transformation to the result to forecast the original returns. Van Bellegem and von Sachs (2004) provide such a development by rescaling time so that the process is "locally" stationary and apply it to forecast daily returns from several markets. Okabe, Matsuura, and Klimek (2002) use another technique to detect the early breakdown of stationarity, claiming that their method can be used to help predict stock-market crashes.[4] Modeling and forecasting allowing for nonstationarity in returns is at an early stage, so it provides, as yet, no reliable guide as to the future evolution of the market price of risk. It does suggest, though, that forecasts from stationary models should be used with circumspection.

Risky assets provide, by definition, an uncertain payoff. Forecasting the equity risk premium must be done in tandem with forecasting the expected course of the riskiness of the asset. But, as this discussion hopes to make clear, the problem is compounded in that we are uncertain of even the riskiness of risky assets.

[4] If the unconditional volatility changes significantly as a quantum jump, then one would expect the future ERP demanded to hold risky assets to also change markedly, which necessitates a considerable change in the price of the assets (which might give economic justification for the stylized fact noted earlier that current return is negatively correlated to future volatility). A glance at Figure 1 shows that the volatility of the Irish equity market increases prior to its major crash in 1974. Hence, the ability to forecast changes in the unconditional volatility could reasonably be expected to help in forecasting significant changes in asset values.

REFERENCES

CONT, RAMA 2001. Empirical Properties of Asset Returns: Stylised Facts and Statistical Issues. *Quantitative Finance* 1: 223–36.

DIMSON, ELROY, PAUL MARSH, AND MIKE STAUNTON. 2002. *Triumph of the Optimists*. Princeton: Princeton University Press.

GOETZMANN, WILLIAM, AND PHILIPPE JORION. 1999. Re-emerging Markets. *Journal of Financial and Quantitative Analysis* 34: 1–32.

IBBOTSON ASSOCIATES. 2003. *Stocks, Bonds, Bills and Inflation Yearbook*. Chicago: Ibbotson Associates.

IBRAHIM, AHAMADA 2003. Non-stationarity Characteristics of the S&P Returns: An Approach Based on the Evolutionary Spectral Density. *Economics Bulletin* 3(32): 1–7.

LORETAN, MICO, AND PETER PHILLIPS. 1994. Testing Covariance Stationarity of Heavy-Tailed Time Series. *Journal of Empirical Finance* 1: 211–48.

OFFICER, R. R. 1973. The Variability of the Market Factor of the New York Stock Exchange. *Journal of Business* 46(3): 434–53.

OKABE, YASUNONRI, MASAYA MATSUURA, AND MACIEJ KLIMEK. 2002. On a Method for Detecting Certain Signs of Stock Market Crashes by Non-linear Stationary Tests. *International Journal of Pure and Applied Mathematics* 3(4): 443–84.

PAGAN, ADRIAN 1996. The Econometrics of Financial Markets. *Journal of Empirical Finance* 3: 15–102.

PLEROU, VASILIKI, ET AL. 1999a. Scaling of the Distribution of Fluctuations of Financial Market Indices. *Physical Review E* 60: 5305–16.

———. 1999b. Scaling of the Distribution of Price Fluctuations of Individual Companies. *Physical Review E* 60: 6519–29.

SIEGEL, JEREMY. J. 2002. *Stocks for the Long Run: The Definitive Guide to Financial Market Returns and Long-Term Investment Strategies*. 3rd ed. New York: McGraw Hill.

VAN BELLEGEM, SEBASTIEN, AND RAINER VON SACHS. 2004. Forecasting

Economic Time Series with Unconditional Time-Varying Variance. *International Journal of Forecasting* 20: 611-27.

WHELAN, SHANE 2004. Measuring Investment Risk in Pension Funds. Paper presented to the Society of Actuaries in Ireland, February 24, 2004. Online at www.actuaries.ie.

WILSON, JACK W., AND CHARLES P. JONES. 2002. An Analysis of the S&P500 Index and Cowles Extensions: Price Indices and Stock Returns, 1870–1999. *Journal of Business* 75(3): 505–33.

AUTHORS' REPLY

We welcome Whelan's discussion for highlighting the role that stationarity plays in discussing the theoretical formation of an equity risk premium (ERP) and for the introduction of data from the Irish capital markets. Our reply consists of three points. First, we do not view the assertion of the absence of weak stationarity as a "devastating critique," as Whelan concludes. Second, the principal thrust of our paper was the wide definitional disparity among the many studies of the ERP puzzle for the U.S. market and their subsequent expectations for the future, most of which avoided or ignored the question of stationarity in any form. Third, our recommendation to practicing actuaries was to use the Ibbotson-Chen building-block method to forecast the ERP, a tool that could be applied equally well to the Irish and other equity markets and that does not depend on stationarity but does depend on replication of the historical mean for each block, absent a rationale *and an estimate* for a change in the block value.

STATIONARITY

Strictly speaking, a time series is *stationary* if all of the statistical properties remain unchanged when the period of observation is shifted forward or backward, or equivalently, if the distribution functions of all consecutive subseries are independent of time (Kruskal and Tanur 1978, pp. 1168–69; Kendall and Stuart 1976, p. 424). Thus the mean, variance, and all other existing moments will remain the same when the period of observation is shortened or lengthened in a stationary series. *Weakly stationary* generally means that only the first two moments, mean and variance, need to be equal. The ERP puzzle literature we reviewed relates only to the expected mean and only incidentally to the other moments. Whelan discusses our test for stationarity of the Ibbotson 1926–2002 series (Derrig and Orr 2004, pp. 51–52), where we informally define stationarity as a mean value unchanging with time (Kendall and Stuart [1976, p. 424] define a separate "stationary in the mean" as the "customary" definition of stationarity of stochastic processes), in line with the ERP puzzle, and test for equal means for the entire series and the 1960–2002 subperiod of the Ibbotson annual data. We find that the *t*-test supports equal means whether or not the variances are assumed equal or not and that there is also some support for unequal variances: that is, the entire Ibbotson series is not *weakly* stationary. This result is due to the large volatility of the depression years of the 1930s (41.6% versus less than 20% for later decades; see Ibbotson Associates 2004, Yearbook, Table 6-1, p. 110), much as the latter years of the Irish market data appear to be more volatile. (Ibbotson Associates' [2004, Valuation Edition, pp. 85–86] graphic shows the large pre–World War II volatility similar to Whelan's post-1970s Irish market.) Table 1 indicates that beginning the Ibbotson series in 1943 (60 years) would give us an annual ERP series with about the same subperiod (30 years) means but equal variances.

Whelan cites studies of monthly return data (not ERPs) that show that U.S. and Irish equity returns are not weakly stationary. But in an equity premium world such as CAPM, one would not expect the total return series to be stationary given the history of wide-ranging nominal and real risk-free rates. That is precisely why the ERP, rather than total returns, is of prime interest. (The modeling of the risk-free rate series has fared no better than that of ERPs, leading to a *risk-free rate puzzle.*) His Figure 1 shows graphically an Irish Capital Market ERP series for 1900–2003 with a changing variance in later years,

Table 1
Equity Risk Premium Variability

Ibbotson Annual Data		
Data Period	Mean	Variance
1926–1959	0.1182	0.0600
1960–2002	0.0527	0.0250
1926–2002	0.0817	0.0410
1943–1972	0.1186	0.0279
1973–2002	0.0527	0.0292
1943–2002	0.0856	0.0292

ruling out weak stationarity but not stationarity in the mean, the object of our paper. Indeed, Finnerty and Leistikow (1993) tested for trend and mean reversion in the Ibbotson ERP series and concluded that the ERP series trended downward over time and, therefore, was not stationary in the mean. That conclusion was later refuted by Ibbotson and Lummer (1994) with the original authors "admit[ing] that the alleged decline is not statistically significant" in a reply (Finnerty and Leistikow 1994). The working assumption of stationarity in the mean was reasserted by Ibbotson and Lummer (1994, p. 99) and continues in the current Ibbotson Associates yearbook (2004, Valuation Edition, p. 75). Additionally, long-run annual ERPs are more valuable to actuaries, as opposed to investment traders, precisely because the annual results smooth the monthly and daily results. Actuarial models of equity and other returns have been built for "scenarios" to be used in sensitivity testing for solvency, pricing, and other actuarial problems. They also adopt practical ERP assumptions, similar in rigor to our stationarity of the mean, to reach meaningful models for practical use such as testing the reasonability of assumptions about the future (see, e.g., Wilkie 1995; Ahlgrim et al. 2003).

ERP Studies

Our review of ERP estimates covered a wide variety of techniques (see Derrig and Orr 2004, Appendix B, for a listing of 25 studies and their methodologies reviewed), most of which were concerned with forecasting a long-run average ERP for some future period as long as 75 years. The majority of the puzzle research studies employed analyses about dividend or earnings series relative to price over different time periods to support various theories about the relation of the future market ERP to the past. Other studies reported surveys of academics and educated guesses by professionals and managers, many concentrating on the next 10 years (2000–2009) and not the long-run average. Whelan asserts that nonstationarity of the U.S. ERP series is "casting doubt on many of the approaches used to forecast it surveyed in the paper." But stationarity was only an implicit consideration for these analyses of the so-called market fundamentals to reach conclusions that the ex ante ERP is *not* equal to the historical average, implying nonstationarity (see Derrig and Orr 2004, Appendix B ERP Estimates, and Appendix C, ERP Estimates adjusted to a common definition). Like many of these studies, Whalen makes a verbal argument about the equity risk. He asserts that risk is measured by the "high volatility of excess returns" and, hence, higher ERP should correspond with higher volatility as in the U.S. depression or the post-1970s Irish data and nonstationarity follows. Absent a theory of ground-up overall returns, this statement is only a plausible working assumption like many of the others. It would be helpful to show whether it conforms to the data. It would support, however, stationarity in the mean for the 1943–2002 series in Table 1.

The Derrig-Orr Recommendation

We recommend for a best estimation methodology for an ex ante ERP the Ibbotson-Chen six building-block methods as described in the paper and laid out in detail in Appendix D. These methods are related to stationarity but do not depend upon it. They specifically allow for the importation of changes in the historical means of the building blocks, such as inflation, growth in earnings, and reinvestment rates, as those changes can be supported or as the judgment of the practicing actuary wills it. Finally, a simple examination of the ERP numerical series for stationarity or any other property would be a misreading of the message of our paper. Rather, one must go beyond the simple numerical values, as those who created the ERP puzzle did, to attempt to understand the process generating the values, including the behavior of investors.

References

AHLGRIM, KEVIN C., STEPHEN P. D'ARCY, AND RICHARD W. GORVETT. 2003. Report on Modeling of Economic Series Coordinated with Interest Rate Scenarios. Research Sponsored by the Casualty Actuarial Society and the Society of Actuaries. www.casact.org/research/econ.

DERRIG, RICHARD A., AND ELISHA D. ORR. 2004. Equity Risk Premium: Expectations Great and Small. *North American Actuarial Journal* 8(1): 45–69.

FINNERTY, JOHN D., AND DEAN LEISTIKOW. 1993. The Behavior of Equity and Debt Risk Premiums. *Journal of Portfolio Management* (summer): 73–84.

———. 1994. "The Behavior of Equity and Debt Risk Premiums": Reply to Comment. *Journal of Portfolio Management* (summer): 101–2.

IBBOTSON ASSOCIATES. 2004. *Stocks, Bonds, Bills, and Inflation*. Yearbook. Chicago: Ibbotson Associates.

——. 2004. *Stocks, Bonds, Bills, and Inflation*. Valuation Edition. Chicago: Ibbotson Associates.

IBBOTSON, ROGER G., AND SCOTT L. LUMMER. 1994. The Behavior of Equity and Debt Risk Premiums: Comment. *Journal of Portfolio Management* (summer): 98–100.

KENDALL, SIR MAURICE, AND ALAN STUART. 1976. *The Advanced Theory of Statistics: Volume 3, Design and Analysis, and Time-Series*. 3rd ed. New York: Hafner Press.

KRUSKAL, WILLIAM H., AND JUDITH M. TANUR, EDS. 1978. *International Encyclopedia of Statistics*. New York: Free Press.

WILKIE, A. D. 1995. More on a Stochastic Model for Actuarial Use. *British Actuarial Journal* 777–964.

"Further Analysis of Future Canadian Health Care Costs," Robert L. Brown and Uma Suresh, April 2004

BEDA CHAN*

THREE-DIMENSIONAL GRADUATION IN POPULATION PROJECTIONS

This excellent paper treats the important topic of health care cost containment in this half century of demographic unfolding in many countries, using Canada as an example.

I view the Lubitz-Scitovsky survivor/decedent costs-split projection (Projection II) as the principal projection. The Denton and Spencer age/sex-specific costs-times-population-pyramid projection (Projection I) is the high variant, since publicly funded health care, when supported, is likely to be utilized as much as it is supported. The advance directive projection (Projection III) is the low variant. The paper by Brown and Suresh, summarized in their Figure 4, is thus a study of a high-principal–low-cost-control environment under a single population projection, which is the best estimate projection for 2001–75 in OSFI's eighteenth actuarial report of 2001. To study the relative importance of population aging scenarios (High Dependency, Best Estimate, Low Dependency of OSFI) versus cost control measures (DS, LS, AD of Brown and Suresh), one can study the 3×3 projections of the high-principal–low-cost-control environment crossed with the High-Dependency–Best-Estimate–Low-Dependency population scenarios.

As variability of health care cost is piggybacked on the variability of population projections, studies on variability of population projections are pertinent to health care cost analyses. In this discussion I would like to offer another high–best estimate–low-population projection over 75 years, in a case where official statistics do not provide high-best-low variants. The case in point is the Hong Kong Special Administrative Region. After 1997 Hong Kong (C&SD 2000, 2002, 2004) caught up with the United Kingdom (ONS 2004, and back every second year for about 40 years) in issuing biannual population projections. The Hong Kong projections, however, do not provide variants in fertility, migration, and mortality and are single-scenario best estimates down 30 years of projections. To keep the discussion brief, I summarize my points in Figure 1.

When the set encompassing high-dependency variant, best estimate, and low-dependency variant is not given in a projection, the later revisions of the projection can be used to construct a high-principal–low-projection band. One can say that based on the June 2004 population projection, 34% of the Hong Kong population will be over age 65 by the year 2063. The high variant would be 39% (mirror image), and the low variant would be 29% (extension of the May 2002 projection). I trusted and used the 2002 projection because it was based on the 2001 full census. The 2000 projection used a high total fertility assumption (1.6 by year 2029) when the observed value was 1.024 for the calendar year 2000. It has since been declining, reaching 0.941 by calendar year 2003. Some technical details are pertinent. The C&SD projections are 30-year projections, and I extended their projections to 75 years by the component method. Since C&SD's fertility, migration, and mortality assumptions stabilized after 15 years, using their last 15 years in five-year intervals allows for graduation and extrapolation

* Beda Chan, ASA, MAAA, PhD, is an Associate Professor in the Department of Statistics and Actuarial Science, University of Hong Kong, Pokfulam Road, Hong Kong, e-mail:bchan@hku.hk.