THE JOURNAL OF FINANCE • VOL. LIV, NO. 3 • JUNE 1999

Global Stock Markets in the Twentieth Century

PHILIPPE JORION and WILLIAM N. GOETZMANN*

ABSTRACT

Long-term estimates of expected return on equities are typically derived from U.S. data only. There are reasons to suspect that these estimates are subject to survivorship, as the United States is arguably the most successful capitalist system in the world. We collect a database of capital appreciation indexes for 39 markets going back to the 1920s. For 1921 to 1996, U.S. equities had the highest real return of all countries, at 4.3 percent, versus a median of 0.8 percent for other countries. The high equity premium obtained for U.S. equities appears to be the exception rather than the rule.

IN A NOW-FAMOUS ARTICLE, Mehra and Prescott (1985) argue that standard general equilibrium models cannot explain the size of the risk premium on U.S. equities, which averages about 6 percent over the 1889–1978 period. They show that one would need a very large coefficient of risk aversion, largely in excess of the usual value of two, to generate such a premium. This unsettling result has sparked a flurry of theoretical research that explores alternative preference structures, including dropping the expected utility assumption and introducing habit formation.[1] Such efforts, however, come at the cost of losing the intuition of standard models.[2]

Rather than searching for preference structures that fit historical data, other explanations focus on the limitations of the data. Rietz (1988) proposes a solution to the puzzle that involves infrequently occurring "crashes." Assuming a crash where output falls by 50 (or 25) percent of its value with a

*Jorion is with the University of California at Irvine; Goetzmann is with the Yale School of Management. We thank seminar participants at the University of California at Los Angeles, Carnegie-Mellon, Indiana University, the London Business School, the Stockholm School of Economics, the University of Houston, the University of Michigan, the University of Notre Dame, the University of Southern California, the 1997 European Finance Association meetings, and the 1997 Western Finance Association meetings for useful comments. The referee and the editor, René Stulz, also provided valuable comments. Able research support was provided by Robin Brooks. George Bittlingmayer kindly provided a copy of the German data. This research received financial support from the Institute for Quantitative Research in Finance, for which we are grateful.

[1] See Epstein and Zin (1991) for nonadditive utility functions and Constantinides (1987) for habit formation. Bansal and Coleman (1996) suggest that liquidity services provided by cash partly explain why returns on cash are so low.

[2] Burnside and McCurdy (1992) provide a good review of the equity premium puzzle.

probability of 0.4 percent (or 1.4 percent), Rietz generates ex ante equity premiums consistent with those observed in the United States and risk aversion of five (or ten).

A related argument is advanced in Brown, Goetzmann, and Ross (1995), who claim that survival of the series imparts a bias to ex post returns. They show that an ex ante equity premium of zero can generate a high ex post positive premium by simply conditioning on the market surviving an absorbing lower bound over the course of a century.[3] The implication is that risk aversion cannot be inferred from the empirical analysis of historical data whose observation is conditional on survival. Although the Rietz (1988) argument leads to higher ex ante equity premiums, the survival argument points to biases in ex post premiums.

Unfortunately, these arguments are nearly impossible to sort out based on a century of U.S. equity data. Consider, for instance, a 0.4 percent annual probability of a large crash. We would then expect one crash to occur every 250 years. Even if we observed such a long sample series, our estimate of the crash probability would still be subject to enormous estimation error.

The only solution to this dilemma is to expand the sample by collecting additional cross-sectional data. In this paper, we reconstruct real capital appreciation series for equity markets in 39 countries over much of the twentieth century. We include not only those markets that survived, but also those markets that experienced both temporary and permanent interruptions. We use this new database to estimate the long-term returns to investing in global markets over the twentieth century.

The first part of our analysis treats each market separately. In effect, it takes all stock market histories as draws from one urn. Under these conditions, we show that the process of discarding markets with interruptions creates serious biases in the measurement of expected returns. Such an experiment assumes that all markets have the same statistical characteristics. This framework is valid when markets are segmented due, for instance, to capital controls. The assumption of constraints on such diversification is not unreasonable for the time period under study.

This paper provides the first comprehensive long-run estimates of return on equity capital across a broad range of markets. To date, virtually the only long-run evidence regarding equity rates of return is derived from the United States, for which we have continuous stock price history going back to 1802. We are able to augment the U.S. experience with a wide range of different global equity market histories.

We find striking evidence in support of the survival explanation for the equity risk premium. Over our sample period, the United States has the highest uninterrupted real rate of appreciation, at 4.3 percent annually. For other countries, the median real appreciation rate is approximately 0.8 per-

[3] A similar argument is advanced by Goetzmann and Jorion (1996). They argue that many so-called "emerging markets" are in fact "reemerging markets" as they have longer histories than commonly believed. Few analysts, however, bother to track the histories of markets that have disappeared.

cent. This strongly suggests that estimates of equity premiums obtained solely from the U.S. market are biased upward by survivorship. An alternative line of explanation is that of fundamentally different risk premiums. With segmented markets, risk premiums are determined by local market conditions. Thus differing expected returns could be due to different investor expectations about risk or to different risk aversion.

Beyond its potential value for shedding light on the equity premium puzzle, this global database allows a broad investigation into the behavior of equity markets over the very long run. We have been able to construct monthly real and dollar-valued capital appreciation indices for virtually all the equity markets that existed during the twentieth century. This enables us to examine markets in crisis and to compare the behavior of losing markets to the behavior of winning markets.

In the second part of the study, we construct a world market appreciation index in order to examine the potential experience of a diversified global investor. This allows us to analyze the benefits of international diversification, comparing return and risk measures across the U.S. and the global portfolios. We estimate the return that such an investor would have earned had it been possible to hold the world market from the early 1920s. Even though one could argue that few investors could have held globally diversified portfolios during these turbulent times, this is still an informative experiment as a guide for future investing.

This paper is organized as follows. Section I motivates the search for differences in return on capital. Section II describes the construction of the global market database. Section III compares the performance of global stock markets and discusses biases affecting the construction of a global stock market index. Section IV contains some concluding comments.

I. The Importance of Compound Growth

In September 1626, Pierre Minuit, the Governor of the West India Company, purchased Manhattan Island from the local Indians for the total sum of 60 guilders, or about 24 dollars. At first sight, this seems like the deal of the century.

Yet, slight differences in the time value of money over long horizons can result in vastly different conclusions. If one compounds this payment at a 5 percent rate of interest, it would have grown in 1995 to about 1.6 billion in current dollars, which seems expensive for 31 square miles of undeveloped land. Compounding at 3 percent, however, results in a much lower current price of $1.3 million—a thousandfold difference! This story shows that differences in rates of return on capital can lead to drastically different numbers when compounded over long horizons.

Our estimates of the rate of return on equity capital are typically based on a century of U.S. data, which reveals an equity premium of about 6 percent. As shown in this example, however, small differences in rates of return can have momentous implications over the long run. How much faith can we have in this number?

Not much, given the volatility of stock returns. Consider, for instance, a market that grows at a 6 percent annual rate with a standard deviation of 20 percent. The question is, how many years do we require to establish that growth is positive with statistical confidence? Using the standard t-test at the 5 percent level, we require that the statistic

$$t = \frac{\hat{\mu}}{\hat{\sigma}/\sqrt{N}} = \frac{0.06}{0.20/\sqrt{N}} \tag{1}$$

be greater than two. This requires N to be at least 44 years. In other words, we need approximately half a century of returns to be confident that this 6 percent equity premium is positive. If the expected return is 3 percent instead, we will need more than 178 years of data to establish statistical significance.

Another problem is that we have reasons to suspect that estimates of return on capital from the United States are affected by survival. At the beginning of the century, active stock markets existed in a number of countries, including Russia, France, Germany, Japan, and Argentina. All of these markets have been interrupted for a number of reasons, including political turmoil, war, and hyperinflation. Assuming there was some probability of disruption for the U.S. market, this probability is not reflected in the observed U.S. data. In turn, this will bias our estimates of the equity premium.

As small differences in estimates on equity capital have dramatic implications for long-term growth, we feel it is important to extend our knowledge of equity premiums to a large cross-sectional sample of long-term data.

II. A Global Stock Market Database

The standard data sources on international stock prices are *Morgan Stanley Capital International Perspectives* (MSCIP) for developed markets and the International Finance Corporation (IFC) for emerging markets. Both are relatively recent.

MSCIP started to construct equity indices in January 1970 for a sample of 19 markets from industrial (developed) countries. These indices are built using a uniform methodology and include income and currency effects. A similar approach was undertaken by the IFC, which in 1980 started to build indices for nine emerging markets, which were expanded to 26 by 1995.

Beyond these databases, unfortunately, there is little systematic information on the long-term performance of global stock markets. The United States is a rare exception, as monthly stock market indices have been constructed

by Standard and Poor's and, prior to 1926, by Alfred Cowles (1939), going back into the 1870s.[4]

For the non-U.S. data, we must turn to a variety of sources. The first is the International Monetary Fund (IMF), which publishes monthly stock price indices as reported by the local authorities in its *International Financial Statistics* (IFS) publication. The published indices generally represent monthly averages, as opposed to the end-of-month MSCIP and IFC data, and do not include dividends.[5] The IMF also publishes price indices and exchange rates, which can be used to compute real returns and dollar returns. We use the Wholesale Price Index (WPI) to deflate nominal returns, whenever available. The WPI measure offers a number of advantages, in that the WPI indices generally have longer histories than consumer indices, are less affected by differences in domestic consumption patterns, and are more responsive to monetary disturbances than other inflation measures.[6]

One drawback of this dataset is that it does not allow us to measure directly the equity premium, usually defined as the difference between the total return on stocks minus the Treasury bill rate. Decomposing the total return on stocks (R_S) into capital return (CR_S) and income return (IR_S), and the Treasury bill rate (R_{TB}) into the inflation component and the real rate, we can write

$$\begin{aligned}\text{Equity Premium} &= R_S - R_{TB} \\ &= [CR_S + IR_S] - [\text{Inflation} + \text{Real Rate}] \\ &= [CR_S - \text{Inflation}] + [IR_S - \text{Real Rate}]. \qquad (2)\end{aligned}$$

Our methodology measures the capital return in excess of inflation, which is the first bracketed term. To the extent that cross-sectional variations in the second bracketed term are small, this allows comparisons of equity premiums across countries. Some evidence on the quality of this approximation is presented later.

[4] For evidence on long-term U.S. data, see Wilson and Jones (1987), Schwert (1990), Siegel (1992) and Goetzmann and Ibbotson (1994). There is some long-term evidence from the U.K. markets; for instance, see Goetzmann (1993), DeLong and Grossman (1993), and Goetzmann and Jorion (1995). Parsons (1974), Mirowski (1981), and Neal (1987, 1990) provide data on the Amsterdam and London exchanges in the eighteenth century.

[5] Relative to more modern data, the IFS data suffer from two drawbacks: possible noncomparability in the construction of the series and use of monthly average instead of end-month price. The Cowles indices, the standard data source before 1926 for U.S. data, however, have similar drawbacks because prices are measured as the average of high and low values during the month.

[6] There are a few instances where we have to use Consumer Price Index data (e.g., post-1947 data for Belgium, France, New Zealand, Peru, and Israel). Because nominal prices in Germany were distorted during the hyperinflation period, we measure nominal prices for 1921–1923 in gold marks.

The first IFS publication was issued in 1948. Prior to the IMF, our source is the *Statistical Yearbooks* of the League of Nations (various issues), which include data on the capital appreciation of market indices in the period from 1929 through 1944. This collection effort was bridged by the United Nations' *Monthly Bulletin of Statistics* from 1945 to 1948. Finally, the *International Abstract of Economic Statistics* publications (ICES 1934, 1938) have stock market data going back to 1919.[7]

By connecting data from these sources, we are able to reconstruct histories for a number of stock markets going back to the early 1920s. This is a challenging effort, because of erratic data reporting.[8] The IMF, for example, provides a CD-ROM with data starting in 1957. Unfortunately, this database suffers from sample selection biases, as a number of markets that were followed in the 1960s are not contained in the CD-ROM. Data for these markets have to be collected from the IFS monthly publications. More recent emerging market data, when not available from the IFS publication, are available from the IFC database.

In order to minimize survivorship biases, we follow all markets that were reported by the League of Nation or the IMF at any point during the 1929 to 1970 period. After 1970, a flurry of new markets opened (or reopened). These emerging markets, however, have relatively short histories and are not included in the database as they have been already extensively analyzed. We obtain a total of 39 markets.[9] All in all, this involves a total of approximately 76,000 data points.

Whenever data sources do not overlap, we attempt to link series by comparing annual averages. This is the case for Austria, for instance, whose price history was interrupted by the Anschluss (German annexation) in April 1938. Fortunately, the United Nations' publications provide annual averages from 1946 on and going back to 1935; allowing us to reconstruct a long-term history for Austria, albeit with an 8-year gap during the war.

[7] Alfred Cowles, founder of the Cowles Commission for Research in Economics, was apparently the first scholar to document time-series data on global stock markets. We learned of the League of Nations data from the appendix to his 1939 publication which lists periodical sources for stock market data in 20 countries. A recent source of global stock market information which uses the League of Nations data, as well as information from other historical sources, is the Global Financial Markets database collected by Bryan Taylor, which we learned of after submission of this paper for publication. Taylor's database covers similar markets to ours; there are, however, some differences in the data sources and in particular during the breaks. For instance, we find the German stock price data collected by Gielen (1994) to be an excellent source for reconstruction of the German markets during the early part of the 20th century.

[8] The measurement of exchange rates also proves quite difficult. The League of Nations, for instance, reports rates in percentage of their 1929 gold parity value, from which current spot rates relative to the dollar have to be reconstructed. Many currencies also changed units or denomination during this century. Around World War II, trading in some currency pairs was either nonexistent or subject to heavy governmental control.

[9] The only market we deliberately omit is Lebanon, for which we cannot find inflation data.

Initially, we begin by collecting annual data. We find, however, that the monthly data create more precise estimates. In particular, we notice discrepancies between returns using monthly and annual data.[10] We also find that monthly data lead to cleaner linkages between various sources, which is particularly important as we sometimes have to patch series together. Finally, the monthly data allow us to perform event studies centered around specific dates.

Note that, despite all our efforts, this database is still not free from selection biases. The first type of bias occurs when backfilling of an index uses only stocks that are in existence at the end of the sample. In the case of Austria, for instance, even though the stock market has recovered, some companies may have fared badly or disappeared during the war. Therefore, a selection bias is induced if these companies are not included in the index.

The second type of remaining bias is much more serious. The UN–IMF data sources do not allow us to link gaps for six countries. In particular, there appears to be no link between stock market prices of Germany and Japan before and after the war in standard data sources. As these two countries did not fare well during these gaps, we can surmise that omitting the gaps misses important negative information. We attempt to correct for this by turning to other data sources for bridging these gaps.[11]

III. Empirical Analysis

A. Performance of Global Stock Markets

We calculate returns using three different numéraires: the local currency, a real price index, and the dollar. Because of wide differences in inflation across time and countries, we primarily focus on WPI-deflated returns. Returns in dollars as a common currency should give similar results over the long run if exchange rates move in line with inflation differentials—that is, if Purchasing Power Parity holds. Differences between real and dollar returns, however, may be induced when exchange rates are pegged by central banks at artificial levels, or when official exchange rates do not reflect the actual rates facing international investors.

[10] The difference can be particularly pronounced over short periods when the data are monthly or annual averages. As an illustration, comparing returns on the S&P index total returns series over 1926–1945, we find the annual growth to be 7.2 percent and 6.6 percent, respectively, for monthly and annual data.

[11] We have permanent gaps in the series for Chile, Germany, Japan, Peru, Portugal, and Argentina. The gap for Chile is filled using data from publications from the Chilean Central Bank. The gap for Germany is covered using data spliced by Gielen (1994). The gap for Japan is bridged using Bank of Japan (1966) data. The gap for Peru is filled using data received by the Lima stock exchange. To cover the gap for Portugal, we use information from the Portuguese Central Bank. Overall, Argentina is the only remaining country with a permanent break over July 1965 to December 1975, which is the first date for which we have data from the IFC. We have been unable to find data to bridge the gap.

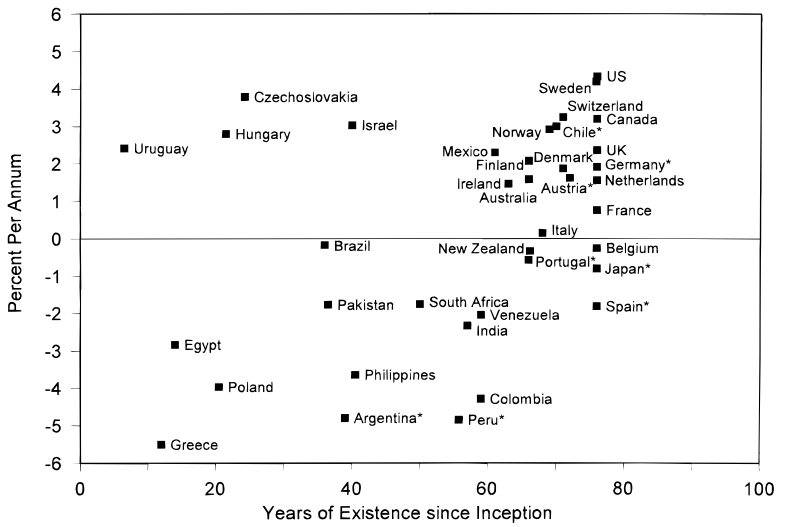

Figure 1. Real returns on global stock markets. The figure displays average real returns for 39 markets over the period 1921 to 1996. Markets are sorted by years of existence. The graph shows that markets with long histories typically have higher returns. An asterisk indicates that the market suffered a long-term break.

Table I presents geometric returns for 39 markets grouped by regions, compounded annually. These results are striking. Of the sample of 39 countries, real returns are the highest for the United States, at 4.32 percent per annum. There is no country with a higher return over the total period. Therefore, the high U.S. equity premium seems to be the exception rather than the rule.

These results are perhaps better visualized in Figure 1, which plots the compound return for each market against its observed "life" since 1921. Longer lives lead to more precise, less volatile, estimates of expected returns. Moving to the right of the figure, we observe that the U.S. market has the highest realized return of all markets.

At the bottom of Table I we show average and median returns for all countries, as well as for a group of countries for which we have data going to the 1920s. The median real returns for all 39 countries is 0.75 percent. By way of contrast, we also analyze countries with continuous histories going back to the 1920s; the median return for this group is also much higher, at 2.35 percent. These results strongly suggest that the 4.3 percent real capital appreciation return for the United States is highly unusual. As it is also one of the few series without any break, this high return could be ascribed to survival.

An alternative explanation is that the United States had a higher level of risk than any other market over the period. In perfectly integrated capital markets, a high equity premium can simply compensate for a high β. Of

course, this is a difficult proposition to test directly because survivorship affects not only returns but also capital weights. Ex post, the most successful index will represent the largest share of the market.

Other high returns, however, are obtained in some cases. Over 1921 to 1996, Swedish equities displayed returns quite close to the 4.32 percent obtained in the United States, perhaps not surprisingly as Sweden also avoided major upheavals in this century. Higher returns are observed over more recent periods. For instance, Germany experienced a steep run-up in prices, 6 percent in real terms, over the period 1950 to 1996. But this high return must be offset against mediocre growth up to July 1944; additionally, during the five-year break in our series, German equities fell by 72 percent in real terms. As a result, the long-term growth of the German market is only 1.91 percent when evaluated over most of this century. The story is similar for Japan, where we observe a sharp difference between the postwar return of 5.52 percent and the prewar return of -0.34 percent. During the 1944 to 1949 break, the market fell by 95 percent in real terms.

Other markets that gapped, such as Portugal, Chile, and Peru, also did well recently, but not so well when going back further in time. These are typical "reemerging markets," whose recent performance appears to be, on the surface, nothing short of stellar. Our analysis shows that the performance of the same markets has also been mediocre at other times.

Table I also reports dollar returns. As expected, rankings for this column are very similar to those obtained with real returns.[12] In general, dollar returns for other currencies are slightly closer to U.S. returns than real returns. For example, the difference between U.S. equities and the median is $4.32 - 0.75 = 3.57$ percent when measured in real terms; the difference is $6.95 - 4.68 = 2.27$ percent in dollar terms. This discrepancy reflects the slight depreciation of the dollar, relative to its Purchasing Power Parity value, over the sample period.

In addition to geometric returns, which represent returns to a buy-and-hold strategy, it is also useful to consider arithmetic averages, which give equal weight to each observation interval. Table II presents conventional measures of annualized average (arithmetic) capital appreciation returns and standard deviations.[13] Data are presented in the local currency, in real terms, and in dollars. The table shows that the 16.2 percent volatility of the U.S. market is not particularly high when compared with other stock markets. Therefore the high return obtained in the United States does not seem to compensate for higher risk as measured by volatility (which would be the appropriate measure of risk under segmented capital markets).

[12] Uruguay and Czechoslovakia had higher returns than U.S. equities, but this was over shorter periods during which currencies were subject to controls; hence, these returns are not representative.

[13] Since price data are monthly averages, it should be noted that the reported standard deviations are lower than those from using month-end data. Additionally, averaging induces spurious positive autocorrelation in the return series.

Table I

Long-Term Performance of Global Equity Markets

(Compound Return in Percentage per Annum)

The table compares the long-term performance of global equity markets with annually compounded data. The sample period varies across country and is reported in the second column. Data for subperiods are reported within brackets. Percentage returns are measured in nominal terms in the local currency, in real terms—deflating by the Wholesale Price Index, and translated into U.S. dollars. The last column reports the inflation rate. * indicates a break in the series that has been bridged; + indicates a permanent discontinuity in the series.

Country	Period	Nominal Return	Real Return	Dollar Return	Inflation
United States	1/21–12/96	6.95	4.32	6.95	2.52
Canada	1/21–12/96	5.78	3.19	5.35	2.51
Austria*	1/25–12/96	5.64	1.62	5.00	3.95
Belgium	1/21–12/96	4.45	−0.26	3.51	4.73
Denmark	1/26–12/96	5.87	1.87	5.19	3.93
Finland	1/31–12/96	10.23	2.07	6.19	7.99
France	1/21–12/96	9.09	0.75	4.29	8.28
Germany*	21–96	4.43	1.91	5.81	2.47
Germany	1/21–7/44	[3.29]	[2.23]	[5.59]	[1.04]
Germany	1/50–12/96	[8.46]	[6.00]	[10.78]	[2.32]
Ireland	1/34–12/96	7.00	1.46	5.14	5.45
Italy	12/28–12/96	10.10	0.15	3.22	9.94
Netherlands	1/21–12/96	3.71	1.55	4.47	2.12
Norway	1/28–12/96	7.13	2.91	6.29	4.10
Portugal*	31–96	6.89	−0.58	3.78	7.51
Portugal	12/30–4/74	[5.21]	[1.16]	[4.96]	[4.00]
Portugal	3/77–12/96	[20.11]	[5.63]	[11.92]	[13.71]
Spain*	1/21–12/96	4.66	−1.82	1.53	6.61
Sweden	1/21–12/96	7.42	4.29	7.00	3.00
Switzerland	1/26–12/96	4.83	3.24	6.84	1.54
United Kingdom	1/21–12/96	6.30	2.35	5.20	3.86
Czechoslovakia	1/21–4/45	4.33	3.79	9.50	0.52
Greece	7/29–9/40	−2.12	−5.50	−8.08	3.58
Hungary	1/25–6/44	6.29	2.80	9.07	3.40
Poland	1/21–6/39	−7.00	−3.97	−4.30	−3.15
Romania	12/37–6/41	−5.36	−28.06	−14.64	31.55
Australia	1/31–12/96	7.06	1.58	6.29	5.39
New Zealand	1/31–12/96	5.69	−0.34	3.63	6.01
Japan*	21–96	7.33	−0.81	1.80	8.21
Japan	1/21–5/44	[1.23]	[−0.34]	[−1.83]	[1.58]
Japan	4/49–12/96	[8.30]	[5.52]	[10.90]	[2.63]
India	12/39–12/96	5.10	−2.33	0.80	7.60
Pakistan	7/60–12/96	7.79	−1.77	0.59	8.57
Philippines	7/54–12/96	5.95	−3.65	−0.30	9.96
Argentina+	47–65,75–96	87.48	−4.80	−1.43	96.92
Argentina	9/47–7/65	[−5.78]	[−25.09]	[−23.64]	[25.78]
Argentina	12/75–12/96	[236.29]	[16.71]	[22.43]	[188.15]
Brazil	2/61–12/96	142.34	−0.17	4.68	147.52
Mexico	12/34–12/96	20.13	2.30	6.12	17.43

Table I—*Continued*

Country	Period	Nominal Return	Real Return	Dollar Return	Inflation
Chile*	27–96	37.12	2.99	6.38	33.16
Chile	1/27–3/71	[12.98]	[−5.37]	[−4.23]	[19.39]
Chile	1/74–12/96	[64.19]	[15.52]	[20.94]	[42.13]
Colombia	12/36–12/96	10.15	−4.29	−0.88	15.09
Peru*	41–96	45.29	−4.85	3.45	52.68
Peru	3/41–1/53	[2.03]	[−12.36]	[2.03]	[16.41]
Peru	1/57–12/77	[1.53]	[−9.88]	[−7.40]	[12.66]
Peru	12/88–12/96	[340.95]	[30.45]	[50.92]	[232.18]
Uruguay	3/38–11/44	6.70	2.42	10.01	4.19
Venezuela	12/37–12/96	9.67	−2.04	0.78	11.95
Egypt	7/50–9/62	−1.46	−2.84	−1.63	1.42
Israel	1/57–12/96	37.05	3.03	7.21	33.02
South Africa	1/47–12/96	6.13	−1.76	1.48	8.03
All 39 countries					
Mean			−0.47	3.11	
Median			0.75	4.68	
11 countries with continuous histories into the 1920s					
Mean			1.88	5.09	
Median			2.35	5.20	

The table also reports the results from standard statistical tests of significance of the real capital appreciation return premium. At the 99 percent level, we can only reject the hypothesis of a zero long-run appreciation return for the United States and Sweden. Over shorter periods, we observe significantly positive returns for Germany and Japan in the postwar period. When averaged with prewar data, however, these returns look less impressive.

B. The Effect of Dividend Omission

The previous section has revealed a striking result: long-term returns on the U.S. stock market appear to be greater than those of any other market during this century. One question that arises is whether this result could be due to the omission of dividends. To shed light on this issue, Table III presents performance numbers for markets for which we have dividend data.

Panel A reports data for the more recent MSCIP indices, which mainly cover industrial countries since 1971. The table displays compound real returns, with and without reinvestment of dividends. The difference due to the omission of dividends is shown in the third column. The fourth column reports the average level of inflation. Presumably, the results in the previous section could simply reflect a bias due to the omission of dividends. For this bias to be effective, other markets must systematically display a higher income component of return than the United States.

Table II

Return and Risk of Global Equity Markets

(Arithmetic Return in Percentage per Annum)

The table compares average stock returns and their standard deviations. Percentage returns are measured in nominal terms in the local currency, in real terms, deflating by the Wholesale Price Index, and translated into U.S. dollars. The arithmetic average return is obtained from the monthly average multiplied by 12; the standard deviation is annualized by multiplying the monthly volatility by the square root of 12. For series with breaks, (1), (2), (3) refer to different subperiods.

		Nominal Return		Real Return		Dollar Return	
Country	Period	Average	(Std.Dev.)	Average	(Std.Dev.)	Average	(Std.Dev.)
United States	1/21–12/96	8.09**	(16.20)	5.48**	(15.84)	8.09**	(16.20)
Canada	1/21–12/96	7.06**	(16.81)	4.54*	(16.65)	6.88**	(18.17)
Austria	1/25–12/96	6.77**	(18.92)	2.32	(19.49)	7.22**	(21.49)
Belgium	1/21–12/96	6.25**	(17.92)	1.49	(18.97)	5.77**	(21.80)
Denmark	1/26–12/96	6.43**	(12.04)	2.65	(12.69)	6.10**	(14.36)
Finland	1/31–12/96	10.74**	(16.56)	3.50	(17.07)	8.18**	(20.49)
France	1/21–12/96	11.19**	(21.57)	3.16	(21.25)	7.76**	(25.50)
Germany (1)	1/21–7/44	10.22	(40.24)	7.62	(34.26)	12.54	(40.49)
Germany (2)	1/50–12/96	9.35**	(15.50)	7.06**	(15.60)	11.75**	(17.19)
Ireland	1/34–12/96	7.88**	(14.85)	2.59	(15.02)	6.43**	(16.73)
Italy	12/28–12/96	12.62**	(26.01)	3.15	(25.66)	3.15	(25.66)
Netherlands	1/21–12/96	4.78**	(15.12)	2.78*	(14.80)	5.85**	(16.50)
Norway	1/28–12/96	8.49**	(17.90)	4.47*	(17.90)	7.97**	(19.33)
Portugal (1)	12/30–4/74	6.50**	(15.15)	2.34	(14.69)	7.40**	(15.03)
Portugal (2)	3/77–12/96	27.08**	(46.38)	14.69	(47.68)	20.42	(47.11)
Spain	1/21–12/96	6.77**	(18.92)	−0.51	(16.00)	2.44	(28.89)
Sweden	1/21–12/96	8.56**	(16.61)	5.60**	(16.65)	8.38**	(17.69)
Switzerland	1/26–12/96	5.83**	(14.79)	4.28*	(14.73)	7.91**	(15.97)
United Kingdom	1/21–12/96	7.25**	(15.43)	3.60*	(15.68)	6.66**	(17.57)
Czechoslovakia	1/21–4/45	5.04*	(12.53)	4.56	(12.84)	10.50**	(17.12)
Greece	7/29–9/40	−0.09	(21.77)	−3.44	(21.61)	−5.31	(25.50)
Hungary	1/25–6/44	9.34	(25.84)	6.20	(26.58)	11.99*	(26.02)
Poland	1/21–6/39	13.60	(71.20)	14.40	(65.69)	16.69	(71.54)
Romania	12/37–6/41	0.14	(33.31)	−27.30	(31.38)	−9.45	(35.06)
Australia	1/31–12/96	7.78**	(13.49)	2.57	(13.94)	7.68**	(18.06)
New Zealand	1/31–12/96	6.20**	(12.12)	0.55	(12.50)	4.98**	(15.97)
Japan (1)	1/21–5/44	2.72	(17.51)	0.89	(15.79)	−0.35	(17.40)
Japan (2)	4/49–12/96	9.79**	(18.78)	7.21**	(18.90)	12.61**	(20.97)
India	12/39–12/96	6.18**	(15.53)	−1.07	(16.13)	2.37	(17.46)
Pakistan	7/60–12/96	7.46**	(14.37)	−0.64	(15.23)	2.39	(17.50)
Philippines	7/54–12/96	10.62	(37.35)	1.21	(37.21)	5.30	(38.91)
Argentina (1)	9/47–7/65	−1.13	(31.91)	−23.32**	(32.73)	−18.17	(40.11)
Argentina (2)	12/75–12/96	179.34	(133.55)	49.68	(87.83)	57.85**	(93.68)
Brazil	2/61–12/96	110.69**	(68.22)	12.92	(51.93)	18.45*	(53.44)
Mexico	12/34–12/96	21.97**	(26.79)	5.37	(24.45)	10.46**	(29.09)
Chile (1)	1/27–3/71	14.51**	(22.45)	−3.91	(21.85)	−0.12	(28.64)
Chile (2)	12/73–12/96	57.19**	(40.34)	20.48**	(36.25)	25.94**	(38.59)
Colombia	12/36–12/96	11.66**	(21.56)	−2.32	(21.78)	1.67	(23.39)
Peru (1)	3/41–1/53	3.02	(12.90)	−12.08**	(14.15)	3.39	(16.58)
Peru (2)	1/57–12/77	1.89	(8.62)	−9.94**	(9.08)	−6.61*	(13.66)
Peru (3)	12/88–12/96	200.64**	(118.38)	55.55	(87.98)	71.95*	(87.18)
Uruguay	12/36–11/44	10.55	(28.98)	6.67	(29.66)	13.80	(29.63)
Venezuela	12/37–12/96	12.03**	(24.65)	0.88	(24.84)	4.85	(28.08)
Egypt	7/50–9/62	−0.83	(11.50)	−2.11	(12.54)	−0.19	(17.33)
Israel	1/57–12/96	35.18**	(26.07)	5.68	(22.96)	10.07*	(24.33)
South Africa	1/47–12/96	7.24**	(15.75)	−0.46	(15.89)	3.34	(18.87)

*, ** Significantly different from zero at the 5 and 1 percent levels, respectively.

Table III

Comparison of Real Returns with and without Dividends

The table compares stock returns with and without dividends. Returns are measured in real terms and are annually compounded. The top part reports Morgan Stanley Capital International Perspective (MSCIP) data; the bottom part presents long-term data, obtained from various sources.

Country		Compound Return with Dividend (% pa)	Compound Return without Dividend (% pa)	Difference due to Dividend	Inflation (% pa)
Panel A: Markets Covered by MSCIP, 1970–1995					
Australia		3.65	−0.71	4.36	6.79
Austria		4.89	2.07	2.82	2.75
Belgium		12.97	4.05	8.92	2.46
Canada		4.34	0.65	3.69	5.78
Denmark		6.54	2.71	3.83	5.62
France		4.45	−0.29	4.74	7.40
Germany		5.52	1.44	4.08	3.09
Italy		−0.26	−2.95	2.69	9.87
Japan		8.59	6.75	1.84	2.18
Netherlands		8.84	3.09	5.74	3.41
Norway		6.03	2.78	3.26	5.90
Spain		2.30	−4.00	6.31	8.40
Sweden		8.79	5.03	3.76	7.42
Switzerland		5.72	3.06	2.66	2.54
United Kingdom		6.39	1.23	5.16	8.35
United States		6.15	2.01	4.14	4.89
Average		5.93	1.68	4.25	5.43
Panel B: Long-Term Markets					
Denmark	1923–95	4.88	0.64	4.24	3.72
Germany	1924–95	4.83	1.21	3.63	2.47
Sweden	1926–95	7.13	3.30	3.83	3.64
Switzerland	1921–95	5.57	2.12	3.45	2.49
United Kingdom	1921–95	8.16	2.99	5.17	3.75
United States	1921–95	8.22	3.38	4.84	2.69
United States	1871–1920	5.43	0.27	5.16	0.59

Table III clearly shows that this is not the case. Over the 1970–1995 period, the dividend effect for the United States was 4.14 percent, which is quite close to the group average of 4.25 percent. Therefore, there is no indication that the high return obtained for U.S. equities in Table I is due to dividend bias. If anything, the bias is in the opposite direction. For example, Japanese equities, which by now constitute the largest market outside the United States, paid an income return of 1.84 percent over the past 25 years, which is much lower than that of U.S. equities.

Panel B of Table III reports the only long-term data with dividends that we are aware of.[14] To maintain comparability with the original data sources, we use the Consumer Price Index (CPI) to deflate returns, except for Denmark where the WPI is employed. Including dividends, the United States displays the highest real equity returns since 1921, at 8.22 percent. Britain, another long-term survivor, is a close second; other markets provide returns that are lower by 109 to 334 basis points. Another way to look at the data is to notice that the ranking of returns is essentially the same with and without dividends. Therefore, there is no evidence that the performance of U.S. equities is artificially high because of relatively low U.S. dividend payments.

C. Evidence on the Equity Premium Puzzle

The data we present thus far do not explicitly solve the equity premium puzzle, as theoretically formulated. Strictly speaking, the equity premium puzzle concerns the spread of expected total return on the market portfolio of equities over the return of a riskless security. Siegel (1994) points out that defaults on "riskless" government securities have often occurred in periods of global stress—which of course raises the question of what the riskless asset might actually be and whether the stylized, single economy, two-asset formulation of the equity premium puzzle is robust.

In the absence of a riskless asset that is immune to the crisis events imagined by Rietz (1988), it seems reasonable to substitute physical storage of goods (i.e., inflation rates for T-bill rates). In this case, using real returns as a proxy for the equity premium clearly supports the hypothesis that the ex post observed U.S. premium is higher because the United States was a winner. This evidence, in turn, is consistent with the "survival" hypothesis suggesting that the magnitude of ex post observed equity returns may be higher than their ex ante expectation.

Is there any evidence in the data supporting the Rietz (1988) hypothesis that the ex ante equity premium is as high as supposed? The issue is whether there was some probability of the U.S. market experiencing a large crash. In fact, this problem is akin to the "peso problem" in the foreign exchange market, where peso forward rates appeared to be biased forecasts of future spot rates over short sample periods, essentially because they account for a nonzero probability of devaluation that is not observed. More generally, peso problems can be interpreted as a failure of the paradigm of rational expectations econometrics, which requires that the ex post distribution of endogenous variables be a good approximation to the ex ante distribution that agents think may happen. The failure may not be that of the economic agent, but that of the econometrician, who only analyzes series with continuous

[14] Data sources are as follows: For the U.S. market, Ibbotson (1995) and prior to that, Cowles (1939); for the U.K., Barclays deZoete Wedd (1993); for Switzerland, Wydler (1989); for Sweden, Frenneberg and Hansson (1992); and for Denmark, Timmerman (1992). All of the data have been updated to 1995 using the MSCIP indices.

histories. Unusual events with a low probability of occurrence but severe effects on prices, such as wars or nationalizations, are not likely to be well represented in samples and may be totally omitted from survived series.

Our cross-sectional data provide evidence about major market crashes not present in U.S. data. We have, for example, 24 markets for which we have data in 1931. Of these, seven experienced no interruption (the United States, Canada, the United Kingdom, Australia, New Zealand, Sweden, and Switzerland), seven experienced a temporary suspension of trading (less than one year), and the remaining 10 markets suffered long-term closure. Even though these events are not independent, they indicate that market failure is not a remote possibility. Under the assumption that market risks are "priced" individually, rather than under the assumption of integration, the frequency of failure would provide clear justification for a peso problem explanation.

Although it is entirely possible that the magnitude of the observed equity premium is due both to survival bias and to the "pricing" of an infrequently occurring crash, it is difficult to believe that the ex ante premium for the United States should be higher than for other markets. The increased probability of a large crash may explain a higher average equity premium, but if past crash frequency is any indication of future crash probability, then the Rietz (1988) hypothesis would suggest that markets with more interruptions should have a higher equity premium. If we believe that the magnitude of the equity premiums for each country is related to the ex post historical real appreciation, then the opposite appears to be the case. Absent survival effects, the Rietz hypothesis is inconsistent with cross-sectional differences in historical global equity market returns. In the next section, we investigate the possibility that markets anticipate major crashes.

Table III provides additional evidence on the equity premium puzzle by comparing the performance of U.S. equities during the recent period with longer term, 1871–1920, Cowles data. The last line in the table shows that the high real capital return obtained since 1921 is much higher than that obtained in the preceding 50 years—3.38 percent during 1921–1995 against 0.27 percent during 1871–1920. Siegel (1992) also points out that the U.S. equity premium is particularly high during this century. Put differently, this large premium seems not only large in a cross-country comparison but also by historical standards. Siegel concludes that "investors in … 1872 did not universally expect the United States to become the greatest economic power in the next century." If so, returns on U.S. equities this century cannot be viewed as representative of global stock markets.

D. Disappearance as an Event

To understand how risk premiums respond to the probability of major market crashes, we can examine the behavior of markets around interruptions. Sample selection of markets will create a bias if the performance of interrupted markets is systematically poor before the break. By the same token,

Table IV

Analysis of Stock Prices around Breaks

The table describes the behavior of stock prices measured in real terms around major breaks. It reports the break date, the return in the year previous to the break, the series restart date, and subsequent change, when available. Real returns are in excess of the Wholesale Price Index for the corresponding countries. * indicates that equities were effectively subject to price controls; + indicates that the subsequent change was obtained from alternative data sources.

Country	Break date	Previous year return	Series restart date	Subsequent change	Comment
Hungary	7/31	−0.222	9/32	0.125	Financial crisis, country in default
Germany	7/31	−0.316	4/32	−0.232	Credit crisis
Greece	10/31	−0.099	12/32	−0.581	Financial crisis, drought
Spain	7/36	−0.113	3/40	−0.147	Civil War
Austria	4/38	−0.179	12/46	0.941	Annexation by Germany
Czechoslovakia	10/38	−0.205	1/40	0.015	Session of land to Germany
Poland	7/39	0.169			Invaded by Germany (Sep 30)
Finland	12/39	−0.192	3/40	−0.101	Invaded by Soviets (Nov 30)
Denmark	4/40	−0.328	6/40	−0.084	Invaded by Germany (Apr 9)
Norway	4/40	−0.274	6/40	−0.154	Invaded by Germany (Apr 11)
Netherlands	5/40	−0.231	9/40	0.105	Invaded by Germany (May 10)
Belgium	5/40	−0.267	12/40	0.850	Invaded by Germany (May 10)
Switzerland	5/40	−0.193	7/40	−0.207	Mobilization
France	6/40	−0.122	4/41	0.824	Invaded by Germany (Jun 14)
Greece	10/40	−0.249	none		Invaded by Germany (Oct 28)
Romania	7/41	−0.396	none		Enters war
Czechoslovakia*	7/43	−0.141	none		War
Japan*	6/44	−0.211	4/49	−0.949+	War
Hungary*	7/44	−0.491	none		War
Belgium*	8/44	0.161	6/45	−0.145	War
Germany*	8/44	−0.013	1/50	−0.838+	Invaded by Allies (Sep 15)
Egypt	10/62	−0.126	none		Arab socialism
Argentina	8/65	−0.692	N/A		Widespread unrest, hyperinflation
Chile	4/71	−0.543	1/74	1.618+	State takes control of economy (Apr 4) Junta reverses policies (Sep 11, 73)
Portugal	4/74	−0.112	3/77	−0.860+	Takeover by leftist junta (Apr 27)

falling stock prices prior to a market break may be indicative of investor assessment of increasing probability that the market will fail.

To test this hypothesis, we adopt the event-study methodology by constructing an equally weighted index in which real returns are aligned on the interruption date. We identify a sample of 25 breaks for which the data series are clearly interrupted. Table IV identifies each of these events. Many are of a global nature, such as the Second World War, or the depression of the early 1930s. A number of events, however, are country-specific, involving a banking crisis or political turmoil.

Figure 2 plots the time-series of the portfolio value, starting one year before the break. It shows prices falling on average by 21 percent relative to their peak. The *t*-test based on the standard deviation of monthly changes in

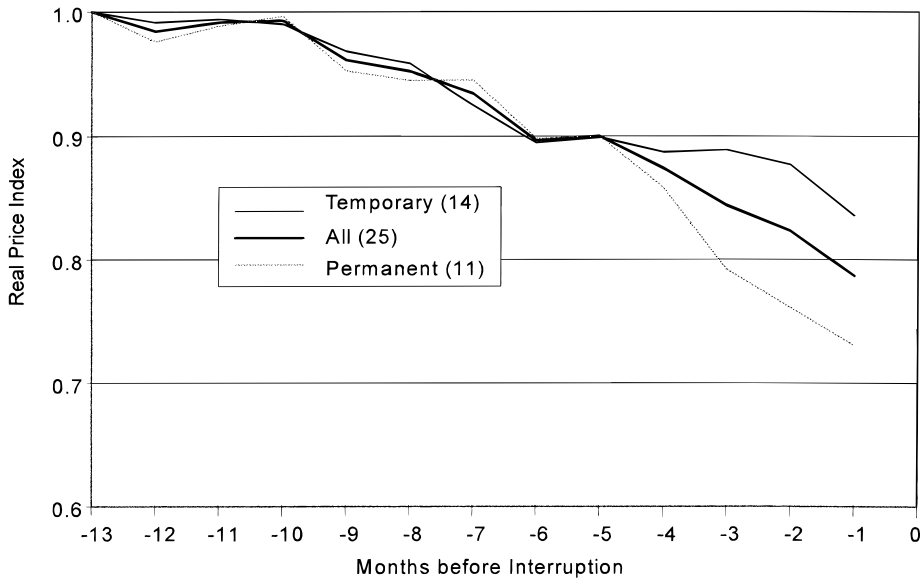

Figure 2. Real stock prices before interruption. The figure displays the performance of an equally weighted index where real returns are aligned on the interruption date. The total sample of 25 is further divided into a sample for which the interruption turns out to be temporary, and a sample for which the interruption is permanent.

the previous year is −4.95 for this number, which is highly significant. However large, this fall of 21 percent in real terms understates the true loss of value to equities. During World War II, in particular, prices were kept artificially high through price controls and do not represent transaction prices as liquidity dried up.[15]

Eventually, reality prevailed. Figure 3 compares the performance of markets sorted by country involvement during the war.[16] As the figure shows, the advent of the war led to a sharp fall of about 20 percent in the value of equities of Allied countries (including the United States, Canada, the United Kingdom, and Commonwealth countries) for the next two months. A similar fall was suffered by neutral countries (Sweden and Switzerland). The index

[15] In Germany, Italy, and German-occupied territories, dealing in shares was subject to strict controls, ranging from taxes on profits and capital gains to the rationing of purchases and to the compulsory declaration of securities holdings. In June 1942, for instance, the sale of German shares became prohibited unless they were first offered to the Reichsbank. The Reichsbank had the option to buy them at December 1941 prices in exchange for bonds that remained in the bank's possession. It is no wonder that this confiscatory system led to a sharp fall in trading activity. There were also rigid price controls in Japan during the war; see for instance Adams and Hoshii (1971). Therefore many of these price indices do not represent market-determined prices.

[16] The index for occupied countries includes Belgium, Czechoslovakia, France, Denmark, Finland, Germany, Hungary, Italy, Netherlands, and Norway.

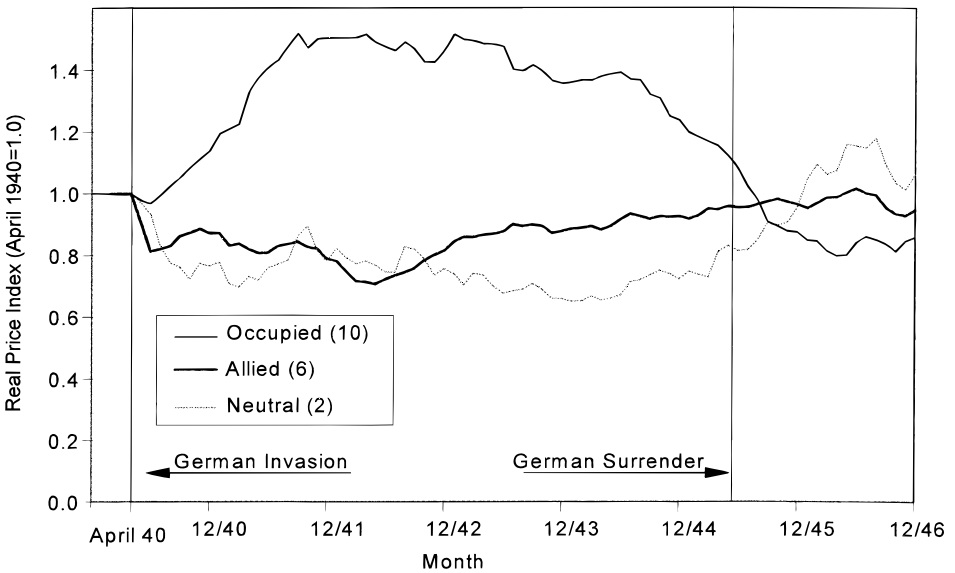

Figure 3. Real stock prices during World War II. The figure displays the performance of portfolios of equities measured in real terms during the war. The sample is divided into occupied, allied, and neutral countries.

for occupied countries, in contrast, registered steady gains, which were only wiped out later as stock prices started to reflect transaction prices and as inflation became apparent. Five years later, the index moved below that of Allied countries, as we would have expected. In reality, the index should have been even lower if we had accounted for those markets that disappeared in the process (such as Germany, Hungary, and Czechoslovakia.)

Table IV also details the performance around each individual break. All markets suffered a substantial drop before the break, reaching 69 percent for Argentina. One exception is Poland, which experienced a slight price increase, possibly because the series was stopped in July, three months before Poland was invaded, or because the advent of the war was unanticipated. As explained before, the price drops in Germany and occupied Europe are also unusual, for artificial reasons. In all other cases, the event creating the market closure was anticipated.

In eleven of these cases, the UN–IMF equity series are interrupted without restarting later (or there are no continuous series spanning the interruption). These cases include Germany, Japan, Eastern European countries taken over by the Soviet Union, Greece, Egypt, Chile, Argentina, and Portugal. Some of these were the result of a foreign occupation and widespread destruction due to war. In Egypt and Chile, the state took control of the economy. The Buenos Aires Stock Exchange, the oldest in Latin America, virtually disappeared as a result of inflation and interest rate policies in the late 1960s; reportedly, investors lost all interest in the market. These are precisely the situations where we would expect equities to fare most badly.

We have to turn to other data sources to bridge these "permanent" breaks. We find that, over the 1944–1949 break in Japan, equities fell by 95 percent in real terms.[17] For Germany, we find that equities fell by 84 percent in real terms over 1944–1950. Another example is the Portuguese stock market, which closed in April 1974 as a military junta took over the country, reopened in March 1977, then traded intermittently. The stock price series suffered a fall of 86 percent in real terms during the interruption in trading. In contrast, most of the loss for the Chilean stock market occurred before the interruption; the market recovered somewhat over the 1971–1974 break, as the military junta reversed the socialist policies of the Allende government.[18] Furthermore, these numbers probably underestimate the true loss in value by ignoring companies that failed during the interruption, as indices are backfilled from companies quoted before and after the break.

Going back to Figure 2, we have separated markets that were temporarily interrupted from those that disappeared, or "died," later. Markets that became extinct dropped by 27 percent the year before the break; markets that subsequently recovered dropped by only 16 percent before the break. To the extent that the event causing the break was anticipated, the market seems to have been able to gauge the gravity of unfolding events. Price declines before breaks are consistent with increasing demand for risk compensation for a catastrophic event.

E. A Global Stock Index

The global equity data provide a unique opportunity to construct a global equity index—an index that for the first time includes defunct as well as surviving countries and extends back 75 years. Because we have no data on market capitalization going back that far, we assign weights based on Gross Domestic Product (GDP). Annual GDP information is obtained from Mitchell (1992, 1993, 1995) and converted to U.S. dollars using annual averages. At the beginning of each decade, we construct a cross section of national GDPs, which are used to construct initial weights.

To minimize rebalancing, we adopt a portfolio value-weighted approach. Our global indices are therefore similar to market capitalization indices, except that the weights are reset to GDP weights at the beginning of each decade. A value-weighted scheme is more appropriate for measurement of investor returns when survival is an issue. As our analysis in the previous section demonstrates, markets that die tend to have less weight when they do so.

[17] The Bank of Japan (1966) estimates that the material damage due to World War II was to reduce national wealth from 253 to 189 billion yen, which is a fall of 64 billion yen (not accounting for human losses), or about $15 billion. For comparison purposes, the market value of equities in 1945 was about 40 billion yen.

[18] The market lost 54 percent in the year to April 1971 during the Allende ascent to power, but then increased by 62 percent later, which is only a partial recovery. Assuming a starting value of 100, the market fell to 46, then recovered to 1.62 times 46, or 74, ending with a net loss in value relative to the starting point.

The indices represent the return an investor would have earned had it been possible to hold the market since the 1920s. This is a hypothetical experiment, however, because it would have been difficult to maintain such a portfolio. Constraints on cross-border capital flows and on liquidation of equity positions were acute during crises—precisely the times when the ability to diversify is most beneficial. In this period, investors were sometimes involuntarily separated from their assets, due to expropriations or nationalizations. As a result, it is not clear whether, for example, a U.S. investor could have continued to hold German or Japanese equities during World War II.

Table V presents the GDP weights at three points in time: 1920, 1950, and 1990. The table reveals a number of interesting observations. The United States accounts for about one-half of the world's output until the 1950s; the proportion then declines to approximately 30 percent. This decline is due to faster growth in other countries such as Japan and Germany. Japan, in particular, zooms from 4 percent of world GDP to 16 percent over this period, even after dipping below 2 percent after the war.

The GDP-based weights can be compared to stock market capitalization-based weights, which are reported in the last column. We observe that the stock market capitalization percentages of the United States, the United Kingdom, Japan, and South Africa are generally greater than those of other countries. Continental Europe, for example, has a history of relying on bank lending rather than raising funds through capital issues. Overall, however, the GDP weights are roughly of the same order of magnitude as market weights.

Biases can be introduced in the measured performance in a number of ways. The first is backfilling, and the second is due to interruptions. There is not much the researcher can do about backfilling if the series are the only ones available. As for interruptions, the problem is that data before the interruption are commonly ignored. Interruptions can be of two types: temporary closure of an exchange, with the series starting again later, or permanent interruption of these series, with no information about the continuity of prices across the interruption.

We take two approaches to the construction of the global index:

(i) Our "survived markets" index includes all markets since the last interruption, which can be a temporary break or a permanent closure; only markets in existence at the end of the sample are considered. As of December 1996, we have a total of 32 markets; of which only 18 had continuous histories to December 1940, for instance.

(ii) Our "all markets" index extends the sample to all markets in existence in our sample, including returns before temporary and permanent closures. As of December 1940, this "comprehensive" series yields 29 markets, adding Austria, Belgium, and France (which suffered a temporary interruption of trading during World War II), Chile, Germany, Japan, Portugal, Uruguay, and three markets that suffered a permanent break during the war: Czechoslovakia, Hungary, and Romania.

Table V

Relative Importance of Economies

(Percentage Weights Based on U.S. Dollar Prices)

The table describes the percentage of each country in the total Gross Domestic Product (GDP) in 1920, 1950, and 1990. The last column shows the percentage weight based on stock market capitalization.

Country	GDP Weights			Stock Market Capitalization
	1920	1950	1990	1995
United States	46.17%	51.52%	30.59%	41.03%
Canada	2.40%	3.16%	3.17%	2.16%
Austria	0.48%	0.47%	0.87%	0.24%
Belgium	0.73%	1.27%	1.09%	0.66%
Denmark	0.55%	0.56%	0.72%	0.37%
Finland	0.17%	0.42%	0.76%	0.26%
France	6.14%	5.19%	6.61%	3.27%
Germany	6.04%	4.19%	8.29%	3.75%
Ireland	0.42%	0.19%	0.24%	0.16%
Italy	1.67%	2.43%	6.07%	1.16%
Netherlands	0.98%	0.89%	1.57%	1.97%
Norway	0.56%	0.38%	0.59%	0.28%
Portugal	0.62%	0.25%	0.33%	0.12%
Spain	2.16%	0.82%	2.72%	0.99%
Sweden	1.22%	1.11%	1.26%	1.14%
Switzerland	0.84%	0.80%	1.25%	2.60%
United Kingdom	10.36%	6.57%	5.41%	8.77%
Czechoslovakia	0.52%	0.31%	0.25%	0.10%
Greece	0.33%	0.39%	0.37%	0.11%
Hungary	0.38%	0.71%	0.18%	0.02%
Poland	1.82%	0.35%	0.03%	
Romania		0.00%	0.21%	—
Australia	2.31%	1.07%	1.63%	1.59%
New Zealand	0.15%	0.35%	0.24%	0.21%
Japan	4.06%	1.96%	16.24%	23.19%
India	6.92%	3.54%	1.68%	0.82%
Pakistan		0.67%	0.22%	0.06%
Philippines		0.63%	0.24%	0.38%
Argentina	1.20%	0.90%	0.78%	0.24%
Brazil	0.75%	2.84%	2.66%	0.96%
Mexico	0.66%	0.85%	1.34%	0.59%
Chile	0.19%	0.75%	0.15%	0.48%
Colombia		0.72%	0.22%	0.12%
Peru		0.19%	0.20%	0.08%
Uruguay		0.18%	0.05%	0.00%
Venezuela		0.57%	0.27%	0.02%
Egypt		0.45%	0.31%	0.05%
Israel		0.24%	0.29%	0.24%
South Africa	1.03%	0.65%	0.56%	1.82%
Memorandum:				
GDP (millions)	$198,200	$556,500	$18,049,700	
Market cap (m)				$15,448,900

Table VI

Performance of Global Stock Index: 1921–1996

(Real Returns in Percentage per Annum)

The table displays the risk and return of real returns on stock market indices, measured in excess of the Wholesale Price Index inflation. Arithmetic return is obtained from the monthly average multiplied by twelve; risk is monthly volatility multiplied by the square root of twelve; Sharpe ratio is the ratio of monthly average to monthly volatility; geometric return uses annual compounding. Ending wealth reports the final value of $1 invested on December 1920 at the end of the sample. "Survived markets" series includes only markets in our sample in existence in 1996, taken since the last interruption (temporary or permanent). "All markets" series accounts for all markets in the sample, imputing a 75 percent loss in the month the series permanently disappears, or the actual loss spread over the period of the break.

Index	Arithmetic Return	Risk	Monthly Sharpe	Geometric Return	Ending Wealth
U.S. index	5.48	15.83	0.0999	4.32	27.3
Global index					
Survived markets	4.98	12.08	0.1190	4.33	27.3
All markets	4.59	11.05	0.1199	4.04	21.9
Non-U.S. index					
Survived markets	4.52	10.02	0.1301	4.09	22.2
All markets	3.84	9.96	0.1114	3.39	13.1

We expect the bias to decrease as we move from (i) to (ii). The difficult part, of course, is to estimate market losses during a permanent interruption such as war or nationalization. We have 11 occurrences of permanent breaks (or "deaths") out of our sample of 39 markets. For some of these, such as Germany, Japan, Portugal, we are able to trace the fall in value, which we evenly spread over the time period of the interruption. This smoothing preserves the geometric return, but induces an artificially low volatility and therefore increases the arithmetic return. We should note, however, that the same problem occurs when reported prices are controlled or do not represent transaction data. For the few remaining markets that suffered a permanent interruption, we assume that the market fell by 75 percent the following month.[19]

Table VI presents the performance of the various global stock indices. We focus on performance data first and discuss volatility later. Over the past 76 years, the U.S. stock market provided an arithmetic capital return of 5.48 percent, measured in real terms. Its geometric growth was 4.32 percent over this period. Figure 4 plots the performance of the U.S., global, and non-U.S. real capital growth indices (using the comprehensive series).

[19] The markets affected were Czechoslovakia, Egypt, Greece, Hungary, Poland, and Romania. The 75 percent imputed drop is in line with the fall in value of markets that suffered a severe breakdown. The arbitrariness of the charge is mitigated by the fact that all of these markets are relatively small.

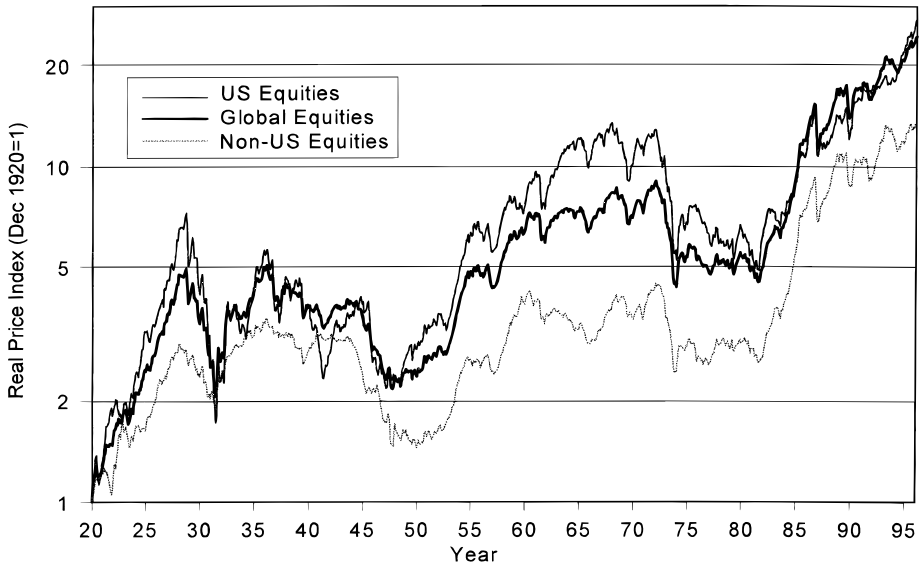

Figure 4. A Global Stock Market Index. The figure displays the performance of the U.S., global, and non-U.S. real capital growth indices. The latter indices are obtained using GDP weights and all existing markets, even if they fail later.

The differences in the performance of the global indices point to the importance of accounting for losing markets. The "survived markets" index has a compound return of 4.33 percent; it accounts only for markets in existence in 1996 and examined since their last break. The "all markets" index has a compound return of 4.04 percent; it accounts for all markets and attempts to interpolate returns over major breaks in the series. Going from the first to the second estimate should move us closer to a true, unbiased measure of long-term return.

At first sight, the difference between the long-term performance of the U.S. index and of the global comprehensive index appears to be small, at only 29 basis points. This result may appear puzzling in light of the evidence in Table I that all non-U.S. markets have lower long-term growth than the United States, often significantly so. One reason for the narrow difference lies in the temporal variation in weights. Consider the Japanese market, for instance. In the first half of the century, the performance of Japanese equities was mediocre. At that time the market carried a weight of less than 4 percent in the global index. In the second half of the century, however, Japanese equities outperformed U.S. equities, precisely at a time when their weight in the index was rising, reaching 16 percent in 1990. Another reason is the large weight in the U.S. market at the beginning of the century. Consider, for example, a $100 investment in global stocks starting in 1921. From the GDP weights in Table V, the amount to allocate to U.S. stocks was $46.17. Over the next 76 years, this amount grew to $1149, using the 4.32 percent

Table VII

Performance of Global Stock Index: 1921–1996

(Nominal Returns in U.S. Dollars, Percentage per Annum)

The table displays the risk and return of dollar returns on stock market indices, translated into U.S. dollars at the official rate. Arithmetic return is obtained from the monthly average multiplied by twelve; risk is monthly volatility multiplied by the square root of twelve; Sharpe ratio is the ratio of monthly average to monthly volatility; geometric return uses annual compounding. Ending wealth reports the final value of $1 invested on December 1920 at the end of the sample. "Survived markets" series includes only markets in our sample in existence in 1996, taken since the last interruption (temporary or permanent). "All markets" series accounts for all markets in the sample, imputing a 75 percent loss in the month the series permanently disappears, or the actual loss spread over the period of the break.

Index	Arithmetic Return	Risk	Monthly Sharpe	Geometric Return	Ending Wealth
U.S. index	8.04	16.19	0.1433	6.95	171.2
Global index					
Survived markets	7.98	13.34	0.1728	7.32	222.9
All markets	7.76	12.14	0.1845	7.25	211.2
Non-U.S. index					
Survived markets	7.53	12.17	0.1785	7.00	176.5
All markets	7.28	12.08	0.1740	6.75	146.2

U.S. growth rate. Let us make now an extreme assumption, which is that all of the money invested outside the United States is lost. Using the $1149-to-$100 ratio, the rate of growth is still 3.26 percent. The large initial size of the U.S. market therefore ensures that the growth on the global index must be within 100 basis points of the U.S. growth number.

The last column in Table VI shows that a difference of 29 basis points can be quite significant over 76 years. Assuming a dollar invested in the U.S. index and in the comprehensive global index, the investments would have grown to 27.3 and 21.9 in real terms, which is a substantial difference.

Table VI also shows that a non-U.S. stock market index, based on our "comprehensive" measure, has grown at the rate of 3.39 percent, which is a full 93 basis points below U.S. equities. If one ignores survivorship issues, however, the return of the non-U.S. index appears to be 4.09 percent. Survival bias therefore induces a difference of 70 basis points in this index, which is quite substantial when accumulated over 76 years.

Table VII presents similar data, measured in nominal U.S. dollars. Over 1921 to 1996, the compound capital return on U.S. equities was 6.95 percent. The return on the global survived index was 7.32 percent; the return on the global comprehensive index was 7.25 percent. Similarly, the average return on the non-U.S. index was 7.00 percent and 6.75 percent. Here the survival bias is on the order of 25 basis points.

As in Table I, we observe that the difference between U.S. and non-U.S. returns is smaller when returns are measured in dollars instead of in real terms. In fact, the return on the unbiased global index is now variation in weights and the real appreciation of most other currencies discussed previously. Also, the return on the value-weighted global index appears not too sensitive to the survivorship issue.

Tables VI and VII also provide estimates of the volatility of the various indices. Using real returns, the volatility of the U.S. index is 15.8 percent. All other indices display lower volatility. For instance, the volatility of the non-U.S. indices is about 10 percent, which is much lower than that of the U.S. market alone, reflecting the fact that the portfolio is spread over a greater number of markets, thus benefiting from imperfect correlations across markets. Next, the risk of our global indices is also driven by correlations. Over the 76 years, the correlation coefficient between returns on the U.S. index and on the comprehensive non-U.S. index is 0.460 in real terms and about the same, 0.452, in dollar terms.[20] As a result of lower volatility for foreign markets and a low correlation coefficient, the risk of the global portfolio is substantially lower than that of U.S. equities. The "comprehensive" global index, for example, displays a volatility of 11.05 percent. Based on these long-term series, the main benefit of going international appears to be risk reduction rather than increased returns.

Taking into account survivorship decreases returns slightly, but also decreases volatility. This is partly due to the (artificial) interpolation of returns when markets are closed, but also because of additional diversification resulting from the inclusion of more markets. We measure the trade-off between risk and return with the Sharpe ratio, defined as average monthly returns divided by their volatility. These are reported in the third columns of Tables VI and VII. With real returns, the Sharpe ratio of the global index is 0.1199, which is higher than that of U.S. equities at 0.0999. With dollar returns, the Sharpe ratio of the global indices is about 0.1845, also higher than that of U.S. equities, at 0.1433. These differences, however, are not statistically significant.[21]

Systematic differences in return can be attributed to two classes of explanations. The first is survivorship, an ex post explanation. The second is rational, ex ante, differences in risk profiles. For example, if markets can be viewed as integrated, a higher return for U.S. equities could be explained by the fact that the U.S. market has a higher world β. Indeed, over the 1921–1996 period, U.S. equities had the highest beta, with a value of 1.24. A regression of real returns on real betas reveals a correlation of 0.53, which is significantly positive.

[20] As for the measurement of volatilities, correlations may be too low because of the smoothing of the series during the breaks. However, the correlation with the survived series is very close, at 0.510 in real terms and 0.520 in dollars. This suggests that the bias is not large.

[21] Using the performance tests developed by Jobson and Korkie (1981).

Testing this proposition is not straightforward because estimation of β with respect to the world index depends on survival issues as well. Had the outcome of the Second World War been different, for example, the β of the United States on the world index would likely have been different. The regression is also afflicted by data and econometric problems. The variables are estimated over different periods and thus have quite different sampling variability. Additionally, the betas that include periods of price controls or infrequent trading are not reliable. Thus it seems difficult to disentangle the higher systematic risk explanation from survivorship to explain the high returns on U.S. equities.

To understand the momentous implications of differences in long-term rates of return reported here, consider the following experiment. First, let us record the current capitalization of non-U.S. equity markets, which was approximately $9,000 billion at the end of 1996. From Table VI, these markets have grown at an average rate of 3.39 percent, which is less than the 4.32 percent growth rate for the United States. Going back to 1921, this implies that the market capitalization of non-U.S. equities was $9,000 billion divided by $(1 + 3.39\%)^{76}$, which amounts to $714 billion in current dollars.

Next, assume that all markets have grown at the U.S. rate of growth. The market value of these equities would then be $714 billion times $(1 + 4.32\%)^{76}$, which amounts to $17,775 billion. In other words, the opportunity cost of growing at about 3.4 percent instead of the 4.3 percent U.S. rate is $8,775 billion in today's dollars. Foreign markets would be double their current size if they had grown only 1 percent faster than they did. Viewed in this context, survival biases of 70 basis points recorded in Table VI are quite significant.

IV. Conclusion

"Financial archaeology" involves digging through reams of financial data in search for answers. Sometimes this involves relying on poor quality data from which to draw inferences about markets in states of crisis. Even so, these data provide invaluable information to help understand long-term histories of capital markets. If one relies on historical data as the basis for estimates of long-term market growth, there is no reason to look at U.S. data only. This is why our paper paints a broad picture of the performance of global stock markets over more than 75 years of a turbulent century for financial markets.

The main lesson from our long-term data is that global capital markets have been systematically subject to dramatic changes over this century. Major disruptions have afflicted nearly all the markets in our sample, with the exception of a few such as the United States. Markets have been closed or suspended due to financial crises, wars, expropriations, or political upheaval.

No doubt this explains our finding that the 4.3 percent real capital appreciation return on U.S. stocks is rather exceptional, as other markets have typically had a median return of only 0.8 percent. These results suggest that the large equity premium obtained in the United States is at least partly to

the result of conditioning estimates on the best performing market. This conditioning may also create time-variation in expected returns; for instance, we expect markets that have done well to exhibit more mean-reversion than others because periods of large losses must be followed by periods of upswings.[22]

This line of analysis treats each market separately. Another approach is to track the hypothetical performance of a diversified global investment. Interestingly, we find that the performance of a globally diversified portfolio is much closer to the performance of U.S. equities, averaging 4.0 percent. This is partly because markets with large capitalization at the beginning of the century performed well. This result also reflects the benefits of diversification, which spreads the risk of dramatic events over a large portfolio.

Whether similar disruptions will happen again is an open question. By now, however, it should be clear that if we fail to account for the "losers" as well as the "winners" in global equity markets, we are providing a biased view of history which ignores important information about actual investment risk.

REFERENCES

Adams, Thomas, and Iwao Hoshii, 1971, *A Financial History of the New Japan* (Kodansha, Tokyo, Japan).

Bank of Japan, 1966, *A Hundred Years of Statistics of Japanese Economy* (Bank of Japan, Tokyo, Japan).

Bansal, Ravi, and Wilbur Coleman, 1996, A monetary explanation of the equity premium, term premium, and risk-free rate puzzles, *Journal of Political Economy* 104, 1135–1171.

Barclays de Zoete Wedd, 1993, The BZW Equity Gilt Study: Investment in the London Stock Market since 1918 (Barclays deZoete Wedd, London, United Kingdom).

Brown, Stephen J., William Goetzmann, and Stephen Ross, 1995, Survival, *Journal of Finance* 50, 853–873.

Burnside, Craig, and Thomas McCurdy, 1992, The equity premium puzzle; in Peter Newman, Murray Milgate, and John Eatwell, eds.: *The New Palgrave Dictionary of Money and Finance* (Stockton Press, New York).

Constantinides, George, 1987, Habit formation: A resolution of the equity premium puzzle, *Journal of Political Economy* 98, 519–543.

Cowles, Alfred, 1939, *Common Stock Indices, 1871–1937* (Cowles Commission for Research in Economics, Monograph no. 3, Principia Press, Bloomington, Ind.).

Epstein, Larry, and Stanley Zin, 1991, Substitution, risk aversion and the temporal behaviour of consumption and asset returns: An empirical investigation, *Journal of Political Economy* 99, 263–286.

Frenneberg, Per, and Bjorn Hansson, 1992, Swedish stocks, bonds, bills, and inflation (1919–1990), *Applied Financial Economics* 2, 79–86.

Gielen, Gregor, 1994, *Konnen Aktienkurse Noch Steigen?* (Gabler, Wiesbaden, Germany).

Goetzmann, William N., 1993, Patterns in three centuries of stock market prices, *Journal of Business* 66, 249–270.

Goetzmann, William N., and Roger Ibbotson, 1994, An emerging market, the New York Stock Exchange 1816–1872, *Journal of Business* 68, 483–508.

[22] Goetzmann and Jorion (1995) also show that survival should induce other effects of interest, such as predictability based on dividend yields.

Goetzmann, William N., and Philippe Jorion, 1995, A longer look at dividend yields, *Journal of Business* 68, 483–508.

Goetzmann, William N., and Philippe Jorion, 1997, Re-emerging markets, {mimeo} University of California at Irvine.

Ibbotson, Roger, 1995, *Stocks, Bonds, Bills, and Inflation: 1995 Yearbook* (Ibbotson Associates, Chicago, Ill.).

International Conference of Economic Services, 1934, *International Abstract of Economic Statistics* (ICES, London, United Kingdom).

International Conference of Economic Services, 1938, *International Abstract of Economic Statistics* (International Statistical Institute, Permanent Office, The Hague, The Netherlands).

International Finance Corporation, 1995, *The IFC Indexes: Methodology, Definitions, and Practices* (International Finance Corporation, Washington, DC).

International Monetary Fund, various issues, *International Financial Statistics* (International Monetary Fund, Washington, D.C.).

Jobson, J. D., and B. Korkie, 1981, Performance hypothesis testing with the Sharpe and Treynor measures, *Journal of Finance* 36, 888–908.

League of Nations, various issues, *Statistical Yearbook* (League of Nations, Geneva, Switzerland).

Mehra, Rajnish, and Edward Prescott, 1985, The equity premium: A puzzle, *Journal of Monetary Economics* 15, 145–161.

Mehra, Rajnish, and Edward Prescott, 1988, The equity premium: A puzzle?, *Journal of Monetary Economics* 22, 133–136.

Mirowski, Philip, 1981, The risk (and retreat) of a market: English joint stock shares in the eighteenth century, *Journal of Economic History* 41, 559–577.

Mitchell, Brian, 1992, *International Historical Statistics: Europe, 1750–1988* (Stockton Press, New York).

Mitchell, Brian, 1993, *International Historical Statistics: The Americas, 1750–1988* (Stockton Press, New York).

Mitchell, Brian, 1995, *International Historical Statistics: Africa, Asia & Oceania, 1750–1988* (Stockton Press, New York).

Neal, Larry, 1987, The integration and efficiency of the London and Amsterdam stock markets in the eighteenth century, *Journal of Economic History* 47, 97–115.

Neal, Larry, 1990, *The Rise of Financial Capitalism: International Capital Markets in the Age of Reason* (Cambridge University Press, Cambridge, Mass.).

Parsons, Brian, 1974, The behavior of prices on the London stock market in the early eighteenth century, Ph.D. dissertation, University of Chicago.

Rietz, Thomas, 1988, The equity premium: A solution, *Journal of Monetary Economics* 22, 117–131.

Schwert, William, 1990, Indexes of U.S. stock prices from 1802 to 1987, *Journal of Business* 63, 399–426.

Siegel, Jeremy, 1992, The equity premium: Stock and bond returns since 1802, *Financial Analysts Journal* 48, 28–38.

Siegel, Jeremy, 1994, *Stocks for the Long Run* (Richard D. Irwin, New York).

Timmerman, Allan, 1992, Changes in Danish stock prices 1914–1990, *Nationalokonomisk Tidsskrift* 130, 473–482.

United Nations, various issues, *Monthly Bulletin of Statistics* (United Nations, New York).

Wilson, Jack, and Charles Jones, 1987, A comparison of annual common stock returns, *The Journal of Business* 60, 239–258.

Wydler, Daniel, 1989, Swiss stocks, bonds, and inflation: 1926–1987, *Journal of Portfolio Management* 15, 27–32.