

DISCUSSION PAPER SERIES

No. 1119

**THE EQUITY PREMIUM AND THE RISK
FREE RATE: A CROSS COUNTRY, CROSS
MATURITY EXAMINATION**

Fabio Canova and Gianni De Nicolò

INTERNATIONAL MACROECONOMICS



Centre for Economic Policy Research

THE EQUITY PREMIUM AND THE RISK FREE RATE: A CROSS COUNTRY, CROSS MATURITY EXAMINATION

Fabio Canova and Gianni De Nicolo

Discussion Paper No. 1119
January 1995

Centre for Economic Policy Research
25–28 Old Burlington Street
London W1X 1LB
Tel: (44 71) 734 9110

This Discussion Paper is issued under the auspices of the Centre's research programme in **International Macroeconomics**. Any opinions expressed here are those of the author(s) and not those of the Centre for Economic Policy Research. Research disseminated by CEPR may include views on policy, but the Centre itself takes no institutional policy positions.

The Centre for Economic Policy Research was established in 1983 as a private educational charity, to promote independent analysis and public discussion of open economies and the relations among them. It is pluralist and non-partisan, bringing economic research to bear on the analysis of medium- and long-run policy questions. Institutional (core) finance for the Centre has been provided through major grants from the Economic and Social Research Council, under which an ESRC Resource Centre operates within CEPR; the Esmée Fairbairn Charitable Trust; and the Bank of England. These organizations do not give prior review to the Centre's publications, nor do they necessarily endorse the views expressed therein.

These Discussion Papers often represent preliminary or incomplete work, circulated to encourage discussion and comment. Citation and use of such a paper should take account of its provisional character.

January 1995

ABSTRACT

The Equity Premium and the Risk Free Rate: A Cross Country, Cross Maturity Examination*

This paper examines the relationship between the equity premium and the risk free rate at three different maturities using post-1973 data for a panel of seven OECD countries. We show the existence of subsample instabilities, of some cross country differences and of inconsistencies with the expectations theory of the term structure. We perform simulations using a standard consumption based CAPM model and demonstrate that the basic features of Mehra and Prescott's (1985) puzzle remain, regardless of the time period, the investment maturity and the country considered. Modifications of the basic set-up are also considered.

JEL Classification: C15, E43, G12

Keywords: equity premium, risk free rate, term structure, consumption based CAPM, calibration, model evaluation

Fabio Canova
Department of Economics
Universitat Pompeu Fabra
Balmes 132
Barcelona 08008
SPAIN
Tel: (34 3) 542 2601

Gianni De Nicolo
Department of Economics
Brandeis University
Waltham
MA 02254-9110
USA
Tel: (1 617) 736 2240

*This paper is produced as part of a CEPR research programme on *Macroeconomics, Politics and Growth in Europe*, supported by a grant from the Commission of the European Communities under its Human Capital and Mobility Programme (no. ERBCHRXCT930234). This research was started when both authors were associated with the European University Institute, Florence. We would like to thank Andy Abel, Silverio Foresi, Rene Garcia, Alan Timmerman and participants of the 1993 SEDC Annual Meeting, Nauplio, Greece, the Second European Summer Symposium in Macroeconomics, and of seminars at the University of Aarhus, Montreal, Birkbeck College, Wharton, and the European University Institute for comments and suggestions.

Submitted 7 December 1995

NON-TECHNICAL SUMMARY

The historical magnitude of the equity premium in the United States has been the object of intense study over the last decade. Since the seminal work of Mehra and Prescott (1985) many authors have modified the basic theoretical model to account for the wide discrepancy between the time series generated by a complete markets Arrow-Debreu economy and the data. The existing literature has, however, disregarded several empirical issues which need to be addressed before one can safely consider the relationship between returns on equities and on risk free assets, an important empirical regularity to be explained by theory.

First, the presence of a large average equity premium and of a small average risk free rate has been documented almost exclusively for the United States and few have investigated whether such a phenomenon exists in other countries. Second, the historical features of the equity premium and of the risk free rate have changed repeatedly over decades. Does the choice of sample period have any influence on our perception of the economic relevance of the phenomenon, and would data which is more homogeneous than typically employed, say, only post-World War II data, or post-1973 data, change the basic features of the relationship? Third, the empirical properties of the equity premium have been documented under the assumption that the sampling period of the data is identical to the holding period of the investment. If the pure expectations theory of the term structure does not hold, restricting attention to a holding maturity which is equal to the sampling period throws away important information. A study of the properties of the equity premium for a number of maturities may therefore indicate the extent of the information waste.

The task of this paper is two-fold. First, we want to provide evidence on these three empirical issues. We characterize the equity premium risk free rate relationship in a number of industrialized countries for the post-1973 era, for buy-and-hold investments of three different maturities. Second, we want to know whether a standard Arrow-Debreu model can account for the time variations, the cross country and the cross maturity heterogeneity present in the data only by means of heterogeneity in the deep parameters of the model. In this respect the term structure of the equity premium-risk free rate pair offers a much more challenging term of comparison to judge whether theory can account for the data and whether some of the suggested modifications are successful or not.

The first part of the paper documents the time series properties of the equity premium-risk free rate pair for seven countries (Canada, Germany, France, Italy, Japan, the United Kingdom and the United States) for three samples (1973–91, 1973–81 and 1982–91) and for three holding maturities (3, 6 and 12 months). We show that within countries there are important statistical differences across subsamples, that these differences emerge both in the slope and in single elements of the term structure, that they are primarily due to the changes in the properties of the risk free rate over time and that they involve both first and second moments. We also show that the term structures of the equity premium risk free rate pairs display important cross-country heterogeneities regardless of the sample period used, that these differences are evident even among countries which are geographically and economically more integrated, that they are due primarily to cross country differences in the time series properties of the risk free rate and that they are more evident in the mean than in the second moment of the distributions. Finally, we document that the expectations theory of the term structure for the equity premium is violated in almost all countries for two samples (the exception is the 1973–81 subsample) and that violations are more important at the shortest end of the term structure.

The second part attempts to confront a standard Arrow-Debreu model with the cross country, cross sample and cross maturity evidence we have discovered. The maintained hypothesis is that countries differ only in preference and technological parameters and not in their market set-ups or institutional arrangements. We outline a formal evaluation procedure which allows us to characterize the discrepancy between the distribution of a vector of statistics of actual and simulated equity premium-risk free rate pairs for each country, maturity and subsample.

The basic results of our simulations suggest that simulated and actual data differ substantially: in most cases the distribution of the moments of actual and simulated equity premium risk free rate pair hardly overlap. Moreover, regardless of the country, the model generates values for the mean and the standard deviation of the equity premium which are too low to be consistent with the cross sectional post Bretton Woods experience. We also demonstrate that, regardless of the maturity, the model is relatively more successful for the first subsample than for the second. This should not come as a surprise, however, since over the period 1973–81 the means of both the risk free rate and the equity premium were negligible or even negative in some countries. In addition, we show that regardless of the country and the time period, the model appears to be slightly more successful for investments lasting 12 months than for shorter investment maturities. The model appears to be less

inadequate in reproducing actual moments of the equity premium-risk free rate pair for Italy than for the other countries.

Finally, we examine how certain modifications of the basic model suggested in the literature are likely to impact on the results. In particular, we discuss the question of heteroskedasticity in the driving forces of the economy the issue of inflation and the problem of leverage. We argue that none of these modifications is likely to improve the performance of the model in all the dimensions examined and that alternative explanations are needed to account for the cross country, cross maturity heterogeneity we have documented.

Capitano, che risolvi con l'astuzia ogni avventura,
ti ricordi di un soldato che ogni volta ha piu' paura?
Ma anche la paura in fondo mi da' sempre un gusto strano
se ci fosse ancora mondo sono pronto, dove andiamo?

Lucio Dalla

1 Introduction

The historical magnitude of the equity premium in the US has been the object of intense study in the last decade. Since the seminal work of Mehra and Prescott (1985) many authors, including Rietz (1988), Weil (1989), Labadie (1989), Epstein and Zin (1990) Constantinides (1991), Mankiw and Zeldes (1991) and Checchetti, Lam and Mark (1993), have modified the basic theoretical model to account for the wide discrepancy between the time series generated by a complete markets Arrow-Debreu economy and the data. Some of these modifications have shed new light on the problem and have contributed to advance our understanding of how simple artificial economies work. The existing literature has however disregarded several empirical issues which need to be addressed before one can safely consider the relationship between returns on equities and on risk free assets an important empirical regularity to be explained by theory. First, the presence of a large average equity premium and of a small average risk free rate has been documented almost exclusively for the US and, to the best of our knowledge, only Siegel (1992) has investigated whether such a phenomenon exists in other countries as well. Second, the historical features of the equity premium and of the risk free rate have changed repeatedly over decades (see e.g. Mehra and Prescott (1985, p.147)). One may wonder whether the choice of sample period has any influence on our perception of the economic relevance of the phenomenon, and whether considering data which is more homogeneous than typically employed, say, only post WWII data, or post 1973 data, may change the basic features of the relationship. Blanchard (1993), for example, claims that for the US, the equity premium has declined substantially since the early 1950's. Third, the empirical properties of the equity premium have been documented under the assumption that the sampling period of the data is identical to the holding period of the investment. If the pure expectations theory of the term structure does not hold, restricting attention to a holding maturity which is equal to the sampling period throws away important information. A study of the properties of the equity premium for a number of maturities may therefore indicate the extent of the information

waste and provide a new perspective on the phenomenon.

The task of this paper is two-fold. First, we want to provide evidence on these three empirical issues. We characterize the equity premium-risk free rate (EP-R) relationship in a number of industrialized countries for the post 1973 era, for buy-and-hold investments of three different maturities. The sample period is selected keeping in mind that the breakdown of the Bretton Wood system has, on one hand, thickened the size of domestic equity markets around the world, and, on the other, increased the integration of capital markets across countries. Therefore, by choosing this sample period, we give the best chance to the phenomenon to have an international dimension. The countries of the panel are chosen because the total value of their stock and T-bill markets constitutes over 70% of the world market and because they are sufficiently homogeneous as a group to make the comparison meaningful. Finally, the choice of investment maturities is dictated by the sample availability. In presenting the empirical evidence we maximize the production of information and organize it in such a way to be useful to those who are interested in confronting theoretical models with data. Second, we want to know whether a standard Arrow-Debreu model can account for the time variations, the cross country and the cross maturity heterogeneity present in the data only by means of heterogeneity in the deep parameters of the model. In this respect, the term structure of the (EP-R) pair offers a much more challenging term of comparison to judge whether theory can account for the data and whether some of the suggested modifications are successful or not.

The paper is organized in six sections. The next section documents the time series properties of the (EP-R) pair for seven countries (US, Canada, UK, Japan, Italy, Germany and France) for three samples (1973-1991, 1973-1981 and 1982-1991) and for three holding maturities (3, 6 and 12 months). For completeness, we also study the time series properties of a pair of equally weighted portfolios, composed of stocks and T-bills for the seven countries over the same samples and for the same three maturities. For each country we present summary statistics describing the term structure of the (EP-R) pair across samples and examine whether these distributions are stable across time, are similar across countries, and satisfy the expectations theory of the term structure. We show that, within countries, there are important statistical differences across subsamples, that these differences emerge both in the slope and in single elements of the term structure, that they are primarily due to the changes in the properties of the risk free rate over time and that they involve both first and second moments. We also show that the term structures of the (EP-R) pairs display important cross-country heterogeneities regardless of the sample period used, that

these differences are evident even among countries which are geographically and economically more integrated, that they are due primarily to cross country differences in the time series properties of the risk free rate and that they are more evident in the mean than in the second moment of the distributions. Finally, we document that the expectations theory of the term structure for EP is violated in almost all countries for two samples (the exception is the 1973-1981 subsample) and that violations are more important at the shortest end of the term structure.

In order to confront a standard Arrow-Debreu model with the cross country, cross sample and cross maturity evidence, we briefly recall in section 3 the theoretical structure employed by Mehra and Prescott and describe the relationship between the deep parameters of the model and the moments of EP and R after imposing simple distributional assumptions on the exogenous processes of the economy. The maintained hypothesis is that countries differ only in preference and technological parameters and not in their market setups or institutional arrangements. We are interested in knowing if there is sufficient variability in these parameters across time, maturity and countries to generate the heterogeneities we see in the data.

In section 4 we outline a formal evaluation procedure which allows us to characterize the discrepancy between the distribution of a vector of statistics of actual and simulated (EP-R) pairs for each country, maturity and subsample. The approach we employ is similar in spirit to the one of Canova (1994) and allows us to draw inference about the quality of the model's approximation to the data in a quasi-bayesian fashion. It is worth stressing that such an approach is more informative than those typically employed in previous work (with the exception of Checchetti, Lam and Mark (1993)) because we explicitly consider parameter uncertainty in the simulations.

Section 5 describes the basic results of our simulations. The probabilistic measures of fit suggest that simulated and actual data differ substantially: in most cases the distribution of the moments of actual and simulated (EP-R) pair hardly overlap. Moreover, regardless of the country, the model generates values for the mean and the standard deviation of the equity premium which are too low to be consistent with the cross sectional post Bretton Wood experience. We also demonstrate that, regardless of the maturity, the model is relatively more successful for the first subsample than for the second one. However, this should not come as a surprise since over the period 1973-1981 the means of both the risk free rate and the equity premium were negligible or even negative in some countries. In addition, we show that regardless of the country and the time period, the model appears to be slightly more successful for investments lasting 12 months than for shorter investment maturities.

Finally, the model appears to be less inadequate in reproducing actual moments of the (EP-R) pair for Italy than for the other countries. Overall, these results suggest that there is an international dimension to the (EP-R) puzzle, that exists regardless of the sample period considered and that, in general, worsens as the investment period shortens. There are however country specific features even in a sample period when financial markets became more integrated. Also, the subsample results suggest that the time path of inflation may be important in determining the economic significance of the phenomenon (as suggested also by Blanchard (1993)).

In section 6 we examine how certain modifications of the basic model suggested in the literature are likely to impact on the results. In particular, we discuss the question of heteroskedasticity in the driving forces of the economy (as in Kandel and Stambaugh (1990) or Canova and Marrinan (1993)), the issue of inflation (as in Labadie (1989)) and the problem of leverage (as in Benninga and Protopapadakis (1990) or Kandel and Stambaugh (1991)). We argue that none of these modifications is likely to improve the performance of the model in all the dimensions examined and that alternative explanations are needed to account for the cross country, cross maturity heterogeneity we have documented. Section 7 concludes indicating avenues of future research.

2 The Term Structure of the Equity Premium-Risk Free Rate

This section documents the quarterly time series properties of the (EP-R) pair for the US, Canada, UK, Japan, Italy, Germany and France for the period 1973,1-1991,4 and for two subperiods (1973,1-1981,4 and 1982,1-1991,4) for buy-and-hold investments of 3, 6, and 12 months and for two equally weighted portfolios, one solely composed of risk free assets and one composed of stocks and risk free assets for the seven countries. The definition of the variables, the data employed and their sources are described in appendix A. Table 1 presents estimates of the mean, standard deviation and AR(1) coefficient for the (EP-R) pair for the whole sample and table B.1 presents the same statistics for the two subsamples. To avoid distortions due to the overlapping nature of investments which last longer than the sampling interval, the second moments of 6 and 12 month investments are computed averaging the corresponding second moments obtained from k nonoverlapping series whose starting date of the investment is moved, successively, by one quarter, where $k=2$ if the investment period is 6 months and $k=4$ if the investment period is 12 months. The subperiod division we employ is somewhat arbitrary but the break point is chosen keeping in mind the behavior of inflation during

the two subperiods (high in the first subsample, low in the second). Garcia and Perron (1993) show that the process for the real risk free rate in the US displays a breaking point at 1981,3 due to changes in Fed policies. Because after that date real interest rates moved to a higher mean level all over the world, it is likely that this date is also crucial for the remaining G-7 countries. It is worth stressing that in calculating the equity premium for maturity k we use the standard procedure of subtracting from the equity return for a k period investment the real rate on a 3 month T-bill compounded $k/3$ times. This is not a completely satisfactory procedure for our purposes because for investments that last longer than 3 months, the calculated equity premium is the sum of a risk premium, due to the fact that stocks are more risky than T-bills for a given maturity, and of a term premium, due to the fact that we use T-bills of a different maturity (see, e.g., Campbell (1986)). However, many countries do not issue T-bills of the required maturity and proxies constructed using either Eurodeposit rates or holding returns on long term bonds are likely to be inadequate as well either because for some currencies euromarkets are very thin or because long term bonds have different risk characteristics than T-bills. Finally, because the CPI bundles we use to convert nominal returns into real differ across countries, term structure differences across countries may be the result of the mismeasurement of real returns. To quantify the extent of this problem, we also computed real returns using GNP deflators, which are more homogeneous across countries, in place of CPI's and found that the reported statistics are insensitive to this change.

There are several aspects of the two tables which deserve some attention. All term structures are upward sloping in the whole sample and in the second subsample, while in the first subsample, which contains the high inflationary period of the 70's, the evidence is mixed. The steepest term structure is for Canada: its risk free rate is, on average and for all maturities, higher than those of other countries but this is due primarily to the 82-91 period. Individual elements of the term structure show large standard deviations with the risk free rate displaying considerably less variability than the equity premium. Also, the standard deviations of the (EP-R) pair increase almost linearly with the holding period. The autocorrelation properties of the two series differ substantially. The equity premium is almost uncorrelated while the AR(1) coefficient for the risk free rate is high. The exceptions to this rule are Italy and Germany for three month investments. In Italy the two variables have the same AR(1) coefficients while in Germany the AR(1) coefficients of EP is larger than the AR(1) coefficients of the risk free rate but both of them are small. It is also worthwhile to note the tendency of the AR(1) coefficient of EP to decrease with the maturity of the investment

and to turn negative for longer investment horizons in all countries, a feature which suggests the presence of mean reversion characteristics in the data (see, e.g., by Fama and French (1988)). The serial correlation properties of the (EP-R) pair are somewhat different across subperiods, with a decrease in the AR(1) coefficient of both variables in the second sample. Finally, the time series properties of the two portfolios are somewhat intermediate between those of the various countries with lower standard deviations and AR(1) coefficients.

To make a direct comparison with Mehra and Prescott calculations it is useful to compute what the mean of the (EP-R) pair would be for the 1973-1991 sample if one had used annual data. Using simple compounding of US returns we obtain that the average equity premium for the period is 5.56 while the risk free rate is 1.62. Therefore, while the EP is lower and the risk free rate is higher than those reported by Mehra and Prescott, they are in line with those reported by Bonomo and Garcia (1993) for the sample 1889-1987.

Next, we examine the stability of the distribution of the term structure of (EP-R) pair over subsamples for each country and their equality across countries for each sample period. We have two goals in mind. First, we would like to know how reliable is a description of the EP-R relationship which uses data from the entire sample. Mehra and Prescott, for example, report that the distribution of the EP-R pair has changed substantially over decades (see table 1, p.147). Siegel (1992) and Blanchard (1993) present similar evidence. Second, we would like to know whether the cross country differences we noted are accidental, in which case restricting the analysis to one country is sufficient, or whether there is additional information in the international cross section of data which can be used to shed new light on the phenomenon.

To examine both hypotheses we use a distance-type test of the form

$$Q = (\mathbf{x}_1 - \mathbf{x}_2)\Sigma^{-1}(\mathbf{x}_1 - \mathbf{x}_2)' \quad (1)$$

where \mathbf{x}_1 and \mathbf{x}_2 are vectors of estimated moments, either across subsamples or across countries, and Σ is the covariance matrix of $\mathbf{x}_1 - \mathbf{x}_2$. Under the null that \mathbf{x}_1 and \mathbf{x}_2 are identical and normally distributed, Q is distributed as a $\chi^2(m)$ where $m = \dim(\mathbf{x}_1) = \dim(\mathbf{x}_2)$. Because the distribution of returns typically displays tails which are fatter than the normal ones, it is necessary to check if the normality assumption for estimates of the means, standard deviations and AR(1) coefficient is appropriate. Using Kendall and Stuart tests for normality, we find that, if we exclude from the data the fourth quarters of 1987 and 1989, there is no evidence of excess kurtosis and only a slight

positive skewness for EP in the US and UK. Therefore, to a first approximation, the asymptotic distribution we use for the tests is appropriate. Table 2 reports the p-values of the tests for the equality of the means, of the standard deviations and of the AR(1) coefficient of EP and of the (EP-R) pair across subperiods jointly at 3, 6 and 12 months maturity for each of the seven countries and for the two portfolios. Together with the single country tests we also report a joint test examining the equality of the mean, standard deviation and AR(1) coefficient for the three maturities for the vector of seven countries.

The table shows very clearly that the term structure is unstable across subperiods for each of the seven countries. When we look in detail at what has changed over time, we find that it is the distribution of the real risk free rate that has been substantially altered in all countries and for all maturities. In contrast, the distribution of EP appears to be more stable and in some countries (e.g. Canada or Italy) the first two moments of the distribution of EP have not significantly changed over time. This result is somewhat surprising, as it indicates that real equity returns and the real risk free rate move together over subsamples, contrary to the characterization offered by Blanchard (1993) for the US. Overall, the mean and the AR(1) coefficient of EP appear to be less stable than its standard deviation and for four countries in the panel (Canada, Italy, US and France) the standard deviations of EP has not significantly changed across subsamples. Looking at single elements of the term structure, we find that at least one moment of the distribution of R has changed at each maturity, while differences across subsamples in EP emerge primarily in the mean for 12 month investments. Very similar results are obtained when we examine the behavior of the two portfolios over the two subsamples.

To complete the characterization of the (EP-R) relationship across sample periods it is useful to look at the slope of the joint (EP-R) mean term structure over subperiods. The slope of the mean term structure is an important ingredient used to test the expectations theory of the term structure (see e.g. Campbell and Shiller (1992)) and our analysis may give a rough idea of the magnitude of the changes in the distribution of the liquidity term across subperiods. Table 3, which reports a joint test for the equality of the slopes of the mean term structure for the 7 countries across subperiods and a test for the portfolios, suggests that not only the levels but also the slopes display statistical changes. Major differences emerge in the slope between 12 and 6 month maturities, while changes in the slope between 6 and 3 month maturities are statistically significant, but smaller in size. Once again differences in the mean slopes of the term structure are due to changes in the

slope of the mean risk free rate. Among the countries in the panel, the slope of the term structure in Canada is the one which shows the largest difference across subperiods. For the portfolios the results are similar except that the slope between 3 and 6 months maturities for the (EP-R) pair are no longer significantly different across subperiods.

In sum, there appear to be subsample instabilities over 1973-1991 period. Because these instabilities may be linked to the properties of inflation and to policy choices (for example, targeting a nominal interest rate or targeting monetary aggregates), and because the first subsample seems at odds with the empirical stylization offered by Mehra and Prescott, we conclude that there are periods when the (EP-R) relationship may strongly deviate from the characterization established in the literature. The discrepancy across subperiods is more evident for investment with horizons longer than 6 months and appears to be due, to a large extent, to the behavior of the risk free rate.

Next, we examine whether the distribution of the term structure of the (EP-R) pair is similar across countries. We conduct tests on single moments of the distribution jointly for 3, 6 and 12 month maturities and on the vector comprising the mean, the standard deviation and the AR(1) coefficient using moments of the term structure in the US as a term of comparison for the entire sample and for each of the two subperiods. The p-values of the tests are presented in table 4. We find that the mean of the three elements of the term structure of the (EP-R) pair displays differences across countries. Differences with the mean US term structure emerge in all three samples but they are more marked in the second subsample. The standard deviations however show no significant differences across countries and the AR(1) coefficients are different primarily for Germany, Japan and Italy. The joint test on the six term structures of the (EP-R) pair rejects the null hypothesis that they are identical to that of the US primarily because the mean of the risk free rate displays substantial differences across countries in each subperiod. The results obtained comparing the term structure of the portfolios with the term structure in the US are similar. The only additional interesting feature is that, overall, the statistical tests reject the hypothesis that the two term structures have very similar moments for all samples.

To check whether the results are sensitive to the choice of the US as benchmark, we also conducted tests using the term structure of Germany as a term of comparison. This alternative test is useful from two different perspectives. First, to examine whether the US term structure has special features while the term structures of other countries display a much higher degree of conformity. Second, to see whether the four European countries in the panel, which experienced

similar government policies and market arrangements over the sample period, display stronger similarities which would allow us to treat them as a block in international comparisons. The results we obtain are very similar to those reported in table 4, confirm the diversity in the moments of the risk free rate across countries, and are omitted for reason of space.

Almost all work we are aware of which tries to account for the (EP-R) relationship via simulation exercises, considers investments whose holding maturity correspond to the sampling frequency of the data used, i.e, if annual data are available, only buy-and-hold investments which last one year are considered. Implicit in this approach is the assumption that the pure expectations theory of the term structure holds. In this case equity premia on longer term investments are simple averages of the equity premium on short term investments and considering one maturity is a sufficient statistic to characterize the dynamics of the entire term structure. Put in another way, if the pure expectations theory of the term structure holds, the rolling premium, defined as the mean excess return on a h -month investments over rolling m times a (k/m) -month investments, is zero, for all h, k, m with $k < h$ (see e.g. Campbell and Clarida (1987)).

The standard approach of considering a holding maturity which is equal to the frequency of the data is partially justified by a result of Mehra and Prescott (1985). They show, in fact, that because consumption growth is approximately uncorrelated over time, the term structure of simulated EP must satisfy the pure expectations theory. However, to the best of our knowledge, this hypothesis has not been tested in the data. Therefore we examine whether the term structure of EP in each of the seven countries is consistent with the assumption that the rolling premium is zero at all maturities. We construct the rolling premium between 6 and 3 month maturities and 12 and 6 month maturities and examine the relationship, for the full sample and for each of the 2 subsamples, for single maturities and for a vector comprising the two rolling premia for each country. The p-values of the tests are reported in table 5. In general, the rolling premium appears to be significantly different from zero. However, there are exceptions. For example, in the first subperiod, the two rolling premia are not statistically different from zero except for Germany at the 6-3 maturity. Similarly, in the second subsample the rolling premium for Canada, UK and Japan at 12-6 maturity is not significantly different from zero. Overall, deviations from the expectations theory of the term structure are more evident for short than for long maturities.

Several conclusions can be drawn at this point. First, the joint distribution of the (EP-R) pair is unstable over time. This is true of single elements of the term structure in each country as

well as of slopes between various maturities. Changes over time in the mean of the risk free rate are the major reason for these instabilities. Hence, by separating historical episodes with different time series characteristics we may have a better chance to understand whether standard theory fails because of incorrect assumptions or intrinsic weaknesses. Second, the joint distribution of the term structure of the (EP-R) pair is significantly different across countries. These differences emerge primarily in the mean and are, to a large extent, due to the differences in the time series properties of the risk free rate across countries. Therefore, there is a scope in trying to confront numerical versions of Arrow-Debreu economies with data from countries other than the US. Third, there is information in the term structure of the (EP-R) pair which is neglected by considering an investment maturity which is equal to the sampling interval of the data. Available data on EP across countries does not seem to satisfy a pure expectations theory which would justify focusing attention on one maturity only. While there are exceptions to the tendencies we have outlined, the results suggest the need, on one hand, for a deeper look into the statistical properties of buy-and-hold strategies and, on the other, to confront numerical versions of Arrow-Debreu models with the cross country, cross maturity, cross sample heterogeneity we have unveiled.

3 A Model of the Equity Premium-Risk Free Rate Relationship

In this section we briefly review the standard consumption based CAPM model used to analyze the (EP-R) relationship. Two maintained hypotheses are made. First, countries can differ in preferences and technological parameters, but not in institutional setups or market arrangements. Although this is a simplifying assumption, it is useful for it provides a benchmark to compare more complicated versions of the model where liquidity or institutional constraints may be introduced. Moreover, it imposes substantial discipline in the simulation exercises since it forces us to account for the heterogeneities we have presented in section 2 only by means of differences in the “deep” parameters of the model. Second, countries’ financial markets are autarkic. We chose this setup because, under the opposite extreme of perfect capital markets integration, the model predicts that investors would hold the world market portfolio of risky assets independently of their country of residence, an implication empirically rejected for most countries in the panel (see French and Poterba (1991)). Therefore, by assuming away financial market integration, we give the model the best chance to explain the heterogeneity previously described by means of differences in countries’

preference and technological parameters. However, for the sake of comparison we also report results for a portfolio which approximates the one a world investor would hold in a perfectly integrated capital market.

After describing the model and because the cross maturity, cross country and cross time heterogeneity we have found appears to be connected with differences across countries and time in the properties of the risk free rate, we will attempt to identify those parameters which affect the risk free rate but not the equity premium and to see whether heterogeneity present in these parameters across time and countries can account for the cross sectional differences.

The model is a standard frictionless pure exchange Arrow-Debreu economy with a single representative agent, one perishable consumption good produced by a single productive unit or "tree" and $K + 1$ assets, an equity share and K risk free assets of maturity $k = 1, \dots, K$, one per maturity. The tree yields a random dividend each period and the equity share entitles its owner to that dividend. The K risk free assets entitle its owner to one unit of the consumption good at maturity. The representative agent maximizes:

$$E_0 \sum_{t=0}^{\infty} \beta^t \left[\frac{c_t^{1-\alpha} - 1}{1-\alpha} \right] \quad (2)$$

subject to :

$$c_t = y_t e_{t-1} + p_t^e (e_{t-1} - e_t) + \sum_{k=1}^K f_{t-k,k} - \sum_{k=1}^K p_t^{f,k} f_{t,k} \quad (3)$$

where c_t is consumption, y_t is the tree's dividend, p_t^e and $p_t^{f,k}$ are the prices of the equity and of the k -th risk free asset, e_t and $f_{t,k}$ are the agent's holdings of equity and of the k -th risk free asset, E_0 is the mathematical expectation operator conditional on information at time zero, β is the discount factor and α is the relative risk aversion parameter. Dividends evolve according to:

$$y_{t+1} = x_{t+1} y_t \quad (4)$$

where x_{t+1} denotes the gross growth rate of dividends. In equilibrium $c_t = y_t$, $e_t = 1$ and $f_{t,k} = 0 \forall t, k$. The returns for the K riskless securities and on k -period buy-and-hold equity investments in each country satisfy the following optimality conditions :

$$1 = \beta^k E_t \left(\prod_{i=1}^k x_{t+i} \right)^{-\alpha} (1 + R_{t,k}) \quad k = 1, \dots, K \quad (5)$$

$$1 = E_t \sum_{i=1}^{k-1} \beta^i \left(\prod_{l=1}^i x_{t+l} \right)^{-\alpha} \frac{y_{t+i}}{p_t^e} + \beta^k \left(\prod_{l=1}^k x_{t+l} \right)^{-\alpha} (1 + R_{t,k}^e) \quad k = 1, \dots, K \quad (6)$$

where $1 + R_{t,k} = 1/p_{t,k}^f$ denotes the risk-free gross return for maturity $k = 1, \dots, K$ and $1 + R_{t,k}^c = (p_{t+k}^c + y_{t+k})/p_t^c$. Closed form expressions for the average risk-free rate of maturity k and the equity premium for investment horizon k in terms of the parameters of agents' preference and technology can be obtained under some distributional assumptions. Following Aiyagary (1993) we let $x_t = \exp(\mu + \epsilon_t)$ and $1 + R_{t,k}^c = (1 + E(R_{t,k}^c)) \exp(u_t - \sigma^2/2)$ where ϵ_t and u_t are i.i.d. normal random variables with 0 mean and variances δ^2 and σ^2 , respectively and μ is the mean of x_t .

Using (5) and (6) and the approximation $\ln(1+z) \approx z$ for small z , the average risk-free rate for maturity k , denoted by \bar{R}_k , and the average equity premium for investment horizon k , denoted by $\bar{E}P_k = \bar{R}_k^c - \bar{R}_k$, are given by:

$$\bar{R}_k = -k \ln(\beta) + k\alpha\mu - 0.5k^2\alpha^2\delta^2 \quad (7)$$

$$\bar{E}P_k = \alpha \text{cov}(\ln(1 + R_{t,k}^c), (\sum_{i=1}^k \ln(x_{t+i}))) + (\alpha - 1) \ln\left(\frac{y_t}{p_t} \sum_{i=1}^{k-1} \beta^i E_t \prod_{l=1}^i x_{t+l}\right) \quad (8)$$

where the second term in equation (8) comes from the fact that dividends may be paid at intermediate dates i of the investment period k . Equation (8) represents a version of the relationships derived by Abel (1988) and Black (1990) when the conditional moments of the dividend process are constrained to be time independent and independent of wealth. Equations (7) and (8) explicitly show the dependence of the average risk-free rate and the average equity premium on the technology parameters (μ, δ) , on the preference parameters (α, β) and on the unconditional covariance between the risky asset return and consumption growth. Note that, given technological parameters and the covariance between consumption growth and asset returns, variations in β affect the mean risk free rate only, while variations in α have a monotonic effect on the mean of the equity premium and a non-monotonic one on the mean of the risk free rate. In other words, there is a range of α which induces opposite movements in $\bar{E}P_k$ and \bar{R}_k .

In general, the expressions for the standard deviations and the AR(1) coefficients depend in a nonlinear way on the differences between the conditional and unconditional distributions of the exogenous driving forces of the model (see also Canova and Marrinan (1992)). For example, the standard deviations of (EP-R) depend on the differences between conditional and unconditional moments of the dividend growth process and on the differences between the conditional and the unconditional covariance between risk returns and dividend growth. For the simple case we are considering here, conditional and unconditional moments are identical so that second moments are

degenerate. In other words, with the log linear approximation we used, the distribution of simulated (EP-R) pair collapses to a point mass and only uncertainty in the parameters generates uncertainty in the outcomes of the model. We will use this feature in designing our evaluation procedure in the next section.

4 Evaluation Procedure

To study the properties of the model for each country, maturity and sample, we use Mehra and Prescott's version of the model where the gross growth rate of dividends x_t follows an ergodic first order Markov chain with probability $P(x_{t+1} = x_j | x_t = x_i) = \phi_{ij}$. As in Mehra and Prescott we specify the process for consumption to have two states of the form $\lambda_1 = 1 + \mu + \delta$; $\lambda_2 = 1 + \mu - \delta$ and restrict the one-period transition matrix to satisfy $\phi_{1,1} = \phi_{2,2} = \phi$ and $\phi_{1,2} = \phi_{2,1} = 1 - \phi$. Several authors, including Abel (1992) and Cecchetti, Lam and Mark (1993) have argued that the assumption that, in equilibrium, consumption equals dividends is a gross misspecification and suggested calibrating the model to a bivariate process for consumption and dividends. We do not consider this possibility here, but we argue in section 6 that it is unlikely to be crucial in reconciling the model and the data.

There are many ways to take into account the uncertainty surrounding the choice of the five parameters of the model $\theta = (\mu, \delta, \phi, \beta, \alpha)$. The standard approach is to calibrate the three technology parameters (μ, δ, ϕ) so that the mean, the standard error and the AR(1) coefficient of the model's consumption match those of the growth rate of annual US consumption over the particular sample period and obtain values for the risk free rate and for the equity premium by grid searching the preference parameters (β, α) over a prespecified interval. Implicit in this procedure is the assumption that technology parameters can be pinned down with sufficient precision, while the sampling uncertainty surrounding point estimates of the preference parameters is so large that a researcher places no confidence in any particular value and simply chooses a reasonable range for (β, α) on the basis of theoretical considerations. In the more formal version of this procedure adopted by Cecchetti, Lam and Mark (1993) one "estimates" the range of (β, α) by choosing those pairs generating (EP-R) which fall within the 95% interval around the point estimates of the mean of (EP-R) found in the data. With this last procedure one can (i) explore the shape of the model's EP to variations in α and β in the chosen range and (ii) construct measures of dispersion for the

statistics of interest using the estimated distribution for (β, α) pair. Here we take a different approach which explicitly considers the uncertainty surrounding estimates of all parameters. Instead of calibrating (μ, δ, ϕ) and searching for the values of (β, α) that best replicate the actual mean of the (EP-R) pair or fall within a prespecified region around the mean of the (EP-R) pair, we estimate all parameters from the data, construct their empirical distribution, perform simulations drawing parameter vectors from these empirical distributions and examine the discrepancy of simulated and actual data using synthetic statistics based on the probability distribution of outcomes (as in Canova (1994)).

Formally, given a density on the parameters $\pi(\theta|\mathcal{I})$ which is data based, where \mathcal{I} represents the information set available to a researcher, the outcomes of the model can be represented with a density $G(W(\theta)|\mathcal{I}, m)$ where $W = W(X_t(\theta))$ is a vector of statistics of the endogenous variables X_t of the model (e.g. moments or regression coefficients) for a given value of θ and m is the particular model specification adopted. Let $\mathcal{F}(W)$ be the empirical distribution of the actual vector of statistics we are interested in. Our task is to compare $G(W(\theta)|\mathcal{I}, m)$ and $\mathcal{F}(W)$, measure how distant they are and what features they have in common. There are two advantages of our approach. First, by examining the *distribution* of a vector of statistics of the model, our evaluation procedure maximizes the use of information present in the data. Second, unlike Cecchetti, Lam and Mark our method allows the formulation of formal statements on the likelihood of the model to reproduce the data.

There is no unique way to obtain $\pi(\theta|\mathcal{I})$. Canova (1994) describes several approaches to the problem. Here to construct $\pi(\theta|\mathcal{I})$ for each of the seven countries we assume that $\pi(\theta|\mathcal{I}) = \prod_{j=1}^5 \pi_j(\theta_j|\mathcal{I})$ where the index j refers to parameters and use a bootstrap algorithm and the implementation method of McCullough and Vinod (1993) to obtain estimates of π_j for each j . Bootstrap distributions are more appropriate than asymptotically normal approximations around the point estimates of the parameters since the sample size is relatively short. In addition, since some parameters must lie in a bounded interval, asymptotic normal approximations allow for values which are unacceptable from an economic point of view. In practical terms, the distribution of technological parameters is obtained by bootstrapping 1000 times the residuals of a regression of the consumption series for each country on a constant, after prefiltering transformations are undertaken to make the regression residuals a homoskedastic white noise. At each replication we generate a new consumption series and collect values of the mean μ , the standard deviation δ and

the AR(1) coefficient, which is used to pin down ϕ . The distribution of the preference parameters is obtained by bootstrapping the residuals of regressions of the risk free rate and of the equity return series on a constant, after a prefiltering transformation is undertaken to make the regression residuals a homoskedastic white noise. At each replication, we generate a new risk free rate and a new equity return series, use Brown and Gibbons' (1985) parametric procedure to obtain values of α and equation (7) to obtain values of β . As a byproduct of this last set of bootstraps, the empirical distribution of the mean, standard deviation and the AR(1) coefficient for the (EP-R) pair is obtained for each country, maturity and sample.

To illustrate the outcomes of this exercise, figure 1 presents $\pi_j(\theta_j|\mathcal{I})$ for each of the five parameters obtained over the 1973-1991 period for the US, and figure 2 the empirical distribution of the mean of the (EP-R) pair in the US for the three maturities. A few features of the figures deserve some attention. First, the empirical distributions of (μ, δ) are slightly skewed, the first to the right, the second to the left, and have a thicker right and left tail respectively, while the one for ϕ is approximately normal. In addition, the distributions of α and β resemble χ^2 with two degrees of freedom, even though for the former some negative values appear. All these facts confirm our previous remark that the uncertainty characterizing the parameters of the model is not well described by multivariate normal distributions. Second, the uncertainty surrounding point estimates is substantial. Hence, calibrating the parameters to the point estimates is restrictive. Finally, the bootstrap distribution of the mean of the EP-R pair at all maturities appears to be approximately normal, even though some slight right and left skewness is present for EP and R respectively. The empirical distributions for the parameters of the other countries are similar in shape even though location measures are slightly different. Because of these features we can anticipate that the model will have a hard time to account for the cross country differences noted in section 2. However, there appears to be sufficient variation for the model to account for the time variations or the maturity differences we have discussed.

To construct the simulated distribution of the statistics of interest we draw with replacement parameter vectors from $\pi(\theta|\mathcal{I})$, perform 1000 simulations for each country, each maturity and each sample period and construct simulated distributions for the mean, the standard deviation and the AR(1) coefficient for (EP-R) pair implied by the model.

To formally evaluate the discrepancy between the actual (bootstrap) and simulated distributions of the (EP-R) pair we use several probabilistic measures of fit. Ideally, one would like to report one

measure for a vector comprising all statistics for all maturities and countries. However, because the model is, at best, a rough approximation to the real world, such an approach produces uninteresting results (probabilities are all 0 or 1). As an alternative, for each country, each maturity and each statistics we report three summary measures of fit: (1) the probability that the model generates statistics for (EP, R) which fall within the $q\%$ bootstrap contour of the actual statistics of (EP, R) where $q=95, 80, 50$, (2) the probability that the bootstrap values for statistics of (EP, R) fall within the $q\%$ contour of the distribution of the simulated statistics for the (EP, R) pair where, again, $q=95, 80, 50$, (3) the probability that the simulated statistics for the $(EP-R)$ pair are in each of the four quadrants of the space delimited by the bootstrap average of the statistics of the actual (EP, R) pair, where quadrant 1 is such that $M^i(EP^S) \leq M^i(EP^B)$ and $M^i(R^S) \leq M^i(R^B)$, quadrant 2 such that $M^i(EP^S) > M^i(EP^B)$ and $M^i(R^S) \leq M^i(R^B)$, quadrant 3 such that $M^i(EP^S) < M^i(EP^B)$ and $M^i(R^S) > M^i(R^B)$, and quadrant 4 such that $M^i(EP^S) > M^i(EP^B)$ and $M^i(R^S) > M^i(R^B)$ where the superscripts "B" and "S" stand for bootstrapped and simulated, $M^i(Y)$ is statistics i of series Y and i is either the mean, the standard deviation or the AR(1) coefficient.

With the first measure we take the bootstrap distribution of the statistics of $(EP - R)$ as the "null" and ask how far is the model from the actual data by calculating the percentage of simulated values in each contour. With the second, we reverse the point of view, take the outcomes of the model as the null and ask how likely are the actual data to be generated by the model. This procedure is entirely analogous to standard hypothesis testing in a classical framework where the null and the alternative are reversed. If a percentage close to the nominal size of 95, 80 and 50% for both measures is found, then this indicates that the distributions of actual data and the model-generated data are sufficiently close since they significantly overlap. In other words, the model generates a distribution of data that reasonably approximates that of actual data. If a percentage close to the nominal size of 95, 80 and 50% for only one measure is found, then this indicates that the distributions of actual data and the data generated by the model differ significantly, even though they may overlap. While with the first measure we compute distance using a "null" distribution as a reference and probability coverings as a synthetic measure of fit, with the third measure we evaluate the model using a simple cell-probability taxonomy on the location of simulated statistics of the simulated $(EP-R)$ pair. Such a measure is particularly useful as it provides information on higher moments of the simulated distribution of each statistic, i.e. whether most of the simulated statistics lie in a particular quadrant, whether there is evidence of bimodality, whether simulated

data all lie on a particular ridge, etc.

To simplify the calculations, and because none of the results depend on this assumption, we assume throughout the exercise that dividends accrue to the equity owner at the end of each period but are available for consumption, cumulatively, only at the end of the investment period. In addition, we assume that the time interval of the model is a quarter so that the maturities of interest for comparison with real data are $k=1, 2$ and 4 .

5 The Results

The results of the simulations appear in tables 6-8 for the entire sample and in tables B.2-B.7 for the two subsamples. Tables 6, B.2 and B.3 contain the results for 3 month investments, tables 7, B.4 and B.5 for 6 month investments and tables 8, B.6 and B.7 for 12 month investments. The tables also report simulations for an equally weighted portfolio and an agent with preference and technology parameters given by the simple averages of country's parameters.

The tables present several interesting facts. First, for the mean of 3 month investments, the probability coverings are low except for Italy when the actual data is taken to be null. The presence of only a marginal overlap between the two distributions clearly emerges when we vary q . For example, when $q=50\%$, the percentage of simulated pairs lying in this contour is close to zero for all countries. The typical outcome is that the model generates both a mean risk free rate and a mean equity premium which are lower than what we see in the data, but there is at least a 6% probability that the mean risk free rate generated by the model is above the bootstrap mean of the actual risk free rate (see the percentage of simulations which fall in quadrants 1 and 3 in the table). In general, the majority of the simulated mean pairs lie on a ridge very close to the R axis, i.e. while the range for the mean of R generated by the model for each country is consistent with the historical experience, the mean of EP is, in general, very close to zero, regardless of the value of the mean of R . Hence parameter variation affects the distribution of the mean risk free rate and of the mean equity return in the same way and the distribution of the mean of EP is almost degenerate. To put this result in another way, changing the location of the univariate distribution for the mean of R will not make the model perform better in the EP dimension. Therefore, attempts to account for the large historical equity premium along the lines of Benninga and Protopapadakis (1990) or Kocherlakota (1990) with a discount factor in excess of 1 in the simulations, will not change the

basic features of the results. The puzzle, as it emerges here, is why the model generates so little independent variation in the mean of the risk free rate and equity return relative to the data (see also Bonomo and Garcia (1993)).

For the standard deviations of three month investments the performance of the model is extremely poor. Only when the model is taken to be the null for Germany and Japan we do find that a substantial portion of the bootstrap distribution falls within the 95% contour of the simulated data. However, even in this case and regardless of the country, most of the simulated values are in the lower portion of the space delimited by the actual estimates of the standard deviations of the (EP-R) pair. The performance for the AR(1) coefficient is different, in the sense that most of the simulations generate a pair of AR(1) coefficients which exceed the values we see in the data, but also in this case the outcomes are not very supportive of the model. The overlap between the bootstrap and the simulated distribution is, in many cases, very small and this is true regardless of the country considered. In general, we find that most of the area of the simulated distributions of the AR(1) coefficient is in the upper tail of the corresponding bootstrap distributions.

The performance of the model for 6 month investments is similar: there is very little overlap between bootstrap and simulated distributions of all three statistics and, relatively speaking, the performance of the model worsens relative to the 3 month investment maturity. Three features of the simulations deserve some attention. First, the model tends to generate mean values of R which are higher and mean values of EP which are lower than what we see in the data. Second, it generates a small percentage of simulations where the standard deviation of EP which is higher and of R which is lower than point estimates of the standard deviations obtained from the actual data. Finally, the overlap of the two distributions for the AR(1) coefficient is close to zero with simulated values always exceeding actual ones.

The performance at the 12 month maturity is slightly better for all countries and all statistics. It is still the case that the simulated mean of EP is too low relative to the data but there is a much more uniform distribution of simulated values in the other three quadrants. In addition, the simulated standard deviation pair are for almost all countries in all four quadrants delimited by actual estimates of the standard deviation of the (EP-R) pair and for Japan, US and Germany the spread around the actual values is roughly appropriate. Also, when we take the model as the null we find a substantial overlap in the distributions for both moments. Finally, the overlap of the two distributions for the AR(1) coefficients is small with the majority of simulations generating a pair

of AR(1) coefficients which are larger than what we see in the data.

For all three investment maturities the performance of the equally weighted portfolio is in general worse than the one of individual countries and changes very little with the investment horizon. Since such a portfolio can be viewed as approximating the one obtained by a world investor in a fully integrated capital market, this result confirms the biasedness of our test toward acceptance of the null.

The results for subsamples confirm previous tendencies. However, for three month investments we find an improvement in the overlap of the simulated and bootstrap distribution for the mean of the (EP-R) pair in the first subsample and a better spread of simulated values around the actual ones. This should not come as a surprise as the means of the (EP-R) pair for this subsample are small and even negative for some countries. For the standard deviation and the AR(1) coefficient the results across subperiods are essentially identical.

The performance of the model for six month investments is very similar across subperiods. Two features need to be noted. First, even in the first subsample, the model generates values for the mean of EP which are too low relative to what we see in the data. Second, except for Italy, the spread of the simulated distribution for the standard deviation of the (EP-R) pair in the second subsample is much larger than the spread of the bootstrap distribution for the same statistic. For 12 month investments, there are some differences across countries and across subsamples, primarily in the mean of the (EP-R) pair. However, because there are only 10 non-overlapping 12 month returns in each sample, differences across countries or subperiods may be due to the large spread in the distribution of estimates of actual statistics.

Overall, we can conclude that, regardless of the country and the time period, the model fails to reproduce the statistics of the data we are considering. It is more successful for 12 month buy-and-hold investments than for the other two maturities, but because of the small sample of non-overlapping 12 month investments, the evidence on the issue is inconclusive. Relatively speaking the model appears to be better suited to explain the bootstrap distribution of the bivariate mean of the (EP-R) pair than the distributions of the bivariate standard deviations or AR(1) coefficients regardless of the country or the maturity. Also, although there are numerical differences, the basic qualitative features of the simulations persist for countries other than the US. However, quantitatively speaking, the performance of the model for these countries is worse. In essence, the model fails to a large extent to capture the differences across countries and across time we have

noted while it can account for some features of the term structure of the (EP-R) pair.

To gain further intuition on the reason why the model fails we have performed a very simple calculation, in the spirit of the original Mehra and Prescott (1985) exercise, using point estimates of the consumption parameters and equations (7) and (8) as approximations to the mean of the (EP-R) pair generated by the model. The exercise is motivated by the following question: given estimates of μ , δ and of the covariance between consumption and risky asset returns, what would be the set of preference parameters needed to match the actual mean of the EP-R pair for each country and for the portfolio? We find that in order to match the observed average equity premium, values of the relative risk aversion parameter much higher in absolute value than those we have estimated from the data for every country and each maturity are required. However, even though the equity premium is increasing in α , it is also the case that the simulated risk free rate is decreasing for high values of α , so that extremely high levels of β are necessary to produce a risk free rate close to the average historical experience, a phenomenon that Weil (1989) termed the risk-free rate puzzle. Constantinides and Duffie (1992) have shown that with the correct amount of within country labor income heterogeneity it is possible to produce the same asset prices obtained in an economy with homogeneous agents while using values of the preference parameters which are lower and values of the discount factor which are higher than those used in the representative agent economy. Their approach appears to be promising to reduce the discrepancy between the model and the data in this dimension. These conclusions persist unaltered when, instead of the mean of the (EP-R) pair for each country, we try to match the mean for the portfolios or when we perform the exercise using data from the two subsamples.

It is legitimate to ask if the results obtained are sensitive to any of the assumptions we have made. Here we want to know whether the measures of fit we employ crucially depend on use of the bootstrap distribution of estimates of the parameters or on our assumption about the timing of dividends. In the next section we study whether alternative economic environments are likely to give a better explanation of the differences across countries, maturities and time we noted in section 2.

To check whether results are sensitive to the use of bootstrap distributions to characterize the dispersion of parameter estimates we have repeated the simulation exercises using two alternative procedures. With the first we obtain recursive estimates of the parameters for each country using rolling samples of 4 years of data. That is, we obtain one value for the 5-parameter tuple using data

from 1973,1 up to 1976,4, another using data from 1973,2 up to 1977,1 and so on up to the last one obtained using data from 1988,1 up to 1991,4. In total we have 61 estimates for each country. Then assuming a uniform distribution on each of the 61 5-parameter tuples, we draw from this empirical distribution and construct the distribution of outcomes of the model for each country, each maturity and each subsample. As compared to the original approach, this procedure has the advantage of better tracking how estimates of the parameters evolve over time and this may give a more real-time description of the uncertainty surrounding the choice of the parameters.

In the second case, we draw replications from the frequency distribution of the estimates of the preference parameters, constructed using previous estimates available in the literature (see El-Gamal (1993)), and use a 4-point uniform distribution on the estimates of technology parameters obtained using 4 year non-overlapping samples (1973-76, 1977-1981, 1982-1986, 1987-1991). This approach has the advantage of condensing information on the parameters coming from both the cross section of experiments and the time series of available data.

Although the quantitative features of the results are altered by each of these modifications, the qualitative message of the exercise is unchanged. It is still true that regardless of the country, the maturity and the sample period employed the model generates values for the moments of EP which are below the moments of EP in the data, that the joint distribution of the moments of the (EP-R) pair is primarily concentrated on a ridge very close to the R axis and that the range of values over which actual and simulated distributions intersect is very small. However, with the second approach the performance of the model is generally very poor.

Next, we study the sensitivity of the results to the assumption that dividends are paid once a year (instead of at quarterly intervals) because the dividend yields series in each country is measured annually and the k -period series is constructed assuming that dividends accrue in equal proportion each period of the year. If we remove this assumption and assume that dividends are paid once a year, say in the quarter j , $j=1, \dots, 4$, no major changes occurs in the properties of the actual (EP-R) pair. From the point of view of the model, this alternative assumption implies that it is impossible to use dividends for consumption at dates other than those corresponding to the j -th quarter. With this modification, the basic message of the simulation exercise remains. The major change concerns the size of the intersection of bootstrap and actual distributions which substantially decreases under the new assumption (and never exceeds 7%).

6 Altering the Basic Model

Many authors have modified the basic setup used in the previous sections and claimed some success in reproducing features of the equity premium in the US. In this section we examine whether three of these modifications have any potential for explaining the cross country, cross maturity, cross time heterogeneities which are unaccounted for by the basic model.

Conceptually these three alternative setups differ from Mehra and Prescott's exercise as they introduce one or more parameters in the basic structure without altering the number of moments (or statistics) to be matched. Therefore, in judging the success of these alternative setups, one should discount the additional degrees of freedom allowed for in the simulations.

As Abel (1988), Black (1990) and Canova and Marrinan (1993) have pointed out, changes in the riskiness of an asset may have direct and indirect effects on a number of asset prices. For example, an increase in the variance of dividends increases the equity premium and reduces the riskless rate of return as portfolio holders substitute away from riskier equities toward the riskless asset. To study the implications of changes in the riskiness of assets, Kandel and Stambaugh (1990), Abel (1992), Checchetti, Lam and Mark (1993) and Bonomo and Garcia (1993) have adopted a Markov switching model for the dividend process and have claimed various degrees of success in matching the first two moments of the equity premium and the autocorrelation function of equity returns at various horizons. It is therefore worthwhile to examine whether the introduction of time variation in the riskiness of dividends is qualitatively and quantitatively important in bringing simulated data closer to the international evidence.

There are several reasons for why time variation in the riskiness of dividends does not help to improve the fit of the model in the dimensions of interest. First, consider the cross country dimensions. If the lack of heteroskedasticity in dividends is the reason for the failure to match the cross country heterogeneity we need substantial differences in the structure of the time variation of second moments across countries. Second, if conditional heteroskedasticity accounts for the cross sample variation, then time variation in the second moments should have substantially increased in the second subsample. None of these features is present in the data. On the one hand, simple ARCH tests for conditional heteroskedasticity do not reject the null of no time variations in the second moments of consumption in all countries but Germany. Therefore, second moments are approximately constant for the period 1973-1991. Second, we do not find any evidence of addi-

tional conditional heteroskedasticity in the second subsample. Finally, note that, because ARCH disappears under time aggregation (see e.g. Diebold (1988)), time variation in the second moment of dividends affects moments of the (EP-R) pair for short maturity but not for long maturity investments (see also Canova and Marrinan (1992)).

Second, we modify the model to account for the lack of a riskless rate of return in the real world. Labadie (1989) has examined the effects of inflation in a monetary version of Mehra and Prescott's model where a cash in advance constraint binds in every state of nature. She shows that there are two channels through which inflation affects the (EP-R) relationship. First, because dividends are paid in money and can be used for consumption only in the next period, changes in the price level generated by random variation in the money supply lead to variations of the purchasing power of dividends and equity returns over time. Second, because there is a correlation between the intertemporal marginal rate of substitution of agents and inflation, the model generates an inflation risk premium which equally affects the risk free rate and equity returns. Labadie argues that the second effect is of minor importance and that once the link between inflation and the purchasing power of dividends is taken into account, the mean equity premium generated by the model is broadly consistent with the historical US experience.

This second modification does not seem to help in the dimensions of interest for two reasons. First of all, recall that the cross country, cross sample differences we noted are due primarily to differences over time and over countries in the risk free rate. Because none of the two effects impact on the risk free rate only, it is unlikely that differences across countries or time in the path of inflation explain the heterogeneities we described. Maturity differences in the inflation risk premium however have potential to explain cross maturity differences. Following Labadie, we define the inflation risk premium at maturity k as the covariance between S_{t+k} and $\frac{1}{\pi_{t+k}}$, where $S_{t+k} = \beta * (\frac{c_{t+k}}{c_t})^{-\alpha}$ and π_{t+k} is the inflation rate. Given the dividend process, we can obtain estimates of this covariance term, after we have pinned down β and α . Using estimates of α obtained using Brown and Gibbons parametric procedure for each country and each maturity and setting $\beta = 0.98$ we find that the largest inflation risk premium generated by the model is at the 12 month maturity, but it is only 0.003 and is not very different from the inflation risk premium generated at the three month maturity. Hence, consistently with Labadie, we conclude that although the presence of an inflation risk premium can qualitatively account for failures of the expectations theory of the term structure to hold, quantitatively, the magnitude of the effect is minor.

Because, as noted, differences across countries and samples appear primarily in the risk free rate, enriching the model by separating the process for consumption and dividends, as suggested by Abel (1992) and Checchetti, Lam and Mark (1993) is not useful since the risk free rate is independent of the dividend process. Therefore calibrating the model to the bivariate consumption-dividends processes may help to match better the time series properties of the equity return across countries, but it is unlikely to reconcile the model to the data in the dimensions of interest.

Finally, we consider the issue of leverage. Mehra and Prescott (1985) examined whether or not leverage was crucial in accounting for the large discrepancy between the EP generated by the model and the one found in the data and concluded that it was not. On the other hand, Benninga and Protopapadakis (1990) and Kandel and Stambaugh (1991) argue that leverage is an important ingredient to consider if one wants to obtain a better match between the model and the data. However, the value of the leverage parameter used by these authors in their simulations is either too high or too unconstrained, as argued by Checchetti, Lam and Mark. Once they restrict the ratio of dividends to consumption to be around the historical average, the model fails to match the data. Can leverage account for the cross country and cross sample differences we noted? We believe it can not, as it can not obviously account for the cross maturity differences. The countries of the panel are the most industrialized of the world and their leverage characteristics are similar except perhaps for Italy, where the percentage of equity financing is slightly lower than in the other countries (0.11 vs. 0.17 on average for the other G-7 countries). Similarly, the sample period is sufficiently homogeneous to doubt that leverage displayed marked differential trends across countries or structural changes across time. Hence, although leverage may be useful to bring simulated data more in line with actual data in each of the seven countries, it can not account for the heterogeneities we have discussed.

7 Conclusions

This paper studied the (EP-R) relationship from two different points of view. First, we were interested in characterizing the relationship empirically in a number of industrialized countries for the post 1973 era for investments of three different maturities (3-6-12 months). We showed that the post 1973 period is not time homogeneous and that important instabilities emerge, primarily in the distribution of the term structure of the risk free rate. In addition we show that there is

independent information in the time series of the (EP-R) pair for the seven countries at different maturities, information which is neglected when we restrict the analysis to the US alone, and that the term structure of the EP-R has characteristics which do not conform to the simple expectations theory of the term structure for many countries. Finally, we show that it is the risk free rate, more than the equity premium, which displays differences across countries and time periods. These results taken together suggest that previous efforts designed to reproduce the features of the US (EP-R) pair for a single maturity equal to the time interval of the model were very limited in scope.

In the second part of the paper we examine the performance of a standard consumption based CAPM when confronted with the richness of the cross-country, cross-maturity, cross-sample evidence. To maintain discipline in the simulations we constrain differences across countries and time periods to appear only in preferences and technological parameters rather than appealing to differences in market arrangements, institutional setups or degree of market incompleteness. We evaluate the discrepancies of a numerical version of the model from the data using measures of distance between simulated and actual distributions of moments of the (EP-R) pair (as in Canova (1994)). We show that the model is inadequate in many dimensions and that it fails to account for the heterogeneities of the data we presented in the first part. We then examined whether three different modifications of the model suggested in the literature help to reconcile the theory with the data. We argued that none of them change the basic flavor of the results even though each of them is helpful in at least one dimension.

The open question that remains to be addressed is: what alternative setups may be useful in bringing theory more in line with the data? We believe that either some form of heterogeneity, which may be individual specific, as in Constantinides and Duffie (1992), maturity specific or country specific, along the lines of Mankiw and Zeldes (1991) or Marcet and Singleton (1992), or the introduction of market incompleteness and transaction costs, along the lines of Aiyagary and Gertler (1991) or Telmer (1993), may be crucial to understand the complexity of asset return characteristics across time, maturity and country.

References

- [1] Abel, A. (1988), "Stock Prices under Time Varying Dividend Risk: An Exact Solution to an Infinite Horizon General Equilibrium Model", *Journal of Monetary Economics*, 22, 375-393.
- [2] Abel, A. (1992), "Exact Solutions for Expected Rates of Return under Markov Regime Switching: Implications for the Equity Premium Puzzle", NBER working paper 4110.
- [3] Aiyagary, R. (1993), "Explaining Financial Market Facts: The Importance of Incomplete Markets and Transaction Costs", *Federal Reserve of Minneapolis, Quarterly Review*, Winter, 17-31.
- [4] Aiyagary, R., Gertler, M. (1991), "Asset Returns With Transaction Costs and Uninsured Individual Risk", *Journal of Monetary Economics*, 27, 311-331.
- [5] Benartzi, S. and Thaler, R. (1993) "Myopic Loss Aversion and the Equity Premium Puzzle", NBER working paper 4369.
- [6] Bonomo, M. and Garcia, R. (1993) "Disappointment Aversion as a Solution to the Equity Premium and the Risk Free Rate Puzzles", CRDE, working paper 2793.
- [7] Black, F. (1990), "Mean Reversion and Consumption Smoothing", *The Review of Financial Studies*, 3, 107-114.
- [8] Benninga, S. and A. Protopapadakis (1990) "Leverage, Time Preference, and the Equity Premium Puzzle", *Journal of Monetary Economics*, 25, 49-58.
- [9] Blanchard, O. (1993), "Movements in the Equity Premium Puzzle", *Brookings Paper on Economic Activity*, 2, 75-138.
- [10] Brown, D. and M. Gibbons, (1985) "A Simple Econometric Approach for Utility Based Asset Pricing Approach", *Journal of Finance*, 359-381.
- [11] Campbell, J. (1986) "Bond and Stock Returns in a Simple Exchange Model", *Quarterly Journal of Economics*, November, 785-803.
- [12] Campbell, J. and Clarida, R. (1987) "The Term Structure of Euromarket Interest Rates: An Empirical Investigation", *Journal of Monetary Economics*, 19, 25-44.
- [13] Campbell, J. and Shiller, R. (1992) "Yield Spread and Interest Rate Movements: A Bird's Eye view", *Review of Economic Studies*, 58, 495-514.
- [14] Canova, F. (1994), "Statistical Inference in Calibrated Models", forthcoming, *Journal of Applied Econometrics*, 9.
- [15] Canova, F. and Marrinan, J. (1992), "Reconciling the Term Structure of Interest Rates with a Consumption Based CAP model", forthcoming, *Journal of Economic Dynamics and Control*.
- [16] Canova, F. and Marrinan, J. (1993), "Risk, Profits and Uncertainty in Foreign Exchange Markets", *Journal of Monetary Economics*, 32, 256-287.
- [17] Cecchetti, S., Lam, P. and Mark, N. (1993), "The Equity Premium and the Risk Free Rate: Matching the Moments", *Journal of Monetary Economics*, 31, 21-46.
- [18] Constantinides, G. (1991), "Habit Formation: A Resolution of the Equity Premium Puzzle", *Journal of Political Economy*, 98, 519-543.

- [19] Constantinides, G. and Duffie, D. (1992) "Asset Prices with Heterogeneous Consumers", University of Chicago, manuscript.
- [20] Diebold, F. (1988), *Empirical Modelling of Exchange Rate Dynamics*, Springer and Verlag, New York, New York.
- [21] El-Gamal, M. (1993), "The Extraction of Information from Multiple Point Estimates", *Journal of Nonparametric Statistics*.
- [22] Epstein, L. and Zin, S. (1991), "First-order Risk Aversion and the Equity Premium Puzzle", *Journal of Monetary Economics*, 26, 387-407.
- [23] Fama, E. and French, K. (1988), "Permanent and Transitory Components in Stock Prices", *Journal of Political Economy*, 96, 246-273.
- [24] French, K. and Poterba, j. (1991) "Investor Diversification and International Equity Markets", *American Economic Review*, Papers and Proceedings, 222-226.
- [25] Garcia, R. and Perron, P. (1993), "An Analysis of the Real Interest Rate Under Regime Shifts", University of Montreal, working paper 2391
- [26] Kandel, S. and Stambaugh, R. (1990), "Expectations and Volatility of Consumption and Asset Returns", *Review of Financial Studies*, 3, 207-232.
- [27] Kandel, S. and Stambaugh, R. (1991), "Asset Returns and Intertemporal Preferences", *Journal of Monetary Economics*, 27, 39-71.
- [28] Kocherlakota, N. (1990), "On the Discount Factor in Growth Economies", *Journal of Monetary Economics*, 25, 45-48
- [29] Labadie, P. (1989), "Stochastic Inflation and the Equity Premium Puzzle", *Journal of Monetary Economics*, 24, 277-298
- [30] Mankiw, G and Zeldes, S. (1991), "The Consumption of Stockholders and Nonstock Holders", *Journal of Financial Economics*, 29, 97-112.
- [31] Marcet, A. and Singleton, K. (1992), "Optimal Consumption -Saving Decisions and Equilibrium Asset Prices in a Model with Heterogeneous Agents Subject to Portfolio Constraints", Manuscript, Stanford University.
- [32] McCullough and Vinod, (1993) "Implementing the Single Bootstrap: Some Computational Considerations", *Computational Economics*, 6, 1-15.
- [33] Mehra, R. and Prescott, E. (1985), 'The Equity Premium: A Puzzle', *Journal of Monetary Economics*, 15, 145-162.
- [34] Mehra, R. and Prescott, E. (1988) 'The Equity Risk Premium: A Solution ?', *Journal of Monetary Economics*, 22, 133-136.
- [35] Reitz, T. (1988) "The Equity Risk Premium: A Solution", *Journal of Monetary Economics*, 22, 117-132.
- [36] Siegel, J. (1992) "The Real Interest Rate from 1800-1990: A Study of the US and of the UK", *Journal of Monetary Economics*, 29, 227-253.

REFERENCES

28

- [37] Telmer, C. (1993), "Asset Pricing Puzzles and Incomplete Markets", *Journal of Finance*, 1893-1930.
- [38] Weil, P. (1989) "The Equity Premium Puzzle and the Risk Free Puzzle ", *Journal of Monetary Economics*, 24, 401-421.

Table 1
Cross Country Statistics: Equity Premium - Risk Free Rate
Sample 1973,1-1991,4

Holding	Period	3 Months		6 Months		12 Months	
		EP	R	EP	R	EP	R
U.S.	Mean	1.39	0.41	3.08	0.85	6.68	1.77
	S.D.	7.31	0.73	11.18	1.39	14.89	2.65
	AR(1)	0.11	0.77	-0.14	0.75	-0.36	0.67
FRANCE	Mean	1.78	0.44	4.35	0.94	9.47	2.06
	S.D.	10.38	0.91	17.16	1.75	26.35	3.38
	AR(1)	0.33	0.75	0.007	0.73	-0.16	0.73
UK	Mean	2.56	0.14	5.47	0.32	11.34	0.81
	S.D.	10.70	1.66	16.17	2.94	18.42	5.38
	AR(1)	0.14	0.56	-0.26	0.64	-0.37	0.56
GERMANY	Mean	2.19	0.21	4.53	0.43	9.15	0.92
	S.D.	7.74	0.65	12.39	0.98	21.14	1.59
	AR(1)	0.18	0.09	0.21	0.25	0.01	0.48
CANADA	Mean	0.56	0.71	1.40	1.49	3.27	3.22
	S.D.	8.07	0.87	13.01	1.62	20.13	3.09
	AR(1)	0.25	0.68	0.09	0.65	-0.20	0.56
ITALY	Mean	0.79	0.09	2.64	0.22	7.65	0.63
	S.D.	12.59	1.19	22.95	2.11	42.07	3.68
	AR(1)	0.48	0.55	0.36	0.56	0.07	0.62
JAPAN	Mean	2.22	0.46	4.47	0.96	9.01	2.11
	S.D.	6.40	0.89	10.27	1.54	17.07	2.47
	AR(1)	0.12	0.50	0.003	0.35	0.09	0.52
PORTFOLIO	Mean	1.64	0.35	3.71	0.39	8.08	1.65
	S.D.	6.68	0.75	10.84	0.71	16.20	2.77
	AR(1)	0.27	0.83	-0.06	0.40	-0.15	0.75

Notes: S.D. is the standard deviation of the series, AR(1) is the first order autoregressive coefficient. Portfolio is an equally weighted portfolio composed of stocks and T-bills (or T-bills only) from the seven countries.

Table 2
Subsample Stability Tests for the Term Structure, P-values
Samples 1973,1-1981,4 and 1982,1-1991,4

		US	France	UK	Germany	Canada	Italy	Japan	Portfolio	Joint
Mean	EP	.003	.001	.97	.00	.70	.33	.00	.002	.00
	EP-R	.00	.00	.00	.00	.00	.00	.00	.00	.00
S.D.	EP	.68	.59	.00	.00	.87	.78	.00	.15	.00
	EP-R	.00	.03	.00	.00	.00	.00	.00	.00	.00
AR(1)	EP	.15	.18	.03	.00	.02	.00	.00	.03	.00
	EP-R	.02	.01	.00	.00	.00	.00	.00	.01	.00
All Moments	EP	.01	.00	.00	.00	.91	.61	.00	.00	.00
	EP-R	.00	.00	.00	.00	.00	.00	.00	.00	.00

Notes: S.D. is the standard deviation of the series, AR(1) the first autoregressive coefficient. Portfolio is an equally weighted portfolio composed of stocks and t-bills (or t-bills only) for the seven countries. The test are joint for 3-6-12 month maturities. Joint refers to a joint test for a vector of 7 countries and 3 maturities in the sample.

Table 3
Subsample Stability Tests for the Mean Slopes, P-values
Samples 1973,1-1981,4 and 1982,1-1991,4

Slope	Joint 7 Countries		Portfolio	
	EP	EP-R	EP	EP-R
6-3	.57	.00	.33	.59
12-6	.57	.00	.38	.00
Joint	.65	.00	.43	.00

Notes: 6-3 refers to the mean slope between 6 and 3 month holding periods. 12-6 refers to the mean slope between 12 and 6 month holding periods. Portfolio is an equally weighted portfolio composed of stocks and T-bills (or T-bills only) for the seven countries. Joint refers to a joint test for the vector of the two mean slopes.

Table 4
Tests for Equality of the Term Structure Across Countries
P-values using US as a baseline

Moment	Variable	France	UK	Germany	Canada	Italy	Japan	Portfolio	Joint
Sample 1973,1-1991,4									
Mean	EP	.45	.08	.27	.50	.90	.63	.07	.46
	EP-R	.002	.08	.01	.001	.01	.54	.00	.00
S.D.	EP	.46	.10	.32	.57	.91	.73	.18	.59
	EP-R	.73	.10	.23	.87	.72	.93	.09	.75
AR(1)	EP	.32	.14	.12	.62	.01	.17	.08	.11
	EP-R	.29	.36	.03	.56	.00	.02	.03	.05
All Moments	EP-R	.02	.04	.01	.01	.07	.87	.00	.00
Sample 1973,1-1981,4									
Mean	EP	.98	.04	.97	.99	.99	.99	.98	.96
	EP-R	.03	.00	.90	.96	.00	.99	.32	.00
S.D.	EP	.99	.99	.99	.99	.99	.99	.97	1.00
	EP-R	.90	.86	.80	.99	.99	.97	.93	.99
AR(1)	EP	.97	.91	.77	.95	.95	.99	.98	.99
	EP-R	.68	.80	.72	.76	.48	.12	.15	.95
All Moments	EP-R	.19	.00	.95	.99	.00	.99	.06	.005
Sample 1982,1-1991,4									
Mean	EP	.41	.99	.59	.17	.97	.41	.17	.80
	EP-R	.00	.00	.00	.00	.85	.39	.09	.00
S.D.	EP	.99	.99	.99	.99	.99	.99	.96	1.00
	EP-R	.99	.73	.98	.99	.99	.99	.90	1.00
AR(1)	EP	.94	.90	.88	.95	.44	.16	.27	.99
	EP-R	.12	.08	.04	.11	.02	.05	.10	.09
All Moments	EP-R	.00	.00	.00	.00	.99	.87	.00	.00

Notes: S.D. is the standard deviation of the series and AR(1) the first autoregressive coefficient. Portfolio is an equally weighted portfolio composed of stocks and t-bills (or t-bills only) for the seven countries. The test are joint for 3-6-12 month maturities. Joint refers to a joint test for a vector of 6 countries and 3 maturities in the sample.

Table 5
Tests for the Existence of Rolling Premia, P-values

Variable	Maturities	US	Canada	Japan	UK	France	Germany	Italy	Portfolio	Joint	Overall
Sample 1973,1-1991,4											
EP	6-3	.007	.00	.00	.02	.003	.00	.00	.004	.00	
	12-6	.00	.00	.009	.002	.00	.33	.16	.00	.00	.00
Sample 1973,1-1981,4											
EP	6-3	.48	.18	.62	.72	.05	.003	.28	.19	.42	
	12-6	.43	.61	.32	.73	.54	.27	.08	.05	.03	.06
Sample 1982,1-1991,4											
EP	6-3	.002	.00	.01	.00	.01	.002	.01	.009	.00	
	12-6	.002	.39	.33	.23	.05	.007	.00	.00	.00	.00

Notes: 6-3 and 12-6 refer to the rolling premia computed using 6 and 3 month and 12 and 6 month maturities. Portfolio is an equally weighted portfolio composed of stocks and T-bills (or T-bills only) for the seven countries. Joint refers to a joint test that both rolling premia are zero. Overall is a joint test that the rolling premia for 2 maturities for the 7 countries is zero.

**Table 6: Probability Measures of Distance
3 Month Holding Period, Sample 1973,1-1991,4**

	US	France	UK	Germany	Canada	Japan	Italy	Portfolio
Mean								
Data is the Null: Probability Coverings								
95%	0.3	31.1	0.1	0.3	28.6	0.5	89.3	9.8
80%	0.2	19.5	0.0	0.0	17.8	0.0	78.4	0.0
50%	0.0	0.0	0.0	0.0	9.8	0.0	47.7	0.0
Model is the Null: Probability Coverings								
95%	0.1	0.2	0.2	0.8	1.6	0.2	0.1	0.1
80%	0.1	0.1	0.1	0.8	0.8	0.1	0.1	0.0
50%	0.0	0.0	0.0	0.0	0.1	0.0	0.1	0.0
Quadrant Probability Coverings								
Q1	89.7	85.7	91.7	91.3	64.9	93.6	55.1	93.3
Q2	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Q3	10.3	14.3	8.3	8.7	35.1	6.1	44.9	6.7
Q4	0.0	0.0	0.0	0.0	0.1	0.0	0.0	0.0
Standard Deviation								
Data is the Null: Probability Coverings								
95%	0.0	0.1	0.1	1.0	0.1	2.5	0.1	0.2
80%	0.0	0.1	0.0	0.6	0.0	1.4	0.0	0.0
50%	0.0	0.0	0.0	0.2	0.0	0.2	0.0	0.0
Model is the Null: Probability Coverings								
95%	0.1	0.2	0.1	80.5	0.1	93.5	0.0	4.1
80%	0.0	0.0	0.0	0.4	0.0	1.4	0.0	2.6
50%	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Quadrant Probability Coverings								
Q1	96.3	99.2	99.5	88.3	98.3	90.3	100.0	89.7
Q2	0.0	0.0	0.1	1.9	0.0	4.1	0.0	0.0
Q3	3.7	0.8	0.4	8.0	1.7	4.2	0.0	10.3
Q4	0.0	0.0	0.0	1.8	0.0	1.4	0.0	0.0
AR(1) Coefficient								
Data is the Null: Probability Coverings								
95%	10.1	7.3	0.6	0.4	3.2	1.3	3.2	1.2
80%	6.2	4.2	0.4	0.4	1.7	0.6	2.3	0.4
50%	3.3	2.2	0.2	0.2	0.9	0.6	1.0	0.0
Model is the Null: Probability Coverings								
95%	23.4	23.3	0.1	0.1	7.2	0.4	0.7	10.9
80%	0.6	0.0	0.0	0.0	0.0	0.0	0.0	1.6
50%	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Quadrant Probability Coverings								
Q1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Q2	8.8	5.8	0.3	0.4	2.5	1.1	2.2	10.8
Q3	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.1
Q4	91.2	94.2	99.7	99.6	97.5	98.9	97.8	89.1

Notes: Q1 is a quadrant where the simulated moments of EP-R are both lower than their actual values, Q2 is the quadrant where the simulated moments of EP are above their actual value and the simulated moments of R are below their actual value, Q3 is the quadrant where the simulated moments of EP are below their actual value and the simulated moments of R are above their actual value, Q4 is the quadrant where the simulated moments of EP and R are both above their actual values. Portfolio is an equally weighted portfolio of stocks and T-bills (or T-bills only) of the seven countries.

**Table 7: Probability Measures of Distance
6 Month Holding Period, Sample 1973,1-1991,4**

	US	France	UK	Germany	Canada	Japan	Italy	Portfolio
Mean								
Data is the Null: Probability Coverings								
95%	0.1	21.5	0.1	0.3	7.8	0.1	55.3	4.8
80%	0.0	11.3	0.0	0.2	5.6	0.0	43.4	0.0
50%	0.0	0.0	0.0	0.0	2.7	0.0	22.5	0.0
Model is the Null: Probability Coverings								
95%	1.9	2.2	7.2	6.8	5.8	4.2	1.1	2.3
80%	1.7	1.6	6.4	6.6	4.9	3.4	1.1	1.1
50%	1.0	0.5	0.0	0.0	0.8	0.0	0.1	0.4
Quadrant Probability Coverings								
Q1	52.8	26.3	84.0	84.9	24.3	88.6	34.1	23.2
Q2	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Q3	47.2	73.7	16.0	15.1	75.7	11.4	65.3	76.8
Q4	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Standard Deviation								
Data is the Null: Probability Coverings								
95%	0.5	0.1	0.2	1.1	0.3	0.9	0.1	0.2
80%	0.2	0.0	0.1	0.6	0.0	0.7	0.0	0.0
50%	0.0	0.0	0.0	0.3	0.0	0.2	0.0	0.0
Model is the Null: Probability Coverings								
95%	6.5	0.1	0.0	99.5	0.1	25.5	0.0	4.1
80%	0.0	0.0	0.0	0.3	0.0	0.2	0.0	2.6
50%	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Quadrant Probability Coverings								
Q1	94.8	100.0	79.5	70.4	98.6	56.6	100.0	98.7
Q2	4.0	0.0	20.2	25.1	1.1	41.8	0.0	0.0
Q3	0.6	0.0	0.0	0.7	0.1	1.0	0.0	1.3
Q4	0.6	0.0	0.3	3.8	0.2	1.5	0.0	0.0
AR(1) Coefficient								
Data is the Null: Probability Coverings								
95%	1.2	0.6	0.4	1.0	0.3	1.0	0.6	1.2
80%	1.0	0.5	0.4	0.9	0.3	0.5	0.2	0.4
50%	0.2	0.4	0.2	0.6	0.0	0.4	0.1	0.0
Model is the Null: Probability Coverings								
95%	0.6	0.1	0.1	0.1	0.0	0.1	0.1	0.1
80%	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
50%	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Quadrant Probability Coverings								
Q1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Q2	1.2	0.5	0.4	0.8	0.4	0.6	0.1	1.8
Q3	0.0	0.0	2.3	0.9	0.0	0.3	0.0	0.1
Q4	98.8	99.5	97.3	98.3	99.6	99.1	99.0	98.1

Notes: Q1 is a quadrant where the simulated moments of EP-R are both lower than their actual values, Q2 is the quadrant where the simulated moments of EP are above their actual value and the simulated moments of R are below their actual value, Q3 is the quadrant where the simulated moments of EP are below their actual value and the simulated moments of R are above their actual value, Q4 is the quadrant where the simulated moments of EP and R are both above their actual values. Portfolio is an equally weighted portfolio of stocks and T-bills (or T-bills only) of the seven countries.

**Table 8: Probability Measures of Distance
12 Month Holding Period, Sample 1973,1-1991,4**

	US	France	UK	Germany	Canada	Japan	Italy	Portfolio
Mean								
Data is the Null: Probability Coverings								
95%	0.1	11.1	0.1	4.3	3.6	0.1	39.3	2.7
80%	0.0	8.2	0.0	2.3	2.1	0.0	25.7	0.0
50%	0.0	0.0	0.0	0.0	0.0	0.0	13.4	0.0
Model is the Null: Probability Coverings								
95%	99.1	14.2	0.1	6.6	35.2	0.1	30.2	2.4
80%	98.2	12.4	0.1	5.8	29.6	0.1	19.8	0.0
50%	49.2	1.0	0.0	0.0	0.0	0.0	10.3	0.0
Quadrant Probability Coverings								
1	33.6	18.3	13.6	54.5	7.3	30.6	53.8	23.3
2	26.5	2.4	85.0	38.4	6.5	56.6	0.1	20.2
3	39.9	79.3	1.4	7.1	86.2	12.8	46.1	56.5
4	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Standard Deviation								
Data is the Null: Probability Coverings								
95%	4.3	0.1	15.5	5.1	7.2	18.0	0.1	0.1
80%	1.8	0.0	3.9	1.8	1.3	9.0	0.0	0.0
50%	0.6	0.0	0.4	0.4	0.1	2.2	0.0	0.0
Model is the Null: Probability Coverings								
95%	99.1	0.2	49.7	99.5	36.5	99.6	0.0	3.4
80%	98.0	0.0	17.9	95.5	5.1	99.4	0.0	2.3
50%	0.0	0.0	4.5	3.2	0.6	87.3	0.0	0.0
Quadrant Probability Coverings								
1	70.1	95.4	31.0	55.8	83.7	48.2	100.0	86.7
2	1.1	0.0	66.6	0.6	0.0	11.3	0.0	0.0
3	16.5	4.3	0.2	29.0	17.1	10.2	0.0	13.3
4	12.3	0.3	2.2	14.6	2.2	30.3	0.0	0.0
AR(1) Coefficient								
Data is the Null: Probability Coverings								
95%	0.7	0.7	0.2	0.9	0.3	0.9	0.7	1.2
80%	0.2	0.1	0.0	0.1	0.1	0.1	0.0	0.4
50%	0.0	0.0	0.0	0.0	0.9	0.0	0.0	0.0
Model is the Null: Probability Coverings								
95%	85.9	0.7	0.1	2.2	0.2	0.1	0.4	2.9
80%	0.0	0.0	0.0	0.0	0.0	0.0	0.0	1.6
50%	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Quadrant Probability Coverings								
1	1.9	0.4	0.1	0.4	0.3	0.4	0.0	0.0
2	0.6	0.6	0.1	1.3	0.2	1.0	1.3	1.8
3	12.6	5.9	1.2	3.4	4.1	3.0	2.0	8.1
4	84.9	93.1	98.6	94.9	95.4	95.6	96.7	90.1

Notes: Q1 is a quadrant where the simulated moments of EP-R are both lower than their actual values, Q2 is the quadrant where the simulated moments of EP are above their actual value and the simulated moments of R are below their actual value, Q3 is the quadrant where the simulated moments of EP are below their actual value and the simulated moments of R are above their actual value, Q4 is the quadrant where the simulated moments of EP and R are both above their actual values. Portfolio is an equally weighted portfolio of stocks and T-bills (or T-bills only) of the seven countries.

Appendix A: Definition of Variables

The basic series employed in the study are: average of nominal stock price indices (p); per capita real seasonally adjusted (SA) consumption of nondurables and services, where quarterly population is obtained from annual data under the assumption of constant quarterly growth (C); consumption of nondurables and services consumption price deflator (PC); nominal yield on three month nominal securities (RF), and dividend yield, constructed using a 12 month moving average of total dividends on the average stock price index (DY).

The derived series are: per capita real seasonally adjusted consumption growth (CG), Real return on 3-month securities (RFR), obtained as

$$RFR_{t,3} = RF_{t,3} - \frac{PC_{t+3} - PC_t}{PC_t}$$

For longer term maturities we use the formula:

$$RFR_{t,k} = \sum_{j=0}^{k-1} RFR_{t+3j,3}$$

Real returns on equity (RT) for holding period k are obtained as:

$$RT_{t,k} = \frac{P_{t+k} - P_t}{P_t} + \frac{DY_{t+k}}{P_t}$$

where $P_t = p_t/PC_t$ and the equity premium (EP) at maturity k , defined as: $EP_{t,k} = RT_{t,k} - RFR_{t,k}$. Because the dividend yield series is annual, DY_{t+k} is obtained by multiplying the original entries of the series by $k/12$ and accumulated the resulting series over k periods.

Data Sources

United States, sample: 1973.1-1991.4

- C Difference between SA total real consumption (USCONEXPD) and SA real consumption on durables (USCONDURD).
- PC Ratio of the difference between SA total value of consumption (USCONEXPB) and SA value of durables consumption (USCONDURB) and the difference between SA total real consumption (USCONEXPD) and SA real consumption on durables (USCONDURD).
- RFR Quarter average of end-of-month rates on 3-month Treasury Bills (USTRSBL)
- RT k -period average of monthly Standard and Poor 500 price index (USSP) deflated by PC
- DY Quarter average of New York-Datastream total market monthly dividend yield (USDY)

Canada, sample: 1973.1-1991.4

- C Difference between SA total real consumption (CNCONEXPD) and SA real consumption on durables (CNCNDURBD).
- PC Ratio of the difference between SA total value of consumption (CNCONEXPB) and SA value of durables consumption (CNCNDURBB) and the difference between SA total real consumption (CNCONEXPD) and SA real consumption on durables (CNCNDURBD).
- RFR Quarter average of end-of-month rates on 3-month Treasury Bills (CNTRSBL)
- RT Quarter average of Toronto SE composite end-of-month price index (CNSHRPRC) deflated by PC
- DY k -period average of monthly Toronto composite SE dividend yield (CNDY)

Japan, sample: 1973.1-1990.4

- C Sum of SA total real nondurables consumption (JPCNNONDD) and SA real services consumption (JPCNSERVD).
- PC Ratio of the sum between SA total value of nondurables consumption (JPCNNONDB) and value of services consumption (JPCNSERVB) and the sum of SA total real non durables consumption (JPCNNONDD) and SA real services consumption (JPCNSERVD).

RFR Quarter average of monthly averages of 3-month Gensaki rates (JPOGGEN)

RT Quarter average of Tokyo New Stock Exchange end-of-month price index (JPTOKYO) deflated by PC
 DY k-period average of monthly Tokyo Datastream total market dividend yield (JPDY)

United Kingdom, sample: 1973.1-1991.4

C Difference between SA total real consumption (UKCONEXPD) and SA real consumption on durables (UKCNDURBD).

PC Ratio of the difference between SA total value of consumption (UKCONEXPB) and SA value of durables consumption (UKCNDURBB) and the difference between SA total real consumption (UKCONEXPD) and SA real consumption on durables (UKCNDURBD)

RFR Quarter average of end-of-month rates on 3-months Treasury Bills (UKTRSBL)

RT Quarter average of FT Actuaries "All Shares" monthly price index (UKFTAALP) deflated by PC

DY k-period average FT Actuaries monthly dividend yield (UKDY)

Germany, sample: 1973.1-1991.4

C SA total real consumption (BDCONEXPD).

PC Ratio of SA total value of consumption (BDCONEXPB) and SA total real consumption (BDCONEXPD).

RFR Quarter average of end-of-month rates on 3-month Treasury Bills (BDTRSBL)

RT Quarter average of Commerzbank end-of-month shares price indices (BDSHRPRC) deflated by PC

DY k-period average Frankfurt total market monthly dividend yield (BDDY)

France, sample: 1973.1-1991.4

C Difference between SA total real consumption (FRCONEXPD) and SA real consumption on durables (FRCNDURBD).

PC Ratio of the difference between SA total value of consumption (FRCONEXPB) and SA value of durables consumption (FRCNDURBB) and Difference between SA total real consumption (FRCONEXPD) and SA real consumption on durables (FRCNDURBD).

RFR Quarter average of monthly average rates on 3-month Treasury Bills (FRTRSBL)

RT Quarter average of end-of-month industrial shares price indexes (FRSHRPRC) deflated by PC

DY k-period average of Paris bourse total market monthly dividend yield (FRDY)

Italy, sample: 1974.1-1991.4

C SA total real nondurables and services consumption (ISTAT)

PC Ratio of SA total value of non durables and services consumption and SA total real non durables and services consumption (ISTAT).

RFR Quarter average of end-of-month rates on 3-month Treasury Bills (ITTRSBL)

RT Quarter average of end-of-month Milan Bourse shares price indices (ITSHRPRC) deflated by PC

DY k-period average of Milan Datastream total market monthly dividend yield (ITDY)

Notes: Datastream codes are in parenthesis. Datastream erroneously reports as seasonally adjusted the consumption series for Japan and UK when they are not. In both cases we seasonally adjusted them with standard methods using TSP procedures. No disaggregated consumption data exists for Germany. The distortions introduced by using total consumption in place of consumption of nondurables and services does not seem to be very serious. For example, in the US and France the difference in the time series properties (mean, standard deviations, autocorrelations and the partial autocorrelations) of total consumption and of consumption of nondurables and services is very small.

Appendix B: Results for Subsamples
Table B.1
Cross Country Statistics: Equity Premium - Risk-Free Rate

Holding	Period	1973,1-1981,4						1982,1-1991,4					
		3 Months		6 Months		12 Months		3 Months		6 Months		12 Months	
		EP	R	EP	R	EP	R	EP	R	EP	R	EP	R
U.S.	Mean	0.02	0.02	0.31	0.13	2.55	0.52	2.76	0.80	5.85	1.56	10.80	3.03
	S.D.	7.15	0.73	11.35	1.48	12.82	3.03	7.31	0.49	10.63	0.83	15.69	1.46
	AR(1)	0.22	0.69	-0.16	0.60	-0.18	0.41	-0.07	0.48	-0.14	0.39	-0.35	0.38
France	Mean	-0.26	-0.29	0.07	-0.54	1.14	-0.78	3.84	1.19	8.62	2.43	17.80	4.91
	S.D.	10.38	0.57	16.23	0.99	22.55	2.09	10.12	0.46	17.50	0.75	28.52	1.14
	AR(1)	0.26	0.32	-0.002	0.19	0.005	0.08	0.36	0.20	-0.11	0.01	-0.45	0.14
UK	Mean	2.29	-0.90	5.26	-1.69	12.21	3.04	2.84	1.18	5.68	2.34	10.46	4.67
	S.D.	13.56	1.70	20.09	2.83	20.21	4.96	6.97	0.69	11.03	1.07	15.14	1.77
	AR(1)	0.16	0.30	-0.22	0.27	-0.42	0.14	0.06	0.12	-0.33	0.19	-0.10	0.04
Germany	Mean	0.54	-0.03	1.10	-0.08	2.53	-0.09	3.84	0.45	7.97	0.96	15.77	1.94
	S.D.	5.17	0.62	7.95	0.77	11.57	1.15	9.45	0.58	15.13	0.89	26.77	1.28
	AR(1)	0.12	-0.20	0.01	-0.12	0.20	0.16	0.15	0.17	0.18	0.02	-0.09	0.09
Canada	Mean	-0.04	0.14	0.10	0.36	2.12	1.00	1.18	1.28	2.70	2.63	4.42	5.44
	S.D.	8.15	0.83	13.43	1.46	21.77	2.59	8.07	0.45	12.86	0.75	18.98	1.28
	AR(1)	0.27	0.55	0.31	0.45	0.24	0.16	0.12	0.23	-0.10	0.13	-0.33	-0.005
Italy	Mean	-0.51	-0.65	-0.32	-1.24	1.98	-2.20	2.10	0.84	5.61	1.69	13.32	3.48
	S.D.	13.25	1.19	24.80	1.90	43.43	2.85	11.95	0.54	21.33	0.95	42.26	1.52
	AR(1)	0.51	0.22	0.31	0.15	0.04	-0.10	0.42	0.47	0.36	0.24	0.05	0.31
Japan	Mean	0.45	0.04	0.76	0.20	2.22	0.75	4.00	0.87	8.17	1.73	15.81	3.46
	S.D.	4.65	1.01	6.80	1.75	7.64	2.74	7.43	0.49	11.91	0.76	21.25	1.16
	AR(1)	0.07	0.43	-0.20	0.20	-0.33	0.28	0.01	0.16	-0.17	0.06	-0.07	0.12
PORTFOLIO	Mean	0.35	-0.23	1.04	-0.17	3.54	-0.54	2.93	0.94	6.37	0.95	12.63	3.85
	S.D.	6.52	0.57	10.23	0.55	12.32	2.16	6.68	0.28	11.04	0.25	18.71	0.86
	AR(1)	0.24	0.59	-0.16	0.26	-0.17	0.32	0.21	0.40	-0.10	0.17	-0.15	0.12

Notes: S.D. is the standard deviation of the series, AR(1) is the first order autoregressive coefficient. Portfolio is an equally weighted portfolio composed of stocks and T-bills (or T-bills only) from the seven countries.

Table B.2: Probability Measures of Distance, 3 Month Holding Period, Sample 73,1-81,4

	US	France	UK	Germany	Canada	Japan	Italy	Portfolio
Mean								
Data is the Null: Probability Coverings								
95%	31.3	33.5	20.8	6.6	11.4	20.5	87.6	11.2
80%	23.8	25.3	16.1	5.0	8.0	14.9	76.4	9.4
50%	13.2	14.7	8.4	2.5	3.5	8.2	46.1	5.5
Model is the Null: Probability Coverings								
95%	0.7	0.6	1.4	9.8	16.2	3.2	10.0	4.8
80%	0.5	0.5	1.2	8.8	10.1	1.1	0.1	3.7
50%	0.2	0.0	0.5	2.1	4.5	0.5	0.0	1.9
Quadrant Probability Coverings								
Q1	0.0	0.0	65.7	0.0	0.6	0.9	0.0	2.5
Q2	35.7	37.7	0.0	33.6	37.2	49.0	48.9	44.8
Q3	0.0	0.0	34.3	4.7	4.9	4.6	0.0	0.0
Q4	64.3	62.3	0.0	61.7	57.3	45.5	51.1	57.5
Standard Deviation								
Data is the Null: Probability Coverings								
95%	0.0	0.1	0.4	5.7	8.3	0.2	0.1	0.5
80%	0.0	0.0	0.3	2.3	3.4	0.1	0.0	0.2
50%	0.0	0.0	0.2	0.7	1.0	0.0	0.0	0.0
Model is the Null: Probability Coverings								
95%	0.1	0.2	0.1	99.5	99.8	0.2	0.0	5.4
80%	0.0	0.0	0.0	99.3	81.3	0.0	0.0	3.1
50%	0.0	0.0	0.0	69.3	10.0	0.0	0.0	0.9
Quadrant Probability Coverings								
Q1	94.3	96.6	98.9	60.8	63.6	88.7	100.0	92.6
Q2	0.0	0.0	0.0	8.2	21.6	0.0	0.0	1.0
Q3	5.7	3.4	1.1	11.2	2.9	11.1	0.0	6.4
Q4	0.0	0.0	0.0	19.8	11.9	0.2	0.0	0.0
AR(1) Coefficient								
Data is the Null: Probability Coverings								
95%	7.3	6.5	1.3	0.1	0.8	3.2	0.1	1.2
80%	3.7	4.3	0.7	0.1	0.6	2.3	0.0	0.4
50%	2.4	2.2	0.4	0.0	0.3	1.2	0.0	0.0
Model is the Null: Probability Coverings								
95%	5.1	9.5	0.1	0.1	0.2	2.9	0.3	0.9
80%	0.2	0.0	0.0	0.0	0.0	0.0	0.0	0.6
50%	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Quadrant Probability Coverings								
Q1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Q2	6.5	4.8	1.1	0.4	0.7	2.8	0.2	12.1
Q3	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Q4	93.5	95.2	98.9	99.6	99.3	97.2	99.8	87.9

Notes: Q1 is a quadrant where the simulated moments of EP-R are both lower than their actual values, Q2 is the quadrant where the simulated moments of EP are above their actual value and the simulated moments of R are below their actual value, Q3 is the quadrant where the simulated moments of EP are below their actual value and the simulated moments of R are above their actual value, Q4 is the quadrant where the simulated moments of EP and R are both above their actual values. Portfolio is an equally weighted portfolio of stocks and T-bills (or T-bills only) of the seven countries.

Table B.3: Probability Measures of Distance, 3 Month Holding Period, Sample 82,1-91,4

	US	France	UK	Germany	Canada	Japan	Italy	Portfolio
Mean								
Data is the Null: Probability Coverings								
95%	0.1	0.0	0.1	0.0	0.2	21.2	80.3	4.4
80%	0.0	0.0	0.0	0.0	0.0	12.8	70.1	3.3
50%	0.0	0.0	0.0	0.0	0.0	7.3	39.8	0.5
Model is the Null: Probability Coverings								
95%	0.0	0.0	0.0	0.0	0.1	1.2	0.3	0.1
80%	0.0	0.0	0.0	0.0	0.0	0.9	0.1	0.0
50%	0.0	0.0	0.0	0.0	0.0	0.5	0.0	0.0
Quadrant Probability Coverings								
Q1	97.1	93.4	97.3	94.9	96.3	70.0	67.1	88.1
Q2	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Q3	2.9	6.6	2.7	5.1	3.7	30.0	32.9	11.9
Q4	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Standard Deviation								
Data is the Null: Probability Coverings								
95%	0.0	0.0	2.1	1.6	0.0	0.1	0.0	0.1
80%	0.0	0.0	0.7	0.6	0.0	0.1	0.0	0.0
50%	0.0	0.0	0.2	0.2	0.0	0.0	0.0	0.0
Model is the Null: Probability Coverings								
95%	0.0	0.0	100.0	93.4	16.5	0.3	0.0	6.6
80%	0.0	0.0	99.4	6.1	0.0	0.0	0.0	4.2
50%	0.0	0.0	67.4	0.0	0.0	0.0	0.0	1.3
Quadrant Probability Coverings								
Q1	67.0	82.0	55.1	72.6	60.6	81.8	99.4	94.2
Q2	0.0	0.0	8.7	1.1	0.1	0.0	0.0	0.5
Q3	32.8	17.9	18.4	22.6	37.8	17.7	0.6	5.1
Q4	0.2	0.1	17.8	3.7	1.5	0.5	0.0	0.2
AR(1) Coefficient								
Data is the Null: Probability Coverings								
95%	8.3	2.7	2.2	2.9	4.9	3.7	0.1	3.3
80%	5.5	2.1	1.0	1.5	2.8	2.3	0.0	1.8
50%	2.8	1.0	0.3	0.7	1.4	1.5	0.0	0.0
Model is the Null: Probability Coverings								
95%	27.2	3.2	13.5	20.9	37.7	16.1	0.3	2.9
80%	1.6	0.0	0.0	0.0	0.0	0.0	0.0	1.6
50%	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Quadrant Probability Coverings								
Q1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Q2	10.8	3.2	2.4	3.5	4.0	7.1	0.0	8.4
Q3	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Q4	89.2	96.8	97.6	96.5	96.0	92.9	100.0	93.6

Notes: Q1 is a quadrant where the simulated moments of EP-R are both lower than their actual values, Q2 is the quadrant where the simulated moments of EP are above their actual value and the simulated moments of R are below their actual value, Q3 is the quadrant where the simulated moments of EP are below their actual value and the simulated moments of R are above their actual value, Q4 is the quadrant where the simulated moments of EP and R are both above their actual values. Portfolio is an equally weighted portfolio of stocks and T-bills (or T-bills only) of the seven countries.

Table B.4: Probability Measures of Distance 6 Month Holding Period, Sample 73,1-81,4

	US	France	UK	Germany	Canada	Japan	Italy	Portfolio
Mean								
Data is the Null: Probability Coverings								
95%	41.3	24.4	14.3	12.4	31.7	15.3	29.3	12.6
80%	30.2	15.9	9.5	7.9	20.3	9.2	15.4	7.8
50%	16.6	7.6	0.2	1.6	5.8	3.5	7.5	0.0
Model is the Null: Probability Coverings								
95%	4.9	6.6	3.2	7.8	11.7	3.2	2.6	4.6
80%	3.3	1.4	1.3	5.0	9.2	0.6	0.9	2.9
50%	0.6	0.1	0.0	0.0	3.6	0.0	0.1	0.4
Quadrant Probability Coverings								
Q1	68.2	14.0	96.3	92.6	66.3	92.6	89.9	77.5
Q2	0.0	78.1	0.2	0.5	0.0	0.4	1.7	0.3
Q3	31.8	0.1	3.5	6.9	33.7	6.7	7.9	22.2
Q4	0.0	7.8	0.0	0.0	0.0	0.0	0.5	0.0
Standard Deviation								
Data is the Null: Probability Coverings								
95%	0.6	0.4	0.7	5.2	2.5	0.8	0.1	0.8
80%	0.3	0.0	0.5	1.2	0.6	0.2	0.0	0.4
50%	0.0	0.0	0.2	0.2	0.1	0.2	0.0	0.0
Model is the Null: Probability Coverings								
95%	13.5	0.4	0.5	99.8	98.6	45.0	0.0	12.4
80%	0.0	0.0	0.0	37.8	1.9	1.1	0.0	9.4
50%	0.0	0.0	0.0	2.9	0.0	0.0	0.0	0.0
Quadrant Probability Coverings								
Q1	95.6	98.3	83.4	48.4	86.3	37.7	99.8	93.5
Q2	3.2	1.2	16.2	42.9	9.0	59.9	0.2	0.0
Q3	0.7	0.3	0.0	1.0	2.2	0.2	0.0	6.5
Q4	0.5	0.2	0.4	7.7	2.5	2.2	0.0	0.0
AR(1) Coefficient								
Data is the Null: Probability Coverings								
95%	1.9	0.1	0.0	0.1	1.7	0.4	0.0	0.8
80%	1.1	0.0	0.0	0.0	1.1	0.2	0.0	0.3
50%	0.7	0.0	0.0	0.0	0.9	0.2	0.0	0.0
Model is the Null: Probability Coverings								
95%	11.6	0.0	0.1	0.0	4.4	0.1	0.0	0.1
80%	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
50%	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Quadrant Probability Coverings								
Q1	0.2	0.0	0.0	0.0	0.3	0.2	0.0	0.0
Q2	1.7	0.1	0.0	0.4	1.3	0.2	0.0	0.5
Q3	1.1	0.0	0.0	2.3	3.3	1.0	0.0	0.4
Q4	97.0	99.9	100.0	97.3	95.1	98.6	100.0	99.1

Notes: Q1 is a quadrant where the simulated moments of EP-R are both lower than their actual values, Q2 is the quadrant where the simulated moments of EP are above their actual value and the simulated moments of R are below their actual value, Q3 is the quadrant where the simulated moments of EP are below their actual value and the simulated moments of R are above their actual value, Q4 is the quadrant where the simulated moments of EP and R are both above their actual values. Portfolio is an equally weighted portfolio of stocks and T-bills (or T-bills only) of the seven countries.

Table B.5: Probability Measures of Distance, 6 Month Holding Period, Sample 82,1-91,4

	US	France	UK	Germany	Canada	Japan	Italy	Portfolio
Mean								
Data is the Null: Probability Coverings								
95%	0.1	0.0	0.0	0.0	4.3	0.0	0.0	0.4
80%	0.0	0.0	0.0	0.0	2.6	0.0	0.0	0.1
50%	0.0	0.0	0.0	0.0	0.7	0.0	0.0	0.0
Model is the Null: Probability Coverings								
95%	1.0	1.2	2.5	5.8	9.1	5.2	0.1	2.2
80%	0.7	0.8	1.4	4.5	5.9	3.7	0.0	0.9
50%	0.0	0.2	0.0	0.1	2.7	0.5	0.0	0.0
Quadrant Probability Coverings								
Q1	77.2	32.3	89.8	79.3	33.3	79.1	2.6	37.9
Q2	0.0	0.0	0.4	0.1	0.0	0.2	0.0	0.0
Q3	22.8	67.7	9.8	20.6	66.4	20.7	97.4	62.1
Q4	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Standard Deviation								
Data is the Null: Probability Coverings								
95%	4.1	2.2	1.2	5.1	0.9	6.9	0.0	3.3
80%	2.0	1.3	0.6	1.7	0.6	2.9	0.0	1.8
50%	0.3	0.3	0.2	0.4	0.0	1.0	0.0	0.0
Model is the Null: Probability Coverings								
95%	100.0	98.3	100.0	100.0	100.0	100.0	0.0	85.4
80%	98.2	3.4	50.0	86.0	0.9	99.3	0.0	56.3
50%	13.4	0.0	1.0	8.1	0.0	32.4	0.0	12.1
Quadrant Probability Coverings								
Q1	61.7	87.4	32.3	65.7	87.1	57.1	99.7	54.5
Q2	19.2	4.8	58.0	20.6	5.1	21.3	0.0	20.3
Q3	6.3	4.5	1.1	3.8	5.0	5.8	0.3	4.5
Q4	12.8	3.3	8.6	9.9	2.8	15.8	0.0	20.7
AR(1) Coefficient								
Data is the Null: Probability Coverings								
95%	1.0	0.3	0.0	0.3	0.8	0.5	0.1	1.3
80%	0.6	0.2	0.0	0.2	0.4	0.3	0.0	0.7
50%	0.3	0.0	0.0	0.0	0.3	0.1	0.0	0.0
Model is the Null: Probability Coverings								
95%	0.1	0.1	0.0	0.1	0.0	0.1	0.3	0.1
80%	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
50%	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Quadrant Probability Coverings								
Q1	0.2	0.0	0.1	0.1	0.2	0.0	0.0	0.1
Q2	1.0	0.6	0.1	0.1	0.6	0.9	0.0	2.0
Q3	0.0	0.0	1.6	1.7	0.5	0.8	0.0	0.5
Q4	98.8	99.4	98.2	97.2	98.5	98.3	100.0	97.4

Notes: Q1 is a quadrant where the simulated moments of EP-R are both lower than their actual values, Q2 is the quadrant where the simulated moments of EP are above their actual value and the simulated moments of R are below their actual value, Q3 is the quadrant where the simulated moments of EP are below their actual value and the simulated moments of R are above their actual value, Q4 is the quadrant where the simulated moments of EP and R are both above their actual values. Portfolio is an equally weighted portfolio of stocks and T-bills (or T-bills only) of the seven countries.

Table B.6: Probability Measures of Distance, 12 Month Holding Period, Sample 73,1-81,4

	US	France	UK	Germany	Canada	Japan	Italy	Portfolio
Mean								
Data is the Null: Probability Coverings								
95%	15.1	8.7	0.1	3.3	23.4	0.1	3.9	1.1
80%	9.4	2.9	0.0	1.7	15.1	0.0	2.8	0.5
50%	4.4	1.6	0.0	0.9	8.1	0.0	1.4	0.0
Model is the Null: Probability Coverings								
95%	57.3	0.2	0.1	0.0	95.2	0.1	0.2	1.4
80%	49.5	0.0	0.1	0.0	77.5	0.1	0.0	0.6
50%	0.7	0.0	0.0	0.0	38.7	0.0	0.0	0.0
Quadrant Probability Coverings								
Q1	0.1	0.0	0.0	1.5	22.8	22.2	0.1	23.3
Q2	81.7	96.4	85.0	96.7	35.1	55.6	97.1	55.5
Q3	11.3	0.5	11.4	1.5	42.1	22.2	1.7	21.2
Q4	6.9	3.1	3.6	0.3	0.0	0.0	0.2	0.0
Standard Deviation								
Data is the Null: Probability Coverings								
95%	16.7	6.1	5.5	17.2	12.6	8.0	1.9	10.2
80%	7.1	2.6	3.9	6.8	7.3	5.2	0.4	6.7
50%	2.2	1.0	0.4	2.2	3.0	1.4	0.2	3.3
Model is the Null: Probability Coverings								
95%	99.8	80.4	55.7	99.5	100.0	99.6	2.3	42.4
80%	94.0	1.0	43.9	99.5	75.2	80.2	0.0	38.9
50%	3.0	0.0	12.5	94.8	1.3	59.4	0.0	15.8
Quadrant Probability Coverings								
Q1	71.4	89.0	61.0	28.0	70.7	28.5	97.5	89.7
Q2	3.3	0.9	26.6	10.5	3.6	13.4	1.1	5.0
Q3	12.4	6.9	2.4	22.7	15.3	22.1	1.2	5.3
Q4	12.9	3.2	0.0	38.8	10.4	46.0	0.2	0.0
AR(1) Coefficient								
Data is the Null: Probability Coverings								
95%	0.5	0.1	0.1	0.1	1.2	0.3	0.2	1.0
80%	0.0	0.1	0.0	0.1	0.2	0.1	0.2	0.2
50%	0.0	0.0	0.0	0.1	0.0	0.0	0.2	0.0
Model is the Null: Probability Coverings								
95%	4.1	0.1	0.0	0.0	3.1	0.1	0.9	1.0
80%	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.2
50%	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Quadrant Probability Coverings								
Q1	1.1	0.1	0.1	0.2	0.7	0.4	0.0	0.0
Q2	0.9	0.1	0.0	0.2	1.9	3.5	0.0	1.8
Q3	5.1	0.8	0.1	0.7	4.1	2.0	0.0	1.1
Q4	92.9	99.0	99.8	98.9	93.2	93.1	100.0	98.1

Notes: Q1 is a quadrant where the simulated moments of EP-R are both lower than their actual values, Q2 is the quadrant where the simulated moments of EP are above their actual value and the simulated moments of R are below their actual value, Q3 is the quadrant where the simulated moments of EP are below their actual value and the simulated moments of R are above their actual value, Q4 is the quadrant where the simulated moments of EP and R are both above their actual values. Portfolio is an equally weighted portfolio of stocks and T-bills (or T-bills only) of the seven countries.

Table B.7: Probability Measures of Distance, 12 Month Holding Period, Sample 82,1-91,4

	US	France	UK	Germany	Canada	Japan	Italy	Portfolio
Mean								
Data is the Null: Probability Coverings								
95%	0.1	0.0	0.1	0.0	0.2	21.2	3.3	2.4
80%	0.0	0.0	0.0	0.0	0.0	12.8	0.4	0.4
50%	0.0	0.0	0.0	0.0	0.0	7.3	0.0	0.0
Model is the Null: Probability Coverings								
95%	0.0	0.0	0.0	0.0	0.1	1.2	0.0	0.1
80%	0.0	0.0	0.0	0.0	0.0	0.9	0.0	0.0
50%	0.0	0.0	0.0	0.0	0.0	0.5	0.0	0.0
Quadrant Probability Coverings								
Q1	97.1	93.4	97.3	94.9	96.3	70.0	1.3	5.3
Q2	0.0	0.0	0.0	0.0	0.0	0.0	0.8	0.0
Q3	2.9	6.6	2.7	5.1	3.7	30.0	97.9	94.6
Q4	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Standard Deviation								
Data is the Null: Probability Coverings								
95%	0.0	0.0	2.1	1.6	0.0	0.1	1.3	2.4
80%	0.0	0.0	0.7	0.6	0.0	0.1	0.0	0.0
50%	0.0	0.0	0.2	0.2	0.0	0.0	0.0	0.0
Model is the Null: Probability Coverings								
95%	0.0	0.0	100.0	93.4	16.5	0.3	2.4	3.1
80%	0.0	0.0	99.4	6.1	0.0	0.0	0.8	1.2
50%	0.0	0.0	67.4	0.0	0.0	0.0	0.1	0.1
Quadrant Probability Coverings								
Q1	67.0	82.0	55.1	72.6	60.6	81.8	94.1	90.1
Q2	0.0	0.0	8.7	1.1	0.1	0.0	0.5	1.5
Q3	32.8	17.9	18.4	22.6	37.8	17.7	5.1	8.1
Q4	0.2	0.1	17.8	3.7	1.5	0.5	0.3	0.3
AR(1) Coefficient								
Data is the Null: Probability Coverings								
95%	8.3	2.7	2.2	2.9	4.9	3.7	0.0	3.3
80%	5.5	2.1	1.0	1.5	2.8	2.3	0.0	1.8
50%	2.8	1.0	0.4	0.7	1.4	1.5	0.0	0.0
Model is the Null: Probability Coverings								
95%	27.2	3.2	13.5	20.9	16.1	37.7	0.0	2.9
80%	1.6	0.0	0.0	0.0	0.0	0.0	0.0	1.6
50%	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Quadrant Probability Coverings								
Q1	0.0	0.0	0.0	0.0	0.0	0.0	0.1	0.0
Q2	10.8	3.2	2.4	3.5	4.0	7.1	0.0	8.4
Q3	0.0	0.0	0.0	0.0	0.0	0.0	0.6	0.0
Q4	89.2	96.8	97.6	96.5	96.0	92.9	99.3	93.6

Notes: Q1 is a quadrant where the simulated moments of EP-R are both lower than their actual values, Q2 is the quadrant where the simulated moments of EP are above their actual value and the simulated moments of R are below their actual value, Q3 is the quadrant where the simulated moments of EP are below their actual value and the simulated moments of R are above their actual value, Q4 is the quadrant where the simulated moments of EP and R are both above their actual values. Portfolio is an equally weighted portfolio of stocks and T-bills (or T-bills only) of the seven countries.



