

BANCA D'ITALIA

Temi di discussione

del Servizio Studi

The Equity Premium Is No Puzzle

by Mordecai Kurz and Andrea Beltratti



Number 282 - October 1996

Temi di discussione

del Servizio Studi

The purpose of the "Temi di discussione" series is to promote the circulation of working papers prepared within the Bank of Italy or presented in Bank seminars by outside economists with the aim of stimulating comments and suggestions.

The views expressed in the articles are those of the authors and do not involve the responsibility of the Bank.

Editorial Board:

MASSIMO ROCCAS, EUGENIO GAIOTTI, DANIELA MONACELLI, DANIELE TERLIZZESE, ORESTE TRISTANI; SILIA MIGLIARUCCI (*Editorial Assistant*).

The Equity Premium Is No Puzzle

by Mordecai Kurz and Andrea Beltratti

Number 282 - October 1996

THE EQUITY PREMIUM IS NO PUZZLE (*)

by Mordecai Kurz (**) and Andrea Beltratti (***)

Abstract

We examine the equity premium puzzle with the perspective of the theory of Rational Beliefs Equilibrium (RBE) and show that from the perspective of this theory there is no puzzle. In an RBE agents need to be compensated for the endogenously propagated price uncertainty which is not permitted under rational expectations. It is then argued that endogenous uncertainty is the predominant uncertainty of asset returns and its presence provides a natural explanation of the observed premium. Utilizing data on the asset allocation of 63 US mutual funds, we test some empirical implications of the theory of rational beliefs and estimate the parameters of risk aversion of mutual fund managers. Our tests show that the predictions of the theory are consistent with the empirical evidence. We then construct a simple two-agent model of the US economy in which the agents hold rational beliefs and calibrate it to the empirical experience in accord with the parameters of the Mehra and Prescott (1985) paper. The results of our calculations show that for a large set of parameter values the model predictions closely fit the historical record.

Contents

1. Introduction	p. 7
2. The equity premium puzzle	p. 8
3. The perspective of RBE and endogenous uncertainty of the equity premium	p. 11
4. Asset allocation in mutual funds: testing some implications of the theory of rational beliefs	p. 18
5. Analysis of the equity premium in a rational belief equilibrium: simulating the economy	p. 23
5.a Rational belief equilibria of the two-agent OLG economy	p. 25
5.b Simulation results	p. 37
6. A final comment	p. 46
Appendix 1	p. 47
Appendix 2	p. 49
Appendix 3	p. 50
References	p. 57

(*) Presentation at a Seminar held by the authors at the Research Department of the Bank of Italy, 2-3 July 1996.

(**) Stanford University, Department of Economics.

(***) Università di Torino, Dipartimento di Economia G. Prato.

1. Introduction¹

The "equity premium puzzle" was introduced in Mehra and Prescott (1985) and we refer to this paper as MP (1985). It arises from the observation that the average real rate of return on equity over the last century has been about 7% while the average rate of return on riskless, short term securities, has been about 1%. Many studies which investigated this 6% premium concluded that it is too large by theoretical yardsticks currently in use in economics and finance. Since the risk premium is the differential between the equilibrium rates of return on stocks and bonds it follows that the equity premium puzzle can be reformulated to become a statement that the optimal share of wealth held in the form of bonds is much lower than we observe in the data. This implies an "asset allocation puzzle" which is equivalent to the equity premium puzzle. The research which reported these puzzles is discussed in section 2, stressing the diverse ways one should look at the puzzle. However, our discussion also indicates that the common theoretical paradigm which links these different approaches is the rational expectations paradigm.

This paper aims to make the case that the equity premium "puzzle" is no puzzle at all. Its diverse manifestations, described in Section 2, are simply a collection of tests of the rational expectations theory and the theory fails the tests in a consistent manner. By any reasonable criterion of

¹ This work was supported by Fondazione ENI Enrico Mattei, Milano, Italy. The authors thank Douglas Bernheim, Darrell Duffie, Carsten Nielsen and Martin Schneider for very useful comments on an earlier draft. They also thank Evren Ergin, Maurizio Motolese and Ron Borzekowski for assistance in compiling the data file and for computer analyses.

scientific evaluation the extensive work on the equity premium should have lead to the rejection of the theory rather than to a declaration that the results constitute "puzzles". Once we reject this theory most of the problems raised by the "puzzle" literature are removed. Our point of departure is the replacement of the theory of rational expectations with the theory of rational beliefs developed by Kurz (1994a), (1994b) (see Appendix 3 for a brief review) which takes rational expectations as a special and unlikely case. Thus, in Section 3 we use the perspective of the theory of rational beliefs to provide an intuitively simple explanation for the data which have given rise to the equity premium puzzles.

The bulk of the work reported here is a positive application of the theory of Rational Belief Equilibrium (RBE) to understand the nature of the premium in asset markets. The theory of rational beliefs predicts that agents will have diverse beliefs and consequently have diverse mistake functions which can be estimated as explained in the paper on asset prices in this volume (see Kurz (1997), Section 2.2.D). In Section 4 of the present paper we use data on the asset composition of 63 U.S. mutual funds to test some empirical implications of the theory and to estimate the parameters of risk aversion of the 63 funds in the study. The range of values of this parameter will then be used in Section 5 of the paper. In that part we formulate a two-agent economy with a stock market and a short term borrowing instrument ("bill") and then construct an RBE for this economy. We assign to the real economy all the parameters used by MP (1985) but select the beliefs of the agents in a manner which allows a relatively simple parametrization. For alternative configurations of the parameters we compute equilibrium prices, long term time average of interest rates and long term variance of returns. We show that the model calculations are entirely compatible with the empirical record.

2. The Equity Premium Puzzle

The debate about the equity premium puzzle has been conducted in three distinct forms. In

order to evaluate this debate from the perspective of the theory of rational beliefs we shall start by reviewing the differences among these three approaches.

We start with MP (1985), whose method of analysis is an adaptation of the Lucas (1978) asset pricing model. It postulates a single, infinitely lived representative agent and a fixed number N of assets which produce a non-storable consumption good with an exogenous stochastic technology. The utility function of the agent is a discounted sum of time invariant utilities of consumption with a constant discount rate. MP (1985) consider the basic case of two assets: a common stock and a risk-free debt. The first order conditions of the optimization with respect to the stock holding are standard:

$$(1) \quad P_t u'(x_t) = \beta E_t u'(x_{t+1}) [P_{t+1} + D_{t+1}]$$

where P_t is the price of the stock at time t , x_t is total consumption at time t , D_{t+1} is the dividend paid by the stock at $t+1$ and $\beta < 1$ is a discount factor. Mehra and Prescott use the specific utility function $u(x) = (1-\gamma)^{-1} x^{1-\gamma}$, $\gamma > 0$ for which the first order conditions become:

$$(2) \quad P_t x_t^{-\gamma} = \beta E_t x_{t+1}^{-\gamma} [P_{t+1} + D_{t+1}].$$

The rate of return on the stock is $R_{t+1} = \frac{P_{t+1} + D_{t+1}}{P_t}$ and the risk premium is the difference between the risky and the riskless rates of return. The "premium" is usually thought of as the time average of the risk premia. In order to compute the theoretical premium implied by the model one must compute the equilibrium prices predicted. The Mehra and Prescott economy has no labor or non-capital resources: there is only one productive activity employing a single asset and the net dividend (which equals consumption) is a stochastic process of the following structure

$$(3) \quad D_{t+1} = d_{t+1} D_t$$

where the *growth rate* of dividends, d_{t+1} , is assumed to be a stationary and ergodic Markov chain. In

the application of the model MP (1985) assume that the process $\{d_t, t = 1, 2, \dots\}$ has only two states. Writing down such a model leads immediately to the calculations of equilibrium prices and consequently to the premium. Given rational expectations, then at any date t the only information needed for price determination is the pair (d_t, D_t) which is the exogenous state for the economy. MP (1985) show that the equilibrium price function is of the form

$$(4) \quad P_t = P(d_t)D_t$$

The function (4) is then calculated from the equilibrium conditions and the specified parameters of the dividend process. The central conclusion of the paper is that for reasonable values of β and γ and calibration of the parameters of the stochastic process of dividend growth to actual data, the model generates a maximal risk premium of the order of .37 percentage point. However, the Mehra-Prescott procedure of selecting reasonable values for the parameters of the economy is then secondary to the fact that the question whether the premium is small or large is entirely determined by *the calculated theoretical time path of asset prices and hence capital gains* which follow from the rational expectations assumption. This methodology of determining the size of the premium is universal to all models which followed the approach of Mehra-Prescott (e.g. Rietz (1988), Weil (1989), Epstein and Zin (1990) and Cecchetti, Lam and Mark (1993)).

An alternative view of the equity premium puzzle is implicit in the results of Grossman and Shiller (1981) who also use a representative agent model with rational expectations in the stock market. They show that in order for the model to reproduce time paths for stock prices with volatility over time which is equal to the observed volatility, they need to assume unreasonably large values of γ . The approach does not lead to a calculation of the theoretical equilibrium asset prices and premium of an artificial general equilibrium model. Instead, it uses a model like (1)-(2) to fit the data to an equation implied by the first order condition (2). The standard assumption made is that all agents know the true

probability distribution of dividends and prices, and the moments of that distribution are the empirical moments of the long run time series of the data in the economy. A similar approach is taken by Mankiw and Zeldes (1991) and many others.

A third view of the puzzle is inherent in the approaches of MaCurdy and Shoven (1992), (1993) and is also frequently encountered in the investment community. MaCurdy and Shoven study the ex-post performance of alternative portfolio compositions over very long horizons of 25-40 years during the period 1876-1990. They examine hypothetical households who could have initiated, *at different starting dates*, investment programs for retirement 25 or 40 years later. Their conclusion is that the "all bonds" portfolios were inferior to the "all equities" portfolios for most planned retirement dates. Therefore they propose that households who own bonds for retirement planning are irrational. Thus, MaCurdy and Shoven interpret the equity premium puzzle to mean what we suggested earlier: that the actual proportion of bonds held in portfolios during the last century appears too large compared with some optimal theoretical portfolio. The judgment of what is optimal is then based, in this case, on the assumption that the realized empirical distribution of returns over the last century is known by the agents to be the true distribution of returns and therefore it is also the distribution which every agent *should have adopted as his belief* and used in his own optimization. MaCurdy and Shoven's suggestion of the "irrationality" of investors implies that they interpret "rational expectations" to mean that agents must adopt the stationary measure as their belief.

3. The Perspective of RBE and Endogenous Uncertainty on the Equity Premium

We shall now use known results about Rational Belief Equilibria (RBE) (see Kurz (1994a), (1994b), (1996), Kurz and Schneider (1996) and Kurz and Wu (1996)) to cast the equity premium debate in a new light. As in the previous section, our discussion will evaluate the three different

perspectives of the equity premium debate in order to clarify how the theory of RBE helps in explaining the puzzle.

Starting again with the general equilibrium perspective, we assume with MP (1985) that the underlying economy functions with spot markets and securities in a sequence of markets as in Lucas (1978) or Arrow (1953). The theory of RBE postulates that agents do not have “Structural Knowledge” about the economy. In the model at hand this means that agents do not know either the map between exogenous variables and asset prices or the true probability distribution of the dividend process. Let Q^k be the probability belief of agent k and denote by Q_t^k the conditional probability of k at date t given information at t . Then equilibrium prices are expressed by an equation like

$$(5) \quad P_t = P(d_t, Q_t^1, Q_t^2, \dots, Q_t^N) D_t.$$

The component of variability of prices which is attributable to the beliefs of the agents is called *Endogenous Uncertainty*, a term introduced by Kurz (1974) to express the idea that these fluctuations of prices are internally propagated. Endogenous uncertainty is then the price uncertainty which the agents face in the market and which is not caused by the variability of the “fundamental” exogenous variables (d_t, D_t). Comparison of equations (4) and (5) shows that since the MP (1985) equity premium was calculated under the assumption of rational expectations of the agents, such “model calculations” exclude all endogenous uncertainty. These calculations insist that only capital gains and losses which can be attributed to the variability of the exogenous variables should be included in the risk faced by the agents. If we allow for the presence of endogenous uncertainty in the model then owners of equities would demand, in equilibrium, compensation for taking the endogenously propagated uncertainty. In that case the risk premium which they actually received in the economy

would have been entirely justified. What the MP (1985) calculations show is that *endogenous uncertainty is the dominant form of uncertainty in the equities markets*. More precisely if we take the coefficient of risk aversion to be between 1 and 10 then the range of equity premium which is compatible with exogenous uncertainty is 0%-35% according to the calculations of MP (1985) but could go as high as 2% according to Mankiw-Zeldes (1991) who restrict the sample to stockholders only. This range of numbers imply that out of the total return on equities of around 7%, most is a risk premium for endogenous uncertainty. This is essentially the conclusion of Kurz (1997) as well.

We now turn to the second approach of analyzing the equity premium. In an RBE where agents have heterogeneous beliefs optimality conditions like (2) are specified for each agent k . That is, rewriting (2), the innovations $(z_{t+1}^k - 1)$ are required to satisfy

$$(6a) \quad \left(\frac{x_{t+1}^k}{x_t^k} \right)^{-\gamma} R_{t+1} \beta = z_{t+1}^k$$

$$(6b) \quad E_{Q_t^k}(z_{t+1}^k) = 1$$

and $\text{Cov}_{Q_t^k}(z_{t+1}^k, u_t) = 0$ for any variable u_t known at date t is the Q^k orthogonality implication of a conditional probability. In an RBE agents hold beliefs Q_t^k which are not equal to the true, equilibrium, probability Π_t . The set function $M_t^k = Q_t^k - \Pi_t$ defined over the relevant random events, is called "the mistake of agent k ." The presence of mistakes implies that

$$(7) \quad E_{\Pi_t}(z_{t+1}^k) \neq 1$$

and this violates the orthogonality conditions of rational expectations. Consequently, the random variables z_{t+1}^k are functions of variables observed at t and this functional dependence is the basis for an econometric testing of the empirical implications of the theory of rational beliefs. The systematic dependence of z_{t+1}^k on market information at date t is created by agents either not knowing that some

information is relevant and consequently not using it or by misinterpreting available information in their decision-making (for details see Kurz (1997) Section 3.3.B). The dependency of z_{t+1}^k on observed market information at date t naturally leads to the failure of partial equilibrium models which estimate (6a) under the rational expectations assumption $E_{\Pi_t}(z_{t+1}^k) = 1$. More specifically, the presence of forecasting mistakes of agents reduces the covariance between realized consumption growth and realized returns on risky assets.

A general equilibrium perspective requires us to think of equilibrium prices and rates of return on assets as functions of the distribution of mistakes. The presence of agent's mistakes in the market generates variability of prices and rates of return which is endogenously propagated. This implies, for example, that the Grossmann-Shiller (1981) model under the perfect foresight assumption is misspecified and leads to biased estimates. Equally so, if one ignores endogenous uncertainty then the covariance between consumption and risky rates of return is too small to justify the observed premium (as in Mankiw and Zeldes (1991)).

We briefly address the MaCurdy-Shoven (1992), (1993) perspective. Although MaCurdy and Shoven did not use such a terminology, we have already indicated that their claim amounts to an "asset allocation puzzle" rather than "an equity premium puzzle". We explain what an "asset allocation puzzle" is with a model of an agent who maximizes over two periods. The reason for this choice is that in Section 4 we view a mutual fund as selecting an optimal portfolio to maximize the expected value of a utility function of second period wealth W of the form

$$(8) \quad u(W) = \frac{1}{1-\gamma} W^{1-\gamma}, \quad \gamma > 0.$$

Assume that the financial assets of the economy consist of only two securities: stocks and bills (or bonds) and an economic agent (a household or a mutual fund) has the utility function (8). Let R_t^f denote 1 plus the risk free interest rate (or bond rate) at t for loans paid at $t+1$ and ζ_t the share of

wealth allocated by the agent to stocks at t . The first order conditions of the optimization are then

$$(9) \quad E_{Q_t} \left\{ (R_t^F + \zeta_t \rho_{t-1})^{-\gamma} \rho_{t-1} \right\} = 0$$

where $\rho_{t-1} = R_{t-1} - R_t^F$ is the risk premium. Assume that the agent takes the empirical distribution of ρ as his belief Q and $(R_t^F - 1) = 1\%$. Then, an asset allocation puzzle is defined by the condition that an optimal solution specifying $\zeta_t = 1$ requires γ to be very large and for smaller γ , $\zeta_t > 1$. This implies that most optimizing agents should not hold bonds in their portfolio. The empirical fact is that in planning for retirement most pensions funds and most financial institutions as well as many individual households hold substantial portions of their portfolios in fixed income instruments and show no sign of facing borrowing constraints. Under Rational Beliefs their behavior is entirely rational and we now turn to explain why.

In an environment which is stationary and in which agents know that it is stationary the MaCurdy-Shoven argument is compelling. On the other hand, a non-stationary environment drastically alters this intuition. To understand why we propose that the reader thinks of the time series of the capital market as a sequence of “regimes” in which the moments of the stochastic process of stock prices change drastically. Each such regime has a random length of, say, 1-20 years and within a regime the parameters of the process are fixed. Hence, within each regime, the process is stationary. In such an environment a decision to substantially change asset allocation could have a dramatic impact on the long term performance of a portfolio. To support such a view note that during the 90 years 1905-1995 there were at least three major phases (i.e. 1905-1920, 1929-1941 and 1966-1981) each lasting more than 10 years in which the real values of equities on the New York Stock Exchange *declined by more than 60%!!* Consequently, a person in 1966 facing retirement in 15 years would have been prudent not to risk his standard of living in retirement and keep a fraction of his savings in fixed income securities of, say, 1-3 years maturity. During the 15 years 1966-1981 the equities portion

of his portfolio would have declined in value by 72% while the fixed income component would have earned a small positive return.

More generally, the long term average rates of return give a deceiving picture of the mean value function which an investor may rationally believe that he faces at any moment of time. Consequently, if an investor believes that the capital market has the non-stationary structure described above, a strategy of switching over time between equities and fixed income securities is optimal. Moreover, if an individual does not wish to take the risk of reduced standard of living in retirement, he may optimally hold a portfolio consisting entirely of fixed income securities from a certain age on if he believes that a sustained phase of low returns on equities is ahead. In order to keep in mind the social consequences of retirement planning recall that the financial collapse of a generation of retirees in the 1930's prompted the creation of the social security system.

A detailed analysis of the different phases of the asset markets in the U.S. during the period 1947-1992 is carried out by Kurz (1997). He shows that a substantial portion of the mistakes of the agents represented by z_{t+1}^k in equations (6a)-(6b) can be explained ex-post by the various regime variables. This last fact leads to his conclusion that endogenous uncertainty is the predominant form of uncertainty in equity markets.

The rest of this paper is devoted to a positive application of the theory of rational beliefs. Our analysis in both Section 4 and 5 is based on modeling the behavior of agents as two-period optimizers in an economy with two financial assets: a stock and a "bill" which is a one period debt instrument. In both Sections agents hold rational beliefs and select optimal portfolios to maximize their expected utility given their probability beliefs about dividends and prices in the second period. Section 4 is a partial equilibrium analysis of the asset allocation of U.S. mutual funds. In Section 5 we solve numerically a general equilibrium model of two infinite sequences of households, each constituting an

overlapping generations (OLG) "dynasty". At each date the two "young" households select optimal consumption and portfolios given their probability beliefs about dividends and prices in the next period.

In addition to the terms defined so far, we employ in both Sections the following notation:

x_t^{1k} - the consumption of k when young at t ;

x_{t+1}^{2k} - the consumption of k when old at $t + 1$. This indicates that k was born at date t ;

θ_t^k - amount of stock purchases of young agent k at t ;

B_t^k - amount of one period debt instrument ("bill") purchased by k at t ;

Ω_t^k - endowment of k when young at t .

P_t - the price of the common stock at t ;

q_t - the price of a one period debt instrument ("bill") at t . This is a discount price;

I_t - information available at t which is the history up to t ;

$u^k(\cdot, \cdot)$ - the utility function of agent k .

In the analysis in both Section 4 and 5 we normalize prices by using consumption as a numeraire.

Given this, the optimization problem of agent k has the following common structure:

$$(10a) \quad \begin{aligned} & \text{Max}_{(x^{1k}, \theta^k, B^k, x^{2k})} E_{Q_t^k} \left\{ u^k(x_t^{1k}, x_{t+1}^{2k}) \mid I_t \right\} \\ & \text{subject to} \end{aligned}$$

$$(10b) \quad x_t^{1k} + P_t \theta_t^k + q_t B_t^k = \Omega_t^k$$

$$(10c) \quad x_{t+1}^{2k} = \theta_t^k (P_{t+1} + D_{t+1}) + B_t^k.$$

In the problem of a mutual fund of section 4 we set $x_t^{1k} \equiv 0$ and make the identification

$\Omega_t^k \equiv W_t^k$ where W_t^k is the value of assets under management by the fund. In that case condition

(10b) is the budget constraint and (10c) is the definition of second period wealth $x_{t+1}^{2k} \equiv W_{t+1}^k$. The

utility function of the fund will be assumed to be $u^k(W_{t+1}^k) = \frac{1}{1-\gamma_k} (W_{t+1}^k)^{1-\gamma_k}$ and it will select the optimal portfolio to maximize expected wealth in the second period.

In Section 5 we assume a standard OLG economy with two young and two old agents; such an economy is different from the infinite horizon single agent economy of MP (1985). In addition to the beliefs of the agents, the presence of the endowment Ω_t^k is the most distinct feature of the difference between the models. In all other respects our assumptions correspond to those made by MP (1985).

Hence, the common utility function of the two agents is of the form

$$(11) \quad u(x^1, x^2) = \frac{1}{1-\gamma}(x^1)^{1-\gamma} + \frac{\beta}{1-\gamma}(x^2)^{1-\gamma} \quad \gamma > 0.$$

We maintain, however, that the portfolio optimization conditions of the finitely lived agents in our model remain the same as in MP (1985) and the significance of the risk aversion coefficient for the equity premium puzzle remains intact. This will be discussed in Section 5.

4. Asset Allocation in Mutual Funds: Testing Some Implications of the theory of Rational Beliefs

This section describes econometric analysis of some implications of the theory of RBE which are relevant to the portfolio allocation of mutual funds. As the previous sections clarified, the main goals of this work are the description of the mistake functions of the optimizing mutual funds, a demonstration of the existence of heterogeneity across them and the estimation of γ_k . The first two are central to the theory of RBE while the third is important for evaluating the economic realism of our model. Also, the simulations in Section 5 will be done with respect to *a relatively narrow range of this parameter* based on the empirical evidence which is available from our sample.

We remark that Hansen and Singleton (1982), using aggregate consumption data, estimated values of γ_k which are not economically plausible and in some cases did not even imply concavity of the utility function. Many other studies, using methods based on rational expectations, obtained similar results regardless of methodologies, time periods and data sets. Kurz (1997), applying the theory of RBE, obtains an economically reasonable estimate of γ_k by taking into account the existence of

structural breaks in the data set. Our econometric study complements Kurz's (1996) analysis by bringing in a truly micro-economic perspective via the study of portfolio allocation of funds.

Before describing our methodology and results we want to stress why such a data set is particularly useful for testing some empirical implications of the theory of RBE. First, funds managers are professionals who make decisions on the basis of extensive analysis of available information. Given the size of these funds it is reasonable to assume that they all have at their disposal every available information and thus they are all approximately equally informed. The theory of rational beliefs explains that the attempt to use existing information in the best possible way cannot avoid mistakes. These mistakes should, therefore, be present in the time series of portfolio asset allocation of the managers. The second reason for the usefulness of this data set is the known objective of the managers of the funds. The 63 funds in the sample are all classified as "Balanced", "Growth" or "Growth and Income" funds when these terms mean that these funds specialize in selecting an optimal allocation among the asset categories of "stocks" "bonds" or "cash". This in contrast with specialized funds which invest in particular industries and seek to maximize return by choosing an *optimal mix of firms* in which to invest. Thus, the equity premium is the direct variable motivating a balanced fund.

We treat the managers of a mutual fund as maximizers of an expected utility of wealth over one period who go through a sequence of portfolio allocation over time. This is a realistic assumption for several reasons. First, managers are evaluated regularly on the basis of the returns which are achieved on initial wealth and in most cases part of the fund's compensation is *proportional* to these periodic rates of return. Second, wage incentive schemes of fund managers (as distinct from the funds themselves), are structured in such a manner that the compensation depends in large part on measures related to the returns achieved on the portfolio. This factor tends to induce a short run perspective based on expected utility defined over short term returns to wealth. Third, although there exist some

reputation effects in the mutual funds industry, there is little evidence for long term relationship between funds and investors in the funds.

In order to set the stage for the econometric analysis it is convenient to recast the constraints (10b) and (10c) in terms of rates of return. This is done by solving (10b) for B_t^k and substituting the result in (10c) to have

$$(12) \quad W_{t+1}^k = P_t \theta_t^k (R_{t+1} - \frac{1}{q_t}) + \frac{W_t^k}{q_t}.$$

Dividing (12) by W_t^k and noting that by definition $R_t^F = \frac{1}{q_t}$ and $\zeta_t^k = \frac{P_t \theta_t^k}{W_t^k}$, one obtains

$$(13) \quad \frac{W_{t+1}^k}{W_t^k} = R_{t+1}^F + \zeta_t^k \rho_{t+1}.$$

The problem of the fund, therefore, is

$$(14) \quad \max_{\zeta_t^k} E_{Q_t^k} \frac{1}{(1 - \gamma_k)} (R_{t+1}^F + \zeta_t^k \rho_{t+1})^{1 - \gamma_k}.$$

The first order condition of the maximization is $E_{Q_t^k} (R_{t+1}^F + \zeta_t^k \rho_{t+1})^{-\gamma_k} \rho_{t+1} = 0$. As emphasized in (6)- (7), this condition may be used to estimate the mistake function of the agent and γ_k . The mistake function is approximated by the following regression model

$$(15) \quad (R_{t+1}^F + \zeta_t^k \rho_{t+1})^{-\gamma_k} \rho_{t+1} = \alpha^k X_t + \epsilon_{t+1}^k$$

$$(16) \quad E_{\Pi_t}(\epsilon_{t+1}^k X_t) = 0$$

where α^k is a vector of coefficients for fund k, X_t is a vector of information variables known at the beginning of time t. This vector will include some fund specific variables. In order to estimate γ_k and α^k we use a GMM procedure based on the following orthogonality conditions

$$(17) \quad E_{Q_t^k} [(R_{t+1}^F + \zeta_t^k \rho_{t+1})^{-\gamma_k} \rho_{t+1} - \alpha^k X_t] Z_t = 0$$

where Z_t is a vector of instruments, some specific to fund k. The number of our instruments is larger than the number of parameters to be estimated, giving rise to a set of overidentifying restrictions. We thus test these restrictions for the joint orthogonality between the residuals of the equation and the

instruments (see Hansen (1982)). By increasing the number of instruments we hope to increase the precision of the estimates although the test becomes more stringent (see Hansen and Singleton (1982)). This is useful as our sample is small.

To estimate (17) we consider the six-month allocation among classes of assets for each of our 63 funds for the period 1982:4-1995:1. This makes available 25 observations per fund. We use the total return on the S&P 500 (dividend yield plus capital gains) to approximate the rate of return on stocks, and the rate on three months Treasury Bills to approximate the rate of return on "bills" (see the Appendix on Data Description for information on the data). We conduct a six-month analysis in order to compute rates of return over six-month intervals avoiding the problems which arise from the use of overlapping data. Our procedure was also motivated by the widely recognized fact that short term stock returns are dominated by noise (see Campbell and Shiller (1988) and Fama and French (1988)). Thus, in order to detect any effect of information variables on the mistake functions, it is desirable to reduce the noise by averaging over some time interval.

Motivated by the reasoning in Kurz (1997), we have chosen the following regressors in (17):

X_1 - the rate of growth of real GDP over 4 past quarters,

X_2 - the lagged rate of growth of real GDP over 4 past quarters,

X_3 - the rate of growth of output per man-hour,

X_4 - the discount rate,

X_5 - the lagged risk premium,

X_6 - the risk premium lagged twice,

X_7 - the fraction of assets of the fund allocated to stocks ($\equiv \zeta^k$ for k); this is a *fund specific* variable.

X_7 may be particularly useful to incorporate the effect of any omitted variable. The list of our instruments is as follows: the regressors, the rate of growth of real GDP over the 4 past quarters lagged

twice, the lagged rate of growth of output per man-hour, the rate of growth of the index of vendors over the past 4 quarters, the rate of growth of the index of basic commodities over the past 4 quarters, the rate of growth of M1 over the past 8 quarters, the lagged discount rate, a dummy which is equal to 1 if the discount rate was increased during the past quarter lagged once and twice, index of vendors lagged once and twice, the risk premium lagged three times, the rate of growth of manufacturing output over the past 4 quarters, the index of capacity utilization and its lag.

The results are reported in Table 1 of Appendix 1, for each of the 63 funds in the sample. We make a few comments on these results.

(1) The test of the overidentifying restrictions *never* rejects the specification. This is notable in light of the relatively large number of instruments used in order to increase precision of the estimates.

(2) The estimates of the coefficient of risk aversion, in the first column of Table 1 (Appendix 1), are mostly within the reasonable range of 2 and 4. More specifically, 15 estimates are below 2.25 and 16 are above 3.75, leaving more than 50% in the range of 2.25 - 3.75.

(3) The mistake functions of the funds exhibit a large number of significant variables. Most parameters corresponding to information variables in the columns 3 - 9 of Table 1 have coefficients which are significantly different from 0. The Euler equations for the funds are therefore not orthogonal to existing public information of macroeconomic nature, as predicted by the theory of rational beliefs.

(4) We find much heterogeneity across funds. Some variables have a uniform impact on the funds while others have a very different impact. For example X_7 , the share of wealth allocated to stocks, is *positive* for 27 funds (significantly different from 0 at the 10% level for 9) and *negative* for 36 (significantly different from 0 at the 10% level for 11). A Wald test of the equality of all coefficients across all funds excluding γ_k yields a statistic equal to 2449.87, against a 5% critical value equal to 483.57. We conclude that there is considerable heterogeneity among the mistake functions of the funds.

In conclusion, our results have two important implications to this paper. First, the presence of significant mistake functions which are heterogenous across agents are consistent with the predictions of the theory of rational beliefs and thus support the basic paradigm employed in this paper. Under the assumptions which we made about the nature and objective of the funds *the only explanation for the above conclusions is that the funds had different probability beliefs about future returns* and adopted investment strategies which reflected these beliefs. Second, the results reported here are important for the simulations in Section 5 in that they demonstrate that a narrow range of, say, 2.25 - 3.75 for γ_k covers more than 50% of the funds.

5. Analysis of the Equity Premium in a Rational Belief Equilibrium: Simulating the Economy

We turn now to the two-agent OLG model discussed in Section 3. In the next section we construct a family of rational belief equilibria for this economy: this family is our central object of analysis. We then calibrate the model to the empirical evidence provided by the long term time series of the U.S. economy as in MP (1985) and compute the moments of the long run distribution of the rates of return and premia implied by the model. We then compare our results with those of MP (1985) and others. As noted in Section 3 both MP (1985) as well as other studies of the equity premium examine the problem with a model of a single, infinite lived, household. Since we compare our results to those in the literature, we clarify the differences between the models used.

We start with the finite life of the agents. We postulate that the utility functions of the two agents are as specified in (11) and, given the discount factors, it is well known that the conditions on the optimal portfolios of our sequence of agents (with the same two period utility functions) are equivalent to the optimality conditions of an infinitely lived agent in a MP (1985) type economy. It then follows that the date t spot security markets of the two economies are entirely comparable. As to the

issue of beliefs, we note that heterogeneity of beliefs is central to our approach and in that sense the two models are very different indeed. However, given our comment about the equivalence of the spot securities markets under the two models, a formulation of our economy as one in which the two heterogeneous agents are infinitely lived would contribute little to the comparability of the two models. Apart from beliefs, the essential difference between the models is the presence of exogenous endowments in our OLG economy.

The MP's (1985) single agent, aggregate, model of the economy aims to study the entire U.S. economy but has the extreme features of excluding all natural resources and all types of labor. As the theoretical model is due to Lucas (1978) the approach taken by MP (1985) must be viewed as a model of the financial sector only. We do not suggest that there is any fault in that. What we do stress is that even within the category of capital income, one must interpret the model as representing only the profits (i.e. the "dividends") of the corporate sector whose shares are traded on public exchanges. This excludes the very large parts of capital income such as the profits of all corporations whose shares are not traded on public exchanges, all non-corporate businesses such as farms and real estate ventures, all owner-occupied housing and all output of consumer durables. In fact, the sector represented in the model is rather small relative to the total economy. The capital categories excluded are often investments where *reported* profits do not reflect all the benefits of ownership. In many of these situations the risk to an owner-investor and to other investors are not symmetric. Moreover, there are many dimensions of risk and liquidity in the capital ownerships of these investment categories which are not equivalent to the risk of owning a publicly traded liquid security.

We stress that if the excluded part of the economy has any effect on the financial sector under study then there is some advantage in representing the rest of economy in the model even if it is in an elementary and exogenous manner. This is the advantage of our strategy to include the endowment

vector (Ω^1, Ω^2) in the model as parameters which represent the effects of the rest of the economy on the financial sector under study. In our view "the rest of the economy" includes many components of national output which are not included in GNP; some were mentioned above but there are others such as household work of female members. In the present study we then take these variables *strictly as calibrating parameters* which can be used as tools of analysis. As it turns out, the level of aggregate endowment has virtually no effect on the equity premium as intuition would suggest. However the endowment has other important effects that will be discussed later.

(5.a) Rational Belief Equilibria of the Two-Agent OLG Economy

Our development of the simulation model uses concepts from the theory of rational beliefs reviewed in the *Editor's Introduction* to this volume and the tools of "generating variables" and Markov RBE developed in Kurz and Schneider (1996). Although we explain below how these tools are used here, the reader who seeks additional details may benefit from these cited papers.

(a) *The Dividend Process and the Budget Constraints.* The simulation model is relatively simple, with a single homogenous consumption good and two agents denoted $k = 1, 2$ who have the same utility function over consumptions (x^1, x^2) of the form specified in (11). Since this is an OLG economy one must think of the model as one of dynasties where each of the two dynasties is characterized by a two period utility function (11) and a rational belief which we shall specify later. As in MP (1985) the dividend process $\{D_t, t = 1, 2, \dots\}$ follows (3) with the growth rate process $\{d_t, t = 1, 2, \dots\}$ specified to be a stationary and ergodic Markov process. Its state space is $\{d^H, d^L\}$ with $d^H = 1.054$ and $d^L = .982$ and a transition matrix

$$(18) \quad \begin{bmatrix} \phi & 1 - \phi \\ 1 - \phi & \phi \end{bmatrix}$$

with $\phi = .43$. This means that over time agents experience a rise in the level of dividends and this requires us to redefine the budget constraints. To do that let ω^k be the endowment/dividend ratio of agent k and b_t^k to be the bill/dividend ratio of that agent at date t . Also, let p_t be the price/dividend ratio of the common stock at t . We assume that ω^k for $k=1,2$ are constant over time in order to accord with the MP's (1985) assumption that the growth rate of the economy *as a whole* is a stationary Markov process with a transition matrix (18). This requires us to assume that there is a constant v such that $(\Omega_t^1 + \Omega_t^2) = vD_t$ for all t . One could introduce into the model a fluctuating endowment of the young and this would contribute to the randomness of the economy. We are not making this assumption in part because we are not modeling this sector of the economy explicitly and in part in order to avoid the complications which arise in a model with incomplete insurance markets implied by the inability of the young to insure against these risks.

Rewriting (10b)-(10c) and using the notation introduced, we conclude that

$$(19b) \quad x_t^{1k} = [\omega^k - p_t \theta_t^k - q_t b_t^k] D_t$$

$$(19c) \quad x_{t+1}^{2k} = [\theta_t^k (p_{t+1} + 1) d_{t+1} + b_t^k] D_t$$

We now write down the first order conditions of the maximization of the agents and note that they are independent of D_t if the beliefs of the agents are not conditioned on D_t since

$$(20a) \quad -(\omega^k - \theta_t^k p_t - b_t^k q_t)^{-\gamma} p_t + \beta E_{Q_t^k}(\theta_t^k (p_{t+1} + 1) d_{t+1} + b_t^k)^{-\gamma} (p_{t+1} + 1) d_{t+1} = 0$$

$$(20b) \quad -(\omega^k - \theta_t^k p_t - b_t^k q_t)^{-\gamma} q_t + \beta E_{Q_t^k}(\theta_t^k (p_{t+1} + 1) d_{t+1} + b_t^k)^{-\gamma} = 0$$

It is then clear that we need to specify what the Q_t^k are.

(b) *Generating Variables and the State Space.* The difficulty in specifying the belief of the agents arises from the fact that beliefs are probabilities over future prices and dividends. In order to state such probabilities we need to specify the state space on which prices are defined and this state space depends upon the beliefs. The formulation of the endogenous state space of prices is one of the central problems addressed in this volume and we refer the reader to Kurz and Schneider (1996) for details on the case of Markov processes. To explain how it is done for the relatively simple case used in this paper we divide the presentation into two steps. In the first step we specify the rule according to which the beliefs of the agents are constructed. In the second step, to be completed in the next subsection, we specify the probabilities themselves to ensure that these are rational beliefs.

The beliefs of the agents are formulated using the method of generating variables (see Kurz and Schneider (1996)). For $k=1,2$ we denote these variables by $\{y_t^k, t=1,2,\dots\}$: they are simply a pair of stochastic processes. In our application they take values in $Y = \{1, 0\}$. The central assumption is that each agent believes that the joint process $\{(p_t, q_t, d_t, y_t^k), t=1,2, \dots\}$ is a Markov process. This means that generating variables are, in general, assumed to be interdependent with the real variables in the economy. Past observable variables can be used to forecast future values of the generating variables and present generating variables are used to forecast future values of the observable variables. Suppose now that the number of possible pairs of equilibrium prices is M then we define $V_p = \{(p_1, q_1), (p_2, q_2), \dots, (p_M, q_M)\}$, $J_d = \{d^H, d^L\}$. Now let $\mathcal{P}_p^- = (V_p \times J_d \times Y)^-$ be the space of all infinite sequences of the variables and denote by $\mathcal{B}$ the Borel σ -field of the appropriate space. Then the belief Q^k of agent k is a probability on $(\mathcal{P}_p^-, \mathcal{B}(\mathcal{P}_p^-))$. However, we can equivalently define the beliefs to be probabilities on the space of infinite sequences of price indices. That is, define $V = \{1, 2, \dots, M\}$ to be *the state space for prices* and let $\mathcal{P} = (V \times J_d \times Y)$. Then beliefs are probabilities on $(\mathcal{P}^-, \mathcal{B}(\mathcal{P}^-))$ which can be constructed from an initial distribution on $(V \times J_d \times Y)$ together with a $4M \times 4M$ transition matrix

(F for agent 1 and G for agent 2) on $(V \times J_d \times Y) \times (V \times J_d \times Y)$.

As a matter of economic interpretation, generating variables can be viewed as parameters indicating how the agent perceives the state of the process and are thus tools for the description of stable and non-stationary processes (see Kurz and Schneider (1996) on this technical point). These variables can also be thought of as private signals with purely subjective meaning to the agent and therefore should not be taken to be objective "information". The variables may be functions of past observed data of prices and dividends hence they can be interpreted as representing the assessment of a "research department" of an organization. Keep in mind that by themselves, *generating variables have no intrinsic meaning*. They gain significance from the way the agent specifies how these variables are to be interpreted within the joint dynamical system with the observed data.

Treating y_t^k at each date t as a *signal*, agent k conditions on the signal jointly with the observed data to derive the conditional probability of $(p_{t+1}, q_{t+1}, d_{t+1}, y_{t+1}^k)$ given (p_t, q_t, d_t, y_t^k) . This, finally, brings us back to the first order conditions (20a)-(20b). It follows from our Markov assumptions that the demands of agent k for stocks and bills are functions of (p_t, q_t, d_t, y_t^k) . Consequently we can write the market clearing conditions as

$$(21a) \quad \theta_t^1(p_t, q_t, d_t, y_t^1) + \theta_t^2(p_t, q_t, d_t, y_t^2) = 1$$

$$(21b) \quad b_t^1(p_t, q_t, d_t, y_t^1) + b_t^2(p_t, q_t, d_t, y_t^2) = 0.$$

The system (21a)-(21b) implies that the equilibrium map of this economy specifies that prices are functions of the form

$$(22) \quad \begin{bmatrix} p_t \\ q_t \end{bmatrix} = \Phi^*(d_t, y_t^1, y_t^2)$$

and the map (22) implies that $M = 8$: at most 8 prices will be observed in this economy under the equilibrium map (22). This solves the problem of the state space for prices; it is $V = \{1, 2, \dots, 8\}$. We then define a new map Φ between the *indices* of prices and the states of dividends and generating variables (which are indexed by a number from 1 to 8 rather than by t) by

$$(23) \quad \begin{bmatrix} 1 \\ 2 \\ 3 \\ 4 \\ 5 \\ 6 \\ 7 \\ 8 \end{bmatrix} = \Phi \begin{bmatrix} d_1 = d^H, y_1^1 = 1, y_1^2 = 1 \\ d_2 = d^H, y_2^1 = 1, y_2^2 = 0 \\ d_3 = d^H, y_3^1 = 0, y_3^2 = 1 \\ d_4 = d^H, y_4^1 = 0, y_4^2 = 0 \\ d_5 = d^L, y_5^1 = 1, y_5^2 = 1 \\ d_6 = d^L, y_6^1 = 1, y_6^2 = 0 \\ d_7 = d^L, y_7^1 = 0, y_7^2 = 1 \\ d_8 = d^L, y_8^1 = 0, y_8^2 = 0 \end{bmatrix}$$

We refer to d^H as the "high dividends" and d^L as the "low dividends" states. The maps (22)-(23) highlight the definition of Endogenous Uncertainty which is the variability of prices at a given state of the exogenous variables (for a formal definition, see Definition 2 of Kurz and Wu (1996)).

We now make the simplifying assumption that the marginal distributions of y^1 and y^2 implied by F and by G are specified to be i.i.d. and we denote the unconditional probabilities by $P\{y_t^k = 1\} = \alpha_k$ for $k=1,2$. This means that the agents have two pairs of matrices (F_1, F_2) and (G_1, G_2) such that the beliefs Q^1 and Q^2 are characterized by the following rule:²

²Note that each of the matrices F and G are 32×32 while F_1, F_2, G_1, G_2 are all 8×8 . This is the consequence of the fact that the agent discovers from the stationary measure all the price-dividend combinations which have zero asymptotic relative frequency. Also, the agents know the asymptotic identity of the high and low dividend states. Since in this paper we are concerned only with the calculations of the long term statistics of the economy we neglect all the rows of the matrices for which the stationary probabilities are 0.

$$(24) \quad \begin{array}{ll} Q^1 \text{ for agent 1: adopt } F_1 & \text{if } y_t^1 = 1 \\ & \text{adopt } F_2 \text{ if } y_t^1 = 0 \end{array} \quad \begin{array}{ll} Q^2 \text{ for agent 2: adopt } G_1 & \text{if } y_t^2 = 1 \\ & \text{adopt } G_2 \text{ if } y_t^2 = 0. \end{array}$$

We finally denote by $Q^k(j|s, y^k)$ agent k 's conditional probability of price state j given price state s and the value of y^k but under the competitive assumption that k neither knows the map (22) nor does he know that he influences prices. The first order conditions (20a)-(20b) are restated for $k = 1, 2$ and $j, s = 1, 2, \dots, 8$:

$$(25a) \quad -(\omega^k - \theta_s^k p_s - b_s^k q_s)^{-\gamma} p_s + \beta \sum_{j=1}^8 (\theta_s^k (p_j + 1) d_j + b_s^k)^{-\gamma} (p_j + 1) d_j Q^k(j|s, y^k) = 0$$

$$(25b) \quad -(\omega^k - \theta_s^k p_s - b_s^k q_s)^{-\gamma} q_s + \beta \sum_{j=1}^8 (\theta_s^k (p_j + 1) d_j + b_s^k)^{-\gamma} Q^k(j|s, y^k) = 0.$$

Once we specify (Q^k, ω^k) for $k = 1, 2$ (recall that $d_1=d_2=d_3=d_4=d^H$ and $d_5=d_6=d_7=d_8=d^L$) we compute the demand functions (θ_s^k, b_s^k) as functions of the 8 prices. In equilibrium

$$(26a) \quad \theta_s^1 + \theta_s^2 = 1 \quad \text{for all } s$$

$$(26b) \quad b_s^1 + b_s^2 = 0 \quad \text{for all } s.$$

(25a)-(25b), (26a)-(26b) is then a system of 48 equations in prices and quantities which are the basis of our simulation work.

(c) *The Stationary Measure*³. We have already specified in (18) the stationary dividend process and the implied marginal probability measure m_D in accordance with the Markov assumptions commonly

³ For a review of the basic concepts of the theory of rational beliefs see Appendix 3.

made in the literature (e.g. MP (1985)). However, in an RBE the driving mechanism is the distribution of the sequences (d_t, y_t^1, y_t^2) , $t = 1, 2, \dots$ which is a stable dynamical system $(\varphi_J^*, \mathcal{CB}(\varphi_J^*), \Pi_{DY}, T)$ where $\varphi_J = (J_d \times Y \times Y)$ with a stationary measure m_{DY} . To understand this point recall that each agent has a marginal distribution on his own generating variables and we have just mentioned the marginal distribution of the dividend process. The marginal distributions specify only what each agent perceives and not what they jointly do. Since in this paper we are only concerned with the long term averages we focus on m_{DY} . It specifies all the interactions among the agents which reflect the structure of communication in society and the manner in which agents influence each other and how the real variables in the economy (i.e. the dividends) affect this interaction. The stationary measure m_{DY} must, however, be compatible with the specification of the dividend process and the assumptions which we have already made on the marginal distributions of the generating variables. More specifically, we require that

- (27a) the marginal measures m_{Y^k} specify y_t^k to be i.i.d. with $P\{y_t^k = 1\} = \alpha_k$;
- (27b) the marginal measure m_D is specified by the stationary dividend process (18);
- (27c) the joint distribution of (y_{t-1}^1, y_{t-1}^2) may depend upon d_t .

There are many matrices which satisfy these conditions but we select one which allows flexibility in the parametrization of the final equilibrium. The following transition matrix Γ defines a stationary probability measure m_{DY} which satisfies these conditions:

$$(28) \quad \Gamma = \begin{bmatrix} \phi A, & (1 - \phi)A \\ (1 - \phi)B, & \phi B \end{bmatrix}$$

where A and B are 4×4 matrices which are characterized by the 10 parameters $(\alpha_1, \alpha_2, a, b)$ and

$a = (a_1, a_2, a_3, a_4)$, $b = (b_1, b_2, b_3, b_4)$:

$$(29) \quad A = \begin{bmatrix} a_1 & \alpha_1 - a_1 & \alpha_2 - a_1 & 1 + a_1 - \alpha_1 - \alpha_2 \\ a_2 & \alpha_1 - a_2 & \alpha_2 - a_2 & 1 + a_2 - \alpha_1 - \alpha_2 \\ a_3 & \alpha_1 - a_3 & \alpha_2 - a_3 & 1 + a_3 - \alpha_1 - \alpha_2 \\ a_4 & \alpha_1 - a_4 & \alpha_2 - a_4 & 1 + a_4 - \alpha_1 - \alpha_2 \end{bmatrix}, \quad B = \begin{bmatrix} b_1 & \alpha_1 - b_1 & \alpha_2 - b_1 & 1 + b_1 - \alpha_1 - \alpha_2 \\ b_2 & \alpha_1 - b_2 & \alpha_2 - b_2 & 1 + b_2 - \alpha_1 - \alpha_2 \\ b_3 & \alpha_1 - b_3 & \alpha_2 - b_3 & 1 + b_3 - \alpha_1 - \alpha_2 \\ b_4 & \alpha_1 - b_4 & \alpha_2 - b_4 & 1 + b_4 - \alpha_1 - \alpha_2 \end{bmatrix}$$

If $A \neq B$ then the distribution of (y_{t+1}^1, y_{t+1}^2) depends upon d_t . Also, (29) implies that $P\{y_t^k = 1\} = \alpha_k$ for $k = 1, 2$ and this is compatible with our individual specifications. Note, however, that although each process $\{y_t^k, t = 1, 2, \dots\}$ for $k = 1, 2$ is very simple, the joint process $\{(d_t, y_t^1, y_t^2), t=1,2,\dots\}$ may be complex: it allows correlation among the three variables over time and we use these effects in the simulations. However, if we set $\alpha_1 = \alpha_2 = .5$ and $a_i = b_i = .25$ for $i = 1, 2, 3, 4$ then all correlations among the three central variables are eliminated. It is easy to see that in this case the stationary distribution $(\pi_1, \pi_2, \dots, \pi_8)$ implied in (28) is $\pi_i = .125$ for all i . If, in addition, we assume that the agents adopt the stationary measure as their belief, then we have exactly the rational expectations equilibrium of MP (1985). This defines an important test of our model: under the conditions of MP (1985) it should replicate the equity premium puzzle!

The central case of our simulation results will exploit the interdependence in the joint distribution of the three variables (d_t, y_t^1, y_t^2) which the model permits. In fact, in search for simplicity we set the following parameter values in all our simulations: $\alpha_1 = \alpha_2 = .5$; $a_1 = a_4, a_2 = a_3$; $b_1 = b_4, b_2 = b_3$ with the following numerical choices: $b_1 = b_4 = .001$; $b_2 = b_3 = .01$; $a_1 = a_4 = .12$; $a_2 = a_3 = .43$. We can see from (28) and (29) that the result of these choices is that whenever price states $\{5,6,7,8\}$ are realized then it is virtually certain that they will be followed by a state in the set $\{2,3,6,7\}$. On the other hand if price states $\{1, 4\}$ are realized then with probability of .24 they will be followed by price states $\{1,4,5,8\}$ and if price states $\{2,3\}$

are realized then with the high probability of .86 they will be followed by one of the price states $\{1,4,5,8\}$. These transitions of (d_t, y_t^1, y_t^2) imply very strong joint movements of these three variables although the marginal distribution of each of the y_t^k is i.i.d and the marginal distribution of d_t is described by (18). We stress, however, that these correlations have no economic meaning without a specification of how the agents interpret the signals provided by their generating variables. Thus, we must then turn to the crucial question of specifying the family of rational beliefs which we shall use in the simulations.

(d) *Rational Beliefs.* The price state space which we selected implies that the agents have two pairs of matrices: (F_1, F_2) for agent 1 and (G_1, G_2) for agent 2. It follows from Nielsen [1994] (Section 4.2) that rationality of beliefs requires

$$(30) \quad \alpha_1 F_1 + (1 - \alpha_1) F_2 = \Gamma, \quad \alpha_2 G_1 + (1 - \alpha_2) G_2 = \Gamma.$$

A word of intuition may be helpful here. The rational agents believe that the price-dividend process is not stationary and their beliefs are parametrized by their private signals (y_t^1, y_t^2) . At different dates they may adopt different Markov matrices and hence different consumptions and portfolios. (30) insists, however, that the sequence of matrices which they adopt is compatible (in the sense of generating the same empirical distribution) with the view that the price-dividend process is a stationary Markov process with transition matrix Γ . Given (30), the selection of the conditional probabilities (where F_1^{sj} is the (s,j) element of F_1)

$$(31) \quad Q_t^1(j | s, y_t^1) = \begin{cases} F_1^{sj} & \text{if } y_t^1 = 1 \\ F_2^{sj} & \text{if } y_t^1 = 0 \end{cases} \quad Q_t^2(j | s, y_t^2) = \begin{cases} G_1^{sj} & \text{if } y_t^2 = 1 \\ G_2^{sj} & \text{if } y_t^2 = 0 \end{cases}$$

defines the beliefs Q^k for $k = 1, 2$. We next select the four matrices (F_1, F_2, G_1, G_2) by using two sets of 8 parameters $\lambda = (\lambda_1, \lambda_2, \dots, \lambda_8)$ and $\mu = (\mu_1, \mu_2, \dots, \mu_8)$ which will be motivated later. To do that

we introduce the notation for the row vectors of A and B :

$$\begin{aligned} A^j &= (a_j, \alpha_1 - a_j, \alpha_2 - a_j, a_{j4}) & a_{j4} &= 1 + a_j - (\alpha_1 + \alpha_2) \\ B^j &= (b_j, \alpha_1 - b_j, \alpha_2 - b_j, b_{j4}) & b_{j4} &= 1 + b_j - (\alpha_1 + \alpha_2). \end{aligned}$$

With this notation we define the 4 matrix functions of $z = (z_1, z_2, \dots, z_8)$ as follows:

$$(32) \quad A_1(z) = \begin{bmatrix} z_1 A^1 \\ z_2 A^2 \\ z_3 A^3 \\ z_4 A^4 \end{bmatrix}, \quad A_2(z) = \begin{bmatrix} (1 - \phi z_1) A^1 \\ (1 - \phi z_2) A^2 \\ (1 - \phi z_3) A^3 \\ (1 - \phi z_4) A^4 \end{bmatrix}, \quad B_1(z) = \begin{bmatrix} z_5 B^1 \\ z_6 B^2 \\ z_7 B^3 \\ z_8 B^4 \end{bmatrix}, \quad B_2(z) = \begin{bmatrix} (1 - (1 - \phi) z_5) B^1 \\ (1 - (1 - \phi) z_6) B^2 \\ (1 - (1 - \phi) z_7) B^3 \\ (1 - (1 - \phi) z_8) B^4 \end{bmatrix}$$

Finally we define

$$(33) \quad F_1 = \begin{bmatrix} \phi A_1(\lambda) & A_2(\lambda) \\ (1 - \phi) B_1(\lambda) & B_2(\lambda) \end{bmatrix} \quad G_1 = \begin{bmatrix} \phi A_1(\mu) & A_2(\mu) \\ (1 - \phi) B_1(\mu) & B_2(\mu) \end{bmatrix}$$

and (F_2, G_2) determined by (30). The motivation for this construction of the matrices F_1 and G_1 is that the parameters λ_i and μ_i are multiplied by the rows of the matrices A and B and hence are interpreted as proportional revisions of the conditional probabilities of the four states $(1, 2, 3, 4)$ and $(5, 6, 7, 8)$ *relative to the stationary measure* represented by Γ . Although $\lambda_i > 1$ and $\mu_i > 1$ imply increased probabilities of states $(1, 2, 3, 4)$ in matrix F_1 of agent 1 and G_1 of agent 2 the interpretation of these parameters is made complicated by the fact that the agent 1 may use F_1 or F_2 at any date and similarly for agent 2. It turns out that the useful concepts are those of "agreement" and "disagreement" between the agents. Thus, suppose that $p_t = p_1$ and that $(\lambda_1 > 1, \mu_1 > 1)$. By the map(23) $y_t^1 = 1$ and $y_t^2 = 1$ hence the matrices in use are (F_1, G_1) . By the construction (33) this means that in this state both agents agree that the probability of states $(1, 2, 3, 4)$ is higher than specified in Γ . If, on the other hand, $(\lambda_1 < 1, \mu_1 > 1)$ then the use of the pair (F_1, G_1) would mean "disagreement" since one agent has an *increased* probability and the other a *decreased* probability of states $\{1, 2, 3, 4\}$. Thus, to determine if a state j is one of agreement one needs to consider

(λ_j, μ_j) as well as the pair of matrices in use.

The central case which we consider in the simulations below is the one where we specify (λ_j, μ_j) by:

$$(34a) \quad \lambda_1 = 1.75, \lambda_2 = .25, \lambda_3 = 1.75, \lambda_4 = .25, \lambda_5 = 1.75, \lambda_6 = .25, \lambda_7 = 1.75, \lambda_8 = .25$$

$$(34b) \quad \mu_1 = .25, \mu_2 = 1.75, \mu_3 = .25, \mu_4 = 1.75, \mu_5 = .25, \mu_6 = 1.75, \mu_7 = .25, \mu_8 = 1.75.$$

(e) The Emergence of Regimes. The interpretation of the specification in (34a)-(34b) in terms of price forecasting by an agent at date t depends upon the values of prices and the matrices used at date t . However, a deeper understanding of the behavioral pattern of the agents may be gained by noting that our model leads to the emergence of "regimes" within the time series of prices in the economy. Formally speaking, a "regime" is simply a set of states. The stochastic regime process which emerges is a Markov process and it takes four values defined as follows:

$$(RG^1, d^H) = \{\text{matrices used at } t \text{ are } (F_1, G_2) \text{ or } (F_2, G_1), d_t = d^H\} = \{p_t = p_2 \text{ or } p_t = p_3\}$$

$$(RG^1, d^L) = \{\text{matrices used at } t \text{ are } (F_1, G_2) \text{ or } (F_2, G_1), d_t = d^L\} = \{p_t = p_6 \text{ or } p_t = p_7\}$$

$$(RG^2, d^H) = \{\text{matrices used at } t \text{ are } (F_1, G_1) \text{ or } (F_2, G_2), d_t = d^H\} = \{p_t = p_1 \text{ or } p_t = p_4\}$$

$$(RG^2, d^L) = \{\text{matrices used at } t \text{ are } (F_1, G_1) \text{ or } (F_2, G_2), d_t = d^L\} = \{p_t = p_5 \text{ or } p_t = p_8\}.$$

Under the chosen parametrization of the price dynamics of the economy then takes the following form:

from (RG^1, d^H) the economy is most likely to move to (RG^2, d^H) or (RG^2, d^L) ;

from (RG^1, d^L) the economy is most likely to move to (RG^1, d^H) or (RG^1, d^L) ;

from (RG^2, d^H) the economy moves to all regimes;

from (RG^2, d^L) the economy is most likely to move to (RG^1, d^H) or (RG^1, d^L) .

We provide in Table 4 detailed information about the regime process but here we focus on the behavior of the agents in the four regimes. Recall that we parametrize the beliefs as *deviations* from the conditional

probabilities of the matrix Γ . The behavior then takes the following form in each of the regimes:

Regime (RG^1, d^H) is an (Agreement, d^H) regime

Both agents have *increased* probability of (RG^2, d^L) and *decreased* probability of (RG^2, d^H) .

Regime (RG^1, d^L) is an (Agreement, d^L) regime

Both agents have *increased* probability of (RG^1, d^L) and *decreased* probability of (RG^1, d^H) .

Regime (RG^2, d^H) is a (Disagreement, d^H) regime

Agent 1 has *increased* probability of (RG^1, d^H) and (RG^1, d^L) but *reduced* probability of the others.

Agent 2 has *reduced* probability of (RG^1, d^H) and (RG^1, d^L) but *increased* probability of the others.

Regime (RG^2, d^L) is a (Disagreement, d^L) regime

Agent 1 has *increased* probability of regime (RG^1, d^H) and *decreased* probability of (RG^1, d^L) .

Agent 2 has *decreased* probability of regime (RG^1, d^H) and *increased* probability of (RG^1, d^L) .

When we compute equilibrium prices these patterns become more concrete. For example, in the crash state (RG^2, d^L) when the price is lowest (22.92) the agents disagree: one has increased probability of a recovery to the "boom" regime (RG^1, d^H) where the price is highest (29.23) while the second agent has increased probability of a weak recovery to regime (RG^1, d^L) where the price is intermediate (27.02). On the other hand, in the boom regime (RG^1, d^H) both agents agree on having increased probability of the low crash price of (RG^2, d^L) .

To integrate the discussion we briefly note that endogenous propagation of price volatility is the result of two forces. On the one hand, the correlation in the joint distribution of the three variables (d_t, y_t^1, y_t^2) has an effect on the frequency in which the agents increase or decrease their demands simultaneously. On the other hand, the selection of rational beliefs which is parametrized here by (λ, μ) specifies how the agents

interpret their private signals. This, in turn, determines on which side of the market they are at each realization of their generating variables (for more details see Kurz and Schneider (1996)).

(e) Non Capital Income. The parameters (ω^1, ω^2) were discussed earlier. We make a selection of $\omega^j = 28$ for $j = 1, 2$ but will explain below how the results change if this scaling parameter changes.

(5.b) Simulation Results

We have now selected all the parameters of the model except for the discount rate β and the risk aversion parameter γ . We present below the results for fixed ranges of these parameters: $.80 \leq \beta \leq .92$ and $2.25 \leq \gamma \leq 3.75$. The motivation for these ranges is rather simple. The range for the discount rate was selected since it is generally viewed as reasonable. The range for the risk aversion parameter is relatively narrow and is motivated by the empirical results in Section 4. Contrary to MP (1985) our results are sensitive to the ranges of these parameters and we shall further comment below on this issue.

The recent literature on the equity premium has sought to explain not only the riskless rate and the premium but also the variance of the risky returns as well as the price/dividend ratio. Consequently we exhibit in each of the tables below the following 4 numbers:

- | | |
|--|---|
| 1. r^F - the average riskless interest rate, | 3. σ_r^2 - the variance of the risky rate, |
| 2. ρ - the average risk premium | 4. p - the average price/dividend ratio. |

The historical values of these variables which have given rise to the debate are:

- | | |
|----------------------------|------------------------|
| 1. $0 \leq r^F \leq 1\%$, | 3. $\sigma_r^2 = .034$ |
| 2. $\rho = 6\%$, | 4. $p = 22.1$. |

Exhibiting the Mehra and Prescott (1985) Puzzle in Our Model. We start by selecting the parameter values which transform our equilibrium into the MP (1985) rational expectations equilibrium. These are: $a_j = b_j = .25$ for $j = 1, 2, 3, 4$ and $\lambda_i = \mu_i = 1$ for $i = 1, 2, \dots, 8$. Our results are reported in Table 2:

Table 2: Recovering the Mehra - Prescott Puzzle

		$\beta=.92$	$\beta=.88$	$\beta=.84$	$\beta=.80$
$\gamma=2.25$	r^F	5.24	5.28	5.32	5.37
	ρ	.33	.33	.33	.33
	σ_r^2	.0016	.0016	.0016	.0016
	p	27.1	26.8	26.5	26.2
$\gamma=2.75$	r^F	5.16	5.19	5.22	5.26
	ρ	.41	.41	.41	.41
	σ_r^2	.0016	.0016	.0016	.0016
	p	27.1	26.9	26.6	26.4
$\gamma=3.25$	r^F	5.08	5.10	5.13	5.16
	ρ	.49	.49	.49	.49
	σ_r^2	.0017	.0017	.0017	.0017
	p	27.1	26.9	26.7	26.5
$\gamma=3.75$	r^F	4.99	5.01	5.04	5.06
	ρ	.57	.57	.57	.57
	σ_r^2	.0017	.0017	.0017	.0017
	p	27.2	27.0	26.8	26.6

Table 2 shows that in the selected ranges the riskless rate is around 5% while the equity premium is between .33% and .57%. These values correspond closely to the order of magnitude of the variables reported in the literature on the equity premium.

The Equity Premium is No Puzzle in Our RBE In Table 3 we present the long term average values of the same key variables calculated by our model under the parametrization specified in Section (5.a) above.

Table 3: Key Variables in the Kurz-Beltratti RBE

		$\beta=.92$	$\beta=.88$	$\beta=.84$	$\beta=.80$
$\gamma=2.25$	r^F	2.00	2.02	2.03	2.04
	ρ	4.12	4.16	4.21	4.25
	σ_r^2	.0125	.0128	.0132	.0136
	p	26.9	26.6	26.3	26.0
$\gamma=2.75$	r^F	1.51	1.50	1.50	1.50
	ρ	4.87	4.93	4.99	5.06
	σ_r^2	.0182	.0187	.0192	.0198
	p	27.0	26.8	26.6	26.3
$\gamma=3.25$	r^F	1.04	1.02	1.01	.99
	ρ	5.60	5.67	5.75	5.83
	σ_r^2	.0237	.0244	.0251	.0259
	p	27.1	26.9	26.7	26.5
$\gamma=3.75$	r^F	.59	.56	.53	.50
	ρ	6.28	6.37	6.46	6.56
	σ_r^2	.0289	.0297	.0306	.0315
	p	27.2	27.1	26.9	26.7

Table 3 is the main result of this paper. It shows that our calibrated model predicts the average values of all four key variables within a reasonable order of magnitude of the historical averages. Moreover, it also shows that *the results are sensitive to the values of the parameters*, as theory suggests that they should. To some extent the most peculiar aspect of the debate on the equity puzzle has been the lack of model sensitivity to changes in the important parameters.

The sensitivity of the model's predictions to parameter changes implies that other configurations of model parameters would lead to different complex configurations of the key variables. Moreover, our results are continuous in the parameters hence there are other configurations of beliefs and model parameters that would produce results of the same order of magnitudes as those reported in Table 3. Hence, instead of focusing on the particular parameters selected, we shall try first to provide some economic intuition for the mechanism which generates the premium and the riskless rates as in Table 3.

The long term average dividend yield in any model provides a basic benchmark for the rate of return

on assets in the model. In our RBE this yield is 3.75% for the case of $\gamma = 3.75$ and $\beta = .84$ which we review below. How do we get to a premium of 6.5% and a riskless rate of .5%? The discussion of the premium is usually focused on the demand side which stresses the need to satisfy the risk preferences of the agents without an explicit explanation of how the economy actually provides the premium in equilibrium. Since in the present case preferences of the agents are held fixed and we wish to explain why our model can generate a large premium, we stress the supply side. Note first the elementary observation that, given the stationary distribution, the riskless rate is determined entirely by the *level* of the discount prices q_t in the different states. The premium, however, is determined by the dividend yield together with the averaging of the *rates of change of prices* across states. Thus, to attain the above configuration of riskless rate and premium, the equilibrium mechanism must keep the average discount price of bonds high so that the average cost of borrowing in equilibrium remains in the .5% range. Correspondingly, it must increase, *via capital gains and losses*, the average rate of return on equity by 3.25% so that the average total return on equity is 7%. In order to understand how this mechanism works in our model the reader needs to focus on the non-stationary nature of the RBE. This non-stationarity is best understood in terms of the regime process of "agreement" and "disagreement" between the agents which we defined earlier by $RG^1 = \{y_t^1 \neq y_t^2\}$ and $RG^2 = \{y_t^1 = y_t^2\}$.⁴

Postponing further comments on the particulars of our model, what are the conditions which typically induce the low riskless rate together with high premium? In our non-stationary model there are three characteristics of an RBE which yields these results:

1. In one of the regimes the demand for assets is high, this regime occurs with relatively high frequency and price volatility within the regime is moderate.
2. In the other regime the demand for assets is low and by implication it occurs with relatively lower

⁴ To highlight the use of generating variables we have defined them in such a way that "agreement" between the agents is implied by $y_t^1 = y_t^2$.

frequency and price volatility within the regime is moderate.

3. Price volatility across regimes is very high and hence dramatic price changes occur at switching points of time.

The reason why this pattern could generate the desired results can now be explained. In the regime of high demand for assets the discount price of bonds is very high and the rate of return on debt is low (it is, in fact, negative in some states within the regime) while the average rate of return on common stocks is moderate and may be lower than the dividend yield. This is so since asset price volatility within the regime is moderate and hence capital gains and losses are moderate and could affect moderately the average premium. In the regime of low demand for capital the prices of common stock and debt instruments are very low and the rate of return on bills can be very high. This contributes to *increase* the time average of the riskless rate. We can then see that the riskless rate is determined by the *relative frequency of the two regimes*: as the stationary probability of the regime of high demand increases relative to the probability of the regime of low demand, the economy will tend to have an RBE with lower average riskless rate as stated in the conditions above. This conclusion is correct regardless of the fluctuations of the price of the common stock. Now note that this argument about the riskless rate *depends only on equilibrium prices of the RBE and on the stationary distribution of the price states*.

The economic interpretation can now be completed since the volatility of equity prices across states together with the transition probabilities provide the final mechanism for determining the equity premium. Price volatility is needed in order to generate capital gains and losses. However, for a given price structure the configuration of capital gains and losses is determined by the dynamics of the economy which is represented by the transition probabilities. It is evident that a given stationary distribution can be induced by many different configurations of transition probabilities and this is exactly how two economies with the same vector of stationary probabilities can have different equity premia which are induced by different transition

probabilities across price states. This argument is not entirely precise since in a general equilibrium framework different transition matrices induce different equilibrium prices and this induces a separate "price effect" and "premium effect". As it turns out, under our parametrization of the RBE prices are relatively insensitive to small changes in transition probabilities and consequently the premia effects can be large.

Our discussion may also help explain why under the MP (1985) parametrization reported in Table 2 the economy cannot generate a large premium. We simply observe that the variance of the price/dividend ratio in our RBE is 5.0585 compared to .0080 in the MP (1985) parametrization (!) and without capital gains and losses one cannot generate a risk premium in a MP (1985) type model. Note, however, that price volatility by itself is not sufficient to generate a premium as our discussion above explains the subtle role of the vectors a and b *in correlating the generating variables of the agents and in allowing this correlation to be dependent upon the dividend states so as to alter the transition probabilities between price states.* This highlights the fact that *the crucial property of our model which generates the premium is endogenous uncertainty* and equally so, the failure of the MP (1985) approach to the premium results from the fact that under rational expectations only exogenous uncertainty is allowed as permissible risk.

Using the above framework we now briefly explain why the above conditions are satisfied by the parametric configurations of our model. In the regime of agreement, RG^1 , the demand for assets is high since both agents agree on an increased likelihood of capital losses and lower dividends and hence on lower income prospects. In addition, since they have a large initial endowment they increase their demand for financial assets to ensure adequate consumption in the second period. The price of stocks is then pushed to 29.233 when $d_t = d^H$ and to 27.021 when $d_t = d^L$. In the regime of "disagreement" their demands are conflicting and this reduced aggregate demand for assets leads to the low prices of 26.352 when $d_t = d^H$ and to 22.917 when $d_t = d^L$. The stationary frequency of the high demand regime is .655 and that of the low demand is .345. Note, however, that within the high demand regime prices fluctuate in the range 27.021-29.233 while in the

low demand regime they move in the range 22.352 - 26.352. To see the effect of transition probabilities on the premium consider the joint process of agreement-dividend which is a Markov process with the four states $\{(RG^1, d^H), (RG^1, d^L), (RG^2, d^H), (RG^2, d^L)\}$. Table 4 reports the transition probabilities ($\pi_{j,\cdot}$), the risky rates of return conditional of each state ($R_{j,\cdot}$), the stationary probabilities of the process (π), the expected returns conditional on each state and the dividend yield.

Table 4: Regime Transition Probabilities and Other Regime Conditional Measures

state		RG^1, d^H	RG^1, d^L	RG^2, d^H	RG^2, d^L	stationary π	Expected Return in state	Dividend Yield
RG^1, d^H	$\pi_{1,\cdot}$ $R_{1,\cdot}$	.060 9.01%	.080 -5.87%	.370 -1.38%	.490 -19.66%	.351	-10.07%	3.42%
RG^1, d^L	$\pi_{2,\cdot}$ $R_{2,\cdot}$	.559 17.93%	.421 1.83%	.011 6.69%	.009 -13.08%	.304	10.75%	3.70%
RG^2, d^H	$\pi_{3,\cdot}$ $R_{3,\cdot}$	.327 20.92%	.433 4.42%	.103 9.40%	.137 -10.88%	.150	8.23%	3.79%
RG^2, d^L	$\pi_{4,\cdot}$ $R_{4,\cdot}$	.569 39.05%	.429 20.07%	.001 25.80%	.001 2.49%	.195	30.86%	4.36%

We have already noted the high price volatility of the model. Table 4 reveals a particular pattern of price variability: given the regime of agreement prices have 49% chance of "crashing" to (RG^2, d^L) if $d=d^H$ but are very likely to return to RG^1 if $d=d^L$. Also, from RG^2 prices are very likely to return to RG^1 and this ensures that the frequency of RG^1 is .655. Given the high frequency of RG^1 the premium is very sensitive to the transition probabilities which ultimately account for the fact that the expected risky rate when transiting from regime (RG^2, d^L) is 30.86. Table 4 also shows the impact of the dividend process as part of the subtle mechanism of correlating the behavior of the agents. In states of agreement the transition probability to the low price regimes is very high if $d_L = d^H$ and very low if $d_L = d^L$. This reveals the very general principle according to which the spontaneous correlation between the agents arises from two sources: on the one hand

the generating variables are correlated reflecting communication in society. On the other hand, agents use the same public information (i.e. prices and dividends) when computing their conditional probabilities and this adds a common dimension of correlation which cannot be attained with the privately observed generating variables.

In closing we make two comments. First, note that in the "agreement" regime RG^1 , agents increase their demand for assets in a state of "pessimism". This conclusion is a result of the two period structure of the model. In a model with a longer horizon the motivation of the agents will change; in such a model we shall need to redefine the circumstances which generate the high demand states. However, the structure of the argument will remain the same and for this reason the use of our simple OLG model captures the phenomenon at hand in full generality.

Secondly, we have stressed that the crucial component in our RBE which accounts for endogenous price volatility is the correlation between the generating variables and the impact of the dividends on the mechanism. This is reflected in the configurations of the parameters a and b . To understand the importance of this factor think of a market with a large number of agents who hold rational beliefs. Suppose now that these beliefs are "independent" in the sense that the law of large numbers applies to their forecasts and demands. This heterogeneity of beliefs will have virtually no effect on equilibrium prices because the endogenous variability will be averaged out and this averaging will occur regardless of what the beliefs are. Aggregate price fluctuations are influenced by beliefs only when some correlation is present among the beliefs of the agents. In our model the beliefs of the agents are relatively simple in that the marginal distribution of their generating variables is i.i.d. The parameter configurations ensure however, that it is the correlation which regulates the frequency of the regimes and the structure of price fluctuations through their impact on the transition between states. Thus, finding fault in our relatively simple rational beliefs of the agents will miss the central point; the focus must be on the structure of correlation as reflected in Table 4

because this is where the propagation of price volatility originates. One may question why we do not model explicitly the communication between the agents and the reasons why they "influence" each other's thinking. This is an important question but it is like asking why humans live in communities and not in isolation.

The Scaling Effect of Non-Capital Income. We have noted the important difference between our model and MP (1985) in terms of the endowment. MP (1985) discuss the effect of non-capital income (see MP (1985) Section 4.2) on the premium and conclude that this effect is negligible. To examine the effect we start with the Mehra-Prescott parametrization of our model as in Table 2. Table 5 shows the variations in the key variables when we vary the endowment vector (ω^1, ω^2) between 20 and 30 while keeping $\omega^1 = \omega^2 = \omega$.

Table 5: The Effect of Variations of Non-Capital Income in the Mehra-Prescott Economy

		$\omega = 20$	$\omega = 22$	$\omega = 24$	$\omega = 26$	$\omega = 28$	$\omega = 30$
$\gamma = 2.25$ $\beta = .92$	r^F	6.77	6.28	5.88	5.54	5.24	4.99
	ρ	.33	.33	.33	.33	.33	.33
	σ_r^2	.0016	.0016	.0016	.0016	.0016	.0016
	p	19.3	21.2	23.2	25.1	27.1	29.0
$\gamma = 3.75$ $\beta = .92$	r^F	6.51	6.03	5.62	5.28	4.99	4.74
	ρ	.58	.58	.58	.57	.57	.57
	σ_r^2	.0018	.0017	.0017	.0017	.0017	.0017
	p	19.3	21.3	23.2	25.2	27.2	29.1

Table 5 reveals that variations in the aggregate non-capital income almost completely scales the equilibrium price/dividend ratio but has virtually no effect on the equity premium and on the variance of risky returns. It also has a subtle non-linear effect on the riskless rate. However, due to obvious non-linearities, the effect of this income on the riskless rate is limited: even if we set $\omega = 100$ the riskless rate falls only to 2.54 when $\gamma = 2.25$ and to 2.30 when $\gamma = 3.75$. The price/dividend ratio rises to 98. Table 6 presents the effect of the non-capital income on our model under the parametrization in Table 3

Table 6: The Effect of Variations of Non-Capital Income on the Key Variables

		$\omega = 20$	$\omega = 22$	$\omega = 24$	$\omega = 26$	$\omega = 28$	$\omega = 30$
$\gamma = 2.25$ $\beta = .92$	r^F	3.52	3.03	2.63	2.29	2.00	1.75
	ρ	4.15	4.14	4.13	4.12	4.12	4.11
	σ_r^2	.0128	.0127	.0126	.0126	.0125	.0125
	p	19.1	21.1	23.0	25.0	26.9	28.9
$\gamma = 3.75$ $\beta = .92$	r^F	2.10	1.62	1.22	.88	.59	.33
	ρ	6.31	6.30	6.29	6.29	6.28	6.28
	σ_r^2	.0296	.0294	.0292	.0291	.0289	.0288
	p	19.3	21.3	23.3	25.3	27.2	29.2

Table 6 shows that the "scaling" results of Table 5 remain in effect. However, the large difference in the riskless rate between the tables should help the reader see the strong effect of endogenous uncertainty in our RBE on the level of this variable. Moreover, if one considers other variables in the economy such as the price/dividend ratio, then the non-capital income plays an important role in the model. For these reasons we think it is useful to show explicitly how it affects the workings of our model.

6. A Final Comment

The equity premium debate is an important debate. This is so, in our view, not only because of the specific economic question which it has raised but also because it has, indirectly, questioned the validity of some of the central ideas in contemporary economic thought. In this paper we have attempted to demonstrate that the equity premium puzzle is only a puzzle from the perspective of the theory of rational expectations. The theory of RBE holds that the dominant form of uncertainty in asset markets is endogenous uncertainty and we have demonstrated in this paper that the theory offers a perspective according to which the equity premium is not a puzzle at all. Hence, apart from the intrinsic interest in a better understanding of the functioning of financial markets, we propose that the reader views the present paper as a stringent test of the validity of the theory of RBE.

Table 1
Estimates of Mistake Functions

(i) Standard errors in parentheses. (ii) The $\chi^2_{0.05}$ critical value for the test of overidentifying restrictions (TOR) is 19.68

Fund	gamma	constant	X1	X2	X3	X4	X5	X6	X7	TOR
1	2.3981 (1.1367)	-0.1516 (0.2121)	-5.2203 (0.7803)	0.7842 (0.4530)	3.4374 (0.9284)	0.0174 (3.22E-3)	-0.5796 (0.1326)	-0.1949 (0.0572)	0.1107 (0.2165)	11.08
2	2.5027 (1.8469)	0.2719 (0.1012)	-5.3495 (0.7190)	1.9940 (0.5839)	3.9885 (0.9283)	0.0453 (0.0129)	-0.6375 (0.1333)	-0.2124 (0.0695)	-0.9366 (0.3017)	9.42
3	2.4204 (1.6136)	-0.0411 (0.0915)	-5.3461 (0.7093)	0.6171 (0.4652)	3.9326 (0.9302)	0.0184 (5.11E-3)	-0.5842 (0.1182)	-0.1818 (0.0555)	-0.0267 (0.1095)	8.24
4	1.9719 (1.4120)	-0.3987 (0.2328)	-5.0739 (0.7537)	0.2441 (0.4564)	3.6052 (0.7983)	0.0216 (4.91E-3)	-0.6012 (0.1259)	-0.2480 (0.0625)	0.3585 (0.2381)	11.04
5	3.2965 (1.8426)	-0.0462 (0.0462)	-4.9970 (0.7778)	0.3953 (0.4954)	3.6728 (0.9295)	0.0196 (4.53E-3)	-0.6105 (0.1062)	-0.2282 (0.0545)	-0.0390 (0.0669)	11.60
6	2.0393 (1.6100)	-0.2633 (0.2151)	-5.6334 (0.7578)	0.8680 (0.5239)	3.7103 (0.8940)	0.0181 (5.03E-3)	-0.6224 (0.1062)	-0.1900 (0.0493)	0.2514 (0.2632)	8.03
7	5.4830 (2.3402)	-0.0806 (0.0562)	-4.9718 (0.8330)	0.9108 (0.6038)	3.2177 (0.6906)	0.0169 (5.02E-3)	-0.5172 (0.0944)	-0.1398 (0.0625)	0.0408 (0.0614)	8.78
8	3.6093 (1.6232)	-0.2300 (0.0936)	-5.1799 (0.7405)	1.1215 (0.5253)	3.1074 (0.4689)	0.0221 (4.99E-3)	-0.5103 (0.0750)	-0.1288 (0.0674)	0.1768 (0.0875)	10.35
9	2.5020 (1.4948)	-0.0501 (0.0427)	-5.2527 (0.7843)	0.5768 (0.5342)	3.7056 (0.6670)	0.0202 (4.10E-3)	-0.5773 (0.0565)	-0.1947 (0.0459)	-0.0230 (0.0608)	8.47
10	1.9757 (1.3653)	-0.2754 (0.1495)	-5.1929 (0.8537)	0.3492 (0.5687)	4.1015 (0.9078)	0.0237 (5.08E-3)	-0.6259 (0.1016)	-0.2378 (0.0771)	0.2009 (0.1492)	8.59
11	2.9851 (2.0025)	-0.0587 (0.0418)	-5.2478 (0.8314)	0.5331 (0.4540)	3.7997 (0.9165)	0.0193 (4.96E-3)	-0.5911 (0.1086)	-0.1913 (0.0644)	-7.33E-3 (0.0534)	8.18
12	2.7824 (1.2775)	-0.0556 (0.0350)	-5.4808 (0.6101)	0.6162 (0.5133)	4.0750 (0.6551)	0.0190 (4.79E-3)	-0.6060 (0.1064)	-0.1739 (0.0607)	-0.0195 (0.0334)	8.01
13	2.6320 (1.5237)	-0.0491 (0.0339)	-5.6026 (0.6560)	0.8361 (0.5037)	4.0174 (0.8421)	0.0201 (4.60E-3)	-0.6120 (0.1082)	-0.1737 (0.0608)	-0.0360 (0.0283)	12.07
14	3.4452 (1.9878)	0.1172 (0.0693)	-4.8281 (0.8129)	0.1897 (0.5569)	4.9743 (0.9359)	0.0280 (5.70E-3)	-0.5549 (0.0937)	-0.1648 (0.0617)	-0.5168 (0.1852)	11.44
15	1.7021 (1.1601)	-0.9593 (0.4125)	-5.7674 (0.9637)	1.5129 (0.6547)	3.3883 (0.6281)	0.0382 (0.0110)	-0.6553 (0.1002)	-0.1259 (0.0538)	0.8503 (0.3829)	11.78
16	2.8865 (1.7524)	-0.0615 (0.0675)	-5.3757 (0.8496)	0.6508 (0.5784)	3.8764 (0.8658)	0.0190 (5.42E-3)	-0.6029 (0.1140)	-0.1645 (0.0676)	-7.71E-3 (0.0593)	8.41
17	2.3344 (1.4634)	-0.0645 (0.0305)	-5.4360 (0.7493)	0.3544 (0.4630)	3.9247 (0.8086)	0.0219 (4.56E-3)	-0.6436 (0.0970)	-0.2074 (0.0599)	-0.0141 (0.0304)	8.77
18	2.6084 (1.9794)	-0.0572 (0.0297)	-5.0971 (0.5999)	0.6007 (0.6179)	3.4935 (0.6099)	0.0197 (4.62E-3)	-0.5559 (0.1040)	-0.2082 (0.0694)	-6.28E-3 (0.0461)	10.12
19	9.4660 (4.1408)	-0.0757 (0.0354)	-4.1948 (0.8246)	0.3118 (0.3926)	2.8876 (0.7074)	0.0137 (2.88E-3)	-0.5462 (0.0997)	-0.1471 (0.0585)	0.1377 (0.0923)	9.59
20	2.1721 (1.1151)	-0.0845 (0.0550)	-5.3365 (0.6083)	0.5428 (0.4353)	3.8790 (0.6915)	0.0195 (4.47E-3)	-0.6160 (0.1055)	-0.1945 (0.0521)	0.0222 (0.0603)	8.36
21	1.7271 (1.1249)	-0.2693 (0.1244)	-6.4310 (0.6686)	0.6622 (0.4860)	4.3780 (0.7891)	0.0176 (4.20E-3)	-0.6374 (0.1091)	-0.1700 (0.0461)	0.2742 (0.1687)	9.60
22	2.7589 (2.4471)	0.1650 (0.0732)	-4.6470 (0.9217)	0.1524 (0.4013)	3.6978 (0.8535)	0.0121 (3.68E-3)	-0.6074 (0.0998)	-0.1845 (0.0564)	-0.3266 (0.1028)	9.96
23	1.5277 (1.4716)	-0.1544 (0.1254)	-5.9324 (0.8809)	0.5942 (0.5542)	4.2343 (0.8198)	0.0203 (5.52E-3)	-0.6782 (0.1313)	-0.2089 (0.0652)	0.1145 (0.1515)	8.61
24	5.5753 (2.6980)	3.93E-3 (0.0607)	-4.5474 (0.6259)	0.4250 (0.4918)	3.0520 (0.5168)	0.0167 (4.00E-3)	-0.5237 (0.0980)	-0.1840 (0.0638)	-0.1021 (0.1153)	8.85
25	2.1685 (2.0491)	0.0814 (0.0750)	-5.3172 (1.0777)	0.7166 (0.4626)	3.2827 (0.9140)	0.0227 (7.43E-3)	-0.5370 (0.1097)	-0.1781 (0.0633)	-0.1960 (0.1184)	9.68
26	2.7628 (1.1037)	0.0250 (0.1021)	-5.1914 (0.4615)	0.4475 (0.4632)	3.6358 (0.5502)	0.0238 (6.18E-3)	-0.6135 (0.0941)	-0.2026 (0.0587)	-0.1396 (0.1531)	8.75
27	3.6039 (2.3810)	-0.0459 (0.0371)	-5.3584 (0.9414)	0.5699 (0.6054)	3.9453 (0.9877)	0.0193 (5.23E-3)	-0.5985 (0.1144)	-0.1857 (0.0664)	-0.0428 (0.0416)	8.65
28	3.2450 (1.9414)	-0.2341 (0.1104)	-4.8127 (0.6946)	0.9372 (0.5185)	3.2615 (0.6419)	0.0279 (8.87E-3)	-0.5089 (0.0766)	-0.1428 (0.0625)	0.1480 (0.0809)	9.55
29	1.7064 (1.1712)	1.0356 (0.6191)	-5.8467 (0.8762)	0.2958 (0.5826)	5.2726 (1.2086)	0.0271 (6.47E-3)	-0.6846 (0.1130)	-0.2499 (0.0799)	-1.2702 (0.7206)	10.44
30	1.8067 (1.2815)	-0.1296 (0.1400)	-5.5568 (0.8166)	0.8782 (0.6048)	4.0094 (0.8359)	0.0172 (3.89E-3)	-0.5723 (0.1030)	-0.1533 (0.0663)	0.0785 (0.1508)	8.67
31	3.1909 (1.8127)	0.0886 (0.0534)	-4.6945 (0.8350)	0.2768 (0.3973)	3.5560 (0.8262)	0.0232 (5.32E-3)	-0.5361 (0.1123)	-0.1460 (0.0436)	-0.2644 (0.0814)	9.61

Fund	gamma	constant	X1	X2	X3	X4	X5	X6	X7	TOR
32	3.0733 (1.9565)	-0.0572 (0.1592)	-5.5745 (0.8463)	0.8423 (0.6769)	4.1658 (0.8495)	0.0179 (6.31E-3)	-0.5957 (0.1379)	-0.1647 (0.0982)	-0.0182 (0.2553)	9.95
33	3.8299 (1.6894)	-0.0298 (0.0360)	-4.6297 (0.5209)	0.4378 (0.4161)	3.0605 (0.4154)	0.0172 (4.21E-3)	-0.5376 (0.0834)	-0.1960 (0.0497)	-0.0254 (0.0372)	9.08
34	4.0995 (1.8242)	-0.0316 (0.0438)	-4.4840 (0.6279)	0.4014 (0.4050)	2.9408 (0.5066)	0.0164 (4.29E-3)	-0.5328 (0.0909)	-0.2006 (0.0496)	-0.0149 (0.0491)	10.03
35	1.9567 (1.3448)	-0.0371 (0.1710)	-5.7145 (0.9836)	0.8306 (0.6422)	4.0488 (0.8985)	0.0194 (6.80E-3)	-0.6334 (0.1256)	-0.1682 (0.0614)	-0.0320 (0.1634)	8.94
36	1.9531 (1.6265)	-0.0281 (0.0594)	-5.7418 (0.7971)	0.8010 (0.5452)	3.9421 (0.8340)	0.0222 (5.83E-3)	-0.5996 (0.1088)	-0.1497 (0.0709)	-0.0654 (0.0921)	9.24
37	3.5860 (1.6065)	-0.0434 (0.0401)	-5.3276 (0.6416)	0.9302 (0.4164)	3.5170 (0.7896)	0.0179 (3.78E-3)	-0.5581 (0.0989)	-0.1439 (0.0575)	-0.0262 (0.0734)	10.73
38	2.3763 (1.8664)	-0.1852 (0.1193)	-5.1073 (0.9310)	0.1032 (0.5786)	3.9195 (0.9034)	0.0246 (7.64E-3)	-0.6530 (0.1247)	-0.2115 (0.0669)	0.1175 (0.1079)	8.58
39	3.0876 (1.5770)	4.63E-4 (0.0477)	-5.3436 (0.5855)	0.6053 (0.5463)	4.1910 (0.6788)	0.0205 (4.43E-3)	-0.6139 (0.0800)	-0.1452 (0.0724)	-0.1458 (0.1120)	8.42
40	2.6807 (1.4368)	-0.0464 (0.0340)	-5.2312 (0.7197)	0.5616 (0.5367)	3.5309 (0.7658)	0.0199 (4.56E-3)	-0.5815 (0.0968)	-0.1948 (0.0520)	-0.0156 (0.0551)	9.45
41	6.4307 (3.6030)	-0.2084 (0.1065)	-4.3179 (0.7196)	-0.0660 (0.7276)	3.0545 (0.7130)	0.0205 (6.59E-3)	-0.5615 (0.0946)	-0.2221 (0.0548)	0.3189 (0.1941)	15.01
42	11.9748 (6.0762)	0.0413 (0.0233)	-4.4551 (0.8472)	0.7997 (0.3304)	3.0911 (0.6345)	8.18E-3 (4.82E-3)	-0.5097 (0.0990)	-0.1434 (0.0493)	-0.2696 (0.0207)	6.59
43	1.3287 (1.4915)	0.3476 (0.2553)	-5.9759 (1.1275)	0.4318 (0.6472)	4.7216 (1.1014)	0.0219 (6.19E-3)	-0.7056 (0.1489)	-0.2397 (0.0699)	-0.4600 (0.2851)	10.42
44	2.6954 (1.3384)	-0.0984 (0.1060)	-5.0721 (0.7106)	0.4538 (0.4558)	3.5376 (0.8410)	0.0210 (5.27E-3)	-0.5889 (0.0906)	-0.1981 (0.0502)	0.0320 (0.1052)	9.26
45	8.5705 (4.3688)	-4.21E-3 (0.0381)	-3.3549 (0.7185)	0.3372 (0.3680)	2.0719 (0.5078)	0.0101 (5.77E-3)	-0.4098 (0.0677)	-0.1768 (0.0463)	-0.0197 (0.0284)	9.59
46	3.5621 (1.4302)	-0.0527 (0.0552)	-4.7401 (0.6146)	0.4638 (0.5804)	3.0906 (0.4896)	0.0182 (3.65E-3)	-0.5439 (0.0876)	-0.2026 (0.0590)	1.70E-3 (0.0647)	9.85
47	3.4367 (2.2534)	0.3960 (0.1565)	-5.0703 (0.7991)	1.7964 (0.7107)	2.4078 (0.9582)	7.37E-3 (5.98E-3)	-0.6319 (0.1115)	-0.1576 (0.0525)	-0.5973 (0.2112)	11.49
48	2.4512 (2.2901)	-0.2678 (0.1327)	-5.1277 (0.9212)	1.0744 (0.4240)	3.5980 (0.8945)	0.0272 (9.57E-3)	-0.5938 (0.1124)	-0.1430 (0.0557)	0.1912 (0.1053)	10.64
49	2.7990 (1.2282)	-0.3849 (0.1205)	-5.1245 (0.7973)	2.5537 (0.6566)	2.0568 (0.4147)	0.0183 (3.77E-3)	-0.2982 (0.0577)	-0.0826 (0.0463)	0.3711 (0.1254)	11.15
50	8.9665 (6.8709)	-0.1016 (0.0799)	-5.0958 (1.4730)	0.7889 (0.4018)	3.0547 (1.0583)	0.0218 (0.0111)	-0.5452 (0.1600)	-0.1155 (0.0585)	0.0574 (0.0777)	10.22
51	3.1295 (1.6167)	-0.0379 (0.0282)	-5.2984 (0.6831)	1.0071 (0.2878)	3.4989 (0.7936)	0.0166 (3.52E-3)	-0.5475 (0.0903)	-0.1358 (0.0391)	-0.0189 (0.0293)	10.24
52	4.1805 (2.5851)	-0.0331 (0.1029)	-4.3786 (0.8774)	0.5141 (0.4033)	2.6321 (0.6385)	0.0166 (3.92E-3)	-0.5073 (0.1229)	-0.1819 (0.0521)	-7.02E-3 (0.1519)	10.97
53	7.3470 (2.3227)	-0.0175 (0.0310)	-4.3697 (0.6986)	0.3993 (0.4529)	2.8985 (0.5648)	0.0144 (4.13E-3)	-0.5292 (0.0926)	-0.1546 (0.0626)	-0.0426 (0.0384)	8.59
54	2.7564 (1.1006)	-0.0312 (0.0326)	-5.2566 (0.5973)	0.8185 (0.4405)	3.3850 (0.7247)	0.0191 (3.96E-3)	-0.5693 (0.0889)	-0.1516 (0.0554)	-0.0394 (0.0251)	11.55
55	5.5766 (6.2393)	-0.0690 (0.0816)	-4.9037 (1.0497)	0.5320 (0.4784)	3.4029 (0.9953)	0.0186 (0.0107)	-0.6683 (0.2958)	-0.1832 (0.1288)	0.0582 (0.1596)	10.82
56	3.7730 (2.9434)	-0.0837 (0.0854)	-4.9656 (0.9823)	0.6443 (0.5408)	3.5362 (0.9883)	0.0212 (9.52E-3)	-0.5967 (0.1539)	-0.1841 (0.0648)	9.52E-3 (0.0687)	9.57
57	118.0283 (34.1962)	-0.0528 (0.0173)	-0.0242 (0.3139)	-0.3765 (0.2338)	0.5310 (0.1536)	6.06E-3 (1.59E-3)	-0.1332 (0.0351)	-0.0722 (0.0203)	0.0675 (0.0585)	9.65
58	4.0932 (2.4704)	-0.1126 (0.0709)	-5.2318 (0.8750)	0.8272 (0.5602)	3.5103 (0.8052)	0.0205 (5.41E-3)	-0.5655 (0.1028)	-0.1984 (0.0678)	0.0589 (0.0795)	10.60
59	1.8437 (1.1337)	0.9631 (0.4914)	-5.2112 (0.6762)	0.5566 (0.4717)	3.2552 (0.7580)	0.0205 (4.75E-3)	-0.5693 (0.1118)	-0.1400 (0.0517)	-1.0264 (0.4989)	10.55
60	6.5078 (4.1423)	-0.0518 (0.0421)	-4.5365 (1.0578)	1.0415 (0.5900)	2.8867 (0.9411)	0.0121 (4.98E-3)	-0.4986 (0.1119)	-0.0860 (0.0645)	0.0399 (0.0669)	10.75
61	3.0083 (1.8614)	-0.2472 (0.4095)	-5.7228 (1.4811)	0.5987 (0.6360)	3.8634 (0.8031)	0.0171 (4.31E-3)	-0.6395 (0.1270)	-0.1896 (0.0672)	0.3258 (0.7021)	8.53
62	2.2518 (1.5744)	-0.3053 (0.1006)	-5.6802 (1.2081)	1.2771 (0.4539)	3.0081 (1.0049)	0.0173 (5.22E-3)	-0.4321 (0.1125)	-0.0621 (0.0587)	0.3298 (0.1227)	11.24
63	1.6578 (1.2863)	-2.2216 (1.0347)	-5.1217 (0.6619)	1.0054 (0.5102)	3.2643 (0.4535)	0.0170 (4.32E-3)	-0.4422 (0.1010)	-0.1423 (0.0555)	2.2674 (1.0793)	9.03

APPENDIX 2

Data Description

Funds variables. We have acquired from CDA Wiesenerger, CDA Investment Technologies, Inc., 1355 Piccard Drive, Rockville, Maryland, 20850 an information file containing end-of-period allocation in the following asset classes: cash, bonds, convertible bonds, preferred stock, stock. In order to reduce this allocation to the level of aggregation that we consider in the theoretical model we group convertible bonds and bonds into one class and stocks and preferred stocks into a second class. The data were provided only at the one year frequency for the period 1982-1990, and at quarterly frequency for the period 1990 -1995. We therefore completed the data file by using The Wiesenerger Investment Companies Service, Current Performance and Dividend Record, by Warren, Gorham & Lamont, Inc. Boston, Mass. The various publication involved contain similar data on portfolio allocation at quarterly frequency. These publications are: *Asset Allocations and Total Asset Values - Management Results: 12/31/82-9/30/90*, and *Dividends, Capital Gains and NAV - Mutual Funds Update 1992-*, *Current Dividend Record 1990-1992*, and *Current Performance and Dividend Record 1982-1990*.

Return variables. S&P500 and Treasury bills from *Stocks, Bonds, Bills and Inflation* by Ibbotson and Associates, 1995 Yearbook, Chicago, Illinois,

Macroeconomic variables. The data source is Citibase, Citibank, New-York, N.A. This includes Quarterly GDP in 1987 dollars, quarterly output per man/hour, monthly index of capacity utilization, monthly index of vendors, monthly M1, monthly discount rate, monthly index of basic commodities, monthly index of industrial production, monthly CPI. From this basic series we constructed the following information variables: growth rate of GDP over 4 quarters, growth rate of output per hour over 8 quarters, growth rate of index of purchasing managers over 4 quarters, growth rate of index of basic commodities over 4 quarters, growth rate of M1 over 8 quarters.

APPENDIX 3

A Brief Review of the Theory of Rational Beliefs

We start with some notation. $x_t \in \mathbb{R}^N$ is a vector of N observables at date t and the sequence $\{x_t, t = 0, 1, \dots\}$ is a stochastic process with true probability Π . Since every $x = (x_0, x_1, \dots)$ is an infinite sequence in $(\mathbb{R}^N)^\infty$ we use the notation $\Omega = (\mathbb{R}^N)^\infty$ and denote by $\mathcal{F}$ the Borel σ -field of Ω . We thus think of the probability space $(\Omega, \mathcal{F}, \Pi)$ as the *true* probability space. A belief of an agent is a probability Q ; such an agent is then adopting the theory that the probability space is $(\Omega, \mathcal{F}, Q)$. An agent who observes the data takes $(\Omega, \mathcal{F}, \Pi)$ as fixed but does not know Π . Using past data he will try to learn as much as possible about Π . The theory of Rational Beliefs aims to characterize the set of all beliefs which are compatible with the available data.

The basic assumption made is that date 1 has occurred a long time ago and at date t , when agents form their beliefs about the future beyond t , they have an *ample* supply of past data. We think of the vector $x = (x_0, x_1, x_2, x_3, \dots)$ as the vector of observations generated by the economy. However, in studying complex joint distributions among the observables, econometricians consider *blocks* of data rather than individual, primitive observations. For example, if we study the distribution of $(x_{\text{today}}, x_{\text{today}+1})$ we would consider the infinite sequence of blocks $(x_0, x_1), (x_1, x_2), (x_2, x_3), \dots$. It is thus useful to think of the data *from the perspective of date 0* as the infinite vector $x = (x_0, x_1, x_2, \dots)$ and the data *from the perspective of date n* as $x^n = (x_n, x_{n+1}, \dots)$ where $x \equiv x^0$ and

$$x^n = T x^{n-1} \quad n = 1, 2, 3, \dots$$

T is known as the *shift transformation*. The stochastic dynamical system at hand is denoted by $(\Omega, \mathcal{F}, \Pi, T)$ where Π is the unknown probability. Now for any $B \in \mathcal{F}$ consider the set $T^{-n}B$ which is the preimage of B under T^n defined by

$$T^{-n}B = \{x \in \Omega : T^n x \in B\}.$$

$T^{-n}B$ is the set in Ω such that if we shift it by n dates we enter B ; $T^{-n}B$ is the event B occurring n dates later. A system $(\Omega, \mathcal{F}, \Pi, T)$ is said to be *stationary* if $\Pi(B) = \Pi(T^{-1}B)$ for all $B \in \mathcal{F}$. A set $S \in \mathcal{F}$ is said to be *invariant* if $S = T^{-1}S$; it is said to be *invariant Π a.e.* if $\Pi(S \Delta T^{-1}S) = 0$, ($S \Delta T^{-1}S = (S \cup T^{-1}S) \setminus (S \cap T^{-1}S)$). The distinction between these two concepts of invariance are minimal and will be disregarded here. A dynamical system is said to be *ergodic* if $\Pi(S) = 1$ or $\Pi(S) = 0$ for any invariant set S . In the discussion below we assume for simplicity of exposition that $(\Omega, \mathcal{F}, \Pi, T)$ is ergodic but this assumption is not needed (see Kurz [1994a] where this assumption is not made).

In order to learn probabilities agents adopt the natural way of studying the *frequencies* of all possible economic events. For example, consider the event B

$$B = \left\{ \begin{array}{l} \text{price of commodity 1 today} \leq \$1, \text{ price of commodity 6 tomorrow} \geq \$3, \\ 2 \leq \text{quantity of commodity 14 consumed two months later} \leq 5 \end{array} \right\}.$$

Now using past data agents can compute for any finite dimensional set B the expression

$$m_n(B)(x) = \frac{1}{n} \sum_{k=0}^{n-1} 1_B(T^k x) = \left\{ \begin{array}{l} \text{The relative frequency that } B \text{ occurred among} \\ n \text{ observations since date } 0 \end{array} \right\}$$

where

$$1_B(y) = \begin{cases} 1 & \text{if } y \in B \\ 0 & \text{if } y \notin B \end{cases}$$

This leads to a definition of the basic property which the system $(\Omega, \mathcal{F}, \Pi, T)$ is assumed to have:

Definition 1: A dynamical system is called *stable* if for any finite dimensional set (i.e. cylinder) B

$$\lim_{n \rightarrow \infty} m_n(B)(x) = \overset{\circ}{m} \text{ exists } \quad \Pi \text{ a.e.}$$

The assumption of ergodicity ensures that the limit in Definition 1 is independent of x . In Kurz [1994a]

it is shown that the set function $\overset{\circ}{m}$ can be uniquely extended to a probability m on $(\Omega, \mathcal{F})$. Moreover, relative to this probability the dynamical system $(\Omega, \mathcal{F}, m, T)$ is stationary. There are two crucial observations to be made at this point.

- (a) Given the property of stability, in trying to learn Π all agents end up learning m which is a *stationary* probability. In general $m \neq \Pi$: the true dynamical system $(\Omega, \mathcal{F}, \Pi, T)$ may not be stationary. Π *cannot be learned*.
- (b) Agents know that m may not be Π but with the data at hand m is the only thing that they can learn and agree upon.

Non-stationarity is a term which we employ to represent the process of *structural change which cannot be explained by the statistical regularity of past data*. Hence, a stable but non-stationary system is a model for an economy with structural change but in which econometric work can still be successfully carried out. If all agents knew that the true system is stationary they would adopt m as their belief. The problem is that they do not know if the environment is stationary and hence even if it was stationary, agents may still not adopt m as their belief.

It is important to see that m summarizes the entire collection of *asymptotic restrictions* imposed by the true system with probability Π on the empirical joint distributions of all the observed variables. It is shown in Kurz [1994a] that for each stable system with probability Π there is an entire set $B(\Pi)$ of stable systems with probabilities Q which generate the same stationary probability m and consequently *impose the same asymptotic restrictions on the data as the true system with Π* . The question is how one can determine *analytically* (i.e. without observing any data generated by a system) if a dynamical system $(\Omega, \mathcal{F}, Q, T)$ generates m as a stationary measure. To examine this question let us return to $(\Omega, \mathcal{F}, \Pi, T)$ and consider, for any cylinder B the set function

$$m_n^\Pi(B) = \frac{1}{n} \sum_{k=0}^{n-1} \Pi(T^{-k}B)$$

It is important to note that $m_n^\Pi(B)$ has nothing to do with data: it is an analytical expression derived from $(\Omega, \mathcal{F}, \Pi, T)$.

Definition 2: A dynamical system $(\Omega, \mathcal{F}, \Pi, T)$ is said to be *weak asymptotically mean stationary* (WAMS) if for all cylinders $S \in \mathcal{F}$ the limit

$$m^\Pi(S) = \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{k=0}^{n-1} \Pi(T^{-k}S) \text{ exists.}$$

It is *strong asymptotically mean stationary* if the limit above holds for all $S \in \mathcal{F}$.

It is shown in Kurz [1994a] that m^Π can be uniquely extended to a probability measure on $(\Omega, \mathcal{F})$.

We then have the important theorem which is the main tool in Kurz [1994a]:

Theorem 1: $(\Omega, \mathcal{F}, \Pi, T)$ is stable if and only if it is WAMS. If m is the stationary measure calculated from the data, then

$$m(S) = m^\Pi(S) \text{ for all } S \in \mathcal{F}.$$

The implication of Theorem 1 is that every stable system $(\Omega, \mathcal{F}, \Pi, T)$ generates a *unique* stationary probability m^Π which is *calculated analytically* from Π . This last fact is crucial since it is the foundation of the following:

Definition 3: A selection of belief Q cannot be contradicted by the data m if

- (i) the system $(\Omega, \mathcal{F}, Q, T)$ is stable,
- (ii) the system $(\Omega, \mathcal{F}, Q, T)$ generates m and hence $m^Q = m$.

We can finally state the two axioms which define the Rationality of beliefs:

Rationality Axioms

A selection Q by an agent is a *Rational Belief* if it satisfies

- (I) *Compatibility with the Data*: Q cannot be contradicted by the data.
- (II) *Non-Degeneracy*: if $m(S) > 0$, then $Q(S) > 0$.

One may interpret condition (B) to say that if a finite dimensional event S is observed infinitely often (generating a positive relative frequency) then one cannot be certain that S cannot occur *from the perspective of today*. Now, to express a belief in the non-stationarity of the environment, an agent may select a probability Q^+ . This probability is said to be *orthogonal with m* if there are events S and S^c such that

- (i) $S \cup S^c = \Omega$, $S \cap S^c = \emptyset$,
- (ii) $m(S) = 1$, $m(S^c) = 0$,
- (iii) $Q^+(S) = 0$, $Q^+(S^c) = 1$.

We are now ready to characterize the set $B(\Pi)$ of all Rational Beliefs when the data is generated by $(\Omega, \mathcal{F}, \Pi, T)$.

Theorem 2 (Kurz [1994a]): Every Rational Belief must satisfy

$$Q = \lambda Q_\bullet + (1 - \lambda) Q^+$$

where $0 < \lambda \leq 1$, $Q_\bullet$ and m are probabilities which are mutually absolutely continuous (i.e. they are equivalent) and Q^+ is orthogonal with m such that

- (i) $(\Omega, \mathcal{F}, Q_\bullet, T)$ and $(\Omega, \mathcal{F}, Q^+, T)$ are both stable,
- (ii) $m^{Q_\bullet} = m^{Q^+} = m$.

Moreover, any Q such that λ , $Q_\bullet$ and Q^+ satisfy the above is a Rational Belief.

The probability Q^+ is central since it represents the theory of the agent of how the probability of an event at any date differs from the stationary probability at that date. This reveals a crucial characteristic of non-stationary systems: the timing of events matters in terms of the probabilities which are attached to them. Thus, the probability Q^+ permits an agent to assign to a given event different probabilities at different dates at which it may occur. Rationality of belief requires that averaging the probabilities assigned to this event over all dates must yield the stationary probability assigned to it by m . However, non-stationary systems can give rise to an unbounded number of such events which are different from each other. Consequently, a Rational Belief Q may induce forecasts which are different from the forecasts of m at all dates and the difference between the forecasts of Q and m need not converge to zero. Q^+ may also place positive probabilities on events on which m places zero probability.

Theorem 2 then says that two economic agents who are equally intelligent and who have *identically the same information* may make two different rational forecasts because they hold two competing theories which are compatible with the data. The agents may disagree on how much weight should be placed on the possibility that the environment is stationary (represented by the subjective parameter λ). They may also disagree on the probabilities of time sequencing of events and on the likelihood of important and rare events. Disagreement among rational agents must, therefore, arise from their having different theories about *the nature of the fluctuations of the system* rather than about the behavior of its long term averages.

An Example.

In order to use the theory of rational belief we need to work with processes which are non stationary but stable. Nielsen [1994] fully characterizes a broad class of such processes called Simple

Independently Distributed Stable (SIDS) processes. This class is analogous to i.i.d. processes and hence extremely useful in applications. A simple example will help clarify the nature of this class. Pick a process $\{y_j, j = 0, 1, \dots\}$ of i.i.d. random variables taking values in $\{0, 1\}$ with probability of 1 being, say, $1/4$ and generate an infinite sequence of observations $y^* = (y_0^*, y_1^*, \dots)$ of the process. The realizations $y_j^* = 1$ or $y_j^* = 0$ are now treated as *parameters* of the non stationary SIDS process $\{x_t, t = 0, 1, \dots\}$ to be defined. That is, there is a set of parameters $\{\alpha, \beta\}$ and a map associating the value $y_j^* = 1$ with the value α and the value $y_j^* = 0$ with β such that the process $\{x_t, t = 0, 1, \dots\}$ where $x_t \in X = \{0, 1\}$ is a sequence of *independent* random variables satisfying

$$(1) \quad P\{x_t = 1\} = \begin{cases} \alpha & \text{if } y_t^* = 1 \\ \beta & \text{if } y_t^* = 0. \end{cases}$$

Suppose that α and β satisfy $(1/4)\alpha + (3/4)\beta = .35$ then the long term average of the x_t 's is .35. Given the specification of y^* and the assumption of independence over time, (1) will determine Π_{y^*} , the probability of infinite sequences $x \in \{0, 1\}^\infty \equiv X^\infty$ given y^* and hence the dynamical system of the x_t 's is $(X^\infty, \mathcal{B}(X^\infty), \Pi_{y^*}, T)$. This SIDS system is stable and has a stationary measure m represented by the i.i.d. process $\{z_t, t = 0, 1, \dots\}$, $z_t \in X$ with $P\{z_t = 1\} = .35$. The stationary measure is independent of the specific realization of the generating parameters y^* .

Theorem 2 above (Kurz [1994a]) implies that the dynamical system $(X^\infty, \mathcal{B}(X^\infty), m, T)$ is a rational belief and, for brevity, we say that "the probability m is a rational belief". At each t there are, however, other rational beliefs Q about $\{x_\tau, \tau = t+1, t+2, \dots\}$ such that $m^Q = m$. For example, an agent may have a "private" generating process $\{z_t, t = 0, 1, \dots\}$ of i.i.d. coin tossing with probability of 1 being π and a generating sequence of parameters z^* in the same way as above with the frequency of $\{z_t^* = 1\}$ being ξ . Define now the perceived process $\{x'_t, t = 0, 1, \dots\}$ by

$$(2) \quad P \{ \mathbf{x}_t' = 1 \} = \begin{cases} \alpha' & \text{if } z_t^* = 1 \\ \beta' & \text{if } z_t^* = 0 \end{cases}$$

independently over time. This independent but non-stationary sequence of random variables defines a stable SIDS measure P with m as its stationary measure if $\xi \alpha' + (1 - \xi)\beta' = .35$. All probabilities $Q = \lambda m + (1 - \lambda)P$ with $\lambda \in [1, 0)$ are Rational Beliefs in accordance with theorem 2. In fact, we often also think of P as a rational belief although it violates Axiom II of Kurz [1994a].

References

- Arrow, K. J. : Le Role des Valeurs Boursieres pour la Repartition la Meillure des Risques, *Econometrie* CNRS 40, 41- 47 1953.
- Campbell, J., Shiller, R. J.: The Dividend-Price Ratio and Expectations of Future Dividends and Discount Factors, *Review of Financial Studies* 1, 195-228 (1988)
- Cecchetti, S.G., Lam, P.S., Mark, N.C.: The Equity Premium and the Risk Free Rate: Matching the Moments, *Journal of Monetary Economics* 31, 21-45 (1993)
- Epstein, L.G., Zin, S.E.: 'First Order' Risk Aversion and the Equity Premium Puzzle, *Journal of Monetary Economics* 26, 387-407 (1990)
- Fama E.F., French, K.R.: Permanent and Temporary Components of Stock Prices, *Journal of Political Economy* 96, 246-273 (1988)
- Grossman, S.J., Shiller, R.J.: The Determinants of the Variability of Stock Market Prices, *American Economic Review* 71, 222-227 (1981)
- Hansen, L.: Large Sample Properties of Generalized Method of Moment Estimators, *Econometrica* 50, 1029-1054 (1982)
- Hansen, L., Singleton, K.: Generalized Instrumental Variables Estimation of Nonlinear Rational Expectations Models, *Econometrica* 50, 1269-1288 (1982)
- Kurz, M.: The Kesten-Stigum Model and the Treatment of Uncertainty in Equilibrium Theory, in *Essays on Economic Behavior Under Uncertainty*, ed. By M.S. Balch, D.L. McFadden, S.Y. Wu, Amsterdam: North Holland, 389-399 1974

- Kurz, M.: On the Structure and Diversity of Rational Beliefs, *Economic Theory* 4, 877-900 (1994a)
- Kurz, M.: On Rational Belief Equilibria, *Economic Theory* 4, 859-876 (1994b)
- Kurz, M.: Asset Prices with Rational Beliefs, mimeo, Stanford University, 1994. (in a forthcoming monograph in the Springer series *Studies in Economic Theory* : Kurz, M. (ed.), *Endogenous Economic Fluctuations : Studies in the Theory of Rational Beliefs*, Berlin and New York: Springer Verlag, 1997)
- Kurz, M., Schneider, M.: Coordination and Correlation in Markov Rational Belief Equilibria, *Economic Theory* 8, forthcoming, November (1996)
- Kurz, M., Wu, H.M.: Endogenous Uncertainty in a General Equilibrium Model with Price Contingent Contracts, *Economic Theory* 8, November (1996)
- Lucas, R.E., Jr.: Asset Prices in an Exchange Economy, *Econometrica* 46, 1429-1445 (1978)
- MaCurdy, T., Shoven, J.B.: Stocks, Bonds, and Pension Wealth, in Wise, D. A.(ed.), *Topics in the Economics of Aging*, The University of Chicago Press, Chicago, 61-75 1992
- MaCurdy, T., Shoven, J.B. : Pension Accumulation with Stocks and Bonds, mimeo, Stanford University 1993
- Mankiw, N.G.: The Equity Premium and the Concentration of Aggregate Shocks, *Journal of Financial Economics* 17, 211-219 (1986)
- Mankiw, M.G., Zeldes, S.P.: The Consumption of Stockholders and Non-Stockholders, *Journal of Financial Economics*, 29, 97-112 (1991)
- Mehra, R., Prescott, E.C.: The Equity Premium: A Puzzle, *Journal of Monetary Economics* 15, 145 - 161 (1985)
- Nielsen, C. K.: Weakly Rational Beliefs, Structural Independence and Rational Belief Structures, a Ph.D dissertation submitted to the Economics Department, Stanford University, CA. 1994
- Rietz, T.A.: The Equity Premium: A Solution, *Journal of Monetary Economics* 22, 117-133 (1988)
- Weil, P.: The Equity Premium Puzzle and the Riskless Rate Puzzle, *Journal of Monetary Economics* 24, 401-422 (1989)

RECENTLY PUBLISHED "TEMI" (*)

- No. 258 — *L'analisi discriminante per la previsione delle crisi delle "micro-banche"*, by L. CANNARI and L. F. SIGNORINI (November 1995).
- No. 259 — *La redditività degli sportelli bancari dopo la liberalizzazione*, by F. CASTELLI, M. MARTINY and P. MARULLO REEDTZ (November 1995).
- No. 260 — *Quanto è grande il mercato dell'usura?*, by L. GUIO (December 1995).
- No. 261 — *Debt Restructuring with Multiple Creditors and the Role of Exchange Offers*, by E. DETRAGIACHE and P. G. GARELLA (December 1995).
- No. 262 — *National Saving and Social Security in Italy (1954-1993)*, by N. ROSSI and I. VISCO (December 1995).
- No. 263 — *Share Prices and Trading Volume: Indications of Stock Exchange Efficiency*, by G. MAJNONI and M. MASSA (January 1996).
- No. 264 — *Stock Prices and Money Velocity: A Multi-Country Analysis*, by M. CARUSO (February 1996).
- No. 265 — *Il recupero dei crediti: costi, tempi e comportamenti delle banche*, by A. GENERALE and G. GOBBI (March 1996).
- No. 266 — *Are Banks Risk-Averse? A Note on the Timing of Operations in the Interbank Market*, by P. ANGELINI (March 1996).
- No. 267 — *Money Demand in Italy: A System Approach*, by R. RINALDI and R. TEDESCHI (May 1996).
- No. 268 — *Asset Pricing Lessons for Modeling Business Cycles*, by M. BOLDRIN, L. J. CHRISTIANO and J. D. M. FISHER (May 1996).
- No. 269 — *Do Measures of Monetary Policy in a VAR Make Sense?*, by G. D. RUDEBUSCH (May 1996).
- No. 270 — *Maximization and the Act of Choice*, by A. SEN (May 1996).
- No. 271 — *Una stima dell'incidenza dell'imposizione diretta sulle imprese negli anni ottanta*, by A. STADERINI (June 1996).
- No. 272 — *Institutions and Labor Reallocation*, by G. BERTOLA and R. ROGERSON (July 1996).
- No. 273 — *Monitoring, Liquidation, and Security Design*, by R. REPULLO and J. SUAREZ (July 1996).
- No. 274 — *Localismo, spirito cooperativo ed efficienza: elementi per un'analisi economica delle banche di credito cooperativo*, by L. CANNARI and L.F. SIGNORINI (July 1996).
- No. 275 — *Intergenerational Transfers, Borrowing Constraints and the Timing of Home Ownership*, by L. GUIO and T. JAPPELLI (July 1996).
- No. 276 — *Monetary Policy Transmission, the Exchange Rate and Long-Term Yields under Different Hypotheses on Expectations*, by E. GIAIOTTI and S. NICOLETTI-ALTIMARI (August 1996).
- No. 277 — *Il fabbisogno finanziario pubblico*, by F. BALASSONE and D. FRANCO (September 1996).
- No. 278 — *Real Interest Rates, Sovereign Risk and Optimal Debt Management*, by F. DRUDI and R. GIORDANO (September 1996).
- No. 279 — *La riscoperta del debito e delle banche: progressi e questioni irrisolte*, by R. DE BONIS (October 1996).
- No. 280 — *Why Banks Have a Future: An Economic Rationale*, by R. G. RAJAN (October 1996).
- No. 281 — *Coordination and Correlation in Markov Rational Belief Equilibria*, by M. KURZ and M. SCHNEIDER (October 1996).

(*) Requests for copies should be sent to:

Banca d'Italia – Servizio Studi – Divisione Biblioteca e pubblicazioni – Via Nazionale, 91 – 00184 Rome
(fax 39 6 47922059)