Diverse Beliefs and Time Variability of Risk Premia

by

Mordecai Kurz[1] Maurizio Motolese[2]

November 29, 2007

Abstract: Why do risk premia vary over time? We examine this problem theoretically and empirically by studying the effect of market belief on risk premia. Individual belief is taken as a fundamental state variables. Market belief is observable, it is central to the empirical evaluation and we show how to extract it from the data. The asset pricing model we use is familiar from the noisy REE literature but we adapt it to an economy with diverse beliefs. We derive the equilibrium asset pricing and the implied risk premium. Our approach permits a closed form solution of prices hence we trace the exact effect of market belief on the time variability of asset prices and risk premia. We test empirically the theoretical conclusions.

Our main result is that, above and beyond the effect of business cycles on risk premia, fluctuations in market belief have significant independent effect on the time variability of risk premia. We study the premia on long positions in Federal Funds Futures, 3-month and 6-month Treasury Bills. The annualized mean risk premium on holding such assets for 1-12 months is about 40-60 basis points and, on average, we find that the component of market belief in the risk premium at a random date exceeds 50% of the mean. Since time variability of market belief is large, this component frequently exceeds 50% of the mean premium. This component is larger the shorter is the holding period of an asset and it dominates the premium for very short holding returns of less than 2 months. As to the structure of the premium we show that when the market holds abnormally favorable belief about the future payoff of an asset the market views the long position as less risky hence the risk premium on that asset declines. More generally, periods of market optimism (i.e. "bull" markets) are shown to be periods when the market risk premium declines while in periods of pessimism (i.e. "bear" markets) the market's risk premium rises. Hence, fluctuations in risk premia are *inversely* related to the degree of market optimism about future prospects of asset payoffs. This effect is strong and economically very significant.

JEL classification: C53, D8, D84, E27, E4, G12, G14.

Keywords: Risk premium; heterogenous beliefs; market state of belief; asset pricing; Bayesian learning; updating beliefs; Rational Beliefs.

For Correspondence:
Mordecai Kurz, Joan Kenney Professor of Economics
Department of Economics, Landau Building
Stanford University
Stanford, CA. 94305 - 6072
Email: mordecai@stanford.edu

---

[1] Department of Economics, Serra Street at Galvez, Stanford University, Stanford, CA. 94305-6072, USA

[2] Istituto di Politica Economica, Università Cattolica di Milano, Via Necchi 5, 20123, Milano, Italy.

Diverse Beliefs and Time Variability of Risk Premia[3]

by

Mordecai Kurz and Maurizio Motolese
November 29, 2007

Market risk premia vary over time and their fluctuations are a major cause of asset price volatility. But what drives changes in risk premia? The standard rational expectations answer relates changes in risk premia to changes in information about *exogenous fundamentals* which correctly alter the market's assessment of future risky events, the most important of which are business cycles. Such a view implies that excess returns are predictable by changes in observed fundamentals which, in turn, explain market volatility. Although there is some empirical support for this view, it cannot be the full explanation. Asset prices are not explained well by fundamental factors and, as Paul Samuelson used to quip, the market has forecasted nine of the last five recessions.

An alternative perspective holds that, in addition to exogenous fundamental conditions, the bulk of asset returns' volatility is caused by fluctuations in market belief. We hold the view that agents do not know the true dynamics of the economy since it is a non-stationary system with time varying structure that changes faster than can be learned with precision from data. With diverse beliefs, a large proportion of price volatility is then endogenously generated. This component is called *Endogenous Uncertainty* (see, Kurz (1974)). Some papers which reflect these ideas includes Harrison and Kreps (1978), Varian (1985), (1989), Harris and Raviv (1993), Detemple and Murthy (1994), Kurz (1974) , (1994), (1997), (2007a),(2007b) Kurz and Beltratti (1997), Kurz and Motolese (2001), Kurz Jin and Motolese (2005a) (2005b), Kurz and Wu (1996), Motolese (2001), (2003), Nielsen (1996), (2003) and Wu and Guo (2003), (2004). In particular Kurz and Motolese (2001) and Kurz Jin and Motolese

[3] This research was supported by a grant of the Smith Richardson Foundation to the Stanford Institute for Economic Policy Research (SIEPR). We are indebted to Randell Moore of the Blue Chip Economic Indicators for providing the detailed BLU data used here. We thank Robert Bliss for providing us with the data, routines and documentation to compute zero coupon rates. John O'Leary made a significant contribution to this work by computing the parameters of the value function used in the optimization described in the text. We thank Marco Aiolfi, Wen Chung Guo, Peter Hammond, Ken Judd, Stephen Morris, Hiroyuki Nakata, Carsten Nielsen, John O'Leary, Ho-Mou Wu for helpful discussions of the ideas of this paper and to Min Fan for extensive technical assistance and for helpful comments on an earlier draft. An earlier version of this paper was circulated under the title"Risk Premia, Diverse Belief and Beauty Contests."

(2005a) demonstrate via simulations that Endogenous Uncertainty explains the equity premium and leads to stochastic volatility. However, these papers study risk premia via simulations of computed equilibria. They do not study the determinants of risk premia either analytically or empirically.

In this paper we study the effect of market belief on the structure of risk premia. Beliefs are diverse but individually rational in a sense of Rational Beliefs. Our problem is to establish the relation between market belief and market risk premia. We derive analytic results which are then tested empirically by using data on the market distribution of beliefs. Observations on market belief are extracted from data on *monthly forecasts of future interest rates and macro economic variables* compiled by the Blue Chip Financial Forecasts (BLUF) since 1983. A *market* state of belief is a distribution of individual beliefs and in the theoretical and empirical analysis we focus on the first two moments. Since an agent's perceived risk premium is the conditional expectation of excess returns of an asset, an economy where agents hold diverse beliefs has many *subjectively* perceived risk premia.

The literature on excess returns and risk premia is large. We mention a few papers which report on convincing evidence gathered in recent years against the expectations hypothesis (e.g. Fama and Bliss (1987), Stambaugh (1988), Campbell and Shiller (1991), Cochrane and Piazzesi (2005) and Piazzesi and Swanson (2004)). They show that investments in Treasury securities generate predictable excess returns. Cochrane and Piazzesi (2005) exhibit predictable excess holding returns in bond markets while Piazzesi and Swanson (2004) find excess returns in two futures markets: Fed Funds futures in 1988:10 - 2003:12 and Eurodollar futures in 1985:Q2-2003:Q4. "Predictability" is used here in the sense of exhibiting *long term* statistical correlation between current information and future excess returns. Cochrane and Piazzesi (2005) and Piazzesi and Swanson (2004) do not estimate structural models to explain the source of excess returns but deduce such returns from estimated reduced form models for forecasting returns. Broadly speaking, they argue that bond excess returns are associated with business cycles and for this reason they use pro-cyclical variables such as current yields or year over year growth rate of Non Farm Payroll (in short NFP) to predict excess returns.

Our results confirm earlier results about the effect of cyclical variables on risk premia. However, using our perspective we show that risk premia contain a large component generated by the dynamics of market belief. This component is orthogonal to the observed fundamental variables used in the above studies where the term "orthogonal" highlights the fact that pure belief is a variable measured *net of all* *observed fundamentals*, and it has its own dynamic law of motion. The market belief is a state variable reflecting investor's *perceived* future returns, net of fundamental information. This state variable functions like any exogenous fundamental variable and may be considered to be an externality taken as given by all. In equilibrium, fluctuations in market belief cause large changes in the risk perception of market participants. Here we study the risk premia on holdings of long positions in Federal Funds Futures, 3-month and 6-month Treasury Bills. The annualized mean risk premium on holding such assets for 1-12 months is about 40-60 basis points and we find that, on average, the component of market belief in the risk premium at a random date exceeds 50% of the mean. Since the time variability of market belief is large, this component is frequently larger than 50% of the mean premium. We find that this component is larger the shorter is the holding period of an asset.

We focus on two sets of results. First we show analytically and empirically that much of the time variability of market risk premium is generated endogenously by the dynamics of beliefs. Second, we show that the effect of market belief on the risk premium takes a specific form. When the market holds abnormally favorable belief about future payoffs of an asset, the long position is taken to be less risky and hence the risk premium on a long position of that asset falls. More generally, market optimism about future economic conditions lowers the risk premium while pessimism about future economic conditions increases the risk premium. This inverse relationship leads to the view that *"bull" markets in an asset class constitute periods of lower risk premia on long positions while "bear" markets constitute periods of high risk premia.* Note that in a rational expectations based asset pricing theory the concepts of "bull" or "bear" markets are not well defined. We test our conclusion empirically in all three markets and find the data supports the theoretical findings.

1. Asset Pricing Under Heterogenous Beliefs

1.1 An Illustrative Decision Model

Consider an asset or a portfolio of assets whose market price is p_t, paying an exogenous risky sequence $\{D_t, t = 1, 2, \ldots\}$ under a true probability $\hat{\Pi}$. Let r_t be the riskless interest rate, $R_t = 1 + r_t$ and hence excess return over the riskless rate is $(1/p_t)(p_{t+1} + D_{t+1} - R_t p_t)$. The *risk premium* over the riskless rate is the conditional expectations of excess returns. Since it is a function of equilibrium prices, a risk premium - as a function of state variables - is best deduced from equilibrium prices. With this in mind, the model below is used to deduce a closed form solution of the asset price map so as to enable a study of the factors determining the risk premium. To obtain closed form solutions we use a model which is very common in the literature on Noisy Rational Expectations Equilibrium (e.g. Brown and Jennings (1989), Grundy and McNichols (1989), Wang (1994), He and Wang (1995), Allen, Morris and Shin (2006) and others cited in Brunnermeier (2001)). Nevertheless, our key results are fully general and do not depend upon the specific model used. We now address a key issue. Our agents do not know the true probability $\hat{\Pi}$ of the process $\{D_t, t = 1, 2, \ldots\}$ and hold diverse probability beliefs about it. The fact that there are many subjective risk premia in the market raises two questions that will be at the basis of our development in the next two sections. First, why do agents not know the probability $\hat{\Pi}$? Second, what is the common knowledge basis of all agents in an economy with diverse beliefs?

Starting with the second question, our answer is *past data on observables.* The economy has a set of observable variables and D_t is one of them. Agents have a long history of the variables, allowing rich statistical analysis which leads all of them to compute the same empirical moments and the same finite dimensional distributions of the observed variables. Using standard extension of measures they deduce from the data a unique empirical probability measure on infinite sequences denoted by $\hat{m}$. It can be shown that $\hat{m}$ is stationary (see Kurz (1994)) and we call it "the stationary measure." This is the *empirical knowledge shared by all agents*. We assume the data reveals that under $\hat{m}$ the $\{D_t, t = 1,2,\ldots\}$ constitutes a Markov process where D_{t+1} is conditionally normally distributed with means $\mu + \lambda_d(D_t - \mu)$ and variance σ_d^2[4]. The unique probability $\hat{m}$ is then known to all. To simplify define $d_t = D_t - \mu$, hence the process $\{d_t, t = 1, 2, \ldots\}$ is zero mean with unknown true probability Π and an empirical probability m. *Why is m not equal to Π?* With this issue in mind we turn to the first question.

Our economy has undergone changes in technology and social organization. These are rapid with major economic effects, making $\{d_t, t = 1, 2, \ldots\}$ a non-stationary process. Although this means that the distributions of the d_t's are time dependent, it is more than viewing $\{d_t, t = 1, 2, \ldots\}$ as a sequence of productivity "regimes." It also means that, although we measure the d_t in a single unit of account, over time the nature of assets and commodities change. Such variability makes it impossible to learn the unknown Π. The probability m is merely an average over an infinite sequence of regimes,

[4] It would be more realistic to assume the values D_t grow and the growth rate of the values has a mean μ rather than the values themselves. This added realism is useful when we motivate the empirical model later but is not essential for the analytic development.

reflecting only long term frequencies. Belief diversity starts with the fact that agents disagree over the meaning of public information. They believe Π is different from m and construct models to express the implications they see in the data. Being common knowledge, the empirical probability m is a reference for any concept of rationality. An agent's model may be viewed as "extreme" but it cannot be declared "irrational" unless proved to contradict the empirical evidence. Thus, belief rationality requires a subjective model not to contradict the empirical evidence m.

Turning now to our infinite horizon model, at date t agent i buys θ_t^i shares of stock and receives the payment $d_t + \mu$ for each of θ_{t-1}^i held. We assume *the riskless rate is constant over time* so that there is a technology by which an agent can invest the amount B_t at date t and receive with certainty the amount $B_t R$ at date t+1. The definition of consumption is then standard

$$c_t^i = \theta_{t-1}^i [p_t + d_t + \mu] + B_{t-1}^i R - \theta_t^i p_t - B_t^i.$$

Equivalently, define wealth $W_t^i = c_t^i + \theta_t^i p_t + B_t^i$ and derive the familiar transition of wealth

(1a) $$W_{t+1}^i = (W_t^i - c_t^i)R + \theta_t^i Q_{t+1} \quad , \quad Q_{t+1} = p_{t+1} + (d_{t+1} + \mu) - R p_t.$$

Q_t are excess returns. Given some initial values (θ_0^i, W_0^i) the agent maximizes the expected utility

(1b) $$U = E_t^i \left[\sum_{s=0}^{\infty} -\beta^{t+s-1} e^{-(\frac{1}{\tau} c_{t+s}^i)} \,|\, H_t\right]$$

subject to a vector of state variables $\psi_t^{(\theta^i, c^i)}$ and their transitions, all specified later. H_t consists of all past observable variables. We recognize the limitations of the exponential utility and use it as a good vehicle to explain the main ideas, hence the term "illustrative" in the title of this Section. After deducing the closed form solution of equilibrium risk premium we show how to generalize the key results.

We now state an assumption and a conjecture. First, we assume the agent believes the payoff $\{d_t, t = 1, 2, ...\}$ is conditionally normally distributed. Second, we conjecture that given the economy's state variables, equilibrium price p_t is also conditionally normal. In the next section we describe the state variables and the structure of belief and Theorem 1 confirms the conjecture. In Appendix A we show that for an optimum of (1a)-(1b), there is a constant vector u so the stock demand function is

(2) $$\theta_t^i(p_t) = \frac{\tau}{R\hat{\sigma}_Q^2}[E_t^i(Q_{t+1}) + u\psi_t^i].$$

$\hat{\sigma}_Q^2$ is an *adjusted* conditional variance (see Appendix A for details) of excess stock returns which is assumed constant and the same for all agents. The term $u\psi_t^i$ is the intertemporal hedging demand which is linear in agent i's state variables. We have earlier assumed the dynamics of payoffs *deduced from the empirical frequencies* is characterized by a first order Markov process with transition

(3) $$d_{t+1} = \lambda_d d_t + \rho^d_{t+1} \quad , \quad \rho^d_{t+1} \sim N(0, \sigma^2_d).$$

Since the implied stationary probability is denoted by m, we write $E^m[d_{t+1}|d_t] = \lambda_d d_t$.

Is the stationary model (3) the true data generating process? Those who believe the economy is stationary accept (3) as the truth. Such belief is rational since there is no empirical evidence against it. However, since $\{ d_t, t=1,2,... \}$ is non-stationary with unknown probability Π, most agents do not believe (3) is adequate to forecast the future. All surveys of forecasters show that subjective judgment about the data contributes more than 50% to the final forecast (e.g. Batchelor and Dua (1991)). Hence, agents form their own beliefs about d_{t+1} and other state variables explored later. With possibly complex beliefs, how do we describe an equilibrium? For such a description do we really need to give a full, detailed, development of the diverse theories of all agents? The structure of belief is our next topic.

1.2 Modeling Heterogeneity of belief I: Individual Belief as a State Variable

The theory of Rational Beliefs (in short, RB due to Kurz (1994), (1997)) defines an agent to be *rational* if his model cannot be falsified by the data and if simulated, its simulated data reproduces the stationary probability m deduced from the actual data. Here we use only the most basic restrictions of this theory but in Section 1.3.1 we review all rationality conditions the theory imposes on our model. One of the theory's aims is to account for the evidence of persistent belief diversity. But this diversity raises a clear methodological question. In formulating an asset pricing theory should we describe *in detail* the subjective model of each agent in the economy? With wide diversity this task is formidable. Also, if the objective is to study dynamics of asset prices, is such a detailed description necessary? An examination of the subject reveals that, although an intriguing question, such a detailed task is not needed. Instead, to describe an equilibrium all we need is to specify how the beliefs of agents affect their subjectively perceived transition functions of state variables. Once specified, the Euler equations are fully defined and market clearing leads to equilibrium pricing. We now explain this observation.

In markets with heterogenous beliefs agents are willing to reveal their forecasts. Consequently, samples of individual forecasts are taken and become publicly available. We thus assume that *data on market distribution of forecasts are public observations*. This fact points to the crucial difference between markets with and without private information. A market with asymmetric private information is secretive: agents do not reveal their forecasts since these are taken by others as real information about unobserved state variables. Such revelation eliminates the small advantage that each agent has relative to others. When an individual's forecasts of a state variable are revealed in our market-*without* private information- others do not view such forecasts as a source of new information or data. They view them as an expression of his opinion and do not update their own beliefs. In this market a forecaster uses the forecasts of state variables by other agents only *to alter his forecasts of endogenous variables* since we shall see that these depend upon market belief (for more details see Kurz (2007a)). But then, how do we describe the individual and market beliefs?

The key analytical step taken (see Kurz (1994), Nielsen (1996), Kurz (1997), Kurz and Motolese (2001), Kurz, Jin and Motolese (2005a),(2005b)) is to treat individual beliefs as state variables, generated by the agents within the economy. Here we use the approach of Kurz, Jin and Motolese (2005a), (2005b) as adapted and applied to the problem of this paper. We outline it now.

An individual belief about an economy's state variable is described with a *personal* state of belief which uniquely pins down the transition function of the agent's belief about next period's *economy's* state variable. This implies that personal state variables and the economy-wide state variables are not the same. A personal state of belief is like any other state variables in the agent's decision problem but is analogous to the concept of a "type" of an agent. A given personal state of belief at t identifies the agent type at date t. However, at t he is not certain of his future belief type which is determined by a transition of his personal state of belief. The *distribution* of individual states of belief, which we call "the market state of belief," is then an economy-wide observable state variable whose moments play an important role. In principle all moments could matter in equilibrium, but due to the exponential utility we use, equilibrium endogenous variables depend only on the *mean market states of belief*. This will be generalized in the empirical work reported later. As noted, the crucial fact is that the market state of belief is observable. In equilibrium, endogenous variables (e.g. prices) are functions of the economy's state variables, including market state of belief. But in a large economy an agent's *"anonymity"* implies that a personal belief state has a negligible effect on prices hence past personal states are not observed. Finally, due to the effect of the market belief on endogenous variables, an agent uses the equilibrium map to forecast all endogenous variables but must forecast future market states of belief. To forecast future endogenous variables an agent must, therefore, *forecast the beliefs of others*. It follows that the main issue we need to discuss next is the dynamics of individual beliefs.

A simple implication of the RB rationality principle says that an individual belief *cannot be a* *constant* unless an agent believes the stationary transition (3) is the truth. To explain this argument suppose agents hold diverse beliefs which are different from (3). If one holds a *constant* belief but not (3) then over time his average belief is different from (3). Since (3) is the time average in the data, this proves he is irrational. Hence, if different from (3) an agent cannot hold a constant belief. But just being wrong is not the issue. Rational agents hold wrong beliefs most of the time when there is no empirical proof they are wrong. This is so since when rational agents hold diverse probability models when there is only one true law of motion, then most are wrong most of the time. It follows the average market forecasting model is often wrong. Such market mistakes are at the heart of endogenous uncertainty.

We now introduce agent i's *state of belief* g_t^i. It describes his perception by pinning down his transition functions. Adding to "anonymity" we assume agent ℓ knows his own g_t^ℓ and the market *distribution* of g_t^i at t across i. In addition he observes past distributions of the g_τ^i for all $\tau < t$ hence he knows past values of all moments of the distributions of g_τ^i. We specify the dynamics of g_t^i by

(4) $$g_{t+1}^i = \lambda_Z g_t^i + \rho_{t+1}^{ig} \quad , \quad \rho_{t+1}^{ig} \sim N(0, \sigma_g^2)$$

where ρ_{t+1}^{ig} are correlated across i reflecting correlation of beliefs across individuals. The concept of an individual state of belief, with dynamics (4), is central to our development and we consider (4) to be a primitive. It is simply a positive description of type heterogeneity but in Appendix B we deduce (4) as a consequence of Bayesian inference. It is postulated in the Appendix that non stationarity is expressed by the sequence of unobserved regime parameters b_t of the dividend process

(5a) $$d_{t+1} - \lambda_d d_t = b_t + \rho_{t+1} \quad , \quad \rho_{t+1} \sim N(0, \frac{1}{\beta}) \text{ i.i.d.}$$

The sequence of parameters b_t has a long term variance of σ_b^2, without empirical long term serial correlation or correlation with ρ_{t+1}^d. Since (3) reflects the empirical record we then must have

(5b) $$\sigma_d^2 = \sigma_b^2 + \frac{1}{\beta} .$$

This condition reflects the fact that $d_{t+1} - \lambda_d d_t$ in the empirical record (3) (i.e. ρ_{t+1}^d) must have the same empirical moments as $d_{t+1} - \lambda_d d_t$ in the true process (5a) (i.e. $b_t + \rho_{t+1}$).

How is g_t^i used by agent i? If d_{t+1}^i denotes agent i's perception of t+1 payoff then g_t^i pins down his expectations $E_t^i[d_{t+1}^i - \lambda_d d_t]$ of the *difference* between his date t forecast of a state variable and the forecast under the empirical distribution m. Hence, agent i's date t *perceived* distribution of d_{t+1} is then

(6a) $$d_{t+1}^i = \lambda_d d_t + \lambda_d^g g_t^i + \rho_{t+1}^{id} \quad , \quad \rho_{t+1}^{id} \sim N(0, \hat{\sigma}_d^2).$$

The assumption that $\hat{\sigma}_d^2$ is the same for all agents is made for simplicity. It follows that g_t^i measures

(6b) $$E^i[d_{t+1}^i | H_t, g_t^i] - E^m[d_{t+1} | H_t] = \lambda_d^g g_t^i .$$

Indeed, (6b) shows how to measure $g_t^{\ i}$ in practice. For a state variable X_t, data on i's forecasts of X_{t+1} (in (6a) it is d_{t+1}) are measured by $E^{\,i}[X_{t+1}^{\ i} | H_t, g_t^{\ i}]$. One then uses standard econometric techniques to construct the stationary forecast $E^{\,m}[X_{t+1} | H_t]$ with which to compute the difference in (6b). This construction and the data it generates are also used by Fan (2006). An agent who believes the empirical distribution is the truth, is described by $g_t^{\ i} = 0$. He believes $d_{t+1} \sim N(\lambda_d d_t, \sigma_d^2)$.Since an agent's belief is about our changing society, the $g_t^{\ i}$ reflect different economies over time. For example, in 1900 the $g_t^{\ i}$ were related to electricity and the combustion engine, while in 2000 the $g_t^{\ i}$ reflected information technology. Hence, success or failure of past $g_\tau^{\ i}$ tell you little about what present day $g_t^{\ i}$ should be.

1.3 Modeling Heterogeneity of belief II: Market Belief and Rationality

1.3.1 Individual and Market Beliefs

Averaging (4) denote by Z_t the mean of the cross sectional distribution of $g_t^{\ i}$ and we refer to it as "the average state of belief." It is observable. Due to correlation across agents, the law of large numbers is not operative and the average of ρ_t^{ig} over i does not vanish. We write it in the form

(4b) $$Z_{t+1} = \lambda_Z Z_t + \rho_{t+1}^Z .$$

The true distribution of ρ_{t+1}^Z is unknown. Correlation across agents exhibits non stationarity and this property is inherited by the { Z_t , t = 1, 2, ...} process. Since Z_t are observable, market participants have data on the joint process $\{(d_t, Z_t), t = 1, 2, \ldots\}$ hence they know the *joint empirical distribution* of these variables. For simplicity we assume that this distribution is described by the system of equations

(7a) $$d_{t+1} = \lambda_d d_t + \rho_{t+1}^d$$

(7b) $$Z_{t+1} = \lambda_Z Z_t + \rho_{t+1}^Z$$

$$\begin{pmatrix} \rho_{t+1}^d \\ \rho_{t+1}^Z \end{pmatrix} \sim N\left(\begin{matrix} 0 \\ 0 \end{matrix}, \begin{bmatrix} \sigma_d^2, & 0, \\ 0, & \sigma_Z^2 \end{bmatrix} = \tilde{\Sigma} \right), \quad \text{i.i.d.}$$

Now, an agent who does not believe that (7a)-(7b) is the truth, formulates his own model\belief. We have seen in (6a) how agent i's belief state $g_t^{\ i}$ pins down his forecast of $d_{t+1}^{\ i}$. We now broaden this idea to an agent's perception model of the two state variables $(d_{t+1}^{\ i}, Z_{t+1}^{\ i})$. Keeping in mind that before observing d_{t+1} agent i knows d_t and Z_t, his belief takes the general form

(8a) $$d_{t+1}^{\ i} = \lambda_d d_t + \lambda_d^g g_t^{\ i} + \rho_{t+1}^{id}$$

(8b) $$Z_{t+1}^{\ i} = \lambda_Z Z_t + \lambda_Z^g g_t^{\ i} + \rho_{t+1}^{iZ}$$

(8c) $$g_{t+1}^{\ i} = \lambda_Z g_t^{\ i} + \rho_{t+1}^{ig}$$

$$\begin{pmatrix} \rho_{t+1}^{id} \\ \rho_{t+1}^{iZ} \\ \rho_{t+1}^{ig} \end{pmatrix} \sim N\left(\begin{matrix} 0 \\ 0 \\ 0 \end{matrix}, \begin{bmatrix} \hat{\sigma}_d^2, & \hat{\sigma}_{Zd}, & 0 \\ \hat{\sigma}_{Zd}, & \hat{\sigma}_Z^2, & 0, \\ 0, & 0, & \sigma_g^2 \end{bmatrix} = \Sigma^i \right)$$

Although g_t^i pins down belief about d_{t+1}, (8a)-(8b) show that we use it also to pin down the transition of Z_{t+1}^i. This simplicity ensures that one state variable pins down agent i's subjective belief of how conditions at date t are different from normal as reflected by the empirical distribution:

$$E_t^i\begin{pmatrix} d_{t+1} \\ Z_{t+1} \end{pmatrix} - E_t^m\begin{pmatrix} d_{t+1} \\ Z_{t+1} \end{pmatrix} = \begin{pmatrix} \lambda_d^g g_t^i \\ \lambda_Z^g g_t^i \end{pmatrix}.$$

The average market expectation operator is defined by $\bar{E}_t(\bullet) = \int E_t^i(\bullet)\,di$. From (8c) it is

$$\bar{E}_t\begin{pmatrix} d_{t+1} \\ Z_{t+1} \end{pmatrix} - E_t^m\begin{pmatrix} d_{t+1} \\ Z_{t+1} \end{pmatrix} = \begin{pmatrix} \lambda_d^g Z_t \\ \lambda_Z^g Z_t \end{pmatrix}.$$

Higher Order Beliefs. One must distinguish between higher order belief which are temporal and those which are contemporaneous. Within our theory the system (8a)-(8c) defines agent i's probability over sequences of (d_t, Z_t, g_t^i) and as is the case for *any probability measure*, it implies temporal higher order beliefs of agent i with regard to future events. For example, we deduce from (8a)-(8c) statement like

$$E_t^i(d_{t+N}) = E_t E_{t+1}^i \ldots E_{t+N-1}^i(d_{t+N}) \quad , \quad E_t^i(Z_{t+N}^i) = E_t E_{t+1}^i \ldots E_{t+N-1}^i(Z_{t+N}^i).$$

It is thus clear that temporal higher order beliefs are properties of conditional expectations. In addition, by (8c) (or equivalently (4)) we have $\bar{E}_t(d_{t+N+1}) = \lambda_d \bar{E}_t(d_{t+N}) + \lambda_d^g \bar{E}_t(Z_{t+N})$. Hence we can also deduce *perceived* higher order *market beliefs* by averaging individual beliefs. For example, we have that

$$\bar{E}_t(Z_{t+N}) = \bar{E}_t \bar{E}_{t+N-1}(d_{t+N}) - \bar{E}_t E_{t+N-1}^m(d_{t+N}).$$

The perception models (8a)-(8c) show that properties of conditional probabilities do not apply to the market belief operator $\bar{E}_t(\bullet)$ since it is *not a proper conditional expectation*. To see why let $X=D\times Z$ be a space where (d_t, Z_t) take values and G^i be the space of g_t^i. Since i conditions on g_t^i, his unconditional probability is a measure on the space $((D\times Z\times G^i)^\infty, \mathscr{F}^i)$ where $\mathscr{F}^i$ is a sigma field. The market conditional belief operator is an average over conditional probabilities, each conditioned on a *different* state variable. Hence, this averaging does not permit one to write a probability space for the market belief. The market belief is neither a probability nor rational and we have the following result:

<u>Theorem 1</u>: The market belief operator violates iterated expectations: $\bar{E}_t(d_{t+2}) \neq \bar{E}_t \bar{E}_{t+1}(d_{t+2})$.

<u>Proof</u>: Since $E_t^i(d_{t+2}) = \lambda_d E_t^i(d_{t+1}) + \lambda_d^g E_t^i(g_{t+1}^i) = \lambda_d[\lambda_d d_t + \lambda_d^g g_t^i] + \lambda_d^g \lambda_Z g_t^i$ it follows that

(9) $$\bar{E}_t(d_{t+2}) = \lambda_d^2 d_t + \lambda_d^g(\lambda_d + \lambda_Z) Z_t.$$

On the other hand we have from (8a) that $\bar{E}_{t+1}(d_{t+2}) = \lambda_d d_{t+1} + \lambda_d^g Z_{t+1}$ hence we have that

$$E_t^i \bar{E}_{t+1}(d_{t+2}) = \lambda_d[\lambda_d d_t + \lambda_d^g g_t^i] + \lambda_d^g[\lambda_Z Z_t + \lambda_Z^g g_t^i].$$

Aggregating now we conclude that

$$(10) \qquad \bar{E}_t \bar{E}_{t+1}(d_{t+2}) = \lambda_d^2 d_t + \lambda_d^g(\lambda_d + \lambda_Z + \lambda_Z^g) Z_t.$$

Comparison of (9) and (10) shows that $\bar{E}_t(d_{t+2}) \neq \bar{E}_t \bar{E}_{t+1}(d_{t+2})$. ■

Belief and Information: Understanding Z_t. For each agent, Z_t is a state variable like any other. News about Z_t are used to forecast prices and assess market risk in the same way macroeconomic data such as GNP growth or NFP are used to assess the risk of a recession. Market belief may be wrong as it forecasts more recessions than are realized. Risk premia may rise or fall just because agents are more optimistic or pessimistic about the future, not necessarily because there is any specific data to convince investors the future is bright or bleak. But then, how do agents update their beliefs when they observe Z_t? In sharp contrast with models of private information, agents do not revise their own beliefs about the state variable d_{t+1}: (8a) specifically *does not* depend upon Z_t. Agents do not view Z_t as information about d_{t+1} since it is not a "signal" about unobserved private information they do not have. Indeed, they know that all use the same public information. However, Z_t is crucial "news" about *what the market thinks* about d_{t+1}! Hence, the importance of Z_t is it's great value in forecasting future *endogenous* variables. Date t endogenous variables depend upon Z_t and future endogenous variables depend upon future Z's. Since market belief exhibits persistence, agents know that today's market belief is useful for forecasting future endogenous variables. How is this equilibrated? This we show in Section 1.4.

1.3.2 Rationality: The Theory of Rational Beliefs

We have seen that the market belief is not necessarily rational hence averaging (8a) -(8c) is not required to imply a consistent probability measure. What about individuals? Since they do not know the true probability Π, we assume (8a) -(8c) may not be the truth. But then, can we rationalize such a belief on its own? That is, what restrictions do (8a) -(8c) need to satisfy in order for them to represent the belief of a rational individual agent? What criteria are used in formulating such restrictions? Note that we have already imposed some rationality conditions. First, we argued that rational agents will exhibit fluctuating beliefs since a *constant* belief which is not in accord with the empirical distribution is irrational. Second, we required g_t^i to have an unconditional zero mean by requiring individual beliefs to be about *deviations from the empirical frequencies.* This, by itself, places restrictions on beliefs. We now explain the additional restrictions imposed by the theory of Rational Beliefs.

The theory of Rational Belief (in short, RB) due to Kurz (1994), (1997) proposes natural restrictions on beliefs with the view to explain the emergence of diverse beliefs and excess volatility. In a sequence of papers since 1994 the theory has been applied to various markets (e.g. Kurz (1997), Kurz and Schneider (1996), Kurz and Wu (1996), Kurz, Jin and Motolese (2005b), Motolese (2001), (2003) Nielsen (1996), (2003), Wu and Guo (2003), (2004)). In relation to the equity risk premium, Kurz and Beltratti (1997), Kurz and Motolese (2001), and Kurz, Jin and Motolese (2005a) explain the equity premium by asymmetry in the distribution of beliefs.

A belief is an RB *if it is a probability model which, if simulated, reproduces the empirical distribution known from the data*. An RB is thus a model which cannot be rejected by the empirical evidence represented by m. Beliefs are specified by the perception models (8a) -(8c) in which the dynamics of g_t^i expresses the subjective belief of an agent. Exactly as in (5b), for (8a) -(8c) to be RB it needs to induce the same empirical distribution as (7a)-(7b). This amounts to the requirement that

(11) The empirical distribution of $\begin{pmatrix} \lambda_d^g g_t^{\,i} + \rho_t^{id} \\ \lambda_Z^g g_t^{\,i} + \rho_{t+1}^{iZ} \end{pmatrix}$ = the distribution of $\begin{pmatrix} \rho_t^d \\ \rho_{t+1}^Z \end{pmatrix} \sim N\left(\begin{matrix} 0 \\ 0 \end{matrix}, \begin{bmatrix} \sigma_d^2, & 0, \\ 0, & \sigma_Z^2 \end{bmatrix}\right)$, i.i.d.

To compute the implied data generated by the model, one treats the $g_t^{\,i}$ symmetrically with other random variables. From (8c), the unconditional variance of $g_t^{\,i}$ is $\mathrm{Var}(g^{\,i}) = \sigma_g^2/(1-\lambda_Z^2)$ Hence, we have the following rationality conditions which follow from (11):

$$\text{(i)}\quad \frac{(\lambda_d^g)^2\sigma_g^2}{1-\lambda_Z^2} + \hat{\sigma}_d^2 = \sigma_d^2 \qquad \text{(ii)}\quad \frac{(\lambda_Z^g)^2\sigma_g^2}{1-\lambda_Z^2} + \hat{\sigma}_Z^2 = \sigma_Z^2 \qquad \text{(iii)}\quad \frac{\lambda_d^g\lambda_Z^g\sigma_g^2}{1-\lambda_Z^2} + \hat{\sigma}_{Zd} = 0$$

$$\text{(iv)}\quad \frac{(\lambda_d^g)^2\lambda_Z\sigma_g^2}{1-\lambda_Z^2} + \mathrm{Cov}(\hat{\rho}_t^{id}, \hat{\rho}_{t-1}^{id}) = 0 \qquad \text{(v)}\quad \frac{(\lambda_Z^g)^2\lambda_Z\sigma_g^2}{1-\lambda_Z^2} + \mathrm{Cov}(\hat{\rho}_t^{iZ}, \hat{\rho}_{t-1}^{iZ}) = 0.$$

The first three conditions pin down the covariance matrix in (8a) -(8c). The last two pin down the serial correlation of the terms $(\hat{\rho}_t^{id}, \hat{\rho}_t^{iZ})$. An inspection of (8a) -(8c) reveals the only choice left for a agent are the two free parameters $(\lambda_d^g, \lambda_Z^g)$. But under the RB theory these are not free either since there natural conditions they must satisfy. First, $\hat{\sigma}_d^2 > 0$, $\hat{\sigma}_Z^2 > 0$ place two strict conditions on $(\lambda_d^g, \lambda_Z^g)$:

(12) $$|\lambda_d^g| < \frac{\sigma_d}{\sigma_g}\sqrt{1-\lambda_Z^2} \quad , \quad |\lambda_Z^g| < \frac{\sigma_Z}{\sigma_g}\sqrt{1-\lambda_Z^2}.$$

Finally, to ensure the covariance matrix in (8a) -(8c) is positive definite one must impose an additional condition. The condition

(13) $$\frac{1-\lambda_Z^2}{\sigma_g^2} > \frac{(\lambda_Z^g)^2}{\sigma_Z^2} + \frac{(\lambda_d^g)^2}{\sigma_d^2}$$

is sufficient. Hence the "free" parameters $(\lambda_d^g, \lambda_Z^g)$ are restricted to a rather narrow range and this shows that the rationality principle of RB permits diverse beliefs which differ in parameter values.

1.4 Combining the Elements: the Implied Asset Pricing Under Diverse Beliefs

We now derive equilibrium prices and the risk premium. For details see Appendix A where we also explain the term $\hat{\sigma}_Q^2$, which is the "adjusted" conditional variance of Q_{t+1}. We have also explained in 1.3 why the state variables in (2) are specified by the vector $\psi_t^i = (1, d_t, Z_t, g_t^i)$. Hence, rewrite (2) as

(14) $$\theta_t^i(p_t) = \frac{\tau}{R\hat{\sigma}_Q^2}[E_t^i(Q_{t+1}) + u\psi_t^i] \ , \ u = (u_0, u_1, u_2, u_3) \ , \ \psi_t^i = (1, d_t, Z_t, g_t^i).$$

For an equilibrium to exist we need some stability conditions. First we require the interest rate r to be positive, $R = 1 + r > 1$ so that $0 < \frac{1}{R} < 1$. Now we add:

(15) **<u>Stability Conditions</u>:** We require that $0 < \lambda_d < 1 \ , \ \lambda_Z < 1 \ , \ 0 < \lambda_Z + \lambda_Z^g < 1$.

The first requires $\{d_t, t = 1, 2, ...\}$ to be stable and have an empirical distribution. The second is *a* stability of *belief* condition. It requires i to believe (d_t, Z_t) is stable. To see why, take expectations of (8b), average over the population and recall that Z_t are market averages of the g_t^i. This implies that

(16) $$\bar{E}_t[Z_{t+1}] = (\lambda_Z + \lambda_Z^g)Z_t.$$

<u>Theorem 2:</u> Consider the model with heterogenous beliefs under the stability conditions specified with supply of shares which equals 1. Then there is a unique equilibrium price function which takes the form $p_t = a_d d_t + a_z Z_t + P_0$.

<u>Proof:</u> Average (14), use the fact that the aggregate stock supply is 1 and rearrange to have

(17) $$\frac{R\hat{\sigma}_Q^2}{\tau} = [\bar{E}_t(p_{t+1} + d_{t+1} + \mu) - Rp_t + (u_0 + u_1 d_t + (u_2 + u_3)Z_t)].$$

Now use the perception models (8a)-(8b) about the state variables, average them over the population and use the definition of Z_t to deduce the following relationships which are the *key implications of treating individual and market beliefs as state variables*

(18a) $$\bar{E}_t(d_{t+1} + \mu) = \lambda_d d_t + \mu + \lambda_d^g Z_t$$

(18b) $$\bar{E}_t[Z_{t+1}] = (\lambda_Z + \lambda_Z^g) Z_t$$

Using these to solve for date t price we deduce

(19) $$p_t = \frac{1}{R}[\bar{E}_t(p_{t+1})] + \frac{1}{R}[(\lambda_d + u_1)d_t + (\lambda_d^g + u_2 + u_3)Z_t] + \frac{1}{R}[\mu + u_0] - \frac{\hat{\sigma}_Q^2}{\tau}$$

(19) shows that equilibrium price is the solution of a linear difference equation in the two state variables (d_t, Z_t). Hence, a standard argument (see Blanchard and Kahn(1980), *Proposition 1*, page 1308) shows that the solution is

(20a) $$p_t = a_d d_t + a_z Z_t + P_0$$

To match coefficients use (20a) to insert (18a) - (18b) into (19) and conclude that

(20b) $$a_d = \frac{\lambda_d + u_1}{R - \lambda_d}.$$

(20c) $$a_z = \frac{(a_d + 1)\lambda_d^g + (u_2 + u_3)}{R - (\lambda_Z + \lambda_Z^g)}$$

(20d) $$P_0 = \frac{(\mu + u_0)}{r} - \frac{\hat{\sigma}_Q^2 R}{\tau r}.$$

The stability conditions ensure that (20a) - (20d) is the unique solution as asserted. ■

Since we do not have a closed form solution for the hedging demand parameters $u = (u_0, u_1, u_2, u_3)$ we computed numerical Monte Carlo solutions. For all feasible values of the model parameters we find that $u_1 = 0$ hence $a_d > 0$ and $(a_d + 1)\lambda_d^g + (u_2 + u_3) > 0$ hence $a_z > 0$. These are reasonable conclusions: p_t *increases with higher* d_t *and with higher* Z_t *- today's market belief in higher future dividends.*

1.5 Equilibrium Risk Premium Under Heterogenous Beliefs

1.5.1 The Main Equilibrium Results

Under heterogenous beliefs we have diverse concepts of risk premia and one chooses a concept which is appropriate for an application. The risk premium on a long position, as a random variable, is

(21) $$\pi_{t+1} = \frac{p_{t+1} + d_{t+1} + \mu - Rp_t}{p_t}.$$

(21) is a random variable measuring actual excess returns of stocks over the riskless bond. The need is to measure the premium as *a known expected quantity*, recognized by participants. We have three such measures. The first is the subjective expected excess returns by agent i, computed by using the equilibrium map (20a) and the perception model (8a) -(8c) to show that

(22) $$\frac{1}{p_t}E_t^i(p_{t+1}+d_{t+1}+\mu-Rp_t) = \frac{1}{p_t}[(a_d+1)(\lambda_d d_t+\lambda_d^g g_t^i)+a_z(\lambda_Z Z_t+\lambda_Z^g g_t^i)+\mu+P_0-Rp_t]$$

Aggregating over i, the market premium is the average market expected excess returns. This perceived premium reflects what the market *expects*, not what it receives. From (22) it is measured by

(23) $$\frac{1}{p_t}\bar{E}_t(p_{t+1} + d_{t+1} + \mu - Rp_t) = = \frac{1}{p_t}[(a_d+1)(\lambda_d d_t+\lambda_d^g Z_t) + a_z(\lambda_Z Z_t + \lambda_Z^g Z_t) + \mu + P_0 - Rp_t]$$

Neither (22) nor (23) are objective risk premia. We thus turn to an objective measure, common to all agents, computed by agents studying the long term time variability of the premium and measuring it by the empirical distribution of (21). Using (20a) and the stationary transition (7a)-(7b) we have

(24) $$E_t^m[\pi_{t+1}] = \frac{1}{p_t}E_t^m[p_{t+1}+d_{t+1}+\mu-Rp_t] = \frac{1}{p_t}[(a_d+1)(\lambda_d d_t)+a_z(\lambda_Z)Z_t+\mu+P_0-Rp_t]$$

Observe that (24) is the way Econometricians and all researchers cited above have measured the risk premium. For this reason we refer to it as “the” risk premium.

We arrive at two conclusions. First, the differences between the premia in (22) and (23) is

(25a) $$\frac{1}{p_t}E_t^i(p_{t+1}+d_{t+1}+\mu-Rp_t) - \frac{1}{p_t}\bar{E}_t(p_{t+1}+d_{t+1}+\mu-Rp_t) = \frac{1}{p_t}[(a_d+1)\lambda_d^g + a_z\lambda_Z^g](g_t^i - Z_t).$$

This results says that from the perspective of trading, all that matters is the difference $g_t^i - Z_t$ of individual from market belief. Also, the following difference is important

(25b) $$\frac{1}{p_t}E_t^m(p_{t+1}+d_{t+1}+\mu-Rp_t) - \frac{1}{p_t}\bar{E}_t(p_{t+1}+d_{t+1}+\mu-Rp_t) = -\frac{1}{p_t}[(a_d+1)\lambda_d^g + a_z\lambda_Z^g]Z_t.$$

The risk premium is thus different from the market perceived premium when $Z \neq 0$. But the second, and more important, conclusion is derived by combining (23) with (25b). Keeping in mind that from (20c) $-(u_2 + u_3) = -a_z(R - \lambda_Z) + [(a_d+1)\lambda_d^g + a_z\lambda_Z^g]$, we deduce an *analytical expression of the risk premium:*

(26a) $$\frac{1}{p_t}E_t^m(p_{t+1}+d_{t+1}+\mu-Rp_t) = \frac{1}{p_t}[(R\frac{\hat{\sigma}_Q^2}{\tau} - u_0 - u_1 d_t) - a_z(R-\lambda_Z)Z_t]$$

Since $a_z > 0$, $R > 1$ and $\lambda_Z < 1$ it follows that the premium per share declines with Z_t. We conclude

(26b) The Risk Premium $E_t^m[\pi_{t+1}]$ is decreasing in the mean market belief Z_t.

Conclusions (26a) -(26b) are important. (26a) and the earlier results exhibit the Endogenous Uncertainty component of the risk premium which we call “The Market Belief Risk Premium.” It shows that market belief has a complex effect on market risk premia. The effect of belief consist of two parts

(I) The first is the direct effect of market beliefs on the permanent mean premium $R\frac{\hat{\sigma}_Q^2}{\tau}$. It is shown in the Appendix that there exist weights $(\omega_1, \omega_{12}, \omega_2)$ such that

$$\hat{\sigma}_Q^2 = \text{Var}_t^{\ i}((\omega_1(\lambda_d d_t + \lambda_d^g g_t^{\ i} + \omega_{12}\rho_{t+1}^{id}) + \omega_2(\lambda_Z Z_t + \lambda_Z^g g_t^{\ i} + \omega_{12}\rho_{t+1}^{iZ})).$$

Volatility of individual and market belief, which we call "Endogenous Uncertainty" contributes directly to the volatility of excess returns and increases permanently the risk premium.

(II) The second is the effect of market belief on the time variability of the risk premium, reflected in $-a_z(R-\lambda_Z)Z_t$ with a negative sign when $Z_t > 0$.

To explain this second result we note that it says that if one runs a regression of excess returns on the observable variables, the effect of the market belief on long term excess return *is negative*. This sign is surprising since when $Z_t > 0$ the market expects *above normal* future dividends but in that case the risk premium on the stock *is lower*. When $Z_t < 0$ the market holds bearish belief about future dividend but the risk premium *is higher*. Since we have data on Z_t and on the distribution of belief the result will be empirically tested. Before proceeding to the empirical test we discuss some ramifications of this result.

1.5.2 The Market Belief Risk Premium is Fully General

The main result (26b) was derived from the assumed exponential utility function. We argue that this result is more general and depends only on the positive coefficient a_z of Z_t in the price map. To show this, assume any additive utility function over consumption and a risky asset which pays a "dividend" or any other random payoff d_t. Denote the price map by $p_t = \Phi(d_t, Z_t)$. We are interested in the slope of the excess return function $E_t^m[\pi_{t+1}]$ with respect to Z_t. Focusing only on the numerator $E_t^m[p_{t+1} + d_{t+1} + \mu - Rp_t]$, linearize the price around 0 and write $p_t = \Phi_d d_t + \Phi_Z Z_t + \Phi_0$. The desired result depends only upon the condition $\Phi_Z > 0$. It is reasonable as it requires current price to increases if the market is more optimistic about the asset's future payoffs. To prove the point note that

$$E_t^m[p_{t+1} + (d_{t+1} + \mu) - Rp_t] \approx E_t^m[\Phi_d d_{t+1} + \Phi_Z Z_{t+1} + \Phi_0 + (d_{t+1} + \mu) - R(\Phi_d d_t + \Phi_Z Z_t + \Phi_0)]$$
$$= [(\Phi_d + 1)\lambda_d - R\Phi_d]d_t - \Phi_Z(R - \lambda_Z)Z_t + [\mu + \Phi_0(1 - R)].$$

The desired result follows from the fact that $\Phi_Z > 0$, $R > 1$ and $\lambda_Z < 1$.

The price map might be more complicated. If we write it as $p_t = \Phi(d_t, Z_t, X_t)$ where X are other state variables (in particular, the distribution of wealth), the analysis is more complicated since we need to specify a complete model for forecasting X_{t+1} but the main result continues to hold.

1.5.3 Interpretation of the Market Belief Risk Premium

Why is the effect of Z_t on the risk premium *negative*? Since this result is general and applicable to any asset with risky payoffs, we offer a general interpretation. Our result shows that when the market holds abnormally favorable belief about future payoffs of an asset the market views the long position as less risky and consequently the risk premium on the long position of the asset falls. Fluctuating market belief implies time variability of risk premia but more specifically, in the long run fluctuations in risk premia are *inversely* related to the degree of market optimism about future prospects of asset payoffs.

To explore the result, it is important to explain *what it does not say*. One could interpret it to confirm a common claim that to maximize excess returns it is optimal to be a "contrarian" to the market consensus. To understand why *this is a false interpretation* note that when an agent holds a belief about future dividends, the market belief Z_t does not offer him new information to alter his belief about dividends. If the agent believes future dividends will be abnormally high but $Z_t < 0$, the agent *does not change his forecast* of d_{t+1}. He uses Z_t only to forecast future prices. Hence, Z_t is a crucial input to forecasting returns without changing the forecast of d_{t+1}. Since given the available information and his probability belief, which is, say, Γ^i an optimizing agent is already on his demand function. He does not just abandon his demand by replacing Γ^i with the empirical measure m. This argument is analogous to the one showing why it is not optimal to adopt the log utility as your utility even though it maximizes the growth rate of your wealth. Yes, it does that, but you dislike the sharp declines which you expect to occur in the value of your assets if you follow the strategy called for by the log utility. By analogy, following a "contrarian" policy implies a high long run average return in accord with m since this is what (26a) says. But if your subjective model disagrees with the probability m you will dislike being short when your optimal position should be long. This argument explains why most people do not systematically bet against the market, as a "contrarian" strategy (26a) would dictate.

Taking a positive view, our results show that fluctuations in market belief are crucial for the time variability of the risk premium and the market pricing of risk. Market optimism in bull markets or pessimism in bear markets have drastic effects on market risk perception. *A bull market is a market in which risk perception is low and a bear market is one in which risk perception is high.* Our result (26a) shows that on average, market optimism induces lower risk premium and market pessimism generate high risk premium. But due to diverse beliefs the individually perceived premia are diverse. To see this use (22) and (24) to show that perceived premia are $E_t^i(\pi_{t+1}) = E_t^m(\pi_{t+1}) + \frac{1}{p_t}[(a_d+1)\lambda_d^g + \lambda_Z^g]g_t^i$. Hence, optimizing agents take into account information about Z_t in calculating their premia. From their perspective the state variable Z_t is used to assess risk in the same way as NFP is used to assess the risk of recessions and hence the market risk premium. We turn now to an empirical test of our theory.

2. Testing of the Endogenous Time Variability of the Risk Premium: The Data

2.1 The Forecast Data

The data we use is on the distribution of commercial forecasts and we take it as a proxy for forecasts made by the general public. The data is circulated monthly by the Blue Chip Financial Forecasts (BLUF). It provides forecasts of over 50 economists at major corporations and financial institutions. The number of forecasters may vary from month to month and, due to mergers and other organizational changes, the list of *potential* forecasters also changes over time. A sample of forecasters includes Moody's Investors Service, Prudential Securities, Inc. Ford Motor Company, Macroeconomic Advisers LLC, Goldman Sachs & Co., DuPont, Deutsche Bank Securities, J. P. Morgan Chase, Merrill Lynch, Fannie Mae, and others. BLUF reports forecasts of U.S. interest rates at all maturities along with forecasts of GDP growth and inflation. Forecasts reported in BLUF are collected on the 24th and 25th of each month and released to subscribers on the first day of the following month.

The BLUF publishes, for each variable, individual and mean ("consensus") forecasts. The mean is taken over all forecasters participating in that month. Forecasts are made for several quarters into the future. For each horizon forecasters are asked to forecast *the average value of that variable during the future quarter in question*. Note, the realized value of any variable for the quarter in which forecasts are released is not known at forecasting time since such data is available only after the quarter ends. As a result, each set of forecasts includes "current quarter" forecast which is denoted by the horizon $h = 0$. Hence, $h = 1$ means "the quarter *following* the quarter in which the forecasts were made." The BLUF publication was initiated in 1983:01 and circulated forecast data with horizons of $h = 0,1,...,4$ quarters. The initial version of the files provided data for the Fed Fund rate, 1-month Commercial Paper rate, 3-month T-Bill rate, 30-year Treasury Bonds rate, AAA long term corporate bonds rate, growth rates of GDP, changes in the GDP deflator and CPI. In 1988:01 the BLUF added individual and market mean forecasts to complete the yield curve on treasury securities covering also maturities of 6 months, 1 year, 2 years, 5 years and 10 years. In 1997:01 the forecast horizon was expanded by one quarter and from that date $h = 0,1,...,5$ quarters. Hence, a uniform panel data set for the entire term structure of interest rates is available starting in 1988:01. The data set has undergone other minor changes since its first release but these are not relevant to this paper and are thus not reported here.

In the empirical work we use a *month* as a unit of time. Hence, our first task was to translate quarterly mean forecasts to monthly forecasts. This was accomplished by an interpolation procedure which selected for each date t and for each variable the B-form of a least squares cubic spline piecewise polynomial which minimized the squared deviations from the given forecasts. When a variable is recorded monthly then all forecasters actually know at each date the realized monthly variable at hand for those months of the present quarter which have already past. This clearly applies to all interest rate data. Hence, it was useful to include in all interpolations past realized data of the variable in question for one quarter *before* date t (hence, three monthly observations). This procedure improves continuity at date t. An optimal polynomial is computed *for each date* and *utilizes no future market data of any kind*. At the end of the interpolation we have monthly data with monthly forecast horizons h=1,2,...,12.

The forecasts reported in BLUF are labeled by their release date, which is the start of each month. Hence, these forecasts are conditional on information available at the moment the forecasts were collected which is the *end of the month previous to release*. For example, data released in 1988:01 is recorded in our "sample period" as 1987:12 since the data released on January 1, 1988 is based on information available to forecasters at a date identified by us as 1987:12. Therefore all dates in this paper should be considered as identified with the end of the month. The data set has been updated in a format suitable for computations up to 2003:11.

2.2 Extracting Market States of Belief

The concepts of individual and market states of belief are central to the empirical work and we now explain how they are constructed. For any variable X denote by $E_t^i\{X_{t+h}\}$ agent i's conditional forecast of X_{t+h} at date t and by $E_t^m\{X_{t+h}\}$ the forecast under the stationary probability m. Agent i's state of belief about X_{t+h} is then defined by

$$Z_t^{(X,h,i)} = E_t^i\{X_{t+h}\} - E_t^m\{X_{t+h}\} \ .$$

The subtraction of $E_t^m\{X_{t+h}\}$ ensures the state of belief is m-orthogonal to date t market information. Also, since $Z_t^{(X,h,i)}$ is the deviation from the stationary forecast, it must be interpreted properly. Thus, suppose y is growth rate of GDP. When $Z_t^{(y,h,i)} > 0$ the agent is "optimistic" about future growth but it does not mean he believes output will necessarily go up. He does believe output will grow faster than "normal," defined by the growth rate expected under m. The market state of belief is defined by

$$Z_t^{(X,h)} = \frac{1}{N}\sum_{i=1}^{N} [\, E_t^i\{X_{t+h}\} - E_t^m\{X_{t+h}\}\,] = \bar{E}_t\{X_{t+h}\} - E_t^m\{X_{t+h}\}$$

and the cross sectional variance of beliefs is

$$(\sigma_t^{(X,h)})^2 = \frac{1}{N}\sum_{i=1}^{N}\left([E_t^i\{X_{t+h}\} - E_t^m\{X_{t+h}\}] - [\bar{E}_t\{X_{t+h}\} - E_t^m\{X_{t+h}\}]\right)^2 = \frac{1}{N}\sum_{i=1}^{N}\left(E_t^i\{X_{t+h}\} - \bar{E}_t\{X_{t+h}\}\right)^2.$$

Since $\bar{E}_t\{X_{t+h}\}$ is the average forecast, $Z_t^{(X,h)}$ reflects the market's views about economic conditions which are different at t from average. These differences are the reason why the market forecasts $\bar{E}_t\{X_{t+h}\}$ and not $E_t^m\{X_{t+h}\}$. "Optimism" or "pessimism" depend upon the context. For example, $Z_t^{(y,h)} > 0$ means the market is optimistic about abnormally high output growth in t+h. If $R^{(j)}$ is j maturity interest rate, then $Z_t^{(j,h)} > 0$ means the market expects this rate to be *higher* than normal at t+h. The market belief about Fed Funds rates is a belief about future monetary policy. Hence, $Z_t^{(F,h)} > 0$ means the market expects an abnormally tight monetary policy. Note that in this paper, *all belief variables are about future interest rates*.

To measure $Z_t^{(X,h)}$ we need data on the two components which define it. BLUF files provide direct data on $E_t^i\{X_{t+h}\}$ and $\bar{E}_t\{X_{t+h}\}$ as discussed. We have monthly forecast data on interest rates at different maturities, GDP growth , change in the CPI and the GDP deflator. The key issue is thus the construction of the stationary forecasts $E_t^m\{X_{t+h}\}$. These forecasts are made with a model that takes into account all data that was available at date t hence we take into account the release date of each variable used in the following analysis. A feature of stationarity is time invariance, implying the model is valid out of sample. This is an idealization which we can only approximate, given the relatively limited data set which we have. We thus compute $E_t^m\{X_{t+h}\}$ employing the Stock and Watson's (1999), (2001), (2002), (2005) method of diffusion indices. We briefly explain this procedure.

We started with the Stock and Watson's data set[5] developed by Data Resources and Global Insight. It contains 215 monthly time series for the US from 1959:01 to 2003:12, covering the main sectors of the economy. As discussed in Stock and Watson (2005), the series are transformed by taking logarithms and/or by differencing. In general, first differences of logarithms (growth rates) are used for real quantity variables, first differences are used for nominal interest rates, and second differences of logarithms (changes in growth rates) for price series. Because of missing data we use (see Stock and Watson (2005)) only 127 series from 1959:01 to 2003:12. These represent ten main categories of economic variables: consumption, employment, exchange rates, housing starts, interest rates, money aggregates, prices, real output, stock prices and the University of Michigan Index of Consumer

[5] The data is publically available on Watson's web page http://www.wws.princeton.edu/mwatson/publi.html

Expectations. Stacking them, we obtain an information matrix of dimension 540 by 127. One of Stock and Watson's (1999) conclusion is that effective time invariant models *need to employ a small number of variables*. The reason for this observation is that linear forecasting models with a large number of variables are unstable and forecast poorly out of sample. The Stock-Watson method reduces the rank of the matrix but keeps as much information as possible by creating diffusion indices constructed via principal component analysis to extract factors that best explain the variance of the information matrix.

For the period at hand the five greatest factors explain 43% of the variation in the information matrix and with twenty factors the variance explained is 74%. However, the marginal contribution of a factor declines rapidly implying that little marginal explanatory power is gained when using more than a few factors. Indeed, since we study interest rates which are rather persistent, nothing in this paper is changed by using more than four factors in the stationary forecasting scheme we adopt below. Stock and Watson (2002) concluded that a *combination* of factors and lags of the forecasted variable is the best information set. For any variable X the objective is to compute forecasts of X_{t+h} using information at time t. In all regressions of Section 3 we need stationary forecasts of market nominal interest rates and for these variables the forecasts are constructed as follows:

(1) let $\Delta x_{T+h} = X_{T+h} - X_T$ denote the stationary h-period change in a nominal interest rate and F_T^i for $i = 1,...,4$ denote the first four factors deduced from date T information matrix;

(2) estimate the parameters $\hat{\alpha}^h, \hat{\beta}^{h,i}, \hat{\gamma}^h$ by the following OLS regression:

$$\Delta x_{T+h} = \alpha^h + \sum_{i=1}^{4} \beta^{h,i} F_T^i + \gamma^h \Delta x_T + \varepsilon_{T+h}, \qquad \text{for } T = 1,...,t-h;$$

(3) the forecasts of $\hat{\Delta} x_{t+h}$ at date t are then given by:

$$\hat{\Delta} x_{t+h,t} = \hat{\alpha}^h + \sum_{i=1}^{4} \hat{\beta}^{h,i} F_t^i + \hat{\gamma}^h \Delta x_t.$$

Finally, the stationary forecasts of the interest rates are $E_t^m\{X_{t+h}\} = X_t + \hat{\Delta} x_{t+h,t}$. A similar procedure is used for the GDP deflator except that $\Delta x_{T+h} = X_{T+h}$.

Real Time vs. A Single Estimate. Had our data set been very long, the stationary forecast $E_t^m\{X_{t+h}\}$ could be constructed from any long time interval. However, since our data set is short and we examine the forecastability of excess returns, we do not use the factor loadings of a *single* model estimated for the entire period 1959:01 to 2003:12 combined. Instead, all our estimates of $E_t^m\{X_{t+h}\}$ and $Z_t^{(X,h)}$ are made by using real time forecasts. For each date in the sample we thus use data from 1959:01 up to the given date in order to recompute the factor loadings, reestimate a stationary model with which we compute $E_t^m\{X_{t+h}\}$ and then deduce the values of $Z_t^{(X,h)}$.

Table 1A: Summary Statistics of Market Beliefs

h = 6 Months or 2 Quarters Ahead	Time Ave.	Std Dev	Skew	Kurt	Autocorr.
Fed Fund rate	0.273	0.528	-0.063	2.945	0.700
1 year T-bill rate	0.238	0.429	0.065	2.701	0.735
GDP deflator	0.365	0.595	0.173	3.054	0.674
h = 12 Months or 4 Quarters Ahead					
Fed Fund rate	0.267	0.763	0.148	2.085	0.632
1 year T-bill rate	0.385	0.681	0.299	2.461	0.841
GDP deflator	0.398	0.798	0.684	4.099	0.740

Tables 1A and 1B provide some summary statistics of a sample of extracted market belief variables $Z_t^{(X,h)}$. The last column ρ_{t-1} in Table 1A reports the first order autocorrelation parameter. Although theory requires each market belief to have a *long term* time average equal to zero, it is clear the means over short time periods are not zero. Indeed, the fact that the belief indices for inflation and nominal interest rates have positive time averages for the period at hand is significant. It reflects the forecasting bias in the US during that era when beliefs in inflation and doubts about the efficacy of monetary policy persisted (see Kurz (2005)) despite the mounting evidence against these beliefs.

Table 1B: Correlation Matrix of Market Beliefs

6 Months or 2 Quarters Ahead	Fed Fund rate	1 year T-bill rate	GDP deflator
Fed Fund rate	1.000		
1 year T-bill rate	0.850	1.000	
GDP deflator	0.363	0.298	1.000
12 Months or 4 Quarters Ahead			
Fed Fund rate	1.000		
1 year T-bill rate	0.856	1.000	
GDP deflator	0.516	0.523	1.000

Figure 1 traces the graph of $Z_t^{(6,h)}$ for the 6-months T-bill rate with the two horizons h = 4, 12. The figure shows the belief index exhibits large fluctuations ranging from -1.5% to +2.5%. which are very significant from the economic point of view. Figure 2 traces the time variability of the cross-sectional standard deviations $\sigma_t^{(6,h)}$ of the $Z_t^{(6,h,i)}$ across i, for horizons h = 4, 12. It is clear from the figure that the dispersion of beliefs increases with the forecasting horizon. This is a common feature of all data on belief distributions.

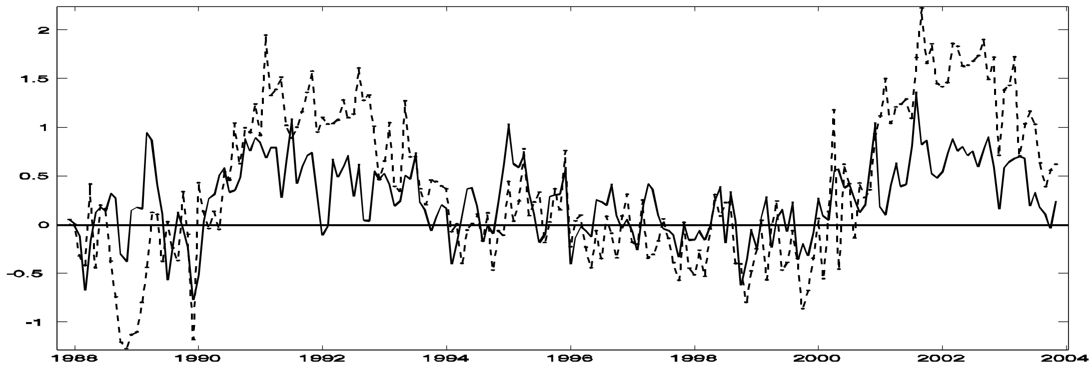

Figure 1: 6-month Treasury Bill rate: 4 and 12(dashed line) month ahead Market Belief $Z_t^{(6)}$

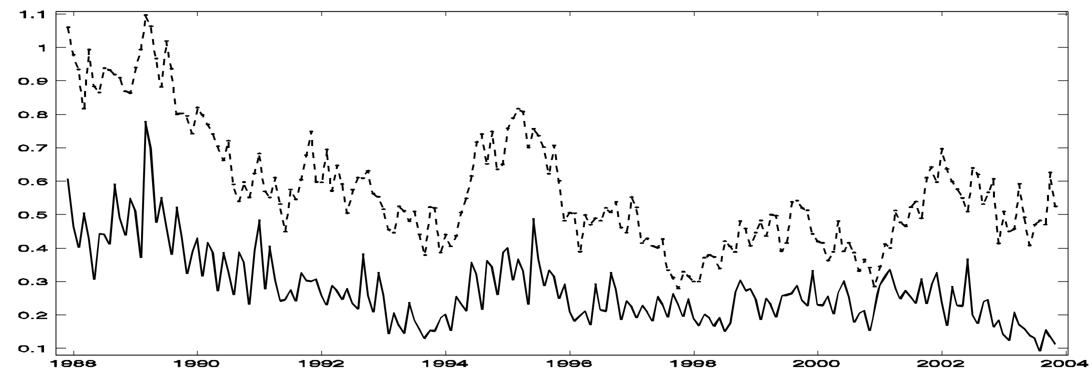

Figure 2: 6-month Treasury Bill rate: 4 and 12(dashed line) month ahead standard deviation of Market Belief $Z_t^{(6)}$

2.3 Data on Realized Market Interest Rates, Rates of Return and Excess Returns

Treasury Bills market. Theory suggests we work with interest rates implied by zero coupon bond prices hence we used data on zero coupon securities with maturities of 1 to 18 months, based on the Fama-Bliss file (see Fama and Bliss (1987)). The data up to 2003:11 was generated by a FORTRAN routines (provided by R.R. Bliss), using a method developed by Bliss for the unsmoothed Fama-Bliss data set (see Bliss (1997)). Let $\pi_{t+h}^{(j,h)}$ be the one period excess holding returns of T Bills with $(j + h)$ maturity held for h periods and sold at maturity j. It can be measured as a monthly or an annualized rate since all we say here about T Bills is independent of the unit of time selected. We study the h - month excess holding returns defined by

$$h\pi_{t+h}^{(j,h)} = (j+h)R_t^{(j+h)} - jR_{t+h}^{(j)} - hR_t^{(h)}$$

where $R_t^{(\tau)}$ is the one period interest rate implied by a zero coupon bond with maturity at τ. We study the two maturities j = 3 and 6 months. All data on the right hand side of the expression are then available in the Fama-Bliss file described above. The limiting factor in the study of this market is the BLUF data hence the period of analysis is 1987:12- 2003:11.

It is useful to clarify the trading mechanics needed to realize h period holding returns earned by selling a specified debt τ dates in the future. For example, to sell a six month Treasury Bill 12 month from now one must buy a Treasury Bond with maturity of 18 months and sell it 12 month from now. Returns on this *long position* consist of interest earned plus capital gains or losses realized.

Federal Fund Futures market. The second set of markets are for non contingent Federal Funds futures contracts with diverse monthly settlement horizons. A Fed Funds futures contract enables buyers and sellers to trade the risk of the Fed Funds rate that would prevail at the time of settlement. Hence this is the risk of the future target of the Fed Fund rate that would be fixed by the Fed's FOMC.

Fed funds futures have traded on the Chicago Board of Trade (CBOT) since October 1988 and settle based on the *mean Fed fund rate that prevails over a specified calendar month*. The mean is computed as a simple average of the daily averages published by the Federal Reserve Bank of New York. Hence, a trader needs to forecast the average federal fund rate during the contract month. The *contract horizon* is the number of months prior to the settlement date when a trader commits to go long or short such a contract. Contracts are settled by cash by the end of the contract month. Keep in mind that traders of such contracts do not invest capital and do not incur any opportunity cost[6]; they commit at t to a contract rate $F_t^{(h)}$ which becomes the contract cost basis at settlement, h months later. h = 3 means a three-month-ahead contract horizon. Data on $F_t^{(h)}$ are then recorded by the exchange and become public information. Some missing observations arise if a contract is not traded. Let us now explain the risks and rewards of a trader in this market.

The trader with a long position (the "buyer") of a Fed Funds futures contract owns a contract under which an interest rate of $F_t^{(h)}$ is paid on a $5 million deposit for a month during month t + h. $F_t^{(h)}$ is quoted as an annual rate. Denote by $R_{t+h}^{(F)}$ the actual average annualized Fed Funds rate during settlement month, h months later. Let n be the number of days in the contract month then at

[6] Traders are required to put up good faith security deposit which is a margin collateral to ensure they honor their pledge for the deposit as agreed. The collateral securities are owned by the parties to the contract who continue to benefit from any return to their investments. Margin cash is often held in the form of T Bills which yield interest to the owner. Hence a buyer or seller of a futures contract do not have any investment or opportunity cost except for the risk they take on the actual Fed Funds rate that would prevail at settlement. In this sense this market permits agents to trade risk of future monetary policy actions.

settlement a seller pays and a buyer receives for each contract the cash amount[7]

$$\$\text{Profits} = [F_t^{(h)} - R_{t+h}^{(F)}] \times \frac{n}{360} \times \$5{,}000{,}000.$$

It is then clear the parties trade the risk of $R_{t+h}^{(F)}$ which is the risk of the rate set by the Open Market Committee. It is reasonable to define the excess return of any gamble in this market to be defined by

$$\pi_{t+h}^{(F,h)} = F_t^{(h)} - R_{t+h}^{(F)}$$

Data on $F_t^{(h)}$ is recorded by CBOT while data on $R_{t+h}^{(F)}$ is reported by the Federal Reserve. Given the data set available the period for analysis of this market is 1988:10- 2003:11.

The problem of serial correlation. Serial correlation in forecast errors is inevitable for a well known reason. Computing excess returns utilizes overlapping data and this fact leads us to report, in the work below, robust standard errors of all estimates. We compute standard error using the heteroskedasticity and autocorrelation (HAC) procedure for robust estimates developed by Hodrick (1992), which generalizes the Hansen-Hodrick (1980) method. This correction places full weight on the lags of serial correlation in excess returns. We thus compute HAC robust standard errors with h-1 lags.

3. Analysis of the Risk Premium in the Bond and Federal Fund Futures Markets

3.1 Estimating Excess Return Functions

We now study the contribution of market belief to long term forecasting of excess returns and test the validity of the theoretical conclusions (26a)-(26b) about the effect of market belief on the time variability of market risk premia. Excess holding returns on three assets are studied: three month Treasury Bills and six month Treasury Bills with holding periods from 1 to 12 moths, and Federal Funds Futures contracts with holding periods of 1 to 6 months. For any asset X we estimate linear excess return functions of the following general form

$$\pi_{t+h}^{(X,h)} = \alpha_0^{(X,h)} + \alpha_1^{(X,h)} M_t + \alpha_2^{(X,h)} B_t + \varepsilon_{t+h}^{(X,h)} \tag{27}$$

where M_t is a vector of macroeconomic variables and B_t is a vector of market belief variables to be specified. We stress, at the outset, that it follows from the definition of $Z_t^{(X,h,i)}$ that belief variables B_t are m-orthogonal to all information in M_t. Since the risk premium is estimated in (27) using the long term statistics, it follows that variables in B_t add something new which is not in the market data M_t.

[7] The CBOT uses the 360 day year as the basic convention for quotation of interest rate and conversion from annual to monthly rates. The CBOT provides more details on its web page.

To specify B_t and M_t note that under an exponential utility the risk premium is a function of the mean market belief only; no other moments matter. For more general utility functions the entire distribution matters and we thus take into account additional moments of this distribution. To that end we study below the following three variables about any asset X:

$-Z_t^{(X,h)}$ – date t mean market belief about X at future date t+h

$\sigma_t^{(X,h)}$ – date t cross sectional standard deviations of individual beliefs about X at future date t+h.

These two variables are clear: they are simply the first two moments of the distribution of individual beliefs. Note the negative sign in $-Z_t^{(X,h)}$. It results from our convention to describe belief as in (8a)-(8c). Belief variables are oriented so that *a positive belief is perceived beneficial to a long position.* Since a belief in a higher future interest rate is a belief in a lower future price of debt, a belief which is beneficial to a long position in debt is a belief in lower rather than higher interest rates.

The macroeconomic variables in M_t are natural and reflect the literature on excess return on debt instruments and futures markets as noted in the introductory section. First, following Piazzesi and Swanson (2004) who concentrated on the cyclical variable, we use the following three macroeconomic variables in estimating risk premium in the Federal Funds futures market:

NFP_{t-1} - lagged year over year growth rate of Non Farm Payroll;

CPI_{t-1} - lagged year over year change in the consumer price index ;

F_t - the Federal Funds rate, reflecting the state of monetary policy at t.

Turning to past yields, recall that Cochrane and Piazzesi (2005) stressed the predictive power of past yields. Thus, we use yield variables to assess the risk premium in markets for 3 month and 6 month Treasury Bills. We introduce data on yields of Treasuries with 18 maturities covering 1970:01 to 2003:11. To reduce the dimension of information we computed principal components in real time (i.e. employ data up to t) and in all estimates we use the first three factors with notation $R_t^{F\upsilon}, \upsilon = 1,2,3$. These three factors account for 98% of the total variance of the yields' information matrix.

Some comments on the time unit are important. Rates of return on holding T Bills are naturally annual rates and hence comparable across different T Bills and horizons. This is not the case of Fed Funds futures. Total returns on such futures are measured in percentage points for the length of time the contracts are held consequently they are not annualized. Returns on short duration contracts are typically smaller than returns on long duration contracts hence excess returns on holding Fed Funds futures are not entirely comparable with returns on holding an asset with clearly defined holding cost.

Table 2A: Federal Fund Futures Market - Time Variability of Excess Returns

	Constant	NFP_{t-1}	CPI_{t-1}	F_t	$\sigma_t^{(F,h)}$	$-Z_t^{(F,h)}$	R^2	Chow Test p-value
h=2	-0.000 (0.056)	-0.016 (0.020)	-0.003 (0.026)	0.006 (0.018)	0.354 * (0.204)	-0.294 † (0.051)	0.313	0.003
h=4	-0.047 (0.101)	-0.142 † (0.040)	-0.037 (0.045)	0.122 † (0.032)	-0.752 * (0.438)	-0.419 † (0.066)	0.380	0.000
h=6	-0.199 (0.136)	-0.284 † (0.047)	-0.056 (0.085)	0.233 † (0.042)	-0.413 (0.455)	-0.397 † (0.130)	0.407	0.000

Table 2B: 3 Months Treasury Bills Market - Time Variability of Excess Returns

	Constant	NFP_{t-1}	CPI_{t-1}	F_t	R_{t-1}^{F1}	R_{t-1}^{F2}	R_{t-1}^{F3}	$\sigma_t^{(3,h)}$	$-Z_t^{(3,h)}$	R^2	Chow Test p-value
h=2	0.724 (0.676)	-0.042 (0.044)	0.005 (0.042)	-0.033 (0.095)	0.203 (0.308)	-0.181 † (0.067)	-0.010 (0.060)	0.065 (0.397)	-0.919 † (0.091)	0.422	0.005
h=4	0.939 * (0.567)	-0.136 † (0.033)	-0.021 (0.045)	0.013 (0.086)	0.398 (0.273)	-0.036 (0.063)	-0.062 (0.052)	-0.721 † (0.321)	-0.542 † (0.082)	0.416	0.067
h=6	1.329 † (0.552)	-0.191 † (0.023)	-0.000 (0.043)	-0.033 (0.079)	0.603 † (0.259)	0.014 (0.057)	-0.062 (0.051)	-0.397 (0.259)	-0.276 † (0.080)	0.434	0.041
h=8	1.964 † (0.535)	-0.181 † (0.025)	-0.003 (0.034)	-0.119 * (0.070)	0.915 † (0.243)	0.028 (0.054)	-0.030 (0.045)	-0.183 (0.193)	-0.250 † (0.035)	0.537	0.003
h=10	1.781 † (0.543)	-0.178 † (0.026)	0.001 (0.023)	-0.042 (0.071)	0.796 † (0.261)	-0.053 (0.035)	-0.023 (0.037)	-0.447 † (0.152)	-0.223 † (0.042)	0.640	0.059
h=12	1.684 † (0.593)	-0.200 † (0.021)	0.004 (0.026)	-0.029 (0.079)	0.749 † (0.295)	-0.068 (0.038)	-0.012 (0.026)	-0.403 † (0.097)	-0.180 † (0.027)	0.673	0.434

Table 2C: 6 Months Treasury Bills Market - Time Variability of Excess Returns

	Constant	NFP_{t-1}	CPI_{t-1}	F_t	R_{t-1}^{F1}	R_{t-1}^{F2}	R_{t-1}^{F3}	$\sigma_t^{(6,h)}$	$-Z_t^{(6,h)}$	R^2	Chow Test p-value
h=2	1.027 (1.414)	-0.111 (0.084)	-0.029 (0.089)	-0.017 (0.204)	0.225 (0.646)	-0.426 † (0.136)	0.095 (0.112)	-0.127 (0.952)	-1.985 † (0.179)	0.444	0.191
h=4	1.713 (1.293)	-0.316 † (0.071)	-0.076 (0.100)	0.056 (0.182)	0.788 (0.594)	-0.062 (0.118)	-0.075 (0.093)	-1.200 (0.747)	-1.182 † (0.148)	0.451	0.447
h=6	2.703 † (1.239)	-0.391 † (0.069)	-0.051 (0.089)	-0.015 (0.161)	1.268 † (0.568)	0.009 (0.116)	-0.081 (0.119)	-1.246 † (0.467)	-0.799 † (0.147)	0.494	0.040
h=8	3.261 † (1.207)	-0.394 † (0.073)	-0.075 (0.082)	-0.052 (0.164)	1.514 † (0.587)	-0.074 (0.106)	-0.028 (0.096)	-0.752 * (0.412)	-0.622 † (0.082)	0.596	0.013
h=10	4.174 † (1.319)	-0.387 † (0.075)	0.005 (0.051)	-0.187 (0.175)	1.852 † (0.640)	-0.170 † (0.084)	-0.020 (0.074)	-0.763 † (0.359)	-0.451 † (0.124)	0.638	0.003
h=12	4.069 † (1.253)	-0.401 † (0.059)	-0.004 (0.050)	-0.173 (0.162)	1.812 † (0.619)	-0.176 † (0.085)	0.024 (0.054)	-0.683 † (0.212)	-0.388 † (0.098)	0.664	0.114

This lack of comparability should be kept in mind in any cross-table comparisons. Tables 2A-2C present parameter estimates of (27) for the three markets and even numbered horizons. (*) denotes significance at 10% level and (†) denotes significance at 5% level or better. All R^2 are adjusted

3.2 Evaluating the Results

Considering Tables (2A)-(2C) combined, we find that the pro-cyclical variable NFP used by Piazzesi and Swanson (2004), and the yield variables used by Cochrane and Piazzesi (2005) are, indeed, important components of the risk premium. We note however that only the first factor of past yields is consistently significant. Our central concern is the size and sign of the belief variables.

3.2.1. Statistical Evaluation.

Starting with a statistical evaluation, the effect of market belief is significant, large and universally compatible with the theoretical predictions. This constitutes an empirical support for the hypothesis that, like society at large, markets are moved by perceptions. Fluctuations of real pro-cyclical variables account for some variability of risk premia but variations in market perceptions, *which may express mistaken interest rate forecasts*, are at least as important. The data supports the Market Risk Premium hypothesis in (26b). Keeping in mind our orientation convention, the data reveals that the parameters of the mean market beliefs $-Z_t^{(X,h)}$ in Tables 2A-2C are *always negative,* they are large, they are always statistically significant and are key contributors to the high R^2.

The parameters of $\sigma_t^{(X,h)}$, which measure market diversity, are not consistently significant across horizons for excess returns in the Fed Fund Futures Market. However, in the 6-month T Bill market the parameters of $\sigma_t^{(X,h)}$ are consistently significant for all horizons $h > 2$ and for 3-month T Bills they are consistently significant at intermediate and longer horizons, h = 4, 8, 10. With one exception, the estimates are always negative and large. This result says that an increase in diversity of market opinions *decreases the risk premium*. We derived this same result in earlier theoretical work from a simulation model (see Kurz and Motolese (2001)). The reason for it is that markets with more diverse beliefs are more stable since beliefs tend to cancel each other out, resulting in reduced volatility. In essence, with increased diversity the effects of the law of large numbers are more pronounced over time. The converse holds as well: markets are more risky the higher is the degree of unanimity in them. In such markets small changes of market belief result in sharp change of prices when too many people try to get through the same door.

Keeping in mind the limitation of R^2 we present in Table 3 the contribution of all belief variables to the R^2. They reveal again that belief variables explain a significant proportion of the risk

Table 3: Contribution of Belief to Excess Returns Predictability

Asset	Horizon	R^2 Without Beliefs	R^2 With Beliefs
Fed Fund Futures	h=2	0.061	0.313
	h=4	0.201	0.380
	h=6	0.345	0.407
3 Months T-Bill	h=2	0.122	0.422
	h=4	0.256	0.416
	h=6	0.367	0.434
	h=8	0.460	0.537
	h=10	0.541	0.640
	h=12	0.595	0.673
6 Months T-Bill	h=2	0.131	0.444
	h=4	0.290	0.451
	h=6	0.389	0.494
	h=8	0.502	0.596
	h=10	0.558	0.638
	h=12	0.600	0.664

premium. To see how belief variables contribute to the risk premium we exhibit in Figures 3-5 the fitted and the realized excess holding returns for a sample of three of our models, in accord with the estimates in Tables 2A-2C. The figures show that the results for Fed Funds futures are less precise than the results for T Bills. However, we note the great success of our estimated model in *predicting the turning points* of the time series. This high accuracy is the crucial contribution of the belief variables in capturing the time variability of the market's risk premia. One may also note that the belief variables enable the fitted values to match the realized data at high frequency within the broader cyclical pattern.

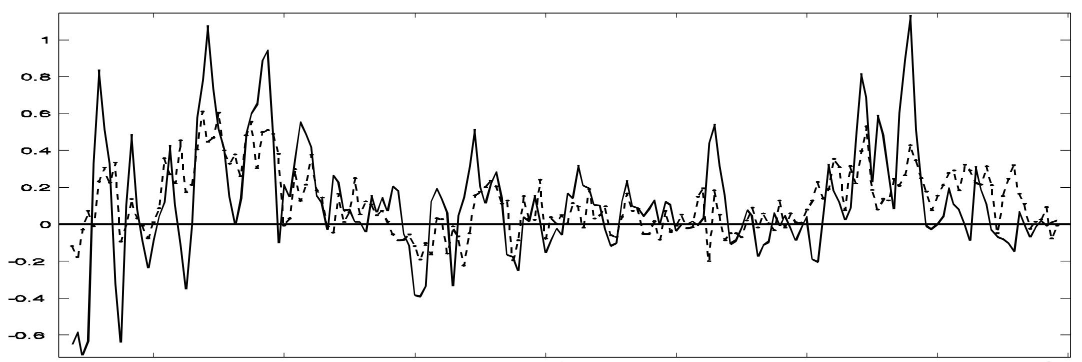

Figure 3: Excess returns on Fed Fund Futures contract 3 months ahead. The dashed line represents the fitted values from regression (27)

To sum up our findings, from the econometric point of view we confirm the result of earlier work which shows that pro-cyclical fundamental variables are important components of the time variability of the risk premium. The new fundamental forces proposed in this paper are the beliefs of agents. These variables make a clear and statistically significant contribution to the risk premium.

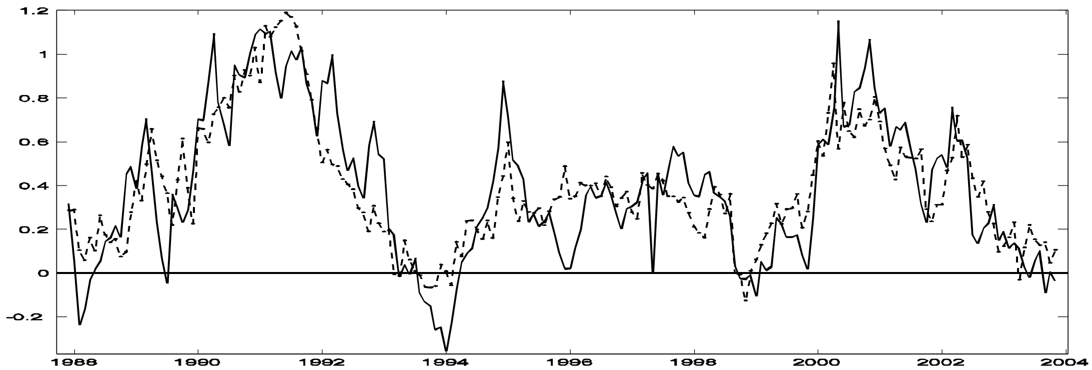

Figure 4: Excess returns on 3 Months TBill 12 months ahead. The dashed line represents the fitted values from regression (27)

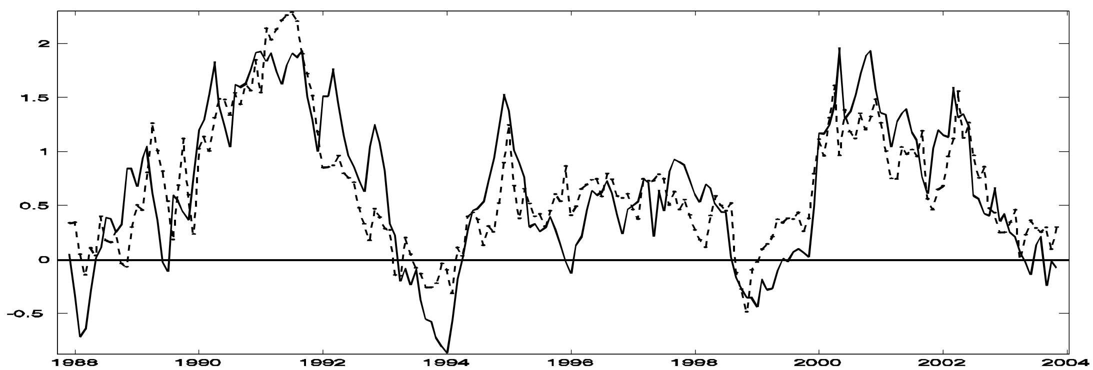

Figure 5: Excess returns on 6 Months TBill 12 months ahead. The dashed line represents the fitted values from regression (27)

Non- Stationarity. Our theory hinges on agents not knowing the true structure of the economy since it exhibits non-stationarity. In that case the risk premium has to exhibit non-stationarity as well. To test for parameter time variability we could select dates when structural changes have been studied by others. Our view is that forecast functions change for many reasons and practically any date will do for a Chow test. Since the periods 1988:10- 2003:11 for Fed Funds and 1987:12-2003:11 for T Bills are relatively short, *we chose the mid-points of 1996:04 and 1995:11 to maximize the number of observations per period.* For these sub - periods we conduct Chow tests of parameter time variability. In Tables 2A-2C, presented earlier, we report parameter estimates for the *entire period* and, in the last columns, p-values of Chow tests for breaks in the two chosen dates. Almost all Chow tests lead to *a rejection of the hypothesis of structural parameter time invariance* *in all markets.* The Chow tests are particularly significant since we have only 91 observations for Fed Futures and 96 for T Bills in each of the sub periods.

3.2.2 Economic Evaluation: Magnitude of the Effect of Market Belief on the Risk Premium

We now evaluate the order of magnitude of the effects of beliefs on risk premia. Note first, that we cannot evaluate the *total* effect since we do not have a measure of the *constant* effect of market beliefs on long term volatility of asset returns, as measured by $[R\frac{\hat{\sigma}_Q^2}{\tau}]$ in (26a). We can only measure the *variable* effect of $(\sigma_t^{(X,h)}, Z_t^{(X,h)})$. Hence, our estimates are lower bounds only. To that end we provide in Table 4 some long term statistics on the belief variables during the period[8] at hand. Together with the estimated parameters in Tables 2A-2C we assess the effects, on market premia, of these variables measured in units of standard deviations. Such computations provide an idea of the order of magnitude of the effect of these changes. To illustrate the effect we consider two cases: decreased optimism and increased diversity of market opinions.

Table 4: Long Term Statistics of Belief Variables (in basis points)

	$\sigma_t^{(F,h)}$		$\sigma_t^{(3,h)}$		$\sigma_t^{(6,h)}$		$Z_t^{(F,h)}$		$Z_t^{(3,h)}$		$Z_t^{(6,h)}$	
	Avg.	Std dev.	Avg.	Std dev.	Avg.	Std dev.	Avg.	Std dev.	Avg.	Std dev.	Avg.	Std dev.
h=2	13.7	7.8	19.0	8.4	20.3	8.4	9.0	34.5	10.8	33.8	15.4	37.7
h=4	21.5	10.6	26.8	11.2	28.0	11.1	22.3	47.9	21.6	35.5	26.3	38.0
h=6	30.4	14.1	35.1	14.4	36.3	13.8	27.3	52.8	25.2	41.6	29.3	42.3
h=8	----	----	43.2	16.8	44.5	16.3	----	----	29.4	52.2	34.1	53.7
h=10	----	----	50.3	18.5	51.4	17.7	----	----	23.7	65.2	30.5	69.2
h=12	----	----	56.7	18.4	57.6	17.9	----	----	33.1	73.1	39.6	75.4

(1) The effect of decreased optimism. If we set $Z_t^{(X,h)}$ equal to two standard deviations above its mean during the studied period as in Table 4, the total effect on the risk premium is as follows:

-0.397×(-(27.3+105.4)) = **+52.68** bp in the Federal Fund Futures Market when h = 6;

-0.180×(-(33.1+146.2)) = **+32.27** bp in the 3 Months Treasury Bills Market when h = 12;

-0.388×(-(39.6+150.8)) = **+73.88** bp in the 6 Months Treasury Bills Market when h = 12.

(2) The effect of diversity of market opinions. If we set $\sigma_t^{(X,h)}$ equal to two standard deviations above its mean during the studied period as in Table 4, the total effect on the risk premium is as follows:

-0.413×(30.4+14.1) = **-18.38** bp in the Federal Fund Futures Market when h = 6;

-0.403×(56.7+36.8) = **-37.68** bp in the 3 Months Treasury Bills Market when h = 12;

[8] As pointed out earlier, theory requires that each market belief have a *long term* time average equal to zero. Due to the short time span of the sample periods we have considered, the time averages of $Z_t^{(X,h)}$ are not zero.

$-0.683 \times (57.6+35.8) =$ **-63.79** bp in the 6 Months Treasury Bills Market when h = 12.

From the above we see that market pessimism can frequently account for an increase in the risk premium of up to about 70 basis points while an increase in the diversity of market opinions can frequently account for a decrease in the risk premium of up to about 60 basis points.

To measure the *joint effect* of the two belief variables $B_t^{(X,h)} = (\sigma_t^{(X,h)}, Z_t^{(X,h)})$ combined, we denote by $J_t^{(X,h)} = \hat{\alpha}_2^{(X,h)} \bullet B_t^{(X,h)}$ the estimated value of the belief component of the risk premium in (27). $J_t^{(X,h)}$ may be positive or negative and could thus increase or decrease the premium at any date. To measure an order of magnitude of the component $J_t^{(X,h)}$ of the risk premium at t relative to the mean premium, let $\overline{|J^{(X,h)}|}$ be the mean of the $|J_t^{(X,h)}|$. Table 5 reports, for each asset, the unconditional annualized mean premium for the sample period and the annualized value of $\overline{|J^{(X,h)}|}$.

Table 5: Component of Belief in the Premium (in basis points, annualized)

	Fed Funds Futures		3 Months Treasury Bills		6 Months Treasury Bills	
	Mean Premium	$\overline{\lvert J^{(F,h)}\rvert}$	Mean Premium	$\overline{\lvert J^{(3,h)}\rvert}$	Mean Premium	$\overline{\lvert J^{(6,h)}\rvert}$
h=2	44.4	61.4	37.2	24.6	60.4	59.2
h=4	57.6	51.3	29.4	17.9	54.2	37.8
h=6	68.8	34.0	27.2	11.5	54.2	34.9
h=8	----	----	27.5	11.1	70.8	32.0
h=10	----	----	40.6	19.7	69.5	35.8
h=12	----	----	38.9	18.5	66.1	33.2

There are two conclusions one can draw from Table 5 about the belief component in the premium:

(1) The $\overline{|J^{(X,h)}|}$ component in the risk premium is large: for the assets at hand it is generally larger than 50% of the mean premium. We remark that both the premium as well as $J_t^{(X,h)}$ are very volatile hence the range of the $J_t^{(X,h)}$ component of the risk premium is wide.

(2) The $\overline{|J^{(X,h)}|}$ component is largest for short holding returns and declines to about 50% at h=12. For very short holding periods of less than 3 months this component may often dominate the premium.

The second result is consistent with the intuition that risk premia are dominated by market beliefs for very short holding periods. This result is also compatible with the results reported in Table 3 that show the R^2 without the belief variables are very small for very short holding periods.

We stress that $J_t^{(X,h)}$ may be negative or positive and in the long run may not contribute much to the mean premium *itself*. We also recall that the average risk premium contains the constant component in (26a) which constitutes an important effect of the market beliefs on the volatility of asset return and hence on the risk premium. We do not measure this effect here.

4. Final Comments: On Bull and Bear Markets

Excess volatility of asset returns above the level accounted by "fundamental" forces is a fact contested by only very few economists. Asset price volatility does not imply time variability of risk premia but the converse does hold true. It follows that the exhibited strong impact of market belief on risk premia teaches us two additional lessons. First, it offers a direct demonstration that market perception should be considered to be as fundamental to asset pricing as the customary exogenous variables. Second, that market belief is actually an observable state variable which can be used for a deeper understanding of the causes of market dynamics. The terms "bull" or "bear" markets have a limited meaning in an REE based asset pricing theory according to which such markets are related to business cycles. Contrast this with the fact that during the last half century business cycles have moderated while volatility of financial markets has not declined and perhaps has increased. Accordingly, we have shown that beyond the standard effect of business cycles "bull" and "bear" markets do have specific meaning. "Bull" markets are periods of low risk premium caused by unusually positive market perception about future asset payoff while "bear" markets are periods of high risk premium caused by unusually negative market perception about future asset payoff.

Turning to the nature of the effect of market belief we have shown that the premium on holding a risky asset over the riskless rate has two components. The first is a direct effect which results from the impact of market belief on increased excess volatility of asset returns. This premium is constant. The second effect, which we call "the market belief risk premium," varies over time. We have shown that the premium is decreasing in the mean market belief Z_t. This means that an optimistic market is a market in which risk perception is low and the risk premium is low. Equipped with a detailed panel data on individual forecasts of interest rates our theory proposes a specific way in which we should deduce the appropriate panel data of market belief. Using such data we then test our theory empirically in the markets for Federal Funds Futures, 3 month Treasury Bills and 6 month Treasury Bills. We show that the data supports the theory and the estimated effect is large.

References

Allen, F., Morris, S., Shin, H.S. (2006) : "Beauty Contests and Iterated Expectations in Asset Markets." *Review of Financial Studies,* **19**, 719-752.

Batchelor, R. and Dua, P. (1991): "Blue Chip Rationality Tests." *Journal of Money, Credit and Banking*, **23**, 692 - 705.

Blanchard, O. J., and Kahn, C. M. (1980): “The Solution of Linear Difference Models Under Rational Expectations.” *Econometrica*, **48**, 1305 - 1311.

Bliss, R. R., (1997): “Testing Term Structure Estimation Methods.” In Boyle, P., Pennacchi, G., Ritchken, P., (Ed.), Vol. 9, *Advances in Futures and Options Research*, JAI Press, Greenwich, Conn. 197 - 231.

Brown, D. and Jennings, R. (1989): “On Technical Analysis.” *Review of Financial Studies*, **2**, 527-551.

Brunnermeier, M.K., (2001): Asset Pricing Under Asymmetric Information. Oxford: Oxford University Press.

Campbell, J. and Shiller, R. J. (1991): “Yield Spreads and Interest Rate Movements: A Bird’s Eye View.” *Review of Economic Studies*, **58**, 495 - 514.

Cochrane, J., and Piazzesi, M., (2005): “Bond Risk Premia.” *American Economic Review* **95**, 138 - 160.

Detemple, J., Murthy S.(1994): “Intertemporal Asset Pricing with Heterogeneous Beliefs.” *Journal of Economic Theory* **62**, 294-320.

Fama, E.F. and Bliss, R.B., (1987): “The Information in Long-Maturity Forward Rates.” *American Economic Review*, **77**, 680-692.

Fan, M., (2006): “Heterogeneous Beliefs, the Term Structure and Time-Varying Risk Premia.” *Annals of Finance*, **2**, 259-285.

Grundy, B. and McNichols, M. (1989): “Trade and Revelation of Information Through Prices and Direct Disclosure.” *Review of Financial Studies*, **2**, 495-526.

Hansen, L.P. and Hodrick, R.J., (1988): “Forward ExchangeRates as Optimal Predictors of Future Spot Rates: An Econometric Analysis.” *Journal of Political Economy,* **88**, 829-853.

Harris, M., Raviv, A. (1993) : “Differences of Opinion Make a Horse Race.” *Review of Financial Studies* **6**, 473-506

Harrison, M. and Kreps, D.(1978): “Speculative Investor Behavior in a Stock Market with Heterogenous Expectations.” *Quarterly Journal of Economics* **92**, 323-336.

He, H. and Wang, J. (1995): “ Differential Information and Dynamic Behavior of Stock Trading Volume.” *Review of Financial Studies*, **8**, 914 - 972.

Hodrick, R.J., (1992): “Dividend Yields and Expected Stock Returns: Alternative Procedures for Influence and Measurement.” *Review of Financial Studies,* **5**, 357-386.

Kurz, M.(1974): “The Kesten-Stigum Model and the Treatment of Uncertainty in Equilibrium Theory.” In Balch, M.S., McFadden, D.L., Wu, S.Y., (ed.) *Essays on Economic Behavior Under Uncertainty*. Amsterdam: North-Holland, 389-399.

Kurz, M. (1994): "On the Structure and Diversity of Rational Beliefs." *Economic Theory* **4**, 877 - 900 . (An edited version appears as Chapter **2** of Kurz, M. (ed.) (1997)).

Kurz, M. (ed) (1997): *Endogenous Economic Fluctuations: Studies in the Theory of Rational Belief*. Studies in Economic Theory, No. **6**, Berlin and New York: Springer-Verlag.

Kurz, M. (2005): "Diverse Beliefs, Forecast Errors and Central Bank Policy." Working Paper, Stanford University, July, 2005.

Kurz, M. (2007a): "Beauty Contests Under Private Information and Diverse Beliefs: How different?" *Journal of Mathematical Economics. (forthcoming*).

Kurz, M. (2007b): "Rational Diverse Beliefs and Economic Volatility." Chapter in Hens, T. and Schenk-Hoppé, K.R. (ed.) *Handbook On Financial Markets: Dynamics and Evolution*." North-Holland, (forthcoming).

Kurz, M., Beltratti, , A. (1997): "The Equity Premium is No Puzzle." Chapter **11** in Kurz, M. (ed.) (1997) *Endogenous Economic Fluctuations: Studies in the Theory of Rational Belief*. Studies in Economic Theory, No. **6**, Berlin and New York: Springer-Verlag, 283 -316.

Kurz, M., Jin, H., Motolese, M. (2005a): "Determinants of Stock Market Volatility and Risk Premia." *Annals of Finance*, **1**, 109-147.

Kurz, M., Jin, H., Motolese, M. (2005b): "The Role of Expectations in Economic Fluctuations and the Efficacy of Monetary Policy." *Journal of Economic Dynamics and Control*, **29**, 2017 - 2065.

Kurz, M., Motolese, M. (2001): "Endogenous Uncertainty and Market Volatility." *Economic Theory*, **17**, 497 - 544.

Kurz, M., Schneider, M.(1996): Coordination and Correlation in Markov Rational Belief Equilibria. *Economic Theory* **8**, 489 - 520. (Appears as Chapter **10** of Kurz, M. (ed.) (1997a))

Kurz, M., Wu, H.M. (1996): "Endogenous Uncertainty in a General Equilibrium Model with Price Contingent Contracts." *Economic Theory*, **8**, 461 -488. (Appears as Chapter **2** of Kurz, M. (ed.) (1997a)).

Motolese, M. (2001): "Money Non-Neutrality in a Rational Belief Equilibrium with Financial Assets." *Economic Theory*, **18**, 97 - 16.

Motolese, M. (2003): "Endogenous Uncertainty and the Non-Neutrality of Money." *Economic Theory*, **21**, 317 - 345.

Nielsen, K. C. (1996): “Rational Belief Structures and Rational Belief Equilibria.” *Economic Theory*, **8**, 339 - 422 .

Nielsen, K. C. (2003): “Floating Exchange Rates vs. A Monetary Union Under Rational Beliefs: The Role of Endogenous Uncertainty.” *Economic Theory*, **21**, 347 - 398.

Piazzesi, M. and Swanson, E., (2004): “Futures Prices as Risk-Adjusted Forecasts of Monetary Policy.” NBER Working Paper 10547, June.

Saari, D. (2006): **“**Parts, Whole, and Evolution”, Lecture notes, University of California at Irvine.

Stambaugh, R.F., (1988): “The Information in Forward Rates: Implications for Models of the Term Structure.” *Journal of Financial Economics*, **21**, 41 -70.

Stock, H. J., Watson, W. M. (1999): “Forecasting Inflation.” Department of Economics, Princeton University.

Stock, H. J., Watson, W. M. (2001): “Forecasting Output and Inflation: The Role of Asset Prices.” Department of Economics, Princeton University, February.

Stock, H. J., Watson, W. M. (2002): “Macroeconomic Forecasting Using Diffusion Indexes.” *Journal of Business and Economic Statistics*, The American Statistical Association, **20**, 147-162.

Stock, H. J., Watson, W. M. (2005): “An Empirical Comparison of Methods for Forecasting Using Many Predictors.” Department of Economics, Harvard University, January.

Toukan, A. (2006): “Privately Held or Publicly Owned? Evolutionary Game Theoretic Analysis.” Working Paper, University of California at Irvine.

Varian, H.R. (1985): “Divergence of Opinion in Complete Markets: A Note.” *Journal of Finance* **40**, 309-317

Varian, H.R. (1989) : “Differences of Opinion in Financial Markets.” In *Financial Risk: Theory, Evidence and Implications*, Proceeding of the 11th Annual Economic Policy Conference of the Federal Reserve Bank of St. Louis, Stone, C.C. ed. Boston: Kluwer Academic Publishers.

Wang, J. (1994): "A Model of Competitive Stock Trading Volume." *Journal of Political Economy*, 102, 127 - 168.
Wu, H.M., Guo, W.C. (2003): "Speculative Trading with Rational Beliefs and Endogenous Uncertainty." *Economic Theory*, **21** , 263 292.
Wu, H.M., Guo, W.C. (2004): "Asset Price Volatility and Trading Volume with Rational Beliefs." *Economic Theory*, **23** , 461- 488.

APPENDIX A: Derivation of the Value Function

For simplicity we ignore in this Appendix the index i identifying the agent who carries out the optimization. Hence, the dynamic programming problem is as follows. Given initial values (θ_0^i , W_0^i), maximize

$$U_t = \underset{(\theta^i, C^i)}{E_t^i} [\sum_{s=0}^{\infty} -\beta^{t+s-1} e^{-(\frac{1}{\tau} C_{t+s}^i)} \mid H_t]$$

subject to the following definitions

$$W_{t+1} = (W_t - C_t)R + \theta_t Q_{t+1}$$
$$Q_{t+1} = p_{t+1} + (d_{t+1} + \mu) - p_t R$$
$$\psi_t = (1, d_t, z_t, g_t)$$

and stochastic transition functions

$$\begin{aligned} d_{t+1} &= \lambda_d d_t + \lambda_g^d g_t + \varepsilon_{t+1}^d \\ Z_{t+1} &= \lambda_z Z_t + \lambda_g^z g_t + \varepsilon_{t+1}^z , \\ g_{t+1} &= \lambda_z g_t + \varepsilon_{t+1}^g \end{aligned} \quad \Lambda_\psi = \begin{pmatrix} 1, 0, 0, 0 \\ 0, \lambda_d, 0, \lambda_g^d \\ 0, 0, \lambda_z, \lambda_g^z \\ 0, 0, 0, \lambda_z \end{pmatrix}, \quad \hat{\varepsilon}_t = (1, \varepsilon_t^d, \varepsilon_t^z, \varepsilon_t^g), \ (\varepsilon_t^d, \varepsilon_t^z, \varepsilon_t^g) \sim N(0, \Sigma).$$

Step 1: simplification. We thus define, for the unknown matrix V

$$\Lambda = \begin{pmatrix} \lambda_d, 0, \lambda_g^d \\ 0, \lambda_z, \lambda_g^z \\ 0, 0, \lambda_z \end{pmatrix} , \quad V = \begin{pmatrix} v_{00}, v_{01}, v_{02}, v_{03} \\ v_{01} & & \\ v_{02} & V_{11} & \\ v_{03} & & \end{pmatrix} = \begin{pmatrix} v_{00}, & \hat{v}_0^T \\ \hat{v}_0 & , V_{11} \end{pmatrix}$$

We now have $\psi_{t+1} = \Lambda_\psi \psi_t + \Lambda_\varepsilon \varepsilon_{t+1}$, where $\Lambda_\varepsilon = \begin{pmatrix} 0, 0 \\ 0, I_{(3\times3)} \end{pmatrix}$ is a 4×4 matrix

We assume that $p_t = a_d d_t + a_z z_t + P_0$ and verify it later when we solve for equilibrium. Using this price map we can compute excess return in terms of the state variables we have that

$$Q_{t+1} = (a_d + 1)[\lambda_d d_t + \lambda_g^d g_t + \varepsilon_{t+1}^d] + a_z[\lambda_z Z_t + \lambda_g^z g_t + \varepsilon_{t+1}^z] + P_0 - [a_d d_t + a_z Z_t + P_0]R + \mu$$

Hence

$$Q_{t+1} = [(a_d + 1)\lambda_d - R a_d] d_t + [a_z \lambda_z - R a_z] Z_t + [(a_d + 1)\lambda_g^d + a_z \lambda_g^z] g_t + [P_0(1 - R) + \mu] + [(a_d + 1)\varepsilon_{t+1}^d + a_z \varepsilon_{t+1}^d]$$

Or,

$$Q_{t+1} = a^T \psi_t + \hat{b}^T \varepsilon_{t+1} , \text{ hence } E_t[Q_{t+1}] = a^T \psi_t$$

where

$$a^T = ([P_0(1 - R) + \mu], [(a_d + 1)\lambda_d - R a_d], [a_z \lambda_z - R a_z], [(a_d + 1)\lambda_g^d + a_z \lambda_g^z]) , \quad \hat{b}^T = (0, (a_d + 1), a_z, 0).$$

Also, we shall use the notation $b^T = ((a_d + 1), a_z, 0)$. Now compute the expression

$$-\alpha W_{t+1} - \frac{1}{2}\psi_{t+1}^T V \psi_{t+1} = -\alpha(W_t - C_t)R - \alpha\theta_t[a^T\psi_t + \hat{b}^T\hat{\varepsilon}_{t+1}] - \frac{1}{2}\psi_t^T\Lambda_\psi^T V\Lambda_\psi\psi_t - \psi_t^T\Lambda_\psi^T V\Lambda_\varepsilon\hat{\varepsilon}_{t+1} - \frac{1}{2}\hat{\varepsilon}_{t+1}^T\Lambda_\varepsilon^T V\Lambda_\varepsilon\hat{\varepsilon}_{t+1}$$

Algebra and simplification leads to the conclusion that we have

$$-\alpha W_{t+1} - \frac{1}{2}\psi_{t+1}^T V\psi_{t+1} = -A_t - e_t^T\varepsilon_{t+1} - \frac{1}{2}\varepsilon_{t+1}^T V_{11}\varepsilon_{t+1}$$

where

$$A_t = \alpha(W_t - C_t)R + \alpha\theta_t a^T\psi_t + \frac{1}{2}\psi_t^T\Lambda_\psi^T V\Lambda_\psi\psi_t$$

$$e_t^T = [\alpha\theta_t b^T + \psi_t^T\Lambda_0^T] \text{ (this is a 3 vector) where } \Lambda_0^T = \begin{pmatrix} \hat{v}_0^T \\ \Lambda^T V_{11} \end{pmatrix} \text{ (3x4) matrix}, \Lambda_0 = (v_0, V_{11}\Lambda)$$

Step 2: The Bellman Equation. It is well known (see, for example, the Appendix of Wang (1994)) that the Bellman Equation for this problem with $\gamma = \frac{1}{\tau}$ is written in the form

$$J_t = \max_{(\theta_t, C_t)}[-\beta^{t-1}\exp\{-\gamma C_t\} - \beta^t E_t\exp\left\{-A_t - e_t^T\varepsilon_{t+1} - \frac{1}{2}\varepsilon_{t+1}^T V_{11}\varepsilon_{t+1}\right\} \quad \text{for some parameter matrix V}$$

But we know that

$$E_t\exp\left\{-A_t - e_t^T\varepsilon_{t+1} - \frac{1}{2}\varepsilon_{t+1}^T V_{11}\varepsilon_{t+1}\right\} = |1 + \Sigma V_{11}|^{-\frac{1}{2}}\exp[\frac{1}{2}e_t^T(I + \Sigma V_{11})^{-1}\Sigma e_t - A_t].$$

Also

$$\frac{1}{2}e_t^T(1 + \Sigma V)^{-1}\Sigma e_t = \frac{1}{2}[\alpha\theta_t b^T + \psi_t^T\Lambda_0^T]^T(I + \Sigma V_{11})^{-1}\Sigma[\alpha\theta_t b + \Lambda_0\psi_t]$$

$$= \frac{1}{2}\alpha^2\theta_t^2 b^T\Omega b + \alpha\theta_t b^T\Omega\Lambda_0\psi_t + \frac{1}{2}\psi_t^T\Lambda_0^T\Omega\Lambda_0\psi_t \quad , \text{ where } \Omega = (I + \Sigma V_{11})^{-1}\Sigma.$$

Hence, we have an expression for the expectations

$$\frac{1}{2}e_t^T(1 + \Sigma V_{11})^{-1}\Sigma e_t - A_t = -\alpha(W_t - C_t)R - \alpha\theta_t[a^T - b^T\Omega\Lambda_0]\psi_t + \frac{1}{2}\alpha^2\theta_t^2 b^T\Omega b - \frac{1}{2}\psi_t^T[\Lambda_\psi^T V\Lambda_\psi - \Lambda_0^T\Omega\Lambda_0]\psi_t.$$

The first order conditions are then stated as follows. Equating the derivative with respect to θ to zero leads to

$$-\alpha[a^T - b^T\Omega\Lambda_0]\psi_t + \alpha^2\theta_t b^T\Omega\psi_t = 0$$

And this proves equation (14) in the text which we can write in the more explicit form (since $E_t[Q_{t+1}] = a^T\psi_t$)

$$\theta_t^i = \frac{1}{\alpha b^T\Omega b}\left\{[a^T - b^T\Omega\Lambda_0]\psi_t\right\} \equiv \frac{1}{\alpha b^T\Omega b}\left\{[E_t(Q_{t+1}) + u^T\psi_t\right\} \quad , \quad u^T = -b^T\Omega\Lambda_0.$$

This last equation determines the parameter vector u. It also shows that this vector is the same for all agents since the assumption made in the text is that all agents are identically the same except for their belief states g_t^i. The last equation shows that the vector u depends only upon parameters of the stochastic structure.

Step 3: The Adjusted Variance and Constants. We can also explain the "adjustment" to the variance in (14) since

$$\hat{\sigma}_Q^2 = b^T\Omega b$$

which is the variance of the excess return function where the covariance matrix used is not Σ but rather Ω.

We now have

$$\alpha^2\theta_t^{i^2} b^T\Omega b = \frac{1}{b^T\Omega b}\left\{\psi_t^T[a^T - b^T\Omega\Lambda_0]^T[a^T - b^T\Omega\Lambda_0]\psi_t\right\}.$$

Hence the optimized value of the exponent is simply

$$\frac{1}{2}e_t^T(1 + \Sigma V_{11})^{-1}\Sigma e_t - A_t = -\alpha(W_t - C_t)R - \frac{1}{2}\psi_t^T M\psi_t$$

Where

$$M = \frac{1}{b^{T}\Omega b}[a^{T}-b^{T}\Omega\Lambda_0]^T[a^{T}-b^{T}\Omega\Lambda_0] + [\Lambda_\psi^T V\Lambda_\psi - \Lambda_0^T\Omega\Lambda_0].$$

Now take the derivative with respect to C and equate to zero to obtain

$$\gamma\exp\{-\gamma C_t\} = \alpha R\beta|1+\Sigma V_{11}|^{-\frac{1}{2}}\exp\left\{-\alpha(W_t - C_t)R - \frac{1}{2}\psi_t^T M\psi_t\right\} \quad , \quad \text{let} \quad G = |1+\Sigma V_{11}|^{-\frac{1}{2}}.$$

Hence the solution for C must satisfy

$$\gamma C_t = -\log[\frac{\beta\alpha RG}{\gamma}] + \alpha(W_t - C_t)R + \frac{1}{2}\psi_t^T M\psi_t$$

hence we finally have

$$C_t = -\frac{1}{\gamma+\alpha R}\log[\frac{\beta\alpha RG}{\gamma}] + \frac{\alpha R}{\gamma+\alpha R}W_t + \frac{1}{2(\gamma+\alpha R)}\psi_t^T M\psi_t.$$

The final details of showing that the value function is indeed the solution of the Bellman Equation leads to the demonstration that the unknown parameter α and matrix V are determined by the conditions

(i) $\alpha = \frac{r\gamma}{R}.$

(ii) $\frac{M}{R} = V.$

APPENDIX B: Deducing (4) $g_{t+1}^{i} = \lambda_Z g_t^{i} + \rho_{t+1}^{ig}$ from Bayesian Inference

We aim to maintain simplicity and analytic tractability and note at the outset that in a rapidly changing environment there is no universal procedures to learn an unknown sequence of parameters. It is then less important to explain why agents disagree and more important to describe their diversity so that equilibrium analysis is *tractable*. Description (4) in the text, $g_{t+1}^{i} = \lambda_Z g_t^{i} + \rho_{t+1}^{ig}$, of the dynamics of belief states leads to a simple and useful description of equilibrium pricing with diverse beliefs as shown in this paper. It does not entail extraction of information from market prices, it needs each agent to have a distinct state space to describe his uncertainty and requires an endogenous expansion of the economy-wide state space for equilibrium pricing. However, we now explore conditions under which the Markov dynamics (4) can be proved as a *consequence of elementary principles of Bayesian inference*.

In a standard Bayesian environment an agent faces data generated under a stationary structure but with an unknown and fixed parameter. The agent starts with a prior on the parameter and then uses Bayesian inference for retrospective updating of his belief. The term "retrospective" stresses that inference is made *after* data is observed. In real time the agent must use the prior to forecast all variables while learning can only improve *future* forecasts of these variables. Our model has some parameters fixed and others that change over time. The fixed parameters are known as they are deduced from the empirical frequencies. The time varying parameters, reflecting the non stationarity of the economy, are modeled by the fact that under the true probability Π the value d_t has a transition function of the form

(B.1) $$d_{t+1} - \lambda_d d_t = b_t + \rho_{t+1}.$$

The sequence of parameters b_t is an exogenous, time varying mean value function. Agents know λ_d but not the "regimes" b_t. This formulation includes economies with slow changing regimes, each lasting a long time or fast changing regimes. The fact that these change restricts the validity of Bayesian updating. To see why observe that at date t our agent has a prior belief about b_t with which he forecasts d_{t+1}. After observing d_{t+1} he updates his prior to have a sharper posterior estimate of b_t. But when date t+1 arrives he needs to forecast d_{t+2} and for that *he needs a prior on* b_{t+1}. Agents do not know if and when a parameter changes. If the b_t change slowly, a sharp posterior estimate of b_t (given d_{t+1}) may serve also as a prior belief about b_{t+1}. Indeed, if the agent knew that $b_t = b_{t+1}$ the *updated posterior of* b_t *is the best prior of* b_{t+1}. In the absence of such knowledge, agents would believe that $b_t = b_{t+1}$ is only one possibility. They would, then, seek additional information and use subjective interpretation of other public data to arrive at alternative subjective estimates of b_{t+1} to supplement the Bayesian posterior they have. Such subjective interpretation of public data arises naturally from the fact that public quantitative data is always provided together with a vast amount of *qualitative* information which is an important source of subjective interpretation of data.

B.I *Qualitative Information and Subjective Interpretation of Public Information*

Bayesian inference is only possible with *quantitative* measures. The fact is that quantitative data like d_t are always accompanied with much *qualitative* public information about usual or unusual conditions. For example, data on inflation are interpreted with reports on normal or abnormal productivity features, conditions of the labor markets, assessment of the price of energy, political environment, etc. If d_t are profits of a firm then d_t is just one number extracted from a detailed financial report of the firm, the industry, the technology or the products involved. If d_t are profits of the S&P500 then qualitative information includes general business conditions, monetary policy, political environment, prospective tax reform, trends in productivity and other macroeconomic conditions. Qualitative information cannot, in general, be compared over time and does not constitute conventional "data." For example, when a firm announces a new research into something that did not exist before, no past data is available for comparison. When a new product alters the nature of an industry, it is a unique event. Financial markets pay a great deal of attention to qualitative announcements which are often the focus of diverse opinions of investors.

There is little modeling of deduction from qualitative information. Saari (2006) uses qualitative information in a competitive model of market shares. Toukan (2006) is a second example. Here we provide a simple formalization (see Kurz (2006), (2007)) of the use of qualitative information. Thus, qualitative information consist of *statements about the future.* Let date t statements be $A_t = (A_{t1}, A_{t2}, ..., A_{tK_t})$, each with quantitative measures in some units. The list may change with t and K_t varies with time. The activity in a statement may turn out to impact $(d_{t+1} - \lambda_d d_t)$ or not. The effects may be desirable or not. A realization at t+1 is a vector $\varphi_{t+1} = (\varphi_{t+1,1}, \varphi_{t+1,2}, ..., \varphi_{t+1,K_t})$ of numbers which are 0 or 1. A 0 means the activity turns out to have no effect and 1 means it has an effect. These can be interpreted as "success" or "failure." There are 2^{K_t} possible vectors of outcomes $\varphi_{t+1}(k)$, $k = 1, 2, ..., 2^{K_t}$. Next, agent i has a subjective map from φ_{t+1} to an expected value $\Phi^i(\varphi_{t+1})$ of $(d_{t+1} - \lambda_d d_t)$. This is an independent estimate by agent i on how different he expects $(d_{t+1} - \lambda_d d_t)$ to be from the stationary forecast conditional upon the success or failure of the statements. But, keep in mind, the quantitative estimate $\Phi^i(\varphi_{t+1})$ depends upon the statement A_t. For example, a research plan with \$1 million budget would be expected to have a smaller impact than a plan with a \$1 billion budget.

Finally, conditional on A_t, agent i attaches probabilities $(a_1^i, a_2^i, \ldots, a_{2^{K_t}}^i)$ to the vectors $\varphi_{t+1}(k)$. The result of this procedure is that agent i has an alternate subjective estimate of $(d_{t+1} - \lambda_d d_t)$ based only on public A_t data:

(B.2) $$\Psi_t^i(A_t) = \sum_{k=1}^{2^{K_t}} a_k^i \Phi^i(\varphi_{t+1}^i(k)).$$

By (A1) or (3) the long term average of $(d_{t+1} - \lambda_d d_t)$ is zero. Hence, rationality requires that the Ψ_t^i are zero mean random variables. Although public *data* consist only of d_t, the procedure outlined shows that in a complex world agents endogenously create subjective quantitative measures which reflect their beliefs. We incorporate such a measure in the Bayesian procedure below.

B.II *A Bayesian Inference: Beliefs are Markov State Variables*

Start by assuming that d_{t+1} has a true transition of the form

(B.3) $$d_{t+1} - \lambda_d d_t = b_t + \rho_{t+1}^d \quad , \quad \rho_{t+1}^d \sim N(0, \frac{1}{\beta}).$$

Agents do not know b_t but β is known. At first decision date t (say, t = 1) an agent has two pieces of information. He knows d_t and observes qualitative information $(A_{(t)1}, A_{(t)2}, ..., A_{(t)K_t})$ with which to assess Ψ_t^i. Assume that without Ψ_t^i the prior subjective mean at t = 1 is b but to start the process he uses *both sources* to form, as yet unspecified, a prior belief $E_t^i(b_t | d_t, \Psi_t^i)$ about b_t (used to forecast d_{t+1}). The changing parameter b_t leads to a problem. When $d_{t+1} - \lambda_d d_t$ is observed, agent i updates his belief about the same parameter b_t to $E_{t+1}^i(b_t | d_{t+1}, \Psi_t^i)$[9] in a standard Bayesian inference but before knowing the assessment Ψ_{t+1}^i. The point is that the agent needs an *estimate of b_{t+1}*, not of b_t. Hence, his problem is how to go from $E_{t+1}^i(b_t | d_{t+1}, \Psi_t^i)$ to a prior of b_{t+1}? Without new information his belief about b_{t+1} is unchanged and he would use $E_{t+1}^i(b_t | d_{t+1}, \Psi_t^i)$ as a prior of b_{t+1}. This is surely true if the b's change very slowly or when $b_{t+1} = b_t$. Hence, if agent i believes $b_{t+1} \neq b_t$ a new prior is needed. To that end he uses the qualitative information $(A_{(t+1)1}, A_{(t+1)2}, \ldots, A_{(t+1)K_{t+1}})$ released publicly before trading at t+1. These lead to an alternate subjective estimate Ψ_{t+1}^i of b_{t+1}. Now the agent has two independent sources for belief about b_{t+1}: the last posterior $E_{t+1}^i(b_t | d_{t+1}, \Psi_t^i)$ is to be used if $b_{t+1} = b_t$ and Ψ_{t+1}^i if $b_{t+1} \neq b_t$. With a Bayesian approach we assume:

Assumption (B): Agent i uses a subjective probability μ to form date t+1 prior belief which is then

(B.4) $$E_{t+1}^i(b_{t+1} | d_{t+1}, \Psi_{t+1}^i) = \mu E_{t+1}^i(b_t | d_{t+1}, \Psi_t^i) + (1 - \mu)\Psi_{t+1}^i \qquad 0 \leq \mu < 1.$$

At t=1 it was assumed an initial posterior b, hence for consistency, if Ψ_1^i is Normal then

(B.5) $$b_1 \sim N(\mu b + (1 - \mu)\Psi_1^i, \frac{1}{\alpha}) \text{ for some } \alpha.$$

This assumption is the new element that permits the posterior $E_{t+1}^i(b_t | d_{t+1}, \Psi_t^i)$ of b_t to be upgraded to a date t+1 prior belief $E_{t+1}^i(b_{t+1} | d_{t+1}, \Psi_{t+1}^i)$ about b_{t+1}, before d_{t+2} is observed. We then have the following:

---

[9] **Note our notation**. We use the notation of $E_t^i(b_t | d_t, \Psi_t^i)$ for date t *prior* belief about the parameter b_t used to forecast d_{t+1}. We then use the notation $E_{t+1}^i(b_t | d_{t+1}, \Psi_t^i)$ for the posterior belief *about the same* b_t given the observation of d_{t+1} but without forming the estimate of Ψ_t^i. Assumption (A) will use this posterior belief as a building block in the formation of the new prior $E_{t+1}^i(b_{t+1} | d_{t+1}, \Psi_{t+1}^i)$ about the new parameter b_{t+1}.

Theorem (Kurz (2007a)): Suppose $\Psi_t^i \sim N(0\,,\frac{1}{\gamma})$, i.i.d. and Assumption (A) holds. Then for large values of t, the prior belief $E_t^i(b_t|d_t,\Psi_t^i)$ is a Markov state variable such that if we define $g_t^i = E_t^i(b_t|d_t,\Psi_t^i)$ and $\mu\kappa = \lambda_Z$ for some $0<\kappa<1$ then the dynamics (4) holds: (B.4) implies (4).

Proof: Pick a starting date t = 1. Data d_t is known and the agent generates a subjective measure of Ψ_t^i. He then forms a prior on b_t, which by Assumption (B) is $b_t \sim N(\mu b + (1-\mu)\Psi_t^i, \frac{1}{\alpha})$. Now we move on to t+1 and d_{t+1} is observed. The agent updates the prior in a standard Bayesian manner:

(B.6) $$E_{t+1}^i(b_t|d_{t+1},\Psi_t^i) = \frac{\alpha(\mu b + (1-\mu)\Psi_t^i) + \beta[d_{t+1} - \lambda_d d_t]}{\alpha + \beta} \quad , \quad 0 \le \mu \le 1.$$

But before date t+1 trading he generates the subjective measure Ψ_{t+1}^i of qualitative data. By Assumption (B) the expected parameter b_{t+1} under the new prior at t+1 is

$$E_t^i(b_{t+1}|d_{t+1},\Psi_{t+1}^i) = \mu\, E_t^i(b_t|d_{t+1},\Psi_t^i) + (1-\mu)\Psi_{t+1}^i \quad , \quad 0 \le \mu \le 1.$$

Denote by $\zeta = \frac{1}{\mu^2}$ and $\xi = \frac{1}{(1-\mu)^2}$. Then the prior is

$$b_{t+1} \sim N(\,E_t^i(b_{t+1}|d_{t+1},\Psi_{t+1}^i)\,,\,\frac{1}{\zeta(\alpha+\beta)+\xi\gamma}).$$

It is used to forecast $d_{t+2} - \lambda_d d_{t+1}$. Moving on to t+2, the agent observes $d_{t+2} - \lambda_d d_{t+1}$ and based on this observation he uses Bayesian inference to deduce a posterior belief about b_{t+1}

$$E_{t+1}^i(b_{t+1}|d_{t+2},\Psi_{t+1}^i) = \frac{(\zeta(\alpha+\beta)+\xi\gamma)[\mu E_t^i(b_t|d_{t+1},\Psi_t^i) + (1-\mu)\Psi_{t+1}^i] + \beta[d_{t+2} - \lambda_d d_{t+1}]}{\zeta(\alpha+\beta)+(\xi\gamma+\beta)}.$$

Before the start of date t+2 trading, the agent generates a new value Ψ_{t+2}^i leading to t+2 prior belief about the unobserved parameter b_{t+2}

$$E_{t+2}^i(b_{t+2}|d_{t+2},\Psi_{t+2}^i) = \mu E_{t+1}^i(b_{t+1}|d_{t+2},\Psi_{t+1}^i) + (1-\mu)\Psi_{t+2}^i \quad , \quad 0 \le \mu < 1.$$

When $d_{t+3} - \lambda_d d_{t+2}$ is observed the posterior belief about b_{t+2} is then

$$E_{t+2}^i(b_{t+2}|d_{t+3},\Psi_{t+2}^i) = \frac{[\zeta^2(\alpha+\beta)+(\xi\gamma+\beta)\sum_{n=0}^{1}\zeta^n - \beta][\mu E_{t+1}^i(b_{t+1}|d_{t+2},\Psi_{t+1}^i)+(1-\mu)\Psi_{t+2}^i] + \beta[d_{t+3}-\lambda_d d_{t+2}]}{\zeta^2(\alpha+\beta)+(\xi\gamma+\beta)\sum_{n=0}^{1}\zeta^n}.$$

Next the agent generates a new value Ψ_{t+3}^i leading to a t+3 prior belief $E_{t+3}^i(b_{t+3}|d_{t+3},\Psi_{t+3}^i)$ about b_{t+3}. By forward iteration we conclude that after N rounds the prior takes the form

$$E_{t+N}^i(b_{t+N}|d_{t+N+1},\Psi_{t+N}^i) = \frac{[\zeta^{N-1}(\alpha+\beta)+(\xi\gamma+\beta)\sum_{n=0}^{N-1}\zeta^n - \beta]}{\zeta^N(\alpha+\beta)+(\xi\gamma+\beta)\sum_{n=0}^{N-1}\zeta^n}[\mu E_t^i(b_{t+N-1}|d_{t+N},\Psi_{t+N-1}^i)+(1-\mu)\Psi_{t+N}^i] + $$

$$+ \frac{\beta[d_{t+N+1}-\lambda_d d_{t+N}]}{\zeta^N(\alpha+\beta)+(\xi\gamma+\beta)\sum_{n=0}^{N-1}\zeta^n}.$$

Now take the limit. Since $\zeta > 1$, as N increases $\zeta^N \to \infty$ hence we find that

$$\lim_{N\to\infty} \frac{[\zeta^{N-1}(\alpha+\beta)+(\xi\gamma+\beta)\sum_{n=0}^{N-1}\zeta^n - \beta]}{\zeta^N(\alpha+\beta)+(\xi\gamma+\beta)\sum_{n=0}^{N-1}\zeta^n} = \frac{(\alpha+\beta)+(\xi\gamma+\beta)(\frac{\zeta}{\zeta-1})}{\zeta(\alpha+\beta)+(\xi\gamma+\beta)(\frac{\zeta}{\zeta-1})} \equiv \kappa$$

Since all terms are positive it is clear that $0 < \kappa < 1$. Hence we have that for large t

$$E_{t+1}^{i}(b_{t+1}|d_{t+2},\Psi_{t+1}^{i}) \;=\; \kappa[\mu E_{t}^{i}(b_{t}|d_{t+1},\Psi_{t}^{i}) + (1-\mu)\Psi_{t+1}^{i}]$$

But by definition we have

(B.7) $$E_{t+1}^{i}(b_{t+1}|d_{t+1},\Psi_{t+1}^{i}) \;=\; \mu E_{t}^{i}(b_{t}|d_{t+1},\Psi_{t}^{i}) + (1-\mu)\Psi_{t+1}^{i}$$

We conclude that for large t, the contribution of each new observation of dividends is negligible hence

$$E_{t}^{i}(b_{t}|d_{t+1},\Psi_{t}^{i}) \;=\; \kappa E_{t}^{i}(b_{t}|d_{t},\Psi_{t}^{i}).$$

Inserting this last equation again into (B.7) we finally have the desired conclusion that for large t

(B.8) $$E_{t+1}^{i}(b_{t+1}|d_{t+1},\Psi_{t+1}^{i}) \;=\; \mu\kappa E_{t}^{i}(b_{t}|d_{t},\Psi_{t}^{i}) + (1-\mu)\Psi_{t+1}^{i}.$$

Now identify $g_t^{\,i} = E_t^{\,i}(b_t|d_t,\Psi_t^i)$, $(1-\mu)\Psi_{t+1}^i = \rho_{t+1}^{ig}$ and $\mu\kappa = \lambda_Z$ to see that (B.8) is actually (4). ■

The Theorem shows that as the d_t data set increases, there is nothing new to learn. The posterior does not converge but *the law of motion* of the posterior converges to a time invariant stochastic law of motion defined by (B.8). The posterior fluctuates forever, providing the basis for the dynamics of market belief but the fluctuations follow a simple Markov transition. New data d_t and Ψ_t^i alter the conditional probability of the agent, but these do not change the dynamic law of motion of $g_t^{\,i}$.