

This work is distributed as a Discussion Paper by the  
**STANFORD INSTITUTE FOR ECONOMIC POLICY RESEARCH**

SIEPR Discussion Paper No. 06-03

**Risk Premia, Diverse Belief and Beauty Contests**

By  
Mordecai Kurz  
Stanford University  
And  
Maurizio Motolese  
Università Cattolica di Milano  
August 2006

Stanford Institute for Economic Policy Research  
Stanford University  
Stanford, CA 94305  
(650) 725-1874

The Stanford Institute for Economic Policy Research at Stanford University supports research bearing on economic and public policy issues. The SIEPR Discussion Paper Series reports on research and policy analysis conducted by researchers affiliated with the Institute. Working papers in this series reflect the views of the authors and not necessarily those of the Stanford Institute for Economic Policy Research or Stanford University.

---

# Risk Premia, Diverse Belief and Beauty Contests

by

Mordecai Kurz<sup>1</sup> Maurizio Motolese<sup>2</sup>

This draft: **September 19, 2006**

**Abstract:** We present a theoretical and empirical evaluation of the role of market belief in the structure of risk premia. To that end we employ a familiar asset pricing model for which we develop in detail the belief structure. The novelty in this development is the treatment of individual and market beliefs as Markov state variables. Moreover, the market belief is observable and the paper explains how we extract it from the data. The advantage of our formulation is that it permits a closed form solution of equilibrium prices hence we can trace the exact effect of market belief on the time variability of equilibrium risk premia. We present a model of asset pricing with diverse beliefs. We then explore the conditions under which diverse beliefs arise. We then derive the equilibrium asset pricing and the risk premium which the model implies. Since asset prices are affected by the dynamics of market belief, the component of market risk which is determined by the belief of agents is thus termed “Endogenous Uncertainty.” The theoretical conclusions are tested empirically for investments in the futures markets, the bond markets.

Our main theoretical and empirical result is that fluctuations in the market belief about state variables are a dominant factor determining the time variability of risk premia. More specifically, we show that when the market holds abnormally favorable belief about future payoffs of an asset the market views the long position as less risky and hence the risk premium on that asset declines. This means that fluctuations in risk premia are *inversely* related to the degree of market optimism about future prospects of asset payoffs. This effect is very strong and empirically very dominant. The strong effect of market belief on market risk premia offers two additional perspectives. First, it offers an additional way of showing (for those who have any doubt) that fundamental factors affect market dynamics but perceptions have equally important effect on volatility. Second, that market belief is actually an observable data which can be used for a deeper understanding of the basic causes of stochastic volatility and time variability of risk premia.

*JEL classification:* D82, D83, D84, G12, G14, E27.

*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

---

<sup>1</sup> Department of Economics, Serra Street at Galvez, Stanford University, Stanford, CA. 94305-6072, USA

<sup>2</sup> Istituto di Politica Economica, Università Cattolica di Milano, Via Necchi 5, 20123, Milano, Italy.

# Risk Premia, Diverse Belief and Beauty Contests<sup>3</sup>

by

Mordecai Kurz and Maurizio Motolese

This draft: **September 19, 2006**

Market risk premia vary over time and their fluctuations are a major cause of market volatility. But what drive the changes in risk premia? The standard rational expectations equilibrium answer relates changes in risk premia to changes in information about fundamental conditions which correctly alter the market's assessment of future risky events. The most important among these are business cycle events. This reasoning implies that excess returns are predictable by changes in observed *fundamental* conditions and market volatility can be explained by such information. This conclusion is rejected by the data. Fluctuations in asset prices cannot be well explained by news about fundamental factors and, as Samuelson used to quip, the market will forecast eleven of the next five recessions.

The alternative perspective holds that, in addition to 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. Hence diverse beliefs is a simple consequence of lack of full knowledge. With diverse beliefs a large proportion of market volatility is endogenously generated. This component is called *Endogenous Uncertainty* (see, Kurz (1974)). A sample of papers includes Harrison and Kreps (1978), Varian (1985), (1989), Harris and Raviv (1993), Detemple and Murthy (1994), Kurz (1974), (1994), (1997a), (1997b), Kurz and Beltratti (1997), Kurz and Motolese (2001), Kurz Jin and Motolese (2005a) (2005b), Kurz and Wu (1996), Motolese (2001), (2003), Nielsen (1996), (2003), Wu and Guo (2003), (2004). In particular Kurz and Motolese (2001) and Kurz Jin and Motolese (2005a) demonstrate via simulations that Endogenous Uncertainty contributes a big

---

<sup>3</sup> 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.

component to the equity premium and leads to stochastic volatility. However, these papers study the structure of volatility and risk premia only via simulations of computed equilibria. They do not study the determinants of market risk premia either analytically or empirically. In this paper we focus on factors which contribute to risk premia. More specifically, we study the relationship between market belief and the structure of risk premia. If belief dynamics cause Endogenous Uncertainty how does the structure of belief affect the equilibrium risk premia? We derive analytical results which we then test empirically by employing new data measuring the market distribution of beliefs. Market belief data are extracted from observations on *monthly forecasts of future interest rates and macro economic variables* compiled by the Blue Chip Financial Forecasts (BLUF) since 1983. Since we study an economy where agents hold diverse probability beliefs and since a risk premium of an asset over the riskless rate is the conditional expectation of excess returns of the asset, there are many subjectively perceived risk premia in such economies. We thus need to sort through the measurement problem of risk premia.

The literature on excess returns and risk premia is large. We mention only 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)). These show that investments in Treasury securities generate large 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 conclusions regarding the effect of cyclical variables on risk premia. However, using our alternative perspective we show that conditional risk premia also contain a large *endogenous component* generated by the dynamics of the market state of belief. We call it "The Market Belief Risk Premium." This component is orthogonal to the observed "fundamental" variables used on

the above studies. The term “orthogonal” highlights the fact that pure belief is a variable which is measured net of all observed variables and has its own dynamic law of motion. This endogenous component reflects investor’s pure *perceived* risks of future returns - net of all fundamental information - and this includes perceived risk of future market’s beliefs itself. Although interest rates fluctuate with business cycles, large fluctuations of asset prices and interest rates are generated endogenously by market *beliefs* about future events, including future market beliefs.

The main results of this paper consist of two parts. First we show analytically and empirically that a large proportion of market fluctuations and risk premia are generated endogenously by the dynamics of market beliefs. These beliefs are entirely rational since in a non-stationary and changing economy investors cannot learn the true dynamics of return and hence often adopt beliefs which are wrong but which cannot be falsified by existing data. Under diverse beliefs the market often moves too high or too low resulting in large time variability of risk premia. Second, our most striking result shows that the market belief has a clear effect on the risk premium. When the market holds abnormally favorable belief about future payoffs of an asset (e.g. future interest rates or dividend payments) the market views the long position as less risky and consequently the *risk premium on holding that asset falls*. Hence, fluctuations in risk premia are inversely related to the degree of market optimism about future prospects of asset payoffs. We test our conclusion empirically in futures and bonds markets and show that this effect is very strong and empirically very dominant.

## **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  $\{D_t, t = 1, 2, \dots\}$ . Let  $r_t$  be the interest rate,  $R_t = 1 + r_t$  and let excess return over the riskless rate be  $(1/p_t)(p_{t+1} + D_{t+1} - R_t p_t)$ . The *risk premium* over the riskless is the conditional expectations of excess return. Since it is a function of equilibrium prices, a risk premium - as a function of the 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 us to study the factors which determine 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)). We thus make strong assumptions but these are justified by the fact that the model helps clarify the main ideas. Once explained, we show that our results are fully general. There is one problem we need to address at the outset. Our agents do not know the true probability  $\hat{\Pi}$  of the payoff process  $\{D_t, t = 1, 2, \dots\}$  and hold diverse probability beliefs about  $\hat{\Pi}$ . The fact that there are many subjective risk premia in the market raises two immediate questions that will be 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 *unequivocal* 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 observables, allowing rich statistical analysis. Given the data, all compute the same finite dimensional distributions of the observations and hence all know the same empirical moments. Using standard extension of measures they deduce from the data a unique empirical probability measure on infinite sequences which is 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 that the long run empirical data reveals that the  $D_t$ ’s constitute a Markov process and they are conditionally normally distributed with mean  $\mu + \lambda_d(D_t - \mu)$  and variance  $\sigma_d^2$ <sup>4</sup>. The unique empirical probability  $\hat{m}$  is common knowledge. To simplify define  $d_t = D_t - \mu$  and note that  $\{d_t, t = 1, 2, \dots\}$  is then a zero mean process with unknown true probability  $\Pi$  and a stationary empirical probability  $m$ . Keeping in mind the question of why  $m$  and  $\Pi$  *are not the same*, we now turn to the second question.

Our economy has undergone changes in technology, social organization and attitudes. These changes are rapid with profound economic effects which induce changes in asset productivity and management. Consequently, the process  $\{d_t, t = 1, 2, \dots\}$  is non-stationary. Although this means the distributions of the  $d_t$ ’s are time dependent, it is far more than saying that  $\{d_t, t = 1, 2, \dots\}$  constitutes a sequence of productivity “regimes.” It requires recognition that although assets pay in some unit of account, assets and commodities used as payoff are different over time. Such time variability makes it impossible to learn the unknown true probability  $\Pi$ . The stationary probability  $m$  is then merely an

---

<sup>4</sup> It would be more realistic to assume the values  $D_t$  grow and the growth rate of the values has a mean  $\mu$  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.

average over an infinite sequence of changing regimes. It reflects long term frequencies but it is not the true probability of the process. Belief diversity starts with the fact that agents disagree over the meaning of public information. They believe  $\Pi$  is different from  $m$  and they construct models to express the implications they see in the data. Being common knowledge, the stationary probability  $m$  is a reference for any concept of rationality. We may disagree with an agent and regard his model as “extreme” but we can declare it irrational only if it is contradicted by the data. Hence, any rationality requirement must insist a model does not contradict the evidence summed up by the probability  $m$ .

Turning now to our infinite horizon model, at date  $t$  agent  $i$  buys  $\theta_t^i$  shares of stock and receives the payment  $d_t + \mu$  for each of  $\theta_{t-1}^i$  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

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

$Q_t$  is excess returns. Given some initial values  $(\theta_0^i, W_0^i)$  the agent maximizes the expected utility

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

subject to a vector of yet unspecified state variables  $\psi_t^i$  and their transitions. Both are specified later.

$H_t$  is date  $t$  information consisting of known values of 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 we deduce the closed form solution of equilibrium risk premium we show how to generalize the main results.

To proceed we state an assumption and a conjecture. First, we assume the agent believes the payoff process  $\{d_t, t = 1, 2, \dots\}$  is conditionally normal. Second, we conjecture that given the state variables of the economy, equilibrium price  $p_t$  is conditionally normally distributed. In the next section we describe the state variables and the structure of belief and in Theorem 2 below we confirm the above conjecture. For an optimum there exists a constant vector  $u$  so the demand functions for the stock is

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

$\hat{\sigma}_Q^2$  is an *adjusted* conditional variance (the “adjustment” is explained later) of excess stock return which is assumed to be 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 stressed that disagreements among agents arises from diverse interpretation of the commonly known empirical record. We made a realistic assumption that the empirical frequencies of past payoffs follow a first order Markov process hence the long term empirical process has the transition

$$(3) \quad d_t = \lambda_d d_{t-1} + \rho_t^d, \quad \rho_t^d \sim N(0, \sigma_d^2).$$

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 mechanism? Those who believe the economy is stationary would accept (3) as the truth. We view such belief as rational since there is no empirical evidence against it. Since  $\{d_t, t=1,2,\dots\}$  is a non-stationary with unknown probability  $\Pi$ , the empirical record (3) is just an average over different regimes. Hence, most agents do not believe past empirical record 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)). Given this environment, each agent forms his own belief about  $d_{t+1}$  and other state variables explored in the next section. With high level of complexity, 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

We start with a methodological comment. We have noted that an agent should not be declared irrational if he does not know what he cannot know and consequently, the concept of rationality must be modified. The theory of Rational Beliefs (in short, RB due to Kurz (1994), (1997a)) defines an agent to be *rational* if his model cannot be falsified by the data and if simulated, it reproduces the empirical distribution. In this paper we use only the most basic restrictions of this theory which are noted later, and in Section 1.3.1 we review all rationality conditions this theory imposes on our model. For the moment we note that under this theory, without a known “true” model any meaningful concept of belief rationality must embrace a wide collection of models without resorting to psychological or behavior principles to explain such diversity. This conclusion raises a clear methodological question. In



formulating an asset pricing theory should we describe in detail the subjective models of each agent in the model? With diverse agents 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 all state variables. Once these are specified, the Euler equations are fully specified and market clearing leads to equilibrium pricing. In the rest of this section we explain this methodology.

We observe that in markets with heterogeneous beliefs agents are willing to reveal their forecasts. Hence, in formulating our theory we now assume that market forecast data are public. The crucial difference between markets with and without private information is that when an individual's forecasts of a state variable are revealed in a market *without* private information, others do not see such forecasts as a source of new data and do not update their own beliefs about any parameter to forecast state variables. In such a market, a forecaster uses knowledge about the forecasts of state variables by other forecasters only *to alter his forecasts of endogenous variables* since these depend upon the market belief. In short, the difference between an equilibrium with asymmetric private information and an equilibrium without private information but with heterogeneous beliefs is that in the latter *agents do not learn from others and do not update their beliefs about state variables based on the opinions of others* (for details see Kurz (2006)). But then, how do we describe the individual and market beliefs?

The key analytical step taken (see Kurz (1994), Nielsen (1996), Kurz (1997a), 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 adapt the ideas of Kurz, Jin and Motolese (2005a), (2005b) to the problem studied in this paper and outline now this adaptation.

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. Note that this implies that personal state variables and the economy-wide state variables are not necessarily the same. A personal state of belief is analogous to any other state variables in the agent's decision problem although it can also be interpreted as defining the more familiar concept of a "type" of that agent. At date  $t$  he is not certain of his future belief type but his behavior model (e.g. Bayesian updating) and interpretation of current information may determine the dynamics of his personal state of belief. The *distribution* of individual states of belief is then an

economy-wide state variable whose moments play an important role in the work below. As we indicated, the crucial fact is that the distribution of beliefs in the market is observable. In equilibrium, endogenous variables (e.g. prices) depend upon the economy's state variables, but in a large economy an agent's "*anonymity*" implies that a personal state of belief has a negligible effect on prices. Thus, as in any equilibrium, endogenous variables are functions of the economy's state variables and here these state variables include the distribution of personal beliefs. Hence, all moments of this distribution could matter in equilibrium. 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. Finally, since endogenous variables are functions of the market beliefs, endogenous variables are forecasted by forecasting the market distribution of beliefs using the known equilibrium map. Hence, to forecast future endogenous variables an agent must forecast the beliefs of others.

A simple principle of rationality implies that an individual state of belief *cannot be a constant* unless an agent believes the stationary measure (3) is the truth and consequently the issue discussed in this section is the dynamics of individual beliefs. To see this argument suppose agents hold diverse beliefs which are different from (3). If an agent 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 is an *empirical proof* that his belief is irrational. Clearly, just being wrong is not the real issue. Rational agents hold wrong beliefs most of the time when there is no empirical proof that they are wrong. To see why note that when rational agents hold diverse probability models while there is only one true law of dynamic motion then most are wrong most of the time. Hence the average market forecasting model is often wrong. This is actually the essence of the market risk we call "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  $\ell$  knows his own  $g_t^\ell$  and the market *distribution* of  $g_t^i$  across  $i$ . In addition he observes past distributions of the  $g_\tau^i$  for all  $\tau < t$  hence he knows past values of all moments of the distributions of  $g_\tau^i$ . We specify the dynamics of  $g_t^i$  by

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

where  $\rho_t^{ig}$  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 Section 1.3 we deduce (4) as a consequence of a Bayesian updating procedure. To motivate our approach we note that  $g_t^i$  is used to

express an agent's assessment of the *difference* between date  $t$  forecast of an observable state variable and the forecast under the empirical distribution  $m$ . In our model agent  $i$ 's perception of date  $t$  distribution of  $d_{t+1}$  (denoted by  $d_{t+1}^i$ ) is described by using the belief state  $g_t^i$  as follows

$$(5a) \quad d_{t+1}^i = \lambda_d d_t + \lambda_d^g g_t^i + \rho_t^{id}, \quad \rho_t^{id} \sim N(0, \hat{\sigma}_d^2).$$

The assumption that  $\hat{\sigma}_Q^2$  is the same for all agents is made for simplicity. It follows that the state of belief  $g_t^i$  measures the deviation of his forecast from the empirical stationary forecast

$$(5b) \quad E^i[d_{t+1}^i | H_t, g_t^i] - E^m[d_{t+1} | H_t] = \lambda_d^g g_t^i.$$

Indeed, (5b) shows how to measure  $g_t^i$  in practice. For a state variable  $X_t$ , data on  $i$ 's forecasts of  $X_t$  (in (5b) it is  $d_t$ ) are measured by  $E^i[X_t^i | H_t, g_t^i]$ . One then uses standard econometric techniques to construct the stationary forecast  $E^m[X_t | H_t]$  with which one constructs the difference in (5b). This construction and the data it generates are the basis of the work of Fan (2006). An agent type 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 belief heterogeneity results from economic dynamic non-stationarity, it should be clear that in 1900 subjective assessments  $g_t^i$  were related to the development of electricity and the combustion engine, while in 2000  $g_t^i$  measured the impact of computers and information technology. Hence, success or failures of past  $g_t^i$  do not tell you anything about what present day  $g_t^i$  should be. We now deduce (4) from Bayesian principles.

### 1.3 Deducing (4) $g_{t+1}^i = \lambda_Z g_t^i + \rho_{t+1}^{ig}$ from Bayesian Updating Procedure

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*. The description (4)  $g_{t+1}^i = \lambda_Z g_t^i + \rho_{t+1}^{ig}$  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 requires each agent to have a distinct state space to describe his uncertainty and dictates the endogenous expansion of the economy wide state space for equilibrium pricing. However, we now explore the conditions under which the Markov dynamics (4) is a consequence of elementary principles of Bayesian inference.

In a standard environment of Bayesian learning an agent faces data generated under an unknown 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 one must use the prior to forecast the variable and learning can only improve *future* forecasts of that variable. Our model has some parameters fixed and others that change over time. The fixed parameters are known since 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  $\Pi$  the value  $d_t$  has a transition function of the form

$$(6) \quad d_{t+1} - \lambda_d d_t = b_t + \rho_t^d.$$

The sequence of parameters  $b_t$  is an exogenous, time varying mean value function. Agents know  $\lambda_d$  but not the sequence  $b_t$ . This formulation includes economies with slow changing regimes, each lasting a long time. Regimes may change rapidly or slowly but the mere fact that they change limits the validity of Bayesian updating. To understand this fact 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 one possibility. They would, however, seek any additional information and use other subjective interpretation of public data to arrive at alternative subjective estimates of  $b_{t+1}$  to supplement the Bayesian procedure. 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 the basis of all subjective interpretation of data.

### 1.3.1 *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 and multitude of reports about 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. Generally, qualitative information cannot 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 formal modeling of deduction from qualitative information. Saari (2006) uses qualitative information in the context of a dynamical model of market shares. The model traces out the equilibrium dynamics of each firm’s market share where qualitative information is represented by the derivative of a firm’s response function at specified points. Such derivatives at isolated points in the space provide a rational player an indication of possible future dynamical evolutions which are consistent with the given derivatives. For an additional application see Toukan (2006).

Here we adopt a very simple formalization of the use of qualitative information. Thus, we consider all qualitative information as *statements about the future*. A statement may turn out to be true or false. Denote date  $t$  statements by  $(C_{t1}, C_{t2}, \dots, C_{tK_t})$ . The list changes over time hence  $K_t$  varies with  $t$ . These may offer *contradictory* perspectives in the sense that if, say,  $C_{t1}$  materializes it would imply bright prospects for  $d_{t+1}$  while  $C_{t2}$  may lead to a negative assessment of  $d_{t+1}$ . A realization at  $t+1$  is a vector  $\varphi_{t+1} = (\varphi_{t+1,1}, \varphi_{t+1,2}, \dots, \varphi_{t+1,K_t})$  of numbers which are 0 or 1: 0 means the statement turns out to be false and 1 means it is true. There are  $2^{K_t}$  possible outcomes, denoted  $\varphi_{t+1}(k)$ ,  $k = 1, 2, \dots, 2^{K_t}$ . We now introduce a subjective map from  $\varphi_{t+1}$  to valuation  $\Phi^i(\varphi_{t+1})$ . Each is a quantitative evaluation by agent  $i$  of the effect of each possible outcome on  $(d_{t+1} - \lambda_d d_t)$ . This is an independent estimate by agent  $i$  on how different  $d_{t+1}$  is from the stationary forecast. Finally, agent  $i$  attaches probabilities  $(a_1^i, a_2^i, \dots, a_{2^{K_t}}^i)$  to each of the qualitative outcomes. This procedure results in agent  $i$  making an alternate subjective estimate of  $(d_{t+1} - \lambda_d d_t)$  based only on the qualitative data at his disposal:

$$(7) \quad \Psi_t^i = \sum_{k=1}^{2^{K_t}} a_k^i \Phi^i(\varphi_{t+1}(k)).$$

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

### 1.3.2 A Bayesian Model: Beliefs are Markov State Variable

As assumed earlier, under the true probability  $\Pi$ , the value  $d_{t+1}$  has a true transition of a form

$$(6a) \quad d_{t+1} - \lambda_d d_t = b_t + \rho_{t+1}^d, \quad \rho_{t+1}^d \sim N(0, \frac{1}{\beta}).$$

$b_t$  is a time varying mean value. Agents do not know  $b_t$  and at first decision date  $t$  (say,  $t = 1$ ) they have two pieces of information. They know  $d_t$  and observe qualitative information  $(C_{(t)1}, C_{(t)2}, \dots, C_{(t)K_t})$  with which they assess  $\Psi_t^i$ . Without  $\Psi_t^i$  the prior mean of the agent at  $t = 1$  is  $b$  but to start the process he uses both sources to form a prior belief  $E_t^i(b_t | d_t, \Psi_t^i)$  about  $b_t$  (used to forecast  $d_{t+1}$ ). However, the changing parameter  $b_t$  leads to a problem. When  $d_{t+1} - \lambda_d d_t$  is observed agent  $i$  updates his belief to  $E_t^i(b_t | d_{t+1}, \Psi_t^i)$ . But agent  $i$  needs an *estimate of  $b_{t+1}$* , not of  $b_t$ . Hence, how does he go from  $E_t^i(b_t | d_t, \Psi_t^i)$  to a new prior of  $b_{t+1}$ ? Without any new information his prior belief of  $b_{t+1}$  will remain the same and he simply takes  $E_t^i(b_t | d_{t+1}, \Psi_t^i)$  as the new prior belief of  $b_{t+1}$ . This is particularly true if the  $b$ 's change very slowly. Indeed, since Bayesian learning draws its inference from the past, it cannot offer a method of updating one's belief about a future value of a changing sequence of parameters. To that end the agent uses the public qualitative information  $(C_{(t+1)1}, C_{(t+1)2}, \dots, C_{(t+1)K_{t+1}})$  released at the start of date  $t+1$  but before trading. These lead to subjective measures  $\Psi_{t+1}^i$  which are, in fact, alternate estimates of  $b_{t+1}$ . Now our agent has two independent sources for belief about  $b_{t+1}$ :  $E_t^i(b_t | d_{t+1}, \Psi_t^i)$  and  $\Psi_{t+1}^i$  which must be reconciled. Under a Bayesian approach we thus assume:

**Assumption (\*):** Agent  $i$  uses a subjective probability  $\mu$  to form date  $t+1$  prior belief which is then

$$(8a) \quad E_{t+1}^i(b_{t+1} | d_{t+1}, \Psi_{t+1}^i) = \mu E_t^i(b_t | d_{t+1}, \Psi_t^i) + (1 - \mu) \Psi_{t+1}^i \quad 0 < \mu < 1.$$

For consistency, if  $\Psi_1^i$  is believed to be Normal then at the initial date  $t=1$  the prior must be

$$(8b) \quad b_1 \sim N(\mu b + (1 - \mu) \Psi_1^i, \frac{1}{\alpha}).$$

This assumption is the new element that permits  $E_{t+1}^i(b_t | d_{t+1}, \Psi_t^i)$  to be upgraded into a prior belief at date  $t+1$ ,  $E_t^i(b_{t+1} | d_{t+1}, \Psi_{t+1}^i)$ , before  $d_{t+2}$  is observed. We can now show the following:

**Theorem 1:** Suppose  $\Psi_t^i \sim N(0, \frac{1}{\gamma})$ , i.i.d. and Assumption (\*) holds. Then for large values of  $t$ , the posterior  $E_{t+1}^i(b_{t+1} | d_{t+1}, \Psi_{t+1}^i)$  is a Markov state variable such that if we define  $g_t^i = E_t^i(b_t | d_t, \Psi_t^i)$  and  $\mu = \lambda_Z$  then the dynamics (4) holds: (8a) implies (4).

**Proof:** Pick a starting date  $t = 1$  when data  $d_t$  is known and the agent generates a subjective measure of qualitative data  $\Psi_t^i$ . He then forms a prior on  $b_t$ , assumed to be  $b_t \sim N(\mu b + (1 - \mu)\Psi_t^i, \frac{1}{\alpha})$ . Now we

move on to  $t+1$  and  $d_{t+1}$  is observed. The agent updates the prior in a standard Bayesian manner:

$$(9) \quad E_{t+1}^i(b_t | d_{t+1}, \Psi_t^i) = \frac{\alpha(\mu b + (1 - \mu)\Psi_t^i) + \beta[d_{t+1} - \lambda_d d_t]}{\alpha + \beta}, \quad 0 \leq \mu \leq 1.$$

But before date  $t+1$  trading he generates the subjective measure  $\Psi_{t+1}^i$  of qualitative data. By the

Assumption (8a) the expected parameter  $b_{t+1}$  under the new prior at  $t+1$  is

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

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 updates his belief to

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

Before the start of date  $t+2$  trading the agent generates a new value  $\Psi_{t+2}^i$  leading to  $t+2$  belief that

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

When  $d_{t+3} - \lambda_d d_{t+2}$  is observed the updated belief 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  $\Psi_{t+3}^i$  leading to  $t+3$  belief  $E_{t+3}^i(b_{t+3} | d_{t+3}, \Psi_{t+3}^i)$ . By induction we iterate forward to conclude that

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

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

$$E_{t+1}^i(b_{t+1}|d_{t+2}, \Psi_{t+1}^i) = \mu E_t^i(b_t|d_{t+1}, \Psi_t^i) + (1-\mu)\Psi_{t+1}^i$$

But by definition we have

$$(10) \quad E_{t+1}^i(b_{t+1}|d_{t+1}, \Psi_{t+1}^i) = \mu E_t^i(b_t|d_{t+1}, \Psi_t^i) + (1-\mu)\Psi_{t+1}^i$$

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

$$E_t^i(b_t|d_{t+1}, \Psi_t^i) = E_t^i(b_t|d_t, \Psi_t^i).$$

Inserting this last equation in (10) we finally have the desired conclusion for large  $t$

$$(11) \quad E_{t+1}^i(b_{t+1}|d_{t+1}, \Psi_{t+1}^i) = \mu E_t^i(b_t|d_t, \Psi_t^i) + (1-\mu)\Psi_{t+1}^i.$$

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  $\mu = \lambda_Z$  to see that (11) is actually (4). ■

Theorem 1 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 (11). The posterior fluctuates forever, providing the foundations for the dynamics of market belief but the fluctuations follow a simple Markov transition. New data  $d_t$  and  $\Psi_t^i$  alter the conditional probability of the agent, but these do not change the dynamic law of motion of  $g_t^i$ .

## 1.4 Modeling Heterogeneity of belief II: Market Belief and Rationality

### 1.4.1 Individual and Market Beliefs

Denote by  $Z_t$  the first moment of the cross sectional distribution of the  $g_t^i$  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  $\rho_{t+1}^{ig}$  over  $i$  does not vanish. We write it in the form

$$(12) \quad Z_{t+1} = \lambda_Z Z_t + \rho_{t+1}^Z.$$

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

$$(13a) \quad d_{t+1} = \lambda_d d_t + \rho_{t+1}^d \quad \left( \begin{array}{c} \rho_{t+1}^d \\ \rho_{t+1}^Z \end{array} \right) \sim N \left( \begin{array}{c} 0 \\ 0 \end{array}, \begin{array}{cc} \left[ \begin{array}{cc} \sigma_d^2 & 0 \\ 0 & \sigma_Z^2 \end{array} \right] & = \tilde{\Sigma} \end{array} \right), \text{ i.i.d.}$$

Now, an agent who does not believe that (13a)-(13b) is the truth, formulates his own model\belief. We



have seen in (5a) 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

$$(14a) \quad d_{t+1}^i = \lambda_d d_t + \lambda_d^g g_t^i + \rho_{t+1}^{id}$$

$$(14b) \quad Z_{t+1}^i = \lambda_Z Z_t + \lambda_Z^g g_t^i + \rho_{t+1}^{iZ}$$

$$(14c) \quad g_{t+1}^i = \lambda_Z g_t^i + \rho_{t+1}^{ig}$$

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

Although the state variable  $g_t^i$  defines belief about future value of  $d_{t+1}$ , (14a)-(14b) 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. Hence,  $g_t^i$  expresses how the agent considers the present conditions to be different from 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 (14c) it is

$$(14d) \quad \bar{E}_t \begin{pmatrix} d_{t+1} \\ Z_{t+1} \end{pmatrix} - E_t^m \begin{pmatrix} d_{t+1} \\ Z_{t+1} \end{pmatrix} = \begin{pmatrix} \lambda_d^g Z_t \\ \lambda_Z^g Z_t \end{pmatrix}.$$

**Higher Order Beliefs.** One must distinguish between higher order belief which are temporal and those which are contemporaneous. Within our theory the system (14a)-(14c) 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 (14a)-(14c) statement like

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

It is thus clear that temporal higher order beliefs are simple properties of conditional expectations. In addition, since (12) is implied by (14c) 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 (14a)-(14c) 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  $((D \times Z \times G^i)^\infty, \mathcal{F}^i)$  where  $\mathcal{F}^i$  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:

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

**Proof:** 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

$$(15a) \quad \bar{E}_t(d_{t+2}) = \lambda_d^2 d_t + \lambda_d^g (\lambda_d + \lambda_Z) Z_t.$$

On the other hand we have from (14a) that

$$\bar{E}_{t+1}(d_{t+2}) = \lambda_d d_{t+1} + \lambda_d^g Z_{t+1}$$

hence

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

Aggregating now we conclude that

$$(15b) \quad \bar{E}_t \bar{E}_{t+1}(d_{t+2}) = \lambda_d^2 d_t + \lambda_d^g (\lambda_d + \lambda_Z + \lambda_Z^g) Z_t.$$

Comparison of (15a) and (15b) shows that  $\bar{E}_t(d_{t+2}) \neq \bar{E}_t \bar{E}_{t+1}(d_{t+2})$ . ■

Contemporaneous higher order beliefs occur naturally in games but not in markets, despite the common and false interpretation of the Keynesian Beauty Contest. To explain the issue observe that, formally speaking, we could incorporate such higher order beliefs in our theory by incorporating belief variables about future distributions of market belief variables. For example, in (14a)-(14b) we could have introduced a *second and separate* variable  $g_t^{i2}$  to express belief about future values of the market belief  $Z$  (which would become  $Z^1$ ). This triggers an infinite regress since the average of  $g_t^{i2}$  is an aggregate market belief  $Z_t^2$  hence we need to introduce  $g_t^{i3}$ , a belief about  $Z_t^2$ , and hence proceed to all higher order beliefs. We did not introduce such a structure for two reasons. First, simplicity is a virtue and infinite number of transitions are not tractable. A much deeper reason is that all higher order market beliefs  $Z_t^j$ , for  $j > 1$  are degenerate. This is so since they are averages of  $g_t^{ij}$  and since for  $j > 1$  the  $Z^j$  are not observable, there is no mechanism for the individual  $g_t^{ij}$  to be correlated as in (12). With independent  $g_t^{ij}$  the averages, defined to be  $Z_t^j$  for  $j > 1$ , are zero and hence degenerate variables.

**Belief and Information: Understanding what  $Z_t$  is.** From the perspective of an agent,  $Z_t$  is a state variable like any other. News about  $Z_t$  are used to forecast prices and assess market risk premia in the same way macroeconomic data such as GNP growth or Non Farm Payroll are used to assess the risk of a recession. Market belief may be wrong as it forecasts more recessions that occur. Market risk premia may fall just because agents are more optimistic about the future, not necessarily because there is any specific data which convinces everybody the future is bright. But then, how do agents update their beliefs when they observe  $Z_t$ ? In sharp contrast with a theory assuming the presence of private information, agents do not revise their own beliefs about the state variable  $d_{t+1}$ : (14a) specifically *does not* depend upon  $Z_t$ . Agents do not consider  $Z_t$  as information about  $d_{t+1}$  because it is not a “signal” about unobserved private information they do not have. This is the case since they know all use the same available information. However, they do consider  $Z_t$  to be “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 is what we show now.

#### 1.4.1 Rationality: The Theory of Rational Beliefs

We have seen that the market belief is not necessarily rational hence averaging (14a) -(14c) is not required to imply a consistent probability measure. What about individuals? Since they do not know the true probability  $\Pi$ , we assume (14a) -(14c) may not be the truth. But then, can we rationalize such a belief on its own? That is, what restrictions do (14a) -(14c) 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), (1997a) 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 (1997a),

(1997b), 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  $m$ . In our model beliefs are specified by perception models (14a) -(14c) in which the dynamics of  $g_t^i$  expresses the subjective belief of an agent. For (14a) -(14c) to be RB it needs to induce the same empirical distribution as (13a)-(13b). But this amounts to the requirement that

$$(16) \quad \text{The empirical distribution of } \begin{pmatrix} \lambda_d^g g_t^i + \rho_t^{id} \\ \lambda_Z^g g_t^i + \rho_{t+1}^{iz} \end{pmatrix} = \text{the distribution of } \begin{pmatrix} \rho_t^d \\ \rho_{t+1}^z \end{pmatrix} \sim N \left( 0, \begin{bmatrix} \sigma_d^2 & 0 \\ 0 & \sigma_Z^2 \end{bmatrix} \right), \text{ i.i.d.}$$

To compute the implied data generated by the model, one treats the  $g_t^i$  symmetrically with other random variables. From (14c), the unconditional variance of  $g_t^i$  is  $\text{Var}(g^i) = \sigma_g^2 / (1 - \lambda_Z^2)$ . Hence, we have the following rationality conditions which follow from (16):

$$\begin{aligned} \text{(i)} \quad & \frac{(\lambda_d^g)^2 \sigma_g^2}{1 - \lambda_Z^2} + \hat{\sigma}_d^2 = \sigma_d^2 & \text{(ii)} \quad & \frac{(\lambda_Z^g)^2 \sigma_g^2}{1 - \lambda_Z^2} + \hat{\sigma}_Z^2 = \sigma_Z^2 & \text{(iii)} \quad & \frac{\lambda_d^g \lambda_Z^g \sigma_g^2}{1 - \lambda_Z^2} + \hat{\sigma}_{Zd} = 0 \\ \text{(iv)} \quad & \frac{(\lambda_d^g)^2 \lambda_Z \sigma_g^2}{1 - \lambda_Z^2} + \text{Cov}(\hat{\rho}_t^{id}, \hat{\rho}_{t-1}^{id}) = 0 & \text{(v)} \quad & \frac{(\lambda_Z^g)^2 \lambda_Z \sigma_g^2}{1 - \lambda_Z^2} + \text{Cov}(\hat{\rho}_t^{iz}, \hat{\rho}_{t-1}^{iz}) = 0. \end{aligned}$$

The first three conditions pin down the covariance matrix in (14a)-(14c). The last two pin down the serial correlation of the two terms  $(\hat{\rho}_t^{id}, \hat{\rho}_t^{iz})$ . An inspection of (14a)-(14c) reveals the only choice left for an agent are the two free parameters  $(\lambda_d^g, \lambda_Z^g)$ . But under the RB theory these are not free either since their natural conditions they must satisfy. First,  $\hat{\sigma}_d^2 > 0$ ,  $\hat{\sigma}_Z^2 > 0$  place two strict conditions on  $(\lambda_d^g, \lambda_Z^g)$ :

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

Finally, to ensure the covariance matrix in (14a)-(14c) is positive definite one must impose an additional condition. The condition

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

is sufficient. Hence the "free" parameters  $(\lambda_d^g, \lambda_Z^g)$  are restricted to a rather narrow range. Since these are

the restrictions on the parameters of (14a)-(14c) they are empirically testable.

### 1.5 Combining the Elements: the Implied Asset Pricing Under Diverse Beliefs

We now derive equilibrium prices and the risk premium. Details about the value function and the demand functions are provided in an Appendix where we explain the term “adjusted” conditional variance of  $Q_{t+1}$ , denoted below by  $\hat{\sigma}_Q^2$ . We first specify the state variables of agent  $i$ 's demand function (2) by  $\psi_t^i = (1, d_t, Z_t, g_t^i)$  and write this function as

$$(17) \quad \theta_t^i(p_t) = \frac{\tau}{R \hat{\sigma}_Q^2} [E_t^i(Q_{t+1}) + u \psi_t^i], \quad u = (u_0, u_1, u_2, u_3), \quad \psi_t^i = (1, d_t, Z_t, g_t^i).$$

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:

**Stability Conditions:** We require that  $0 < \lambda_d < 1$ ,  $0 < \lambda_Z + \lambda_Z^g < 1$ .

The first requires  $\{d_t, t = 1, 2, \dots\}$  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 (14b), average over the population and recall that  $Z_t$  are market averages of the  $g_t^i$ . This implies that

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

**Theorem 3:** Consider the model with heterogenous beliefs under the stability conditions specified with supply of shares equals 1. Then there is a unique equilibrium price function which takes the form

$$p_t = a_d d_t + a_z Z_t + P_0.$$

**Proof:** Average (17) to have

$$(18) \quad \frac{R \hat{\sigma}_Q^2}{\tau} = [\bar{E}_t(p_{t+1} + d_{t+1} + \mu) - R p_t + (u_0 + u_1 d_t + (u_2 + u_3) Z_t)].$$

Now use the perception models (14a)-(14b) 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*

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

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

Using these to solve for date t price we deduce

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

(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*, page 1308) shows that the solution is

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

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

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

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

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

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 entirely reasonable conclusions: *today's price of the asset increases if  $d_t$  rises and today's price of the asset is higher when  $Z_t$ - the present day market belief in higher future dividend rate  $d_{t+1}$ - increases.*

## 1.6 Equilibrium Risk Premium

### 1.6.1 The Main Equilibrium Results

We now explore the often misunderstood problem of measuring market risk premium under heterogenous beliefs. We shall see in a moment that there are many different concepts involved here and the main issue is one of choosing the concept which is most appropriate to any application. The definition of the realized risk premium on a long position, as a random variable, is clearly defined by

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

(21) is a random variable measuring the actual excess return of stocks over the riskless bond. The need is to measure the premium as *a known expected quantity* which is recognized by market participants. We have three such measures. The first is the subjective expected excess returns by agent i which is

computed from (18) to be

$$(22a) \quad \frac{1}{p_t} E_t^i(p_{t+1} + d_{t+1} + \mu - Rp_t) = \frac{1}{p_t} \left[ R \frac{\hat{\sigma}_Q^2}{\tau} - (u_0 + u_1 d_t + u_2 Z_t + u_3 g_t^i) \right].$$

Alternatively, we use the equilibrium map (20a) and the perception model (14a)-(14c) to show that

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

Diverse beliefs implies the perceived premia vary across agents and this diversity is the crucial cause of trade and volatility. Aggregating over  $i$  we define the market premium as the average market expected excess returns. This perceived market premium reflects what the market expects, not necessarily what the market gets. From (22a)-(22b) we deduce that it is measured by

$$(23) \quad \begin{aligned} \frac{1}{p_t} \bar{E}_t(p_{t+1} + d_{t+1} + \mu - Rp_t) &= \frac{1}{p_t} \left[ R \frac{\hat{\sigma}_Q^2}{\tau} - u_0 - u_1 d_t - (u_2 + u_3) Z_t \right] \\ &= \frac{1}{p_t} [(a_d + 1)(\lambda_d d_t + \lambda_d^g Z_t) + a_z (\lambda_z Z_t + \lambda_z^g Z_t) + \mu + P_0 - Rp_t]. \end{aligned}$$

Neither the individual perceived premium nor the market perceived premium are necessarily correct. Hence we focus on the third premium which is an objective measure, common to all agents. Agents who study the long term time variability of the premium would measure it by the empirical distribution of (21). Using the equilibrium map (18) and the stationary transition (13a)-(13b) we have

$$(24) \quad \begin{aligned} E_t^m[\pi_{t+1}] &= \frac{1}{p_t} E_t^m[p_{t+1} + d_{t+1} + \mu - Rp_t] \\ &= \frac{1}{p_t} [(a_d + 1)(\lambda_d d_t) + a_z (\lambda_z) Z_t] + \mu + P_0 - Rp_t. \end{aligned}$$

We stress that (24) is the way researchers cited above have measured the risk premium and therefore we refer to it as “the” risk premium.

We thus arrive at two important conclusions. First, the differences between the individual perceived premium and the market perceived premium is

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

It is thus clear that from the perspective of trading, all that matters is the difference  $g_t^i - Z_t$  of individual from market belief. In addition, the following difference is important

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

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 arrive at an analytical expression of the risk premium:

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

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

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

Conclusions (26a) -(26b) are, perhaps, the most important results of this paper. (26a) and the earlier results exhibit the Endogenous Uncertainty component of the risk premium (see Kurz (1997a)) which we call “The Market Belief Risk Premium.” It shows that market risk premia inherently depend upon market belief. 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 the adjusted variance follows from the existence of weights  $(\omega_1, \omega_{12}, \omega_2)$  such that

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

Hence, the direct effect of belief is in  $\rho_{t+1}^{iz}$ . Changed volatility of the market belief changes the volatility of excess return with a direct impact on 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  $Z_t > 0$  which is very revealing.

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 excess return *is negative*. This sign is surprising since when  $Z_t > 0$  the market expects *above normal* future dividend and in that case the risk premium on the stock *declines*. When the market holds bearish belief about future dividend ( $Z_t < 0$ ) the risk premium *rises*. The importance of this result is not only related to its theoretical implications but also because it is empirically testable since we have data on  $Z_t$  and we shall test this result empirically. However, before we proceed to discuss the empirical test it would be useful to discuss some ramifications of this result.



### 1.6.2 The Market Belief Risk Premium is Fully General

The main result (26b), says the market belief  $Z_t$  has a negative effect on the risk premium. It was derived from the assumed exponential utility function. We now argue that this result is far more general and depends only on the positive coefficient  $a_z$  of  $Z_t$  in the price map. To show this, suppose we assume any additive utility function over consumption and a risky asset which pays a “dividend” or any other random payoff denoted by  $d_t$ . 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 on the numerator only  $E_t^m[p_{t+1} + d_{t+1} + \mu - Rp_t]$  we linearize the price map around 0 and write  $p_t = \Phi_d d_t + \Phi_Z Z_t + \Phi_0$ . We now show that the desired result depends only upon the condition  $\Phi_Z > 0$ . This condition is entirely reasonable as it requires the current price to increase if the market is more optimistic about the asset’s future payoffs. To prove the point note that

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

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.

### 1.6.3 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, fluctuations in risk premia are inversely related to the degree of market optimism about future prospects of asset payoffs.

To further explore the main result, it is important to explain *what it does not say*. One may interpret our result as confirming a common claim that in order to maximize excess returns it is an optimal strategy to be a “contrarian” to the market consensus by betting against it. To understand why this is a false interpretation of our result note that when an agent holds a belief about future payments, the market belief does not offer any new information to alter the individual’s belief about the exogenous

variable. If the agent believes that future dividends will be abnormally high but  $Z_t < 0$ , the agent does not change his forecast of  $d_{t+1}$ . He uses the market belief information only to forecast future prices of an asset. Hence,  $Z_t$  is a crucial input to forecasting returns without changing the forecast of  $d_{t+1}$ . Since given the available information an optimizing agent is already on his demand function, he does not just abandon his demand and becomes a contrarian. Our argument is the same as the one showing why it is not optimal to adopt the log utility as your own utility even though it maximizes the growth rate of your wealth. Yes it does that but you dislike the sharp declines which are expected in the value of your assets. By analogy, following a “contrarian” policy may imply a high long run average return in accord with the empirical probability  $m$ . However, if you disagree with this probability you will dislike being short when your optimal position should be long. Indeed, this argument explains the fact that *most people hold positions which are in agreement with the market belief most of the time* instead of betting against it as a “contrarian” strategy would dictate. Taking a positive perspective, our result shows that if your belief leads to an optimal long position in an asset, the market value of  $Z_t$  will enable you to make a more precise estimate of your excess returns. Even if  $Z_t < 0$  and you disagree with the market, you may still maintain your long position but alter your estimated risk premium. The crucial observation we make is that a maximizing agent has his own belief about future events, and he does not select a new belief when he learns the market belief. From his point of view the market belief is an important state variable used to forecast future prices just like other state variables such as Non Farm Payroll which also changes the risk premium on investments in the bond and stock markets. We turn 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**

Our basic data 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, financial institutions, and consulting firms. 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 24<sup>th</sup> and 25<sup>th</sup> 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 that 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  $h = 0, 1, \dots, 4$  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  $h = 0, 1, \dots, 5$  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 estimation of the effect of market belief on risk premia we use the *month* as a unit of time. Hence, our first task was to translate quarterly mean market 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 available on a monthly basis then all forecasters actually know at each date  $t$  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  $h=1, 2, \dots, 12$ .

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 at the foundation of the theory developed in Section 1. These 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

$$Z_t^{(X,h,i)} = E_t^i\{X_{t+h}\} - E_t^m\{X_{t+h}\}^5.$$

The subtraction of  $E_t^m\{X_{t+h}\}$  ensures the state of belief is  $m$ -orthogonal to information in the market at  $t$ . Also, since  $Z_t^{(X,h,i)}$  is the deviation from the stationary forecast, it must be interpreted properly. For example suppose  $y$  is the 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 under the stationary forecast. Since here we study risk premia, *all belief variables examined in this paper are beliefs about future interest rates*. 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 of economic conditions which are different at  $t$  from past 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,

---

<sup>5</sup> A similar definition was used by Fan (2006) who constructed similar belief data for the study of the dynamics of the term structure.

$Z_t^{(y,h)} > 0$  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.

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<sup>6</sup> 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. Series of real variables are transformed by taking the monthly first difference of their logarithms. Time series of prices are defined to be the second difference of the logarithms of the initial prices. Because of missing data we use (see Stock and Watson (2005)) only 126<sup>7</sup> series from 1959:01 to 2003:12. These 126 time series represent nine main categories of economic variables: consumption, employment, exchange rates, housing starts, interest rates, money aggregates, prices, real output, and stock prices. Stacking them, we obtain an information matrix of dimension 540 by 126. One of Stock and Watson's (1999) conclusion is that effective time invariant models *needs 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.

---

<sup>6</sup> The data is publically available on Watson's webpage <http://www.wws.princeton.edu/mwatson/publi.html>

<sup>7</sup> We excluded 4 housing permit time series due to missing data. For natural reasons we also excluded the University of Michigan Consumer Sentiment data.

For the entire period combined, 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 one factor in the stationary forecasting scheme we adopt below. Stock and Watson (2002) concluded that a *combination* of factors and lagged macro variables is the best information set. To keep the number of variables small, our stationary forecasts of interest rates are derived from a linear regression of a future variable on explanatory variables consisting of the following: (i) one factor deduced from date  $t$  information matrix, (ii) lags of date  $t$  and  $t-1$  values of the variable in question, and (iii) dates  $t-1$  and  $t-2$  values of each of the year over year rates of change of *industrial production* and the *CPI*. The use of a factor also ensures that the information includes all spreads in the yield curve since spreads are just linear combinations of interest rates and such spreads receive significant weight in the factor employed.

*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  $Z_t^{(X,h)}$ .

Tables 1A and 1B provide some summary statistics of a sample of extracted market belief variables  $Z_t^{(X,h)}$ <sup>8</sup>. The last column  $\rho_{t-1}$  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 growing evidence against these beliefs.

---

<sup>8</sup> Since the table reports the belief index about GDP growth we remark that to estimate belief variables  $Z_t^{(X,h)}$  about real variables we use stationary forecasts which employ three factors.

**Table 1A: Summary Statistics of Market Beliefs**

h = 6 Months or 2 Quarters Ahead	Time Ave.	Std Dev	Skew	Kurt	$\rho_{t-1}$
Fed Fund rate	0.215	0.451	-0.195	2.554	0.459
1 year T-bill rate	0.177	0.294	0.224	2.661	0.648
CPI	0.080	0.744	0.452	3.573	0.637
Real GDP % ch.	-1.501	0.609	0.002	2.817	0.801
h = 12 Months or 4 Quarters Ahead					
Fed Fund rate	0.343	0.615	0.028	2.304	0.564
1 year T-bill rate	0.337	0.423	0.388	2.651	0.701
CPI	0.060	0.878	0.380	2.970	0.728
Real GDP % ch.	-1.583	0.564	0.370	2.462	0.837

**Table 1B: Correlation Matrix of Market Beliefs**

6 Months or 2 Quarters Ahead	Fed Fund rate	1 year T-bill rate	CPI	Real GDP % ch.
Fed Fund rate	1.000			
1 year T-bill rate	0.896	1.000		
CPI	0.472	0.543	1.000	
Real GDP % ch.	0.033	0.032	-0.077	1.000
12 Months or 4 Quarters Ahead				
Fed Fund rate	1.000			
1 year T-bill rate	0.919	1.000		
CPI	0.707	0.773	1.000	
Real GDP % ch.	-0.331	-0.223	-0.388	1.000

Figure 1 traces the graph of the extracted  $Z_t^{(6,h)}$  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  $-0.5\%$  to  $+1.5\%$ , which are very significant from the economic point of view.

### FIGURE 1

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 2

## 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  $Q_{t+h}^{(j,h)}$  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

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

where  $R_t^{(\tau)}$  is the one period interest rate implied by a zero coupon bond with maturity at  $\tau$ . 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.

*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 be realized at the time of settlement. Hence this is actually 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 calculated as the simple average of the daily averages published by the Federal Reserve Bank of New York. Hence, a trader in this market 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 futures 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<sup>9</sup>; they commit at date  $t$  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

---

<sup>9</sup> 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.



recorded by the exchange and become public information. 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 settlement a seller pays and a buyer receives for each contract the cash amount<sup>10</sup>

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

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

$$Q_{t+h}^{(F,h)} = F_t^{(h)} - R_{t+h}^{(F)}.$$

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.** The presence of serial correlation in the forecast errors is inevitable for a well known reason. Computation of excess returns entail overlapping data and this fact leads us to report, in all 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), and which generalizes the Hansen-Hodrick (1980) method. This correction places full weight over the lags of serial correlation in excess returns. Hence we 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 Returns

In this Section we study and measure of the contribution of market belief to long term forecasting excess returns and hence to market risk premia. More specifically, we test the validity of the

---

<sup>10</sup> 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.

theoretical conclusions (26a)-(26b) about the effect of market belief on the time variability of 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 months, and Federal Funds Futures contracts with holding periods of 1 to 6 months. We thus estimate linear excess return functions of the following general form

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

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 individual state of belief  $Z_t^{(X,h,i)}$  that belief variables are m-orthogonal to all information in  $M_t$ . Since the risk premium is estimated in (27) using the long term statistics, under which  $B_t$  and  $M_t$  are orthogonal, it follows that variables in  $B_t$  add something new which is not in the market data  $M_t$ .

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 section standard deviations of individual belief about X at future date t+h

-  $SZ_t^{(6-F,h)} = -(Z_t^{(6,h)} - Z_t^{(F,h)})$  – date t mean market belief about the slope of the yield curve at t+h.

The first and second 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 (14a)-(14c). All belief variables are oriented so their sign in perception of future asset payoff is *positive*. 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 a debt instrument, a belief which is beneficial to the long position is a belief in lower rather than higher interest rates.

The inclusion of the third variable is motivated by two considerations. First, the risk premia on holding interest bearing assets are interdependent. Hence, even with fixed market belief about the riskiness of three month Bills the risk premium on holding this asset is likely to be affected by beliefs about change in the slope of the yield curve. Second, although we study the holding return of say, a six month bill, a speculative motive to trade the asset earlier could be affected by the possibility of changes

in the slope of the yield curve. Given that the current Federal Funds rate will be a variable in all equations, a steeper yield curve is not beneficial to the long positions in any debt instrument hence again the negative sign in  $-SZ_t^{(6-F,h)}$ .

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 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  $t$ .

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

Up to now the time unit chosen did not matter. However, for the equations in (27) the time over which excess return are measured does matter. Rates of return on holding T Bills are naturally annual rates and hence comparable across different T Bills and horizons. As to Fed Funds futures, we measure total returns on such futures in percentage points for the length of time the contracts are held. Naturally, returns on short duration contracts are typically smaller than returns on long duration contracts. It is then clear that returns on a gamble of buying Fed Funds futures are not entirely comparable with returns on holding an asset with clearly defined holding cost. This lack of comparability should be kept in mind in assessing the results reported below. Tables 2A-2C present the estimates of equation (27) for the three market at hand and for selected horizons. (\*) denotes significance at 10% level and (†) denotes significance at 5% level. We report adjusted  $R^2$ .

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

	Constant	NFP <sub>t-1</sub>	CPI <sub>t-1</sub>	F <sub>t</sub>	$\sigma_t^{(F,h)}$	Z <sub>t</sub> <sup>(F,h)</sup>	SZ <sub>t</sub> <sup>(6-F,h)</sup>	R <sup>2</sup>	Chow Test p-value
h=2	0.009 (0.063)	-0.035 * (0.020)	0.013 (0.033)	-0.001 (0.018)	0.454 (0.277)	-0.134 * (0.078)	-0.063 (0.109)	0.095	0.001
h=4	0.031 (0.121)	-0.147 † (0.043)	0.089 * (0.053)	0.042 (0.041)	-1.006 * (0.511)	-0.582 † (0.126)	-0.821 † (0.231)	0.279	0.001
h=6	-0.173 (0.129)	-0.224 † (0.043)	0.155 * (0.080)	0.115 † (0.047)	-1.164 † (0.483)	-0.936 † (0.224)	-1.674 † (0.562)	0.441	0.000

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

	Constant	NFP <sub>t-1</sub>	CPI <sub>t-1</sub>	F <sub>t</sub>	R <sub>t-1</sub> <sup>F1</sup>	R <sub>t-1</sub> <sup>F2</sup>	R <sub>t-1</sub> <sup>F3</sup>	$\sigma_t^{(3,h)}$	Z <sub>t</sub> <sup>(3,h)</sup>	SZ <sub>t</sub> <sup>(6-F,h)</sup>	R <sup>2</sup>	Chow Test p-value
h=2	1.695 * (0.868)	-0.127 † (0.050)	0.075 (0.060)	-0.182 (0.122)	0.681 * (0.397)	-0.056 (0.087)	-0.088 (0.075)	0.628 (0.504)	-0.536 † (0.176)	0.008 (0.183)	0.186	0.004
h=4	1.785 † (0.665)	-0.181 † (0.032)	0.071 (0.050)	-0.132 (0.104)	0.716 † (0.319)	-0.030 (0.068)	-0.071 (0.058)	-0.662 (0.421)	-0.390 † (0.138)	-0.213 (0.157)	0.288	0.013
h=6	1.244 † (0.503)	-0.160 † (0.028)	0.111 † (0.051)	-0.081 (0.076)	0.567 † (0.237)	0.057 (0.060)	-0.090 * (0.052)	-0.786 † (0.270)	-0.571 † (0.146)	-0.684 † (0.199)	0.461	0.035
h=8	1.891 † (0.578)	-0.146 † (0.030)	0.093 † (0.042)	-0.162 † (0.077)	0.885 † (0.272)	0.059 (0.058)	-0.039 (0.047)	-0.499 † (0.220)	-0.470 † (0.109)	-0.430 † (0.161)	0.553	0.025
h=10	1.571 † (0.486)	-0.155 † (0.026)	0.068 † (0.033)	-0.048 (0.066)	0.690 † (0.248)	-0.019 (0.037)	-0.034 (0.041)	-0.779 † (0.211)	-0.369 † (0.104)	-0.323 † (0.150)	0.643	0.048
h=12	1.609 † (0.568)	-0.174 † (0.023)	0.072 † (0.032)	-0.058 (0.084)	0.659 † (0.292)	-0.051 (0.037)	-0.018 (0.030)	-0.623 † (0.066)	-0.257 † (0.081)	-0.135 (0.111)	0.666	0.005

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

	Constant	NFP <sub>t-1</sub>	CPI <sub>t-1</sub>	F <sub>t</sub>	R <sub>t-1</sub> <sup>F1</sup>	R <sub>t-1</sub> <sup>F2</sup>	R <sub>t-1</sub> <sup>F3</sup>	$\sigma_t^{(6,h)}$	Z <sub>t</sub> <sup>(6,h)</sup>	SZ <sub>t</sub> <sup>(6-F,h)</sup>	R <sup>2</sup>	Chow Test p-value
h=2	3.525 * (1.902)	-0.353 † (0.105)	0.108 (0.125)	-0.361 (0.273)	1.510 * (0.879)	-0.100 (0.191)	-0.121 (0.161)	1.705 (1.221)	-0.945 † (0.431)	-0.021 (0.416)	0.179	0.144
h=4	3.766 † (1.474)	-0.445 † (0.079)	0.122 (0.113)	-0.280 (0.218)	1.567 † (0.690)	-0.064 (0.121)	-0.100 (0.110)	-0.850 (0.919)	-0.856 † (0.317)	-0.522 (0.354)	0.317	0.215
h=6	2.717 † (1.185)	-0.338 † (0.079)	0.235 † (0.095)	-0.145 (0.162)	1.197 † (0.546)	0.036 (0.114)	-0.167 (0.123)	-2.430 † (0.548)	-1.504 † (0.290)	-1.481 † (0.411)	0.521	0.006
h=8	3.388 † (1.358)	-0.375 † (0.076)	0.125 (0.086)	-0.150 (0.188)	1.509 † (0.683)	-0.047 (0.104)	-0.059 (0.099)	-1.424 † (0.503)	-0.913 † (0.221)	-0.692 † (0.306)	0.586	0.003
h=10	3.858 † (1.234)	-0.365 † (0.065)	0.121 * (0.063)	-0.195 (0.171)	1.677 † (0.633)	-0.100 (0.078)	-0.059 (0.080)	-1.368 † (0.493)	-0.708 † (0.206)	-0.505 (0.332)	0.644	0.003
h=12	3.992 † (1.166)	-0.371 † (0.058)	0.129 † (0.064)	-0.233 (0.162)	1.648 † (0.588)	-0.140 * (0.073)	-0.001 (0.057)	-1.094 † (0.277)	-0.520 † (0.196)	-0.157 (0.267)	0.667	0.010

### 3.2 Evaluating the Results

Looking at Tables (2A)-(2C) together we find that the pro-cyclical variable NFP advocated by Piazzesi and Swanson (2004), and the yield variables used by Cochrane and Piazzesi (2005) are, indeed, important for long term forecasts of excess returns. We note however that only the first factor of past yields is consistently significant. The central question is the size and sign of the belief variables. We find that the effect of market belief is large, significant 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 are partly responsible for variability of risk premia but variations in market perceptions, which may express mistaken forecasts of future interest rates, are equally important. The data supports the Market Risk Premium hypothesis in (26b). Keeping in mind the orientation convention the data reveals the parameters of the mean market beliefs  $(-Z_t^{(X,h)}, -SZ_t^{(6-F,h)})$  in Tables 2A-2C are *always negative* and are key contributors to the high  $R^2$ . Parameters of  $-Z_t^{(X,h)}$  are large and always statistically significant. Those of  $-SZ_t^{(6-F,h)}$  are significant mostly for longer holding periods when an investor holds bonds of longer maturities. For example, to sell a six month Treasury Bill 12 month from now you must buy a Treasury Bond with maturity of 18 months which you see 12 month from now. Such investments are more sensitive to changes in the slope of the yield curve.

The parameters of  $\sigma_t^{(X,h)}$ , which measure market diversity, are significant and negative for longer horizons and we now explore the effect of  $\sigma_t^{(X,h)}$ . For all  $h > 2$  the coefficients of  $\sigma_t^{(X,h)}$  are negative and large. To interpret this result we observe it says that an increase in diversity of market opinions *decreases the risk premium*. This same conclusion was derived earlier in theoretical work using simulations (see Kurz and Motolese (2001)). It reveals that markets with large diversity of beliefs are more stable since beliefs tend to cancel each other, resulting in reduced volatility and lower market risk. In essence, with increased diversity the effects of the law of large numbers is more pronounced over time. The converse is also true: markets are more risky the higher is the degree of unanimity in them. This is so since any change of market belief results in sharp change of prices when too many people try to get through the same door. The negative coefficient of  $\sigma_t^{(X,h)}$  says that lower risk premia are priced into markets with more diverse beliefs. Finally, keeping in mind the limitation of  $R^2$  we present in Table 3 the contribution of belief variables to the  $R^2$ .

**Table 3: Contribution of Belief Data to Excess Returns Predictability**

Asset	Horizon	R <sup>2</sup> Without Beliefs	R <sup>2</sup> With Beliefs
Fed Fund Futures	h=2	0.061	0.095
	h=4	0.201	0.279
	h=6	0.345	0.441
3 Months T-Bill	h=2	0.122	0.186
	h=4	0.256	0.288
	h=6	0.367	0.461
	h=8	0.460	0.553
	h=10	0.541	0.643
	h=12	0.595	0.666
6 Months T-Bill	h=2	0.131	0.179
	h=4	0.290	0.317
	h=6	0.389	0.521
	h=8	0.502	0.586
	h=10	0.558	0.644
	h=12	0.600	0.667

To sum up our findings, pro-cyclical fundamental variables are important since they are used, to some extent, to forecast the onset of recessions. But, as Samuelson liked to quip, markets will forecasted 11 of the next 5 recessions! Our theory captures some of the 6 other recessions which the market will predict but that will not happen.

Figures 3-5 exhibit the fitted and 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 perceived 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 – FIGURE 5**

**What is the order of magnitude of these premia?** To measure the magnitude of the market belief premium in basis points we provide the following information. For T Bills with short horizons  $h = 2$  the values of  $Z_t^{(j,h)}$  during the sample period range from -35 to +142 basis points. For  $h > 2$  it rises, so that at  $h = 6$  it is from -68 to +117 and when  $h = 12$  the ranges is about -68 to +160. For  $Z_t^{(F,h)}$  the range at  $h = 2$  is from -48 to +95 basis points and for  $h > 2$  it rises, reaching at  $h = 6$  the range of -90 to +123. In general, for any given horizon the volatility of  $Z_t^{(F,h)}$  is greater than the volatility of  $Z_t^{(j,h)}$  for T Bills. As to the effect of diversity, the standard deviation of  $\sigma_t^{(X,h)}$  is about 8 -14 basis points for Fed Funds and 8 - 18 for T Bills. To illustrate the marginal effects of the belief variables

we compute the following examples:

If  $Z_t^{(F,6)}$  takes 2/3 of its maximal value in the sample the premium changes by +76 basis points

If  $Z_t^{(F,6)}$  takes 2/3 of its minimal value in the sample the premium changes by -56 basis points

If  $Z_t^{(3,6)}$  takes 2/3 of its maximal value in the sample the premium changes by +44 basis points

If  $Z_t^{(3,6)}$  takes 2/3 of its minimal value in the sample the premium changes by -26 basis points

If  $Z_t^{(6,6)}$  takes 2/3 of its maximal value in the sample the premium changes by +117 basis points

If  $Z_t^{(6,6)}$  takes 2/3 of its minimal value in the sample the premium changes by -68 basis points.

The distributions of the belief variables as realized over time is far from symmetric; they all tend to have large tails. For example  $Z_t^{(3,6)}$  has skewness of 1.14 and kurtosis of 6.04 and  $Z_t^{(F,6)}$  has skewness of -0.20 and kurtosis of 2.68. Nevertheless we present in Table 4 their measured standard deviations during the applicable sample period. Together with the results in Tables 2A-2C the reader can assess the effects of marginal changes in these variables in terms of their standard deviations. For example, suppose for  $h = 6$  we have an increase of  $Z_t^{(6,6)}$  by two standard deviations. In that case the risk premium rises by (see Table 2C)  $1.504 \times 64 = + 96.26$  basis points.

**Table 4: Standard Deviations of Belief Variables (in basis points)**

	$\sigma_t^{(F,h)}$	$\sigma_t^{(3,h)}$	$\sigma_t^{(6,h)}$	$Z_t^{(F,h)}$	$Z_t^{(3,h)}$	$Z_t^{(6,h)}$	$SZ_t^{(6-F,h)}$
h=2	7.8	8.4	8.4	23.9	26.1	27.1	18.4
h=4	10.6	11.2	11.1	38.8	28.8	28.5	15.9
h=6	14.1	14.4	13.8	45.1	33.6	32.0	17.4
h=8	----	16.8	16.3	----	35.2	35.4	19.5
h=10	----	18.5	17.7	----	42.7	44.1	20.3
h=12	----	18.4	17.9	----	46.7	48.1	22.2

**Non- Stationarity.** Our theory hinges on the fact that agents do not know the true structure of the economy since the economy exhibits non-stationarity. In that case the risk the excess return function will have to exhibit non-stationarity as well. To test for parameter time variability we could select dates when structural changes are considered. 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 periods we estimate (27) and conduct Chow tests of parameter time variability. In Tables 2A-2C we report parameter estimates for the entire periods and p-values of Chow tests for parameter time variability. 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.

#### 4. Conclusions and Final Comments

Excess volatility of asset returns, above and beyond the level warranted by fundamental forces, is a fact contested by only very few economists. In earlier work cited above we have shown via simulations how the dynamics of belief impacts the dynamics of asset pricing. In this paper we focus on market risk premia. We first set up a model of asset pricing with heterogenous beliefs and derive an analytical expression for the risk premium of a risky asset over the riskless rate. We find two effects. A direct effect on risk premia, via the effect of market belief on market volatility, which is a constant premium. A second effect which we call “the market belief risk premium” varies over time and is rather surprising in nature. We show that the risk premium  $E_t^m[\pi_{t+1}]$  is decreasing in the mean market belief  $Z_t$ . This result means that when the market holds abnormally favorable belief about future payoffs of an asset, the market views the long position in that asset as less risky. In that case the long term risk premium awarded the long position in that asset is reduced. Fluctuating market beliefs thus imply time variability of risk premia but more important, fluctuations in risk premia are inversely related to the degree of market optimism about future prospects of the asset in question. 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 on 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 varies across markets and holding periods but is, generally, very large.

The strong effects of market belief on market risk premia has thus two important implications. First, *it offers an alternative way of showing (for those who have any doubt) that fundamental factors affect market dynamics but perception is equally important for market volatility. Second, that market belief is actually an observable data which can be used for a deeper understanding of the basic causes of stochastic volatility and time variability of risk premia.*

Although the theory is framed in terms of beliefs in state variables which impact the price of an asset, our data on interest rates forecasts, are actually superior. This is so since such forecasts sum up all state variables which could have affected interest rates and which we could have missed. That is, suppose we had a list of state variables which could impact interest rates and suppose we deduced



all market beliefs about these state variables. We would then find that these market beliefs are heavily correlated and this fact would actually make it harder for us to test our theory. In addition, the data would never be able to capture all the relevant state variables affecting interest rates. We thus conclude that good quality data on the distribution of market forecasts of prices is needed for any future research in this area. Indeed, BLUF has not collected forecast data on stock prices and we have not been able to find a consistent, satisfactory panel data on stock price forecasts. As a result, it has not been possible for us to carry out on stock market returns the test done in this paper for Treasury Securities and Federal Funds futures. Nevertheless, in future research we hope to explore other dimensions of the data.

## **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 Exchange Rates 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) (1997a): *Endogenous Economic Fluctuations: Studies in the Theory of Rational Belief*. Studies in Economic Theory, No. **6**, Berlin and New York: Springer-Verlag.
- Kurz, M. (1997b): "On the Volatility of Foreign Exchange Rates." Chapter **12** in Kurz, M. (ed.) (1997a) *Endogenous Economic Fluctuations: Studies in the Theory of Rational Belief*. Studies in Economic Theory, No. **6**, Berlin and New York: Springer-Verlag, 317 -352.
- Kurz, M. (2005): "Diverse Beliefs, Forecast Errors and Central Bank Policy." Working Paper, Stanford University, July, 2005.
- Kurz, M. (2006): "Beauty contests under private information and diverse beliefs: How different?" Working Paper, Stanford University, August, 2006. (Forthcoming in the *Journal of Mathematical Economics*).
- Kurz, M., Beltratti, , A. (1997): "The Equity Premium is No Puzzle." Chapter **11** in Kurz, M. (ed.) (1997a) *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

The problem is formulated as follows:

$$\begin{aligned}
 W_{t+1} &= (W_t - C_t)R + \theta_t Q_{t+1} \\
 Q_{t+1} &= p_{t+1} + d_{t+1} - p_t R + \mu \quad , \quad d_t = D_t - \mu \quad , \quad \psi_t = (1, d_t, z_t, g_t) \\
 d_{t+1} &= \lambda_d d_t + \lambda_g^d g_t + \varepsilon_{t+1}^d \\
 z_{t+1} &= \lambda_z z_t + \lambda_g^z g_t + \varepsilon_{t+1}^z, \quad \Lambda_\psi = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & \lambda_d & 0 & \lambda_g^d \\ 0 & 0 & \lambda_z & \lambda_g^z \\ 0 & 0 & 0 & \lambda_z \end{pmatrix}, \quad \hat{\varepsilon}_t = (1, \varepsilon_t^d, \varepsilon_t^z, \varepsilon_t^g), \quad (\varepsilon_t^d, \varepsilon_t^z, \varepsilon_t^g) \sim N(0, \Sigma) \\
 g_{t+1} &= \lambda_z g_t + \varepsilon_{t+1}^g
 \end{aligned}$$

We also keep in mind the simpler notation

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

Hence we have  $\psi_{t+1} = \Lambda_\psi \psi_t + \Lambda_\varepsilon \varepsilon_{t+1}$ , where  $\Lambda_\varepsilon = \begin{pmatrix} 0, 0 \\ 0, \mathbf{I}_{(3 \times 3)} \end{pmatrix}$  is a 4x4 matrix

We assume that

$$p_t = a_d d_t + a_z z_t + P_0$$

Computing 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

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

Or,

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

where

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

and also we shall use the notation  $\mathbf{b}^T = ((a_d + 1), a_z, 0)$ . Now compute the expression

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

Algebra and simplification leads to the conclusion that we have

$$-\alpha W_{t+1} - \frac{1}{2} \psi_{t+1}^T \mathbf{V} \psi_{t+1} = -A_t - \mathbf{e}_t^T \varepsilon_{t+1} - \frac{1}{2} \varepsilon_{t+1}^T \mathbf{V}_{11} \varepsilon_{t+1}$$

where

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

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

It is now well known that the Bellman Equation for this problem with  $\gamma = \frac{1}{\tau}$  is

$$J_t = \text{Max}_{(\theta_t, C_t)} [-\beta^{t-1} \exp\{-\gamma C_t\} - \beta^t E_t \exp\left\{-A_t - \mathbf{e}_t^T \varepsilon_{t+1} - \frac{1}{2} \varepsilon_{t+1}^T \mathbf{V}_{11} \varepsilon_{t+1}\right\}]$$

But we know that

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

and also

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

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

Hence,

$$\begin{aligned} \frac{1}{2} \mathbf{e}_t^T (\mathbf{I} + \Sigma V_{11})^{-1} \Sigma \mathbf{e}_t - A_t = & -\alpha (\mathbf{W}_t - \mathbf{C}_t) \mathbf{R} - \alpha \theta_t [\mathbf{a}^T - \mathbf{b}^T \Omega \Lambda_0] \psi_t + \\ & + \frac{1}{2} \alpha^2 \theta_t^2 \mathbf{b}^T \Omega \mathbf{b} - \frac{1}{2} \psi_t^T [\Lambda_\psi^T \mathbf{V} \Lambda_\psi - \Lambda_0^T \Omega \Lambda_0] \psi_t. \end{aligned}$$

The first order conditions are then stated as follows. The derivative with respect to  $\theta$  is

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

And this proves equation (17) in the text

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

We can also explain the “adjustment” to the variance in (17) since

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

which is the variance of the excess return function where the covariance matrix used is not  $\Sigma$  but rather  $\Omega$ .

We now have

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

Hence the optimized value of the exponent is simply

$$\frac{1}{2} \mathbf{e}_t^T (\mathbf{I} + \Sigma V_{11})^{-1} \Sigma \mathbf{e}_t - A_t = -\alpha (\mathbf{W}_t - \mathbf{C}_t) \mathbf{R} - \frac{1}{2} \psi_t^T \mathbf{M} \psi_t$$

Where

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

The derivative with respect to C

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

Hence the solution for C is

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

or

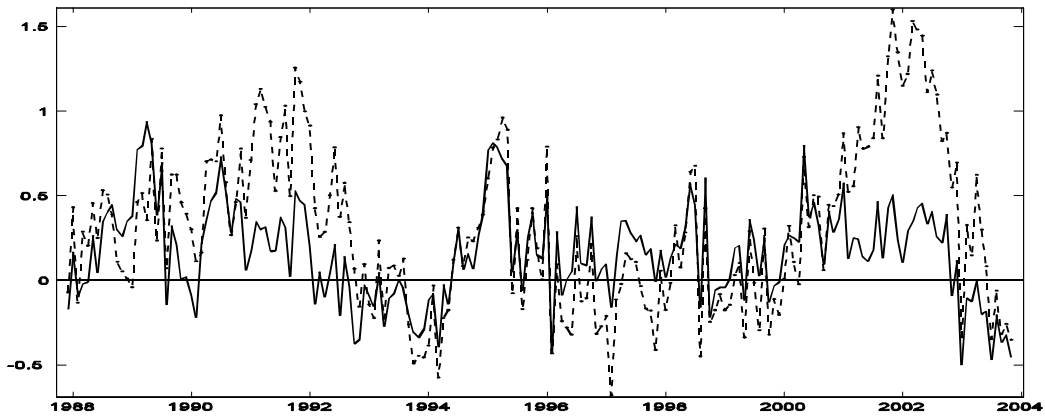
$$C_t = -\frac{1}{\gamma + \alpha R} \log\left[\frac{\beta\alpha R G}{\gamma}\right] + \frac{\alpha R}{\gamma + \alpha R} W_t + \frac{1}{2(\gamma + \alpha R)} \psi_t^T M \psi_t.$$

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  $\alpha$  and matrix  $V$  are determined by the conditions

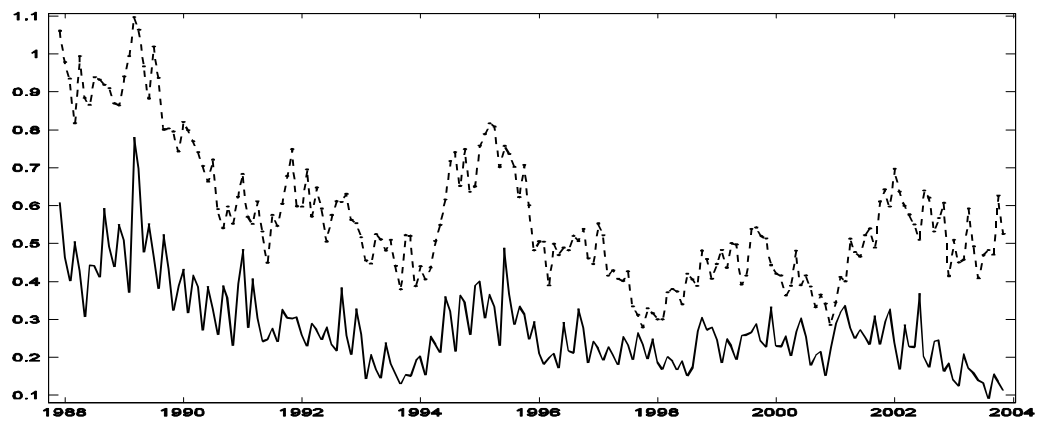
$$(i) \quad \alpha = \frac{r\gamma}{R}.$$

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

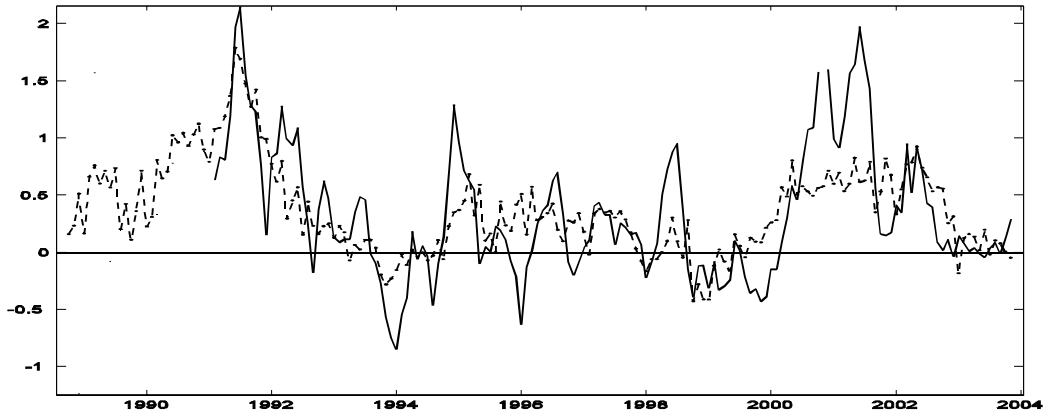
# FIGURES



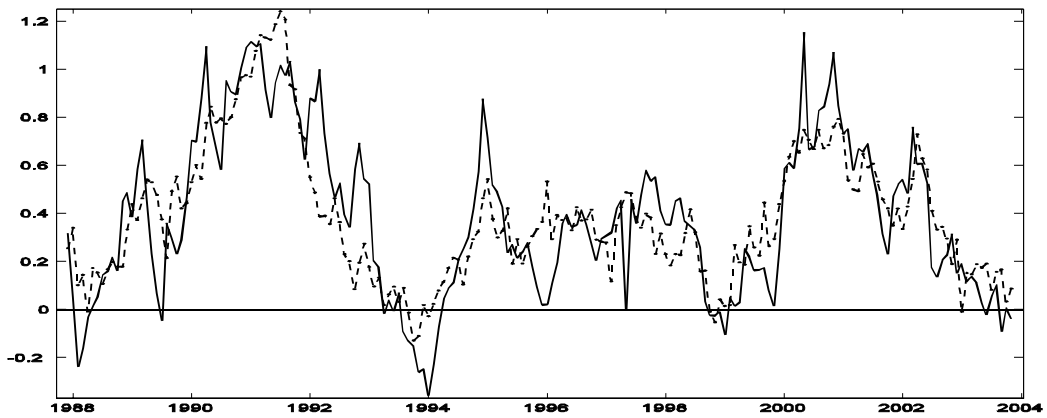
**Figure 1:** 4 and 12 (dashed line) months ahead market beliefs of 6 months T-bill rate ( $Z_t^{(6)}$ )



**Figure 2:** 4 and 12 (dashed line) months ahead standard deviations of market beliefs of 6 months T-bill rate ( $Z_t^{(6)}$ )

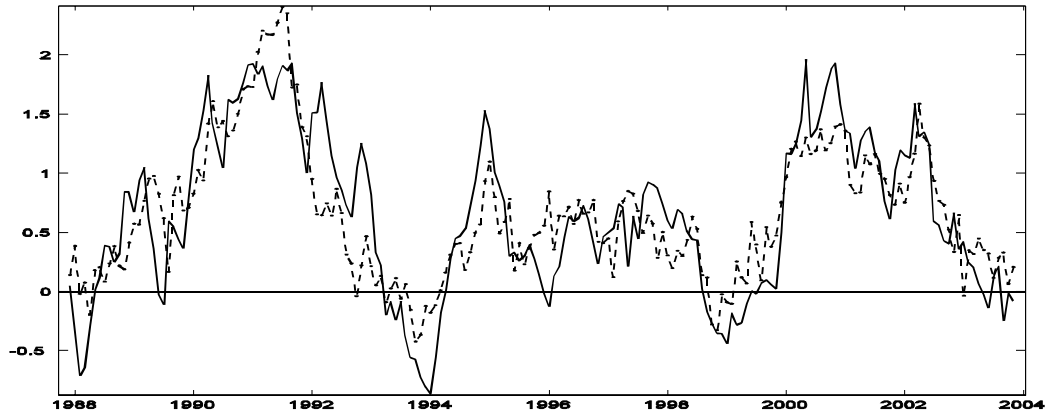


**Figure 3:** Excess returns on Fed Fund Futures contract 6 months ahead. The dashed line represents the fitted values from regression (27)



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





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