This work is distributed as a Discussion Paper by the

STANFORD INSTITUTE FOR ECONOMIC POLICY RESEARCH

Rational Diverse Beliefs and Economic Volatility

By
Mordecai Kurz
Stanford University

August 2007

Rational Diverse Beliefs and Economic Volatility

by

Mordecai Kurz, Stanford University
August 14, 2007

Prepared for Chapter in Handbook in Financial Economics
Entitled
Dynamics and Evolution of Financial Markets

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

Table of Content

Rational Diverse Beliefs and Economic Volatility[1]

by

Mordecai Kurz, Stanford University [2], August 14, 2007

1.   Introduction: Why Diversity is Needed for Understanding Market Dynamics

1.1   Belief Diversity as An Empirical Fact

Diversity of beliefs is an undisputed fact and the idea that allocations and prices are affected by agents' perceptions of the future is not new. Diverse expectations are central to Thornton's (1802) views of paper money and financial markets. Expectations are crucial to Keynes (1936). Chapter 12 of *The General Theory* examines the "state of confidence" and the importance of investors' expectations to asset pricing. Expectations are key to the "cumulative movements" in Pigou (see, Pigou's (1941), Chapter VI) and constitute the mechanism of deviations from a stationary equilibrium in the Swedish school (e.g. see Myrdal's 1939 views in Myrdal [1962], chapter III). Also, "subjective values" based on diverse agents' expectations are cornerstones of Lindahl [1939] theory of money and capital.

In the post world war II era large data bases on heterogenous forecasts of various variables have been assembled. Early work was done in Holland, Germany and Sweden. In the U.S. the Survey of Professional Forecasts, reporting *quarterly* forecasts of private forecasters, was started in 1968. It was first conducted by the American Statistical Association\National Bureau but has since been taken over by the Federal Reserve Bank of Philadelphia. Since 1980 the Blue Chip Economic Indicators (BLUE) reports *monthly* forecast distributions of economic variables by over 50 financial institutions. BLUE was expanded, under the title of Blue Chip Financial Forecasts (BLUF), to include forecasts of interest rates and other variables. To illustrate we report in Table 1 the forecast distribution of Real GDP and GDP price deflator in May of 2000 for 2000. Actual GDP growth in 2000 was 4.1% and the

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

inflation rate 2.3%. Note that in May of 2000, *five months into the year*, large heterogeneity persisted. Also, almost all GDP forecasts were wrong! To understand this correlated error place yourself in May 2000 and make a stationary econometric forecast of GDP growth but make no special judgment about the unique conditions in May 2000. An example of such a model was developed by Stock and Watson (2001),(2002),(2005). They estimate it by using a combination of diffusion indexes and bivariate VAR forecasts and employing a large number of U.S. time series. In May of 2000 the non-judgmental stationary forecast of GNP growth was *lower than most private forecasts*.

Table 1: May 2000 BLUE Forecasts of GNP Growth and Inflation for the full year 2000

May 2000 Forecasted Percent Change In	Real GDP	GDP Price Deflator
Forecast for 2000		
First Union Corp.	5.3H	2.0
Turning Points (Micrometrics)	5.2	2.1
J P Morgan	5.2	2.1
Evans, Carroll & Assoc.	5.1	2.2
Mortgage Bankers Assn. of Amer.	5.1	2.1
Goldman Sachs & Co.	5.1	2.1
U.S. Trust Co.	5.1	2.0
US Chamber of Commerce	5.1	2.0
Banc of America Corp.	5.1	2.0
Morgan Stanley Dean Witter	5.1	1.9
Wayne Hummer Investments LLC	5.0	2.3
Bank One	5.0	2.1
Nomura Securities	5.0	1.9
Merrill Lynch	5.0	1.9
Perna Associates	4.9	2.3
National Assn. of Home Builders	4.9	2.1
Macroeconomic Advisers, LLC	4.9	2.1
Prudential Securities, Inc.	4.9	2.0
LaSalle National Bank	4.8	2.3
Conference Board	4.8	2.3
Wells Capital Management	4.8	2.2
DuPont	4.8	2.1
Northern Trust Company	4.8	2.1
Chicago Capital, Inc.	4.8	2.0
Deutsche Bank Securities	4.8	1.8
Chase Securities, Inc.	4.8	1.8
Credit Suisse First Boston	4.8	1.8
Comerica	4.7	2.4
Moody's Investors Service	4.7	2.2
Fannie Mae	4.7	2.0
Federal Express Corp.	4.7	2.0
SOM Economics, Inc.	4.7	1.9
National Assn. of Realtors	4.7	1.9
National City Corporation	4.7	1.9
Clear View Economics	4.7	1.9
Eggert Economic Enterprises, Inc.	4.6	2.1
WEFA Group	4.6	1.9
Eaton Corporation	4.6	1.9
Bear Stearns & Co., Inc.	4.6	1.2 L
Ford Motor Company	4.5	1.8
Motorola	4.5	1.7
Standard & Poors Corp.	4.5	1.7
UCLA Business Forecasting Proj.	4.4	2.1
Inforum - Univ. of Maryland	4.4	2.0
Prudential Insurance Co.	4.4	1.9
Weyerhaeuser Company	4.3	2.2
DaimlerChrysler AG	4.3	2.0
Georgia State University	4.2	2.2
Kellner Economic Advisers	4.2	2.0
Econoclast	4.1	2.0
Naroff Economic Advisors	4.0 L	2.5 H

Repeating the experiment over time you will find the distribution of forecasts fluctuates in two ways. First, it exhibits large changes in the *cross sectional variance* of the forecasts, reflecting changes in degree of disagreement. Second, it exhibits time fluctuations *in relation to the stationary forecasts* reflecting correlation in forecasters' views about unusual conditions of the time. Sometimes forecast distributions are below the stationary forecast while in May of 2000 the distribution was above the

stationary forecast. It is important to recognize that large data banks of market forecast distributions are publically available for many variables since forecasters are willing to reveal their forecasts. The data shows that for any variable, *individual forecasts are correlated* and both the average market forecast as well as the cross sectional distribution of forecasts, *both fluctuate over time*.

Despite the impact of the Rational Expectations paradigm, belief heterogeneity has been used to explain many phenomena, including asset price volatility, risk premia, volume of trade, money non-neutrality etc. There are two theories of rational behavior motivated by the observed diversity. One follows the Harsanyii doctrine, viewing people as Bayesians who hold the same prior probability but with asymmetric private information (in short PI) used in forecasting. Examples of papers includes Grossman and Stiglitz (1980), Phelps (1970), Lucas (1972), Diamond and Verrecchia (1981), Townsend (1978), (1983), Singleton (1987), Brown and Jennings (1989), Grundy and McNichols (1989), Wang (1994), He and Wang (1995), Judd and Bernardo (1996), (2000), Morris and Shin (2002), (2005), Woodford (2003), Hellwig (2002), (2005), Angeletos and Pavan (2006), Angeletos and Werning (2006), Allen, Morris and Shin (2006), Hellwig, Mukherji and Tsyvinski (2006).

An alternative view finds no evidence for the use of PI in forecasts of market returns or economic aggregates. It insists there is no justification to a common prior hence diverse probability models of state variables is inevitable in a complex world. A sample of papers taking this approach includes Harrison and Kreps (1978), Varian (1985), (1989), Harris and Raviv (1993), Kurz (1994), (1996), (1997a), (1997b), (1997c), (2007) Detemple and Murthy (1994), Frankel and Rose (1995), Kandel and Pearson (1995), Cabrales and Hoshi (1996), Kurz and Beltratti (1997), Kurz and Motolese (2001), (2006), Kurz and Schneider (1996), Kurz and Wu (1996), Kurz, Jin and Motolese (2005a), (2005b), Nielsen (1996), (2003),(2005),(2006), Motolese (2001), (2003),Wu and Guo (2003), (2004), Fan (2006), Nakata (2007), Guo and Wu (2007). Kurz's (1994), (1997a) theory of Rational Beliefs stresses the impossibility of perfect learning. It holds that our environment is non-stationary with structural changes occurring faster than we can learn them hence no learning can ever converge.

In this Chapter we review the role of diverse beliefs in explaining market dynamics, but we also clarify the differences between the above two theories of diversity. Since the range of issues is wide, we exclude two types of models. The first are Behavioral Finance models where belief diversity arises from psychological but irrational motives. Second, learning models with common information are

typically models of convergence to Rational Expectations hence belief diversity is not persistent.

1.2   A Basic Principle: Rational Diversity of Belief Implies Volatility

The term "Market Dynamics" refers to the dynamic characteristics of financial markets but we focus here on dynamic phenomena that have attracted attention in the literature. Examples include excess volatility of asset returns, high and time varying risk premia, high volume of trade, money non-neutrality, etc. Many are termed "anomalies" as they contradict predictions of a Rational Expectations Equilibrium (in short, REE) with full information. Hence, they are outstanding problems in financial economics and we review two theories which aim solve them. All studies show that fundamental exogenous factors cannot explain the observed dynamics, leading Paul Samuelson to quip that "the stock market predicted 9 out of the last 5 recessions." We aim to understand the 4 recessions the market predicted but which did not occur. Since standard models explain dynamics with exogenous shocks, we want to study mechanisms of *endogenous amplification* which explain excess volatility. Most researchers agree that belief heterogeneity is needed to explain market dynamics but the question is what heterogeneity? We thus first discuss criteria we use in the evaluation of any theory.

The data reveals heterogeneity *persists* hence the two criteria for acceptable diversity in any theory is that *it has aggregate effects* and these are *non-vanishing*. The non-vanishing condition is challenging to asymmetric PI models under REE since information revelation of prices leads back to a common belief and diversity cannot persist. Asymmetric PI in REE is therefore supplemented with a "noisy" mechanism to avoid revelation. Is such a mechanism natural or just an artificial construct? Is it testable? Requiring diversity to have persistent aggregate effects implies that heterogeneity by itself is not sufficient and it must be supplemented with dynamic features. To understand the importance of this fact consider two examples. (1) Beliefs are heterogeneous, independently distributed across agents and with a fixed distribution over time. The i.i.d. distribution causes a cancellation of the effects of beliefs hence there is a constant, typically small, aggregate effect. Such a distribution implies diversity is irrelevant. (2) Beliefs are heterogeneous *with a fixed distribution of beliefs around some norm* over time. That is, a specific agent always has the same superior or inferior information or else, specific agents are always more optimistic or more pessimistic than others. A fixed belief distribution also implies that prices, volume of trade and risk premia fluctuate only in response to exogenous shocks

hence we are back to a theory, rejected by the data, that publicly observed exogenous shocks are the only cause of fluctuations. Such distributions of beliefs do not generate the desired endogenous amplification to explain excess volatility. These examples show the *dynamics* of beliefs over time is essential and the question is what is its source. One approach assumes a random sequence of irrational noise traders. As we focus on rational behavior this approach is unsatisfactory. Under asymmetric PI such dynamics could be generated by an exogenous flow of private information which entails a process of belief updating. How effective or plausible is such an assumption must be carefully weighed and we discuss it in detail later on. The situation is drastically different under diverse beliefs with common information since dynamics and rationality are linked. To explain why, we now briefly explain the simple principle *that rational diversity of beliefs without private information implies market volatility!*

Diversity of beliefs exists only when something in the economy is not known. Hence, any model discussed here assumes that agents do not know a true probability or a true parameter and they hold diverse beliefs about what they do not know. This induces two basic questions. First, why do agents not know what they do not know? Second, what is their common knowledge basis?

Starting with the second question, note that although assumptions about what is common knowledge vary across models, one answer is general: it is *past data on publicly observed data.* That is, the economy has a vector of observable variables x_t over time with true unknown probability law Π on infinite sequences. Agents have a long history $\{ x_t, t = 1, 2, ..., N \}$, allowing rich statistical analysis. Given the data, all agents compute the same finite dimensional distributions of the data hence they all know the same empirical moments. They then deduce from the data an empirical probability on infinite sequences denoted by m. It can be shown that m is unique, time invariant or stationary (see Kurz (1994)). This is the *empirical knowledge shared by all agents.*[3] Turning now to the first question, the basic cause for diverse beliefs is the fact that m and Π *are not the same*. We briefly explain why.

Our economy has undergone changes in technology and social institutions. These changes have

---

[3] Since in reality there are only finite data, the computed empirical probability is m_t which depends upon the history $\{ x_j, j = 1, 2, ..., t \}$. With time, m_t converges to a limit probability m and the assumption in the text is that the data sequence is long so that m_t is a good approximation for m. The assumption that all agents know the limit m is very strong. It is made for simplicity and we explain later how it is used. We stress that *we are not assuming that agents have an infinite data set* since such an assumption would lead to different perspective on what is rational to believe. It is legitimate to assume that at any t the common knowledge empirical probability is only m_t which is then re-computed each date. The resulting theory is far more complicated since it entails approximations which vary over time. Please note that such approximations would actually lead to increased diversity of beliefs as there would be an additional dimension of disagreement about what the probability m is.

had deep economic effects causing the process $\{x_t, t = 1, 2, \ldots\}$ to be non-stationary. Although this means the distributions of the x_t's are time dependent, it is far more than saying that $\{x_t, t = 1, 2, \ldots\}$ constitutes a sequence of productivity "regimes." It requires recognition that although we assume assets pay in a unit of account, assets and commodities used as payoff vary over time. However, each "regime" is short and hence it is impossible to learn the unknown true probability Π. Hence, the stationary probability m is just an average over an infinite sequence of changing regimes. It reflects long term frequencies but it is not the process' true probability. Belief diversity arises when agents believe m is not the truth hence past empirical record is not adequate to forecast the future. All surveys of forecasters show that subjective judgment contributes more than 50% to the final forecast (e.g. Batchelor and Dua (1991)). Individual subjective models are, therefore, the way agents express their interpretation of the data. Being common knowledge, the stationary probability m thus becomes a reference for any concept of rationality. That is, any rationality requirement must insist that a model does not contradict the evidence summed up by the probability m.

Is it rational, without any private information, to believe m is the truth? Those who believe the economy is stationary hold this belief. 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 m. This idea of rationality embraces a wide range of models without resorting to psychological or behavioral principles to explain diversity. The RB theory has a simple and powerful implication that addresses the question of dynamics. It says an agent's date t state of belief (i.e. his date t conditional probabilities) *cannot be constant, time invariant,* unless he believes the stationary probability m is the truth. To see why, consider an agent who holds a constant belief different from m. It has a time average belief which is not m. Since m is the time average in the data, this *proves* the agent is irrational. It is thus irrational to be permanently optimistic (or pessimistic) relative to m. By implication, if a rational agent's belief disagrees with m *then such a belief must fluctuate over time around m.* By adding the verified fact (see Section 3.2.1) that individual beliefs are correlated, we complete the argument to deduce the conclusion that *rational diversity implies aggregate dynamics.*

Diversity of beliefs without PI is often questioned by asking how agents can be wrong and rational at the same time. This is exactly the central idea of the RB theory. Indeed, when rational agents hold diverse probability beliefs while there is only one true dynamic law of motion *then most*

agents are wrong most of the time. Since agents' beliefs are correlated, the average market belief is also often wrong. This fact is the source of endogenous propagation of market risk and volatility, called "Endogenous Uncertainty" by Kurz (1974) and Kurz and Wu (1996), and explained later.

2.   Explaining Market Dynamics with Asymmetric Private Information

The literature on "noisy" REE asset pricing under asymmetric PI is large and Brunnermeier (2001) provides a good survey. We discuss it in three stages. In Section 2.1 we present a universally used model which follows Grossman and Stiglitz (1980). In Section 2.2 we discuss dynamic versions of the model and applications. In Section 2.3 we evaluate the ideas developed.

2.1   A General Model of Asset Pricing Under Asymmetric Information

The model reviewed here is an adaptation of the short lived trader model of Brown and Jennings (1989). Similar models were used by Grossman and Stiglitz (1980), Diamond and Verrecchia (1981), Singleton (1987), Grundy and McNichols (1989), Wang (1994), He and Wang (1995), Bacchetta and van Wincoop (2005), (2006), Allen, Morris and Shin (2006) and others.

There is a unit mass of traders, indexed by the [0, 1] interval and a single aggregate asset with unknown intrinsic unit value Q. The economy is static with one period divided into three trading dates (no discounting): at date 1 traders first receive public information y and private signals x^i about the asset value and then they trade. At date 2 they trade again. At date 3 (or end of date 2) uncertainty is resolved, the true liquidation value Q is revealed and traders receive it for their holdings. Public information is that Q is distributed in accord with $Q \sim N(y, \frac{1}{\alpha})$. The private signal x^i about Q is $x^i = Q + \varepsilon^i$ where ε^i satisfy $\varepsilon^i \sim N(0, \frac{1}{\beta})$ *independently across i.* (α, β) are known. Since these facts are common knowledge, agents know the true unknown value Q is "in the market" since by the law of large numbers the mean of all private signals is the true value Q. Trader i starts with S^i units of the aggregate asset and can borrow at zero interest rate to finance trading. (D_1^i, D_2^i) are i's demands in the first and second rounds and (p_1, p_2) are market prices in the two rounds. Ending wealth is thus $W^i = S^i p_1 + D_1^i(p_2 - p_1) + D_2^i(Q - p_2)$. All traders are assumed to have the same utility over wealth $u(W^i) = -e^{-(W^i/\tau)}$ with constant absolute coefficient of risk aversion, and they maximize expected utility. Aggregate supplies (S_1, S_2) of shares, each representing an asset unit, traded in each of the

rounds are *random, unobserved and independently normally distributed with mean zero*. This noise is crucial to ensures that traders cannot deduce from prices the true value of Q. In a noisy REE traders maximize expected utility of final wealth while markets clear after traders deduce from prices all possible information. Indeed, Brown and Jennings (1989) show that equilibrium price at date 1 is

(1a) $$p_1 = \kappa_1(\lambda_1 y + \mu_1 Q - S_1)$$

and since S_1 is normally distributed, p_1 is also normally distributed. Keep in mind that Q and S_1 are unknown, hence (1a) shows that prices are not fully revealing. Since over trading rounds Q remains fixed, more rounds of trading generate more price data from which traders deduce added information about Q. But with additional supply shocks the inference problem becomes more complicated. That is, at date 2 the price p_2 contains more information about Q but it depends upon *two* unobserved noise shocks (S_1, S_2). Hence, the price function is shown to be time dependent and at date 2 takes the form

(1b) $$p_2 = \hat{\kappa}_2(\hat{\lambda}_2 y + \hat{\mu}_2 Q - S_2 + \psi S_1).$$

Since the realized noise S_1 is not observed, traders condition on the known price p_1 to infer what they can about S_1. They thus use a date 2 price function which takes an equivalent form

(1c) $$p_2 = \kappa_2(\lambda_2 y + \mu_2 Q - S_2 + \xi_{21} p_1).$$

By (1a), equivalence means $\kappa_2 = \hat{\kappa}_2$, $\lambda_2 = (\hat{\lambda}_2 + \lambda_1 \psi)$, $\mu_2 = (\hat{\mu}_2 + \mu_1 \psi)$ and $\xi_{21} = -\dfrac{\psi}{\kappa_1}$. Denote by (H_1^i, H_2^i) the information of i in the two rounds. The linearity of the equilibrium price map implies that the payoff is normally distributed. Brown and Jennings (1989) show in Appendix A that there exist constants (G_1, G_2), determined by the covariance matrix of the model's random variables and assumed by most writers to be the same for all agents, such that the demand functions of i are

(2a) $$D_2^i(p_2) = \frac{\tau}{\mathrm{Var}^i(Q|H_2^i)}[E^i(Q|H_2^i) - p_2].$$

(2b) $$D_1^i(p_1) = \frac{\tau}{G_1}[E^i(p_2|H_1^i) - p_1] + \frac{(G_2 - G_1)}{G_1}[E^i(D_2^i|H_1^i)].$$

Most writers assume $\mathrm{Var}^i(Q|H_2^i) = \sigma_Q^2$ independent of i. The second term in (2b) is the usual "hedging demand" arising from a trader's date 1 perceived risk of price change between date 1 and date 2. The hedging demand in a noisy REE complicates the inference problem and raises existence of equilibrium problems. As a result, most writers ignore this demand and study the *myopic-investor economy* in which there are only "short lived" traders. A "short lived" trader lives one period only. He first trades in date 1, gains utility from p_2 and leaves the economy. He is replaced by a new short lived

trader who receives the information of the first trader but trades in date 2 only and gains utility from the revealed Q. Neither trader has a hedging demand. A "long lived" trader lives through *both* periods, trades in dates 1 and 2 hence has a hedging demand. For simplicity we follow here the common practice and ignore the second term in (2b). We now average on i, equate to supply and conclude that

(3) $$p_2 = \overline{E}_2(Q) - (\frac{1}{\tau})\sigma_Q^2(S_1 + S_2) \ , \ p_1 = \overline{E}_1(p_2) - \frac{G_1}{\tau}S_1 \ .$$

$\overline{E}_2(Q)$ is date 2 average market forecast of Q and $\overline{E}_1(p_2)$ is average market forecast of p_2. In this case $G_1 = Var_1^i(p_2)$ and it is assumed this variance is the same for all i. Hence, the proof of (1a)-(1b) amounts to exhibiting a closed form solution of $\overline{E}_2(Q)$ and solving the joint system in (3).

The derivation of (2a)-(2b) used a *general* form of conditional expectations and required only that prices are normally distributed. It is thus a general solution for any informational structure used in the conditioning and *it does not depend upon the private character of information used*. This means that it is applicable to models with diverse beliefs and common information as long as their implied prices are normally distributed. Moreover, differences among theories of diverse beliefs are expressed entirely by differences in their implications to the conditional expectations in (2a)-(2b). In the case of asymmetric PI discussed here, (2a) shows that D_2^i depend upon date 2 expectations which are updated based on the information deduced from p_2 and p_1. This is different from date 1 information which consists of public signal, private signals and inference from p_1 only. This explains why equilibrium price maps are time dependent. Allen et al. (2006) present in Appendix A computations of the closed form solution. To get an idea of the inference involved we briefly review the steps they take.

What does a trader learn in round 1? Given a prior belief $Q \sim N(y, \frac{1}{\alpha})$ trader i observes $p_1 = \kappa_1(\lambda_1 y + \mu_1 Q - S_1)$. Since $S_1 \sim N(0, 1/\gamma_1)$ all he infers from date 1 price is that

$$[1/(\kappa_1\mu_1)](p_1 - \kappa_1\lambda_1 y) = Q - [S_1/\mu_1] \ \sim \ N(Q, 1/(\mu_1^2\gamma_1)) \ .$$

But now, his added piece of information is the private signal $x^i = Q + \varepsilon^i$, $\varepsilon^i \sim N(0, \frac{1}{\beta})$. Using a standard Bayesian inference from these three sources, his posterior belief becomes

(4a) $$E_1^i(Q|H_1^i) = \frac{\alpha y + \beta x^i + \mu_1^2\gamma_1\frac{1}{\kappa_1\mu_1}(p_1 - \kappa_1\lambda_1 y)}{\alpha + \beta + \mu_1^2\gamma_1} = \frac{(\alpha - \mu_1\gamma_1\lambda_1)y + \beta x^i + \frac{\mu_1\gamma_1}{\kappa_1}p_1}{\alpha + \beta + \mu_1^2\gamma_1}$$

(4b) with precision $\alpha + \beta + \mu_1^2\gamma_1$.

Averaging (4a) over the population we can see that the average market forecast at date 1 is then

$$\bar{E}_1(Q|H_1) = \frac{(\alpha - \mu_1\gamma_1\lambda_1)y + \beta Q + \dfrac{\mu_1\gamma_1}{\kappa_1}p_1}{\alpha + \beta + \mu_1^2\gamma_1} \equiv \frac{\alpha y + (\beta + \mu_1^2\gamma_1)Q}{\alpha + \beta + \mu_1^2\gamma_1} - \frac{\mu_1\gamma_1 S_1}{\alpha + \beta + \mu_1^2\gamma_1}.$$

In round 2 a trader observes p_2 which, as seen in (1c), is a function of p_1. Given p_1 and the fact that $S_2 \sim N(0, \frac{1}{\gamma_2})$, he infers from $p_2 = \kappa_2(\lambda_2 y + \mu_2 Q - S_2 + \xi_{21}p_1)$ that

$$[1/(\kappa_2\mu_2)](p_2 - \kappa_2\lambda_2 y - \kappa_2\xi_{21}p_1) = Q - [S_2/\mu_2] \sim N(Q, 1/(\mu_2^2\gamma_2)).$$

He now updates (4a)-(4b). Since supply shocks are i.i.d. the updated posterior is standard

$$E_2^i(Q|H_2^i) = \frac{[\dfrac{(\alpha - \mu_1\gamma_1\lambda_1)y + \beta x^i + \dfrac{\mu_1\gamma_1}{\kappa_1}p_1}{\alpha + \beta + \mu_1^2\gamma_1}](\alpha + \beta + \mu_1^2\gamma_1) + \dfrac{1}{\kappa_2\mu_2}(p_2 - \kappa_2\lambda_2 y - \kappa_2\xi_{21}p_1)(\mu_2^2\gamma_2)}{\alpha + \beta + \mu_1^2\gamma_1 + \mu_2^2\gamma_2}.$$

Simplification leads to

(5a) $$E_2^i(Q|H_2^i) = \frac{[\alpha - \mu_1\gamma_1\lambda_1 - \mu_2\gamma_2\lambda_2]y + \beta x^i + [\dfrac{\mu_1\gamma_1}{\kappa_1}p_1 + \dfrac{\mu_2\gamma_2}{\kappa_2}p_2 - \mu_2\gamma_2\xi_{21}p_1]}{\alpha + \beta + \mu_1^2\gamma_1 + \mu_2^2\gamma_2}$$

(5b) $$Var(Q|H_2^i) = \frac{1}{\alpha + \beta + \mu_1^2\gamma_1 + \mu_2^2\gamma_2}.$$

Finally, to compute (1c) we average (5a) to conclude that

(6a) $$\bar{E}_2(Q) = \frac{[\alpha - \mu_1\gamma_1\lambda_1 - \mu_2\gamma_2\lambda_2]y + \beta Q + [\dfrac{\mu_1\gamma_1}{\kappa_1}p_1 + \dfrac{\mu_2\gamma_2}{\kappa_2}p_2 - \mu_2\gamma_2\xi_{21}p_1]}{\alpha + \beta + \mu_1^2\gamma_1 + \mu_2^2\gamma_2}$$

(6b) $$\bar{E}_1(p_2) = \kappa_2(\lambda_2 y + \mu_2\bar{E}_1(Q) + \xi_{21}p_1).$$

We now solve for prices by inserting (6a)-(6b) into (3). The final step is to match coefficients of the price functions and identify $(\kappa_1, \lambda_1, \mu_1, \kappa_2, \lambda_2, \mu_2, \xi_{21})$. For details of these computations see Allen et al. (2006), Appendix A. This verifies that prices are indeed normally distributed as in (1a)-(1b).

What is the length of memory in prices? Multiple trading rounds provide opportunities to deduce more information from prices about Q. As trading continues, information of *all past prices is used* since prices depend upon all past unobserved supply shocks. In such a case the price system is not a finite memory Markov process. The model has been extended to multi period trading where Q is revealed N periods later (see Brown and Jennings (1989), Grundy and McNichols (1989), He and Wang (1995) and Allen et al. (2006)). In these models the complexity of inference depends upon the

presence of a hedging demand of long lived traders. However, for both long and short lived traders *the number of trading rounds is an arbitrary modeling construct*. It would thus be instructive to examine the limit behavior. In a third round of trading by short lived traders the price map becomes

$$p_3 = \kappa_3(\lambda_3 y + \mu_3 Q - S_3 + \xi_{31} p_1 + \xi_{32} p_2).$$

Hence, the independent supply shock leads to an updating rule which is again standard

$$E_3^i(Q|H_3^i) = \frac{E^i(Q|H_2^i)(\alpha + \beta + \mu_1^2\gamma_1 + \mu_2^2\gamma_2) + \frac{1}{\kappa_3\mu_3}(p_3 - \kappa_3\lambda_3 y - \kappa_3\xi_{31}p_1 - \kappa_3\xi_{32}p_2)(\mu_3^2\gamma_3)}{\alpha + \beta + \mu_1^2\gamma_1 + \mu_2^2\gamma_2 + \mu_3^2\gamma_3}.$$

Simplification and averaging over the population leads to the market forecast

$$\bar{E}_3(Q) = \frac{[\alpha - \mu_1\gamma_1\lambda_1 - \mu_2\gamma_2\lambda_2 - \mu_3\gamma_3\lambda_3]y + \beta Q}{\alpha + \beta + \mu_1^2\gamma_1 + \mu_2^2\gamma_2} +$$

$$+ \frac{\frac{\mu_1\gamma_1}{\kappa_1}p_1 + \frac{\mu_2\gamma_2}{\kappa_2}p_2 + \frac{\mu_3\gamma_3}{\kappa_3}p_3 - \mu_3\gamma_3\xi_{31}p_1 - \mu_3\gamma_3\xi_{32}p_2 - \mu_2\gamma_2\xi_{21}p_1}{\alpha + \beta + \mu_1^2\gamma_1 + \mu_2^2\gamma_2}.$$

Individual and market forecasts can be expressed in terms of the unobserved variables. They can easily be extended to N rounds of trade and it can be shown that they take the general form

$$(7) \qquad E_N^i(Q|H_N^i) = \frac{\alpha y + \beta x^i + \sum_{j=1}^{N} \mu_j^2\gamma_j Q}{\alpha + \beta + \sum_{j=1}^{N} \mu_j^2\gamma_j} - \frac{\sum_{j=1}^{N} \mu_j\gamma_j S_j}{\alpha + \beta + \sum_{j=1}^{N} \mu_j^2\gamma_j}$$

A standard argument shows the μ_j converge. For simplicity assume $\gamma_j = \gamma$. The independence of the noise S_j together with (7) and the law of large numbers imply that with probability 1 the first term converges to Q and the second converges to 0. Hence, in the limit, with probability 1 all forecasts converge to the true Q and the effect of the public signal y disappears. This proves that *repeated trading leads to a full revelation of the true value,* and that in the limit p = Q hence traders do not forecast prices at all. With repeated trading the effect of y disappears. If the unit of time is short, like a month, trading rounds are not really limited. Hence this result contradicts the claim (e.g. see Allen et al. (2006)) that the effect of the public signal y on prices lingers on forever.

What have we learned so far? The key conclusions of the PI paradigm are that equilibrium prices vary with each date's true intrinsic value of the asset and with the random supply shock of that

11

date, both of which are fundamental factors. Private information, as such, has no separate direct effect on price volatility since the effect of private information is averaged out by the law of large numbers. Hence, the model does not posses the endogenous amplification which we seek. In addition, since supply shocks are never observed , the repeated inference causes prices to have infinite memory. In applications such as Brown and Jennings (1989) and Grundy and McNichols (1989), this property was used to explain the phenomenon of "Technical Analysis" defined as the traders' use of past prices, in addition to today's price, to form their demand. More broadly, the static model of asymmetric PI was used in widely diverse applications. The purely theoretic motive was to explain the persistence of diverse beliefs in an REE which is true in the model since the equilibrium is not fully revealing. The range of applications is very wide and, to illustrate, we briefly review two different applications.

The Phelps (1970) and Lucas (1972) island models are macroeconomic examples of models where asymmetric PI plays a central role in generating belief diversity with real effects. In these models agents receive private signals which do not fully reveal the state of an aggregate variable such as rate of growth of money or of nominal GNP. In equilibrium prices reveal some information and agents make the best possible inference from prices in order to improve their knowledge of the state of the economy. Such inference is the same as in the model above. In the resulting noisy REE prices are not fully revealing with real economic results. In the case of Phelps (1970), the book reviews several examples of friction generated by asymmetric information where slow price adjustment results in real short term effects in labor markets, capital markets and others. In Lucas' (1972) island model agents receive private signals about money growth and in the resulting noisy REE money is not neutral.

A second example is Allen et al. (2006) who use the model to explain the Keynes (1936) Beauty Contest. To see how, recall $\overline{E}(S_i) = 0$. Then (3) implies that if there are N rounds of trade then

$$(8) \qquad\qquad p_1 = \overline{E}_1 \overline{E}_2 \dots \overline{E}_N(Q) - [\mathrm{Var}_1(p_2)/(\tau)] S_1.$$

Allen et al. (2006) propose that (8) represent the Beauty Contest metaphor since, the price is not equal to the market expectations of Q but rather, to the average expectation of what the market expects the average expected value of Q will be in the future. We comment on this interpretation in Section 4.4.

2.2   Applications with Dynamic Infinite Horizon Models

In the model of Section 2.1 trades can occur but it is not a truly dynamic model. Extensions to infinite horizon were developed for many applications and to get a sense of the issues involved we

shall review two very different applications. We start with Wang's (1994) study of trade volume.

Wang's (1994) objective is to overcome the no trade theorems of REE and explain the vast volume of trade in asset markets. With a Rational Expectations perspective, his hypothesis is that trade is the result of asymmetric PI. Using his notation, he assumes agent i maximizes expected utility over consumption flows $-E_t^i[\sum_{s=0}^{\infty} \beta^s e^{-\gamma c_{t+s}^i} | H_t^i]$ where expectations are conditioned on information of i. Wang (1994) assumes there are two assets with payoff in consumption units. A riskless asset with infinitely elastic supply which pays a constant rate r and where R = 1+ r. The second asset is a risky stock with a fixed supply set at 1 which pays a dividend D_t at date t. The law of motion of dividend is

$$D_t = F_t + \varepsilon_{D,t} \qquad \text{where} \qquad F_t = a_F F_{t-1} + \varepsilon_{F,t}.$$

$(\varepsilon_{D,t}, \varepsilon_{F,t})$ are i.i.d. normally distributed, zero mean shocks. Here F_t is the persistent component of the dividend process and $\varepsilon_{D,t}$ is the transitory component. The structure of information is intended to ensure that a closed form solution is possible. To that end Wang (1994) assumes there are two types of investors. I-investors have perfect private information and observe F_t. The U-investors receive only a noisy signal about F_t in the form $S_t = F_t + \varepsilon_{S,t}$ where $\varepsilon_{S,t}$ are i.i.d. normal, zero mean shocks. Since all investors observe the dividends, the I-investors observe the persistent as well as the transitory components of dividends while the U-investors observe neither.

In addition to the public asset, Wang (1994) assumes the I-investors have a private production technology which is risky and constant returns to scale. If they invest at t the amount I_t they receive at t+1 the amount $I_t(1 + r + q_{t+1})$ where excess return on the private technology q_{t+1} follows

$$q_{t+1} = \Xi_t + \varepsilon_{q,t+1} \qquad \text{and} \qquad \Xi_{t+1} = a_\Xi \Xi_t + \varepsilon_{\Xi,t+1}.$$

$(\varepsilon_{q,t+1}, \varepsilon_{\Xi,t+1})$ are i.i.d. normal, zero mean shocks. Expected excess return Ξ_t is observed only by the I-investors. This sharp information structure is called "Hierarchical" since it requires one class of investors to permanently have inferior information. The economy's structure is common knowledge and all agents are Bayesians with normal priors about parameters they do not know.

Two forces are used to explain the volume of trade. First, asymmetric information between the U and the I-investors, measured by $\sigma_S^2 = \text{var}(\varepsilon_{S,t})$. If $S_t = F_t$, information about the stock is *symmetric* and $\text{var}(\varepsilon_{S,t}) = 0$. When $\sigma_S^2 > 0$ we have $S_t \neq F_t$ and information is asymmetric. Second, the private technology of the I-investors is unavailable to the U-investors. The effect of this factor on asset demands operate via the correlation between private excess returns q_{t+1} and dividends, measured by $\sigma_{D,q}$ - the covariance of $\varepsilon_{D,t+1}$ with $\varepsilon_{q,t+1}$. To understand how this correlation impacts asset demands

and trade, suppose $\sigma_{D,q} \neq 0$ and Ξ_t increases leading I-investors to increase investments in private technology since expected return on such investments increased. But due to $\text{Cov}[\varepsilon_{D,t+1}, \varepsilon_{q,t+1}] \neq 0$, such increased investment changes the risk posture of their portfolio, calling for control of the risk by changing their investments in the publically traded stock. If $\sigma_{D,q} > 0$, control of risk leads to lower investments in the stock and if $\sigma_{D,q} < 0$, it leads to increased demand for the stock. The effect of asymmetric information about the private technology is thus due to the need of the I-investors to control their risk while the U-investors are unable to distinguish between changes in F_t and Ξ_t. It shows that the set-up of public and private technologies is crucial for Wang's (1994) results. Without private technology, just the asymmetry $S_t \neq F_t$ does not lead to trade since in this case the uninformed investors deduce F_t from prices, Wang's (1994) REE becomes fully revealing and we are back to no trade. With $\sigma_{D,q} \neq 0$ the price is linear in F_t and Ξ_t, uninformed investors are "confused" and cannot deduce either one from the price. This confusion of the U-investors makes it impossible for them to determine the cause of price changes. The U-investors now use the history of the process to conduct a Kalman Filtering in order to form expectation of their unobserved F_t. In sum, Wang (1994) shows that the model generates trade due to the exogenous shocks F_t and Ξ_t which cause time variability in the investment composition of the I-investors.

We pause briefly to examine the *causes* of price and volume volatility in models of noisy REE. In the earlier models equilibrium prices, such as (1a)-(1b), only vary with exogenous shocks to supply. In Wang (1994) prices vary only in response to exogenous shocks F_t and Ξ_t. This result continues to hold in all other dynamic models of PI such as He and Wang (1995). What is the effect of increased *diversity* of private information? In the earlier models private information was so diverse that the law of large numbers was invoked so that PI had no effect on prices. If diverse PI is to be the cause of trade one would expect that *increased* diversity of PI should *increase* the volume of trade. Wang (1994, page 145) shows that *increased* diversity of private information leads to a *decrease* in the volume of trade. This conclusion is a consequence of the fact that with increased diversity of PI uninformed traders have increased difficulties of deducing from prices information which will induce them to trade. In short, with diverse beliefs we find that in noisy REE the volatility of prices and the volume of trade are caused by exogenous shocks *but diverse PI does not cause or explains them.* We are thus back to the standard model without any amplification effect.

A second example is Woodford (2003) who revisits the Lucas (1972) model. It is motivated by

the fact that Lucas (1972) explains transitory effects of monetary policy but fails to account for the observed fact that monetary disturbances have persistent real effects. Woodford assumes agents are Dixit-Stiglitz monopolistic competitive price setters who select the nominal price of their product but cannot set the real price since they do not observe the aggregate price level and aggregate output. Although producers cannot observe the real price, their own output is determined by the real price. Aggregate *nominal* GNP is the exogenous state variable (e.g. determined by monetary policy). In equilibrium date t aggregate price level and aggregate output are functions of date t nominal GNP and of all higher order market expectations (i.e. market expectations of market expectations of ...) about it. As in Lucas (1972) agents cannot observe nominal GNP and receive private signals about it. Being rational they learn from the available information and as in Wang (1994) use a Kalman Filtering procedure to learn about the unobserved state variable. With incomplete learning of the persistent exogenous nominal GNP, Woodford (2003) demonstrates persistent money non-neutrality.

Limited space prevents our discussing other applications. Examples include Townsend (1978), (1983), Amato and Shin (2003), Hellwig (2002),(2005) and Angeletos and Werning (2006) who study business cycles, Morris and Shin (2002),(2005) who study the transparency of monetary policy, Singleton (1987) who study bond markets, Bacchetta and van Wincoop's (2006), and Hellwig, Mukherji and Tsyvinski (2006) who study the volatility of foreign exchange rates and many other policy oriented papers using Global Coordination games. These applications use persistent belief heterogeneity to explain the behavior of market *aggregates*. However, why is it asymmetric PI that should provide a basis for heterogeneity? With asymmetric PI agents clearly make different forecasts. Hence, there is the temptation to assume private information in order to model diversity and many authors have done just that. This is so common that for some, agents with different opinions are *synonymous* to agents with different private information. Such identification should be rejected. Private information is a very sharp sword that must be used with care. As we have seen, deducing information from prices is complicated and should be employed only when well justified. The virtual equivalence between belief diversity and private information is particularly wrong in macroeconomic applications when agents are assumed to have asymmetric PI with respect to aggregate variables such as interest rates or growth rate of GNP. We have serious doubts about the applicability of the PI paradigm to study asset market dynamics and will now pause to evaluate the results derived so far.

2.3   Is Asymmetric Private Information a Satisfactory Theory of Market Dynamics?

In questioning the use of asymmetric PI assumption we recall that phenomena studied with PI include market dynamics and volatility, aggregate risk premia, foreign exchange dynamics, business cycles, etc. In such models individual agents forecast aggregate variables. We thus break our query into two questions. First, is it reasonable to assume that economic agents have PI about such aggregate market variables? Second, are the *explanations* of these phenomena with PI valid and is asymmetric PI a persuasive model of market dynamics? We examine the first question by outlining 5 issues raised with noisy REE. We then turn to the second question.

(i) *What is the data that constitutes "private" information*? For the case of individual firms the nature of PI is clear and we discuss it under (ii). Now, if forecasters of GNP growth or future interest rates use PI, one must specify the data to which a forecaster has an exclusive access. Without it one cannot interpret a model's implications since all empirical implications are deduced from restrictions imposed by PI. In reality it is difficult to imagine the data which constitute private information.

(ii) *Without correlation private information explains little.* Even if some agents have some PI about some firms, an aggregate model may have no implications to market dynamics. To deduce any implications PI has to be repetitive over time, correlated and widespread. There is no empirical evidence for that. Indeed, all noisy REE models assume PI to be i.i.d. distributed and in that case private information has no effect on volatility, asset pricing or on any other dynamic characteristics.

(iii) *Asymmetric information implies a Secretive Economy.* Forecasters take pride in their models and are eager to make their forecasts public. Consequently, there are vast data files on market forecasts of many variables. In discussing public information forecasters explain their interpretation of such information which is often being framed as "their thesis." In contrast, an equilibrium with PI is secretive. Individuals are careful not to divulge their PI since it would deprive them of the advantage they have. In such an equilibrium private forecast data are treated as sources of *new information* used by all to update their own forecasts. The fact that forecasters are willing to reveal their forecasts is not compatible with PI being the cause of persistent divergence of forecasts.

(iv) *If private signals and noise are unobserved, how could common knowledge of the structure be attained and how can we falsify the theory?* To deduce PI from public data the *structure* of private signals x^i must be common knowledge. For example, in Section 2.1 $x^i = Q + \varepsilon^i$ where ε^i are pure

noise, independent across traders. One then asks a simple question: if these signals are not publically observed, how does common knowledge come about? How does agent i know his own signal takes the form $x^i = Q + \varepsilon^i$ and that the signal of k is $x^k = Q + \varepsilon^k$? Also, if the crucial data of a theory are not observable, how can one falsify the theory? What are then the true restrictions of the theory?

(v) *Why are private signals more informative than audited public statements?* Most results of models with PI are based on the assumption that private signals are more informative than public information. For example, in the model of Section 2.1 the public signal is $y = E(Q)$ where Q is unknown. Knowing y is inferior to knowing Q. The private signals are $x^i = Q + \varepsilon^i$ with ε^i i.i.d. and there is a continuum of traders. It is then assumed "the market" aggregates the private x^i and learns Q hence equilibrium price is a function of the unknown Q. This procedure raises two questions.

(a) Why are private signals more precise than the professionally audited public statements?

(b) Who does the aggregation and knows the i.i.d. structure needed for aggregation? If that agent is a neutral agent why does he not announce Q? Or, if not neutral, he should be part of the model.

One must conclude that asymmetric PI is not a persuasive assumption for modeling market dynamics. The answer to the second question is derived from the central conclusion that *diverse private information, by itself, has virtually no effect on market dynamics.* In all noisy REE with diverse and independent private signals, asymmetric PI has no impact either on price volatility or on volume of trade. In general, *increased* diversity of private information *decreases* volatility and volume of trade. In any noisy REE all dynamic characteristics are fully determined either by the standard exogenous shocks such as dividends or by exogenous "noise" which is often questionable if it is unobserved. Since we aim to explain excess volatility of markets with mechanisms of endogenous *amplification*, it follows that the asymmetric PI paradigm does not offer such amplification but rather, it leads us back to the traditional causes of market dynamics.

To close this discussion we ignore, for the moment, issues of dynamics and focus on the *economic explanation* given by models with asymmetric PI to known problems. By way of introducing the next Section we pick two standard problems and contrast the proposed solutions of then by noisy REE models with asymmetric PI and the solution of them by models with diverse belief but without PI.

(I) Asset pricing. The noisy REE asset pricing theory (e.g. Allen et al. (2006)) proposes that an asset has a well defined, commonly recognized, intrinsic value and agents have private information about it.

They cannot deduce that value from prices since unobservable supply shocks make prices "noisy." Models with diverse belief and no PI insist an asset which produces future payoffs *does not have a well defined intrinsic value*. Market price merely expresses the traders' beliefs of the present value of future payoffs hence asset prices may exhibit excess volatility explained by the dynamics of beliefs.

(II) Volume of trade. Wang's (1994) explanation for the volume of trade follows from the assumption that informed agents observe the permanent component of the dividend process and have private technologies for investments with returns which are correlated with the returns on the publically traded assets. This technology is unavailable to uninformed investors who do not observe the permanent component of dividends with precision. Uniformed investors cannot distinguish between changes to returns on public assets and on the private technologies. Informed investors manage their portfolios' risk by shifting assets between public assets and investments in their private technology and these "rebalancing" shifts cause trade of public assets. Models of diverse beliefs without PI explain the volume of trade very differently. In such models agents do not conceal their beliefs since other agents take them as "opinions" rather than new information requiring updating of their own beliefs. Hence, all trades result naturally from agents' disagreements about future prospects of assets. Consequently, *asset markets are the arena where agents trade their diverse beliefs* and the volume of trade reflects the intensity of these differences and the dynamics of the distribution of market beliefs over time.

3.   Diverse Beliefs with Common Information: the General Theory.

Rational Expectations and Behavioral Economics have staked out two extreme positions in contemporary thinking. Under Rational Expectations people know all structural details needed for perfect forecasting while under Behavioral Economics they are motivated by psychological impulses that lead to irrational behavior. The theory of Rational Belief proposes an intermediate concept of rationality that begins with the imperfection of human knowledge. It continues with the idea that people optimize given the limited knowledge they have and concludes with a recognition that without perfect knowledge, rational errors do take place. Rational errors may be magnified to a point where changing perceptions dominate public life and asset markets. This is the road to rational diverse beliefs and endogenous amplification which we explore. Before proceeding we mention early papers such as Harrison and Kreps (1978), Varian (1985), (1989), Harris and Raviv (1993), Detemple and Murthy (1994), Kandel and Pearson (1995), Cabrales and Hoshi (1996). These, together with the early writers

mentioned in the Introduction have recognized the importance of diverse expectations. The reason we do not review them here in detail is that they did not anchor the theory with a concept of rationality.

3.1   The Essential Idea: Stability But Not Invariance

In a stationary economy structural relations are fixed and joint probabilities are time invariant. In such an economy the Ergodic Theorem holds: time averages equal expected values under the true probability that is deduced from the relative frequencies of events. In short, the empirical distribution reveals the truth and since human history is long, agents learn the true structure from the empirical distributions of observed variables. The fact is that what we learn from data depends upon the period covered. This is so since the economy has undergone changes in all aspects of life. These changes are rapid with major effects on the nature of commodities, productivity and asset returns. Consequently, if we denote by $\{ d_t, t = 1, 2, ... \}$ the vectors of observed data on asset payoffs then the process is non-stationary, which requires the distributions of the d_t's to be time dependent[4]. But non-stationarity is more than considering the d_t's as a sequence of "regimes." In addition to changing distribution of returns, assets are different over time, reflecting the technology they embody. Such variability makes it impossible to learn the true and unknown probability Π on infinite sequences $\{ d_t, t = 1, 2, ... \}$. With such changes, what is the regularity one uses for scientific analysis? The answer is simple but central. Kurz (1994) postulates that although the economy is not stationary, it exhibits statistical stability so that empirical distributions of observed variables exist. That is, the long term statistics exhibit empirical regularity so that moments and relative frequencies of events exist. We give a formal definition.

Suppose there are N observable variables and let $x_t \in X \subseteq \mathbb{R}^N$ be a vector of observations. Let $x = (x_0, x_1, x_2, ..)$ be a history from date 0 to infinity and $x^t = (x_t, x_{t+1}, x_{t+2}, ...)$ a history from t on. To define *Statistical Stability* let X^∞ be the space of infinite sequences x and $\mathcal{B}(X^\infty)$ be the Borel σ-field of X^∞. For a finite dimensional set (cylinder set) $B \in \mathcal{B}(X^\infty)$ define

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

---

[4] The technical definition of "non-stationary" which we use requires the process to be time dependent, and this is the customary terminology in Ergodic Theory and Stochastic Processes. It is different from the use of this term in the Time Series literature which requires the process to have infinite variance.

where
$$1_B(v) = \begin{cases} 1 & \text{if} \quad v \in B \\ 0 & \text{if} \quad v \notin B . \end{cases}$$

Although the set B is finite, it can be a very complicated time dimensional set.

Definition 1:   A stochastic process $\{x_t , t = 0, 1, 2, ...\}$ with a true probability Π on $(X^\infty, \mathcal{B}(X^\infty))$ is said to be *Statistically Stable* if for each finite dimensional cylinder set $B \in \mathcal{B}(X^\infty)$

$$\lim_{n \to \infty} m_n(B)(x) = m(B)(x) \text{ exists} \qquad \Pi \text{ a.e.}$$

If it is Ergodic then we have

$$m(B)(x) = m(B) \text{ independent of } x, \quad \Pi \text{ a.e.}$$

The restriction to finite dimensional sets results from the fact that we have only finite data. However, in the ergodic case probabilities on the collection $\{ m(B) , B$ finite dimensional$\}$ can be extended to a full probability measure m on the space $(X^\infty, \mathcal{B}(X^\infty))$. Indeed, we know (see Kurz (1994)) that (i) m is unique, and (ii) m is stationary and hence is called "*the stationary measure of Π.*"

We thus make a fundamental observation. We assume our dynamic economy is non-stationary but statistically stable and ergodic. In this economy agents do not know the true probability Π. They discover from the data[5] the probability m induced by the dynamics under Π. If the economy is stationary then $m = \Pi$ *but in general* $\Pi \neq m$. Since m is deduced from public data, there is no disagreement about it: m is *the common empirical knowledge*. The basis for diverse beliefs is that Π is not known. The discipline for rationality of belief is the common knowledge of m. Basic rationality will then naturally dictate that a belief should not contradict the empirical evidence m.

An Example   Agents observe a Black Box generating numbers x_t , t = 0, 1,... in $\{0, 1\}$ without serial correlation between x_t and x_{t+k} all k. Using a long data set they find the mean is ½. The probability measure m is the i.i.d. measure on infinite sequences of $\{0, 1\}$. If the box contains a single coin the

---

[5] As noted earlier, although we have finite data which enable agents to compute $m_n(B)(x)$, the data are assumed long enough and the probability m is assumed simple enough (i.e. Markov with short memory) so that with finite data agents can get a good approximation for the limit measure m. The probabilities $m(B)$ are independent of x due to ergodicity. The assumption that agents know the limits should not be interpreted to mean we assume agents have an infinite sequence of observations since in that case agents will consider not only limits on sequences but also limits on all infinitely many possible subsequences. With finite data one can observe only a finite number of subsequences at all dates and for this reason we do not incorporate restrictions that would be implied by limits on subsequences. We also note that in the non-ergodic case the data requirement is greater, since then we need data for many alternative data sequences x with different starting points, but the basic theory remains unchanged. For details see Kurz (1994).

observed sequence has an empirical distribution of an i.i.d. fair coin, which would be the truth. What other processes generate the same empirical measure? As an example consider a *two coin* family using realizations of a process $\{g_j, j = 0,1, ...\}$ of i.i.d. random variables in $\{1, 2\}$ with probability of 1 being, say, 1/3. Pick an infinite sequence $g^* = (g_0^*, g_1^*,...)$. These realizations $g_j^* = 1$ or $g_j^* = 2$ are treated as *parameters* of a new process. It is defined by a process $\{v_t, t = 0, 1,...\}$ with two i.i.d. coins in the Box which appear at different times depending upon the $g^* = (g_0^*, g_1^*,...)$. $\{v_t, t = 0, 1, ...\}$ where $v_t \in \{0, 1\}$ is a sequence of *independent* random variables satisfying

$$
(9) \qquad P\{v_t = 1\} = \begin{cases} 0.60 & \text{if } g_t^* = 1 \quad (\text{coin type } 1) \\ 0.45 & \text{if } g_t^* = 2 \quad (\text{coin type } 2). \end{cases}
$$

Since $(1/3)(0.60) + (2/3)(0.45) = 0.50$, the empirical distribution is the same as m for almost all g^* and v. It is easy to see that instead of two possible "regimes" we could have an infinite number of regimes. Note that we have not even specified the true probability Π.

The stationary probability m is then merely an average over an infinite sequence of changing regimes. It is not the true probability of the process. Belief diversity starts with the disagreement over the meaning of public information. Agents know the economy is a complex process and believe $\Pi \neq m$. It leads them to construct models which express the different circumstances that prevail at date t from the circumstances specified under m. Beliefs are then statements on how an agent's assessments deviate from the empirical frequencies. But rationality requires more. It recognizes that statistical stability implies an agent cannot deviate from the empirical frequencies forever, hence an agent's long run average deviations from the invariant forecasts must be zero.

We next seek to describe beliefs in a tractable analytical manner. We carry out this program with a specific model used in most applications reported here. Thus, consider an asset or a portfolio of assets paying an exogenous risky payoff sequence $\{D_t, t = 1, 2,...\}$ with a non-stationary and unknown true probability. To keep exposition simple we assume that with a long history of data the empirical distribution of the D_t's is shown to constitute a Markov process with the following transition

$$
D_{t+1} = \mu + \lambda_d(D_t - \mu) + \rho_t^d \qquad , \qquad \rho_t^d \sim N(0, \sigma_d^2)^{[6]}
$$

---

[6] In many applications it is assumed the D_t grow without bound, do not have a finite mean and d_t is the growth rate of dividends with an empirical distribution of a Markov process with a mean μ. The same applies to other statistically stable processes with trends, in which cases the concept of stability is applied to growth rate data.

Defining $d_t = D_t - \mu$ we note that $\{d_t, t = 1, 2, ...\}$ is a zero mean, non-stationary process with unknown true probability Π and a stationary empirical probability m. That is, the process $\{d_t, t = 1, 2, ...\}$ has an empirical distribution which implies a transition function defined by the first order Markov process

(10a) $$d_t = \lambda_d d_{t-1} + \rho_t^d \qquad , \qquad \rho_t^d \sim N(0, \sigma_d^2).$$

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

Many studies of the equity risk premium, discussed later, assume that

$$D_{t+1} = d_{t+1} D_t \text{ - where } d_t \text{ is the random } \textit{growth rate} \text{ of dividends.}$$

For the growth case, Mehra and Prescott (1985) estimate the empirical distribution of $\{d_t, t = 1, 2, ...\}$ is represented by a stationary and ergodic Markov process which takes only two values. Hence, the state space is $J_D = \{d^H, d^L\}$. The two values reflect exogenous business cycle environments. They estimate the two values are $d^H = 1.054$ and $d^L = 0.982$ with a transition matrix

(10b) $$\begin{bmatrix} \varphi, \ 1 - \varphi \\ 1 - \varphi, \ \varphi \end{bmatrix} \quad , \quad \varphi = .43.$$

Is the stationary model (10a) the true process? Those who believe the economy is stationary would accept (10a) as the truth. Most agents do not believe past empirical record is adequate to forecast the future and this leads to non-stationary and diverse beliefs. The problem is then how do we describe an equilibrium in such an economy? The structure of belief is thus our next topic.

3.2 Modeling Heterogeneity of belief I: Individual Beliefs as Markov State Variables

Our approach to diversity raises a methodological question. In formulating an asset pricing theory do we need to describe in detail each agent's model? Also, are detailed descriptions necessary for the study of asset price dynamics? Although an intriguing question, we suggest that such details are not needed. To describe an equilibrium all we need is to specify how beliefs affect agents' perceived stochastic transition of state variables. Once specified, Euler equations are well defined and market clearing leads to equilibrium pricing. The key step (see Nielsen (1996), Kurz (1997a), Kurz and Wu (1996) and Kurz and Motolese (2001) (2006)) is to treat individual beliefs as state variables, *generated within the economy*. Here we adapt the original concept of Nielsen (1996), ideas of Kurz, Jin and Motolese (2005a), (2005b) and Kurz and Motolese (2006) to the problem and explain it now.

Start with the fact that agents who hold heterogenous beliefs willingly reveal their forecasts. We thus assume that distributions of market forecast data are publicly observable. An individual's belief is described with a *personal* state of belief which uniquely pins down his perception of the

stochastic transition of next period's state variables. It follows that personal state variables and economy-wide state variables are not the same. A personal state of belief is analogous to other state variables in the agent's decision problem but can also be interpreted as defining the familiar notion of a "type" of an agent who is uncertain of his future belief type but knows the dynamics of his personal belief state. The *distribution* of individual belief states is then an economy-wide state variable which, as indicated earlier, is observable. Endogenous variables depend upon the economy's state variables and these include the distribution of beliefs. Hence, moments of the market distribution of beliefs have an effect on endogenous variables such as prices. Also, in a large economy an agent's *"anonymity"* implies that a personal state of belief is perceived to have a negligible effect on prices. Many papers assume an exponential utility without income effects. In that case, equilibrium endogenous variables depend only on the *mean market state of belief and we have full aggregation of beliefs*. Finally, to forecast future endogenous variables an agent must forecast the beliefs of other agents tomorrow.

3.2.1   *Individual States of Belief*

We thus introduce agent i's *state of belief* g_t^i. It describes his perception by pinning down his transition functions. Apart from "anonymity" we assume agent ℓ knows his own g_t^ℓ and the market *distribution* of g_t^i across i. As for the past, he observes only past *distributions* of g_τ^i for $\tau < t$ hence he knows past values of all moments of the distributions of g_τ^i. This last assumption is justified by the fact that an infinite horizon economy consists of a sequence of decision makers. An agent knows his states of beliefs but does not know the states of belief of *all* his own specific predecessors. Past belief distributions are public information available to all since samples of g_t^i are made public at each date t[7]. We specify the dynamics of g_t^i by

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

where ρ_t^{ig} are correlated across i reflecting correlation of beliefs across individuals. In short, the *state of belief* is a central concept and (11) is taken as a primitive description of type heterogeneity. One

---

[7] Individual states of belief are described either with finite state models or with continuous state models. Continuous state models tend to be more complex than discreet state models which are more tractable. Using the provided sources the reader can implement the computational programs of either type of model, but the results are essentially the same. To avoid repetition we report here detailed results deduced only for the continuous state models. Since a reader may find either one of these two more suitable for his\her application, we describe in the text the basic structure of *both* models. This results in a somewhat discontinues presentation style. Hence, we recommend that upon first examination of the material *the reader skips the sections on finite state modeling* and study these only after covering the full development of the continuous state model together with the numerical results of the simulations derived for that model.

can, however, deduce (11) from more elementary principles (see Section 3.2.2 below).

In Example (9), finite belief states are described by a parameter sequence $\{g_j^* , j = 0,1, ...\}$ in $G=\{1, 2\}$. Hence in (9) a belief is a Dirichlet distribution in $(G^\infty, \mathcal{B}(G^\infty))$, an assumption first used by Nielsen (1996). In other applications belief states are random, assuming agents do not know their own future types with certainty. Kurz (1997c), Kurz and Beltratti (1997), Kurz and Schneider (1996), Kurz and Motolese (2001), (2006), Motolese (2003), Nielsen (1996),(2003), Nakata (2007), and Wu and Guo (2003) use finite lived agents (OLG models) where a belief state y^i of i is either i.i.d. or Markov with a 2×2 transition matrix. The marginal probability $Q^i\{g_t^i = 1\} = \alpha_i$ plays a central role in the model since it is the frequency at which agent i is in belief state 1.

How does g_t^i pin down the stochastic transition? In various models 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

(12a) $$ d_{t+1}^i = \lambda_d d_t + \lambda_d^g g_t^i + \rho_{t+1}^{id} \qquad , \qquad \rho_{t+1}^{id} \sim N(0, \hat{\sigma}_d^2). $$

The assumption that $\hat{\sigma}_d^2$ is the same for all i is made only for simplicity. An agent who believes the empirical distribution is the truth expresses it by $g_t^i = 0$. It follows that the state of belief g_t^i measures the deviation of his forecast from the empirical stationary forecast

(12b) $$ E^i[d_{t+1}^i | H_t, g_t^i] - E^m[d_{t+1} | H_t] = \lambda_d^g g_t^i. $$

(12b) shows how g_t^i is measured in practice. For any X_t, publicly available data on i's forecasts of X_t measure $E^i[X_{t+h} | H_t, g_t^i]$. One then uses standard techniques to construct the stationary forecast $E^m[X_{t+h} | H_t]$ used to constructs the difference in (12b). Fan (2006) and Kurz and Motolese (2006) study markets with data constructed this way. Figures 1, 2 and 3 illustrate the time series of variables as in (12b), averaged over the population with horizon of h = 6 month for the 6 month Treasury Bill rate, change in the GDP deflator measuring the inflation rate and growth rate of Industrial Production.

Figures 1, 2 , 3 Place Here

In Figures 1 and 3 market belief fluctuates around zero as predicted by the theory *even during the short period at hand*. In Figure 2 this pattern is not exactly maintained by market belief about inflation. This was studied by Kurz (2005) who argues the persistent deviation of inflation forecasts during the 1980's and early 1990's from the normal pattern was due to persistent money neutrality ideas and erroneous conviction in academic circles and some Fed members that monetary policy cannot control

inflation. Forecasted inflation was consistently higher than actual inflation for a fairly long period of time. All three figures are compatible with the Markov property assumed in (11) (and later in (20)).

We add the observation that since belief variables arise from structural change, g_t^i estimated for 1900 has nothing to do with g_t^i in 2000: they are caused by different factors and reflect different societies and technological environments. The ability of past g_τ^i to forecast then tell you nothing about the predictive value of future g_t^i. To sharpen the contrast with models of asymmetric PI observe that g_t^i describes the *opinion* of agent i. Hence, other agents do not treat g_t^i as information about unknown structural parameters. However, the g_t^i are correlated across agents and their mean is observable but is not treated as information about a parameter, as is the case with asymmetric PI.

In the case of a finite state space an individual belief state pins down a transition function in a more complex way. Thus, consider models with two dividend states and two agent types. Within each type they are identical. Such an economy has 8 social states defined by the following equivalence:

(13)
$$
\begin{bmatrix} 1 \\ 2 \\ 3 \\ 4 \\ 5 \\ 6 \\ 7 \\ 8 \end{bmatrix} \Leftrightarrow
\begin{bmatrix}
d_1 = d^H, \; g^{11} = 1, \; g^{12} = 1 \\
d_2 = d^H, \; g^{21} = 1, \; g^{22} = 2 \\
d_3 = d^H, \; g^{31} = 2, \; g^{32} = 1 \\
d_4 = d^H, \; g^{41} = 2, \; g^{42} = 2 \\
d_5 = d^L, \; g^{51} = 1, \; g^{52} = 1 \\
d_6 = d^L, \; g^{61} = 1, \; g^{62} = 2 \\
d_7 = d^L, \; g^{71} = 2, \; g^{72} = 1 \\
d_8 = d^L, \; g^{81} = 2, \; g^{82} = 2
\end{bmatrix}.
$$

d^H is the "high" and d^L is the "low" dividend states. (13) shows that Endogenous Uncertainty[8] requires an expansion of the state space and beliefs play a central role in price volatility (see Kurz and Wu (1996)). Under the Markov assumption, the empirical distribution with 8 states is described with an 8×8 stationary transition matrix denoted Γ. It is required to have the dividend transition matrix (10b) as a marginal probability. It turns out that Γ must take the form

(14)
$$
\Gamma = \begin{bmatrix} \varphi A, \; (1 - \varphi) A \\ (1 - \varphi) B, \; \varphi B \end{bmatrix}.
$$

---

[8] Although "Endogenous Uncertainty" is defined here as the excess price volatility due to the effect of beliefs, Kurz and Wu (1996) introduce a precise definition of this term in the context of a General Equilibrium model. Exactly as in (13), Kurz and Wu (1996) define the term as a property of the price map which has multiple prices for the same exogenous state.

where A and B are 4×4 matrices that must take the form

(14a) $A = \begin{bmatrix} a_1, & \alpha_1 - a_1, & \alpha_2 - a_1, & 1 + a_1 - \alpha_1 - \alpha_2 \\ a_2, & \alpha_1 - a_2, & \alpha_2 - a_2, & 1 + a_2 - \alpha_1 - \alpha_2 \\ a_3, & \alpha_1 - a_3, & \alpha_2 - a_3, & 1 + a_3 - \alpha_1 - \alpha_2 \\ a_4, & \alpha_1 - a_4, & \alpha_2 - a_4, & 1 + a_4 - \alpha_1 - \alpha_2 \end{bmatrix}$, $B = \begin{bmatrix} b_1, & \alpha_1 - b_1, & \alpha_2 - b_1, & 1 + b_1 - \alpha_1 - \alpha_2 \\ b_2, & \alpha_1 - b_2, & \alpha_2 - b_2, & 1 + b_2 - \alpha_1 - \alpha_2 \\ b_3, & \alpha_1 - b_3, & \alpha_2 - b_3, & 1 + b_3 - \alpha_1 - \alpha_2 \\ b_4, & \alpha_1 - b_4, & \alpha_2 - b_4, & 1 + b_4 - \alpha_1 - \alpha_2 \end{bmatrix}$

If $A \neq B$ the distribution of (g_{t+1}^1, g_{t+1}^2) depends upon d_t . (14a) implies the marginal probabilities $Q^i\{g_t^i = 1\} = \alpha_i$ for $i = 1, 2$. Now, a belief of i is a conditional probability, or a transition function, $Q^i((\bullet)| g^i)$. Since g_t^i takes two values, i must be using two different transition matrices with the rule

Agent 1:Use $F_1 = Q^1((\bullet)| g_t^1 = 1)$ if $g_t^1 = 1$ **Agent 2**: Use $G_1 = Q^2((\bullet)| g_t^2 = 1)$ if $g_t^2 = 1$

Use $F_2 = Q^1((\bullet)| g_t^1 = 2)$ if $g_t^1 = 2$ Use $G_2 = Q^2((\bullet)| g_t^2 = 2)$ if $g_t^2 = 2$.

Hence, $Q^i\{g_t^i = 1\} = \alpha_i$, measure the frequencies at which the agents use matrix F_1 or G_1. But what are the matrices? We later explain that, analogous to the Example in **3.1**, Rationality of Belief requires

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

but this leaves some room for choice. In the simulation models of asset volatility Kurz (1997c), Kurz and Beltratti (1997), Kurz and Schneider (1996), Kurz and Motolese (2001), (2006), Motolese (2001), (2003), Nakata (2007) and Wu and Guo (2003) use the simple parametrization

(16) $F_1 = \begin{bmatrix} \varphi A_1(\lambda) & , & A_2(\lambda) \\ (1 - \varphi)B_1(\lambda), & B_2(\lambda) \end{bmatrix}$ $G_1 = \begin{bmatrix} \varphi A_1(\mu) & , & A_2(\mu) \\ (1 - \varphi)B_1(\mu), & B_2(\mu) \end{bmatrix}$.

(λ, μ) are parameters that shift probabilities of high dividend states. When $\lambda > 1$ an agent using F_1 is optimistic about d_{t+1}: his conditional probability of $d_{t+1} = d^H$ is higher than the stationary probability implied by Γ. If $\lambda < 1$ his conditional probabilities of $d_{t+1} = d^H$ are lower than the probability implied by Γ. Once (F_1, G_1) are given, these authors use rationality (15) to determines the second matrix since by (15) $F_2 = \dfrac{1}{1 - \alpha_1}(\Gamma - \alpha_1 F_1)$, $G_2 = \dfrac{1}{1 - \alpha_2}(\Gamma - \alpha_2 G_1)$. The results of these studies are reported later.

3.2.2   Can One Deduce (11) $g_{t+1}^i = \lambda_Z g_t^i + \rho_{t+1}^{ig}$ *from Bayesian Inference*?

The non-stationary model at hand has some fixed parameters, deduced from the empirical frequencies, and others that change over time. The time varying parameters are modeled by specifying that under the true probability Π the value d_t has a stochastic transition function such as

(17)
$$d_{t+1} = b_t + \lambda_d d_t + \rho_t^d.$$

The parameters b_t are the exogenous, time varying mean values of $d_{t+1} - \lambda_d d_t$ and g_t^i are beliefs about b_t. The problem is that in a changing environment there is no universal method to learn an unknown sequence of parameters. The advantage of the description (11) $g_{t+1}^i = \lambda_z g_t^i + \rho_{t+1}^{ig}$ is that it leads to a simple analysis of asset pricing. It requires each agent to have a distinct state space to describe his own perceived uncertainty and it dictates an endogenous expansion of the economy's state space. (11) is thus viewed as a primitive fundamental, like utility functions. But can (11) be deduced from more elementary and familiar principles of Bayesian inference?

In a traditional Bayesian environment an agent faces data generated by a stationary structure but with an unknown fixed parameter. The agent starts with a prior belief 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. If learning by an optimizing agent takes place in real time the agent must use the prior belief to forecast all variables but learning can improve only *future* forecasts. In the case at hand agents know λ_d but not b_t hence a Bayesian inference about the unknown varying parameters b_t in $d_{t+1} = b_t + \lambda_d d_t + \rho_t^d$ limits the validity of Bayesian updating.

To understand the problem observe that at date t an 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 when parameters change. If b_t change slowly, the posterior estimate of b_t (given d_{t+1}) may serve as a prior belief about b_{t+1}. Indeed, if the agent knew that $b_t = b_{t+1}$ the *posterior of b_t is the best prior of b_{t+1}.* In the absence of such knowledge, agents know that $b_t = b_{t+1}$ is only one possibility. They would then seek any additional information and use subjective interpretation of public data to arrive at alternative sharper subjective prior of b_{t+1}. Kurz (2007) shows that 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 for diverse interpretation. In the Appendix we explain the argument which derives (11) from a Bayesian inference. In such inference an agent inserts some judgment into the Bayesian inference about b_{t+1}, based on subjective interpretation of qualitative public information. The general point made by Kurz (2007) (explained in the Appendix) is that a Bayesian inference is very sensitive to any small deviations based on subjective assessment of public data. Any such persistent subjective deviations lead the posterior to fluctuate forever.

3.3 Modeling Heterogeneous beliefs II: Market Belief and Rationality of Beliefs

3.3.1 Market Beliefs and Individual Perception Models

We assume the market is large and anonymous and the distribution of beliefs is observable hence its moments are known. Let Z_t be the first moment and refer to it as "average market belief" computed by aggregating (11). Due to correlation across agents, the law of large numbers is not operative and the average of ρ_t^{ig} over i does not vanish. Hence we have

(20)
$$Z_{t+1} = \lambda_Z Z_t + \tilde{\rho}_{t+1}^Z.$$

Correlation across agents may exhibit non stationarity, a property that would be inherited by the $\{\tilde{\rho}_{t+1}^Z, t = 1, 2, ...\}$ process. Since the Z_t are observable, market participants have data on the *joint process* $\{(d_t, Z_{t+1}), t = 1, 2, ... \}$ hence they know their *joint empirical distribution*. We assume this distribution is described by the system of equations

(21a) $\quad d_{t+1} = \lambda_d d_t + \rho_{t+1}^d$

(21b) $\quad Z_{t+1} = \lambda_Z Z_t + \rho_{t+1}^Z$

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

(21a)-(21b) is the continuous state space analogue of the matrix Γ in the finite case.

We now explain agent i's perception model, which is analogous to (16) in the finite case. In (12a) g_t^i pins down agent i's forecast of d_{t+1}^i. We now broaden this idea to a perception model of the two state variables (d_{t+1}^i, Z_{t+1}^i) given d_t and Z_t. His belief then takes the general form

(22a) $\quad d_{t+1}^i = \lambda_d d_t + \lambda_d^g g_t^i + \rho_{t+1}^{id}$

(22b) $\quad Z_{t+1}^i = \lambda_Z Z_t + \lambda_Z^g g_t^i + \rho_{t+1}^{iZ}$

(22c) $\quad g_{t+1}^i = \lambda_Z g_t^i + \rho_{t+1}^{ig}$

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

Although g_t^i defines belief about future d_{t+1}, (22a)-(22b) show that we use it also to pin down the transition of Z_{t+1}^i. This simplicity ensures one state variable pins down agent i's subjective belief. Hence, g_t^i expresses how agent i views conditions at t as different from the empirical distribution

(22d)
$$E_t^i \begin{pmatrix} d_{t+1} \\ Z_{t+1} \end{pmatrix} - E_t^m \begin{pmatrix} d_{t+1} \\ Z_{t+1} \end{pmatrix} = \begin{pmatrix} \lambda_d^g g_t^i \\ \lambda_Z^g g_t^i \end{pmatrix}.$$

The average market belief operator is defined by $\bar{E}_t(\bullet) = \int E_t^i(\bullet) di$. From (22c) it is

(22e)
$$\bar{E}_t \begin{pmatrix} d_{t+1} \\ Z_{t+1} \end{pmatrix} - E_t^m \begin{pmatrix} d_{t+1} \\ Z_{t+1} \end{pmatrix} = \begin{pmatrix} \lambda_d^g Z_t \\ \lambda_Z^g Z_t \end{pmatrix}.$$

A comment on market belief. From the perspective of an agent, Z_t is a state variable like any other and used to forecast prices like any other macro data used to assess risk of a recession. But market belief may be wrong; as we have seen it forecasted more recessions than actually occurred. Market risk premia may fall just because agents are more optimistic about the future, not necessarily because there is any objective data which convinces everybody that the future is bright. But now, how do agents update beliefs about d_{t+1} given Z_t? In contrast with PI theories, agents do not use Z_t to update beliefs about d_{t+1}: (22a) specifically *does not* depend upon Z_t. Agents do not consider Z_t information about d_{t+1} since it is not a "signal" about unobserved private information of others. All agents know that all market participants use the same public information. They do consider Z_t as "news" about what *the market* thinks about d_{t+1}! Z_t is thus important for 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 Z_t is useful for forecasting future endogenous variables. We shall later examine how this process works in equilibrium.

3.3.2 Higher Order Market Beliefs.

One must distinguish between higher order beliefs which are *temporal* and those which are *contemporaneous*. (22a)-(22c) defines agent i's probability over future sequences of (d_t, Z_t, g_t^i) and as is the case for any probability, it implies i's temporal higher order beliefs with regard to future events. For example, one deduces from (22a)-(22c) statement like

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

Temporal higher order beliefs are thus simple properties of conditional expectations. In addition, since (20) is implied by (22c), average market belief operator satisfies $\bar{E}_t(d_{t+N+1}) = \lambda_d \bar{E}_t(d_{t+N}) + \lambda_d^g \bar{E}_t(Z_{t+N})$. One deduces perceived higher order temporal market beliefs by averaging over i. For example,

$$\lambda_d^g \lambda_Z \bar{E}_t(Z_{t+N}) = \bar{E}_t \bar{E}_{t+N+1}(d_{t+N+2}) - \bar{E}_t E_{t+N+1}^m(d_{t+N+2}).$$

(22a)-(22c) also show that properties of conditional probabilities do not apply to the market belief operator $\bar{E}_t(\bullet)$ since it is *not a proper conditional expectation*. To see why let $X = D \times Z$ be a space where (d_t, Z_t) take values and G^i be the space of g_t^i. Since i conditions on g_t^i, his unconditional probability is a measure on the space $((D \times Z \times G^i)^\infty, \mathscr{F}^i)$ where $\mathscr{F}^i$ is a sigma field. The market conditional belief operator is just 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*! This is formulated as:

Theorem 2 : The market belief operator violates iterated expectations: $\overline{E}_t(d_{t+2}) \neq \overline{E}_t\overline{E}_{t+1}(d_{t+2})$.
Proof: See Kurz (2007).

Contemporaneous higher order beliefs have attracted attention (e.g. Allen, Morris and Shin (2006), Bacchetta, and van Wincoop (2005) and Woodford (2003)) despite being unobservable. They occur naturally in strategic situations. In a market context they can *formally* arise in (22a)-(22c) as follows. Let Z_t in (20) be defined as Z_t^1. Uncertainty about this variable may lead agents to form belief about its future by using a *second* belief index g_t^{i2} about Z_{t+1}^1 whose transition is deduced from the transition of g_t^{i2}. Now $Z_t^2 = \int g_t^{i2} di$ is a second order aggregate belief for which a *third* belief index g_t^{i3} can be introduced, whose average is Z_t^3, etc. Such infinite regress is problematic and leads us to reject contemporaneous higher order beliefs in markets for two reasons. First, higher order beliefs are degenerate in (22a)-(22c) *because* the single belief index g_t^j fully pins down agent i's belief. Moreover, since agents know the beliefs of others and all variables in the price map (which embody the beliefs of others), there is nothing else to form beliefs about. In response to the formal construction above we now offer a second and more general reason for why in markets *all higher order 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, they can only exist in the minds of the agents and hence there is no possible mechanism for individual g_t^{ij} to be correlated as in (20). Hence, higher order beliefs cannot have an aggregate effect since with independent g_t^{ij} the averages Z_t^j for j > 1 are zero at all t.

3.3.3   Belief Rationality and the Conditional Stability Theorem

Statistical stability of the data implies that individual rationality requires a belief to be stable. But then, is the belief (22a)-(22c) stable? Is there a general principle to connect The Example in 3.1 with (22a)-(22c) and offer a general outlook on stable but non stationary systems? The conditional stability theorem (Kurz and Schneider (1996)) provides the answer. We now explain this theorem.

Although the theorem holds for dynamical systems we avoid formalities and discuss stochastic processes only. Consider an economy with N observables $x_t \in X \subseteq \mathbb{R}^N$. We characterize a general family of non-stationary processes described by sequences of parameters in G. To that end consider a

joint process $(x_t, g_t) \in X \times G$. Define $\Omega = (X \times G)^\infty$ and let $B = B((X \times G)^\infty)$ be its Borel σ-field, and consider a joint process $\{(x_t, g_t), t = 1, 2, \ldots\}$ with true probability Ξ on the probability space (Ω, B). We assume the process $\{(x_t, g_t), t = 1, 2, \ldots\}$ is stationary although it is sufficient for it to be stable, hence the *primitive assumption* is that the joint system is either stable or stationary. It expresses the idea that the statistical properties of the parameters g are interrelated with the statistical properties of the data x but both are derived as a consequence of the stability of the joint system. In relation to the true data generating process, the joint system can be either the true unobserved law of motion of non-stationarity, in which case the $g_t \in G$ are unobserved parameters. Or else, which is the way we use it here, the joint process is a *model* which a rational agent use to formulate his belief. To proceed we need two technical definitions. First, let Ξ_g be the *conditional probability* of x given g. That is

$$\Xi_g(\bullet) : G^\infty \times B(X^\infty) \to [0, 1]$$

such that for each $A \in B(X^\infty)$, $\Xi_g(A)$ is a measurable function of g and for each g, $\Xi_g(\bullet)$ is a probability on $(X^\infty, B(X^\infty))$. Second, Ξ_X is the *marginal measure* of Ξ on X and is defined by

$$\Xi_X(A) = \Xi(A \times G^\infty) \quad \text{for all} \quad A \in B(X^\infty).$$

The joint is thus a global process under Ξ. The data $\{x_t, t = 1, 2, \ldots\}$ is then considered as generated under the conditional probability Ξ_g parametrized by g while the marginal Ξ_X is the average over g.

Theorem 3: (Conditional Stability Theorem, (Kurz and Schneider (1996))) Suppose the joint process $\{(x_t, g_t), t = 1, 2, \ldots\}$ with probability Ξ on (Ω, B) is stationary and ergodic then

   (a) Under the conditional probability Ξ_g the process $\{x_t, t = 1, 2, \ldots\}$ is stable and ergodic for
   Ξ a.a. g. The stationary measure of Ξ_g is denoted by m_g.
   (b) $m_g = \Xi_X$ is independent of g.

A sufficient condition for stability and ergodicity of Ξ_g is then stability and ergodicity of Ξ. Hence, the theorem characterizes a large family of beliefs on $\{x_t, t = 1, 2, \ldots\}$ *which have the same empirical implications*. Thus suppose $\{x_t, t = 1, 2, \ldots\}$ is generated under a true unknown stable, ergodic probability Π with an observed empirical stationary measure m. Now consider a joint process $\{(x_t, g_t), t = 1, 2, \ldots\}$ under probability Ξ which induces a belief Ξ_g on $\{x_t, t = 1, 2, \ldots\}$ with parameters g. The question are then: under what conditions is the belief Ξ_g stable and ergodic with an empirical stationary measure denoted by m_g and when would it have the same empirical implications

as Π? Theorem 3 shows that these conclusion hold if two conditions are satisfied

      (i) Ξ is stationary and ergodic,

      (ii) $m_g = m = \Xi_X$.

Hence, all measures Ξ satisfying (i)-(ii) induce conditional beliefs Ξ_g with the same empirical implications as the true measure Π under which the data is generated.

      In applying the theorem to (22a)-(22c) note this *joint* system is stationary and ergodic hence stable by the Ergodic theorem. We now think of the agent's belief as a conditional probability given values of $g_t \in G$, then we conclude that it is also stable with an empirical stationary measure which is simply computed by integrating $g_t \in G$ out of (22a)-(22c). We next explain that the theory of Rational Beliefs imposes the restriction $m_g = m$ and we illustrate, in the next Section 3.3.4, these restrictions on the continuous time model which we have been discussion all along. To complete the exposition we deduce in Section 4.2.2 similar rationality restrictions on the finite state model.

3.3.4 Belief Rationality: The Theory of Rational Beliefs

      We have seen the market belief is not necessarily rational. What about individuals? Since the true probability Π is not known, the belief (22a)-(22c) may not be the truth. What restrictions do (22a)-(22c) need to satisfy for them to be a belief of a rational agent? Our approach has already imposed some rationality conditions. First, we have argued that rational agents will exhibit fluctuating beliefs since a *constant* belief which is not the empirical distribution is irrational. Second, we required g_t^i to have an unconditional zero mean by requiring beliefs to be about *deviations* from empirical frequencies. Third, we described belief as a conditional probability of a stationary joint system (22a)-(22c). We now explain the additional restrictions imposed by the theory of Rational Beliefs.

      A Rational Belief (in short, RB. See, Kurz (1994), (1997a)) *is a probability model which, if simulated, reproduces the empirical distribution known from the data*. An RB is thus a probability model which cannot be rejected by the empirical evidence m[9]. For (22a) -(22c) to be RB it needs to

---

[9] It may appear that the empirical evidence consists of more than the moments of the data series as stipulated in Section 3.1 and Definition 1. That is, one should look not only at the full data series but also at subsequences. Kurz (1994) argues that economic time series have deterministic patterns in seasonal and cyclical frequencies and hence if these are cleaned out so that we look at seasonally and cyclically adjusted data, then under Ergodicity, with probability 1 the empirical distribution along any randomly selected subsequence is the same as the distribution along the entire sequence of data. Hence, there is no new restrictions that can be deduced from looking at subsequences.

induce the same empirical distribution of the observables (d_t, Z_t) as (21a)-(21b). Comparing these two systems, this condition amounts to requiring that if g_t^i is treated symmetrically with other random variables

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

To compute the implied statistics of the model we first compute the moment of the g_t^i. From (22c), the unconditional variance of g_t^i is $Var(g^i) = \sigma_g^2 / (1 - \lambda_Z^2)$ Hence, we have two sets of rationality conditions which follow from (23). The first arises from equating the covariance matrix

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

The second set arises from equating the serial correlations of the two systems

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

(i) -(iii) fix the covariance matrix in (22a)-(22c) and (vi)-(v) fix the serial correlation of $(\hat{\rho}_t^{id}, \hat{\rho}_t^{iZ})$. An inspection of (22a)-(22c) reveals the only choice left for a agent are the two free parameters $(\lambda_d^g, \lambda_Z^g)$. But under the RB theory these are not free either since there are natural conditions they must satisfy. First, $\hat{\sigma}_d^2 > 0$, $\hat{\sigma}_Z^2 > 0$ place two strict conditions on $(\lambda_d^g, \lambda_Z^g)$:

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

Finally, one need to ensure the covariance matrix in (22a)-(22c) is positive definite. The following is a sufficient condition

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

The "free" parameters $(\lambda_d^g, \lambda_Z^g)$ are thus restricted to a narrow range which is empirically testable.

This discussion applies the RB principle to the beliefs (22a)-(22c). What about more general systems? Kurz (ed.) (1997a) generalizes the construction of rational beliefs. It assumes the exogenous

33

process $\{x_t, t = 1, 2, \ldots\}$ is ergodic and stable under a true probability Π with stationary measure m. By the Conditional Stability theorem any joint process $\{(x_t, g_t), t = 1, 2, \ldots\}$ under a stationary and ergodic probability Ξ induces a belief which is a conditional probability Ξ_g on $\{x_t, t = 1, 2, \ldots\}$ with a stationary measure m_g. Ξ_g is then a Rational Belief if $m_g = m = \Xi_X$ where Ξ_X is the marginal of Ξ on X. An equilibrium where agents hold rational beliefs is a *Rational Belief Equilibrium* (in short, RBE).

3.4   Asset Pricing with Heterogenous Beliefs: An Illustrative Model

We now complete the infinite horizon model we started earlier, assuming a continuum of agents and an exogenous risky payoff process $\{d_t, t = 1, 2, \ldots\}$ with an unknown true probability Π and empirical distribution described by the stochastic transition $d_{t+1} = \lambda_d d_t + \rho_{t+1}^d$, $\rho_{t+1}^d \sim N(0, \sigma_d^2)$ i.i.d. The asset structure of the economy consists of an aggregate stock index (think of it as the S&P500) and a risk free bond. 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. At date t agent i buys θ_t^i shares of stock and receives the payment $d_t + \mu$ for each of θ_{t-1}^i held. 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, we define wealth $W_t^i = c_t^i + \theta_t^i p_t + B_t^i$ and derive the familiar transition for wealth

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

Q_t is excess returns per share. For initial values (θ_0^i, W_0^i) the agent maximizes the expected utility

(25) $$U = \mathop{E_t^i}_{(\theta^i, c^i)} [\sum_{s=0}^{\infty} -\beta^{t+s} e^{-(\frac{1}{\tau} c_{t+s}^i)} | H_t]$$

subject to transitions (22a)-(22c) of the state variables $\psi_t^i = (1, d_t, Z_t, g_t^i)$. H_t is date t information.

Our assumptions are restrictive. Constant R means an exogenous riskless rate and exponential utility exhibits no income effects. Nevertheless, these assumptions have the great advantage of leading to closed form solutions which are helpful vehicles to explain the main ideas. Hence the term "illustrative" in this Section's title. To seek a closed form solution we conjecture prices are linear in the economy's state variables hence equilibrium price p_t is conditionally normally distributed. In Theorem 4 below we confirm this conjecture. For an optimum (for details see the Appendix of Kurz and Motolese (2006)) there exists a constant vector u so the demand functions for the stock is

34

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

$\hat{\sigma}_Q^2$ is an *adjusted* conditional variance (the "adjustment" is explained in Section 4.3) of excess stock return Q_{t+1} 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.

For an equilibrium to exist we impose stability conditions on the dynamics of the economy. First we require the interest rate r to be positive, $R = 1 + r > 1$ so that $0 < \dfrac{1}{R} < 1$. Now we add:

Stability Conditions: We require that   (i) $0 < \lambda_d < 1$, (ii) $0 < \lambda_Z + \lambda_Z^g < 1$.

(i) requires that $\{d_t , t = 1, 2, ...\}$ is stable and has an empirical distribution. (ii) is *a* stability of *belief* condition. It requires i to believe (d_t , Z_t) is stable. To see why take expectations of (22b), average over the population and recall that Z_t are averages of g_t^i. This implies that

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

Kurz (2007) and Kurz and Motolese (2006) then demonstrate the following results:

Theorem 4: 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

(27a) $\qquad\qquad\qquad\qquad p_t = a_d d_t + a_z Z_t + P_0$

with coefficients

(27b) $\qquad\qquad\qquad\qquad a_d = \dfrac{\lambda_d + u_1}{R - \lambda_d}$.

(27c) $\qquad\qquad\qquad\qquad a_z = \dfrac{(a_d + 1)\lambda_d^g + (u_2 + u_3)}{R - (\lambda_Z + \lambda_Z^g)}$

(27d) $\qquad\qquad\qquad\qquad P_0 = \dfrac{(\mu + u_0)}{r} - \dfrac{\hat{\sigma}_Q^2 R}{\tau r}$.

Closed form solutions for the hedging demand parameters $u = (u_0, u_1, u_2, u_3)$ are not available hence Kurz and Motolese (2006) compute numerical Monte Carlo solutions. For all values of the model parameters they find $u_1 = 0$ which imply (i) $a_d > 0$, (ii) $(a_d + 1)\lambda_d^g + (u_2 + u_3) > 0$ and (iii) $a_z > 0$. These conclusions are reasonable: *today's asset price increases if* d_t *or* Z_t *rise*. The conclusion $u_1 = 0$ says individuals have no hedging demand for dividends and this result is reasonable since $u_1 = 0$ is

35

compatible with the fact that individual expected excess returns are independent of dividend states. To see why compute i's perceived excess returns, keeping in mind that perceived excess return depends upon the belief of the agent. From (27a) and (22a)-(22c) we have

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

By (27a) $Rp_t = R[a_d d_t + a_z Z_t + P_0]$ hence, collecting the dividend term in excess return we find it is $(a_d + 1)\lambda_d - Ra_d$. But by (27b) we also find that $(a_d + 1)\lambda_d - Ra_d = u_1$. These two results together show that $u_1 = 0$ implies $(a_d + 1)\lambda_d - Ra_d = 0$ hence there is no covariance between dividend and excess returns. In short, this argument shows that the combination of $u_1 = 0$ and no covariance between dividend and excess return is a solution. What does it say about the hedging demand?

Standard thinking about hedging demand leads one to focus on the dividend state as the only relevant variable. In contrast, we have constructed a model where hedging demand due to dividend information is zero and the key hedging demand is in response to market states of belief. The point of the model is that agents do not believe excess returns are i.i.d. and this leads to a hedging demand for the stock. But the non i.i.d. of excess return is not due to the dynamics of the dividend states but to market beliefs which also have a drastic effect on the risk premium. In more general models there may be a hedging demand due to dividend states but in the present case this demand is absent because of the absence of income effects. Here, investors hedge only against future market beliefs.

One now confirms the earlier conjecture that the price is conditionally normal. Also, once we have the solution above then all implications of the model can be deduced analytically.

4.   Explaining Asset Market Dynamics with Diverse Beliefs and Common Information

The most fundamental implication of the theory says that asset markets are subject to Endogenous Uncertainty (Kurz (1974), Kurz and Wu (1996)). Risk of asset returns is, in part, due to the risk of future market belief. In our illustrative model it is seen in the price map $p_t = a_d d_t + a_z Z_t + P_0$ which, by (21a)-(21b), implies that in the long run $\sigma_p^2 = a_d^2 \sigma_d^2 + a_z^2 \sigma_Z^2$. Price volatility is caused by exogenous as well as endogenous forces and this result has far reaching implications. We note a few:

•   An asset's price is not equal to a unique fundamental value determined by the flow of future payoffs. Moreover, market beliefs about exogenous states matter since they are often wrong, generating an independent and dominant component of asset price volatility.

- Moral hazard and the large dimension of market belief make it impossible for markets to trade contracts contingent on market belief, hence markets are fundamentally incomplete.

- In time scales of days or weeks changes in productivity, growth and profits are slow. Hence, most volume of asset trading results from fluctuations in the market distribution of beliefs.

- Expected individual excess returns $\frac{1}{p_t}E_t^i(p_{t+1}+d_{t+1}+\mu-Rp_t)$ and "efficient frontiers" are both subjective concepts. Hence, in markets with diverse beliefs most predictions of CAPM theory do not hold.

- By anonymity of individuals the market belief is a public externality and hence subject to the effect of coordination and public policy. Stabilization policy can thus have a strong effect on market volatility and this carries over to monetary economies as well. We specify this issue explicitly in Section 4.6 when we review Kurz, Jin and Motolese (2005b).

We now turn to review applications of the theory of diverse beliefs to market dynamics

4.1   Rational Overconfidence

Evidence from the psychological and behavioral literature (e.g. Svenson (1981), Camerer and Lovallo (1999), Russo and Schoemaker (1992)) shows a majority of individuals assess their own probability of success in performing a task (investment, economic decisions, driving etc) above the empirical frequency of success in a population. Hence a majority of people often expect to outperform the empirical frequency measured by the median or mean. In a Rational Expectations paradigm individuals know the true probability of success hence the observed inconsistency is taken to be a demonstration of irrational behavior. Indeed, inconsistency between individual assessments and empirical frequencies has been cited extensively as a "proof" of irrational behavior and in support of behavioral\psychological impulses for belief and forecasting. This phenomenon has thus been called "Overconfidence." We reject this conclusion and show it reveals a fundamental flaw.

The work cited above (and other empirical and experimental work) provides strong evidence against Rational Expectations. But REE is an extreme theory in demanding agents to know the full structure of the economy and make exact probability assessments. Behavioral Economics takes the other extreme view and assumes people are irrational and motivated by psychological impulses. Hence, a rejection of Rational Expectations does not imply acceptance of agents' irrationality. Indeed,

one may reject these two extreme perspectives by observing the fact that most people do the best they can, given the limited knowledge they have. Rational people do not know everything and make "mistakes" relative to a true model which they do not know. The theory of Rational Belief rejects both extremes in favor of an intermediate concept of rationality. We then show that "overconfidence" is compatible with the rationality principle of Rational Belief. *Indeed, agents who hold Rational Belief will universally exhibit "Rational Overconfidence"* hence the cited empirical evidence is no proof people are irrational and motivated by pure psychological factors.

We proceed with a simple example which reinterprets the Example in 3.1 (see also Nielsen (2006)). A group of gamblers look at a black box which generates numbers x_t, $t = 0, 1,...$ in $\{0, 1\}$ where long term past data reveals no serial correlation between x_t and x_{t+k} and mean of ½. Each gambler will play only a small number of periods and be replaced by a new gambler. A gambler will thus know the belief distributions of past gamblers but not their individual beliefs. As in (9) the belief of a gambler is defined by a process $\{g_j, j = 0,1, ...\}$ of i.i.d. random variables in $\{1, 2\}$ with probability of 1 being 1/3 and realizations $g^* = (g_0^*, g_1^*,...)$. Values $g_j^* = 1$ or $g_j^* = 2$ are treated as *parameters* of a gambler's belief. He believes the true process is $\{v_t, t = 0, 1,...\}$ with two i.i.d. coins appearing at different times depending upon the $g^* = (g_0^*, g_1^*,...)$. His belief is then defined by $\{v_t, t = 0, 1, ...\}$ where $v_t \in \{0, 1\}$ is a sequence of *independent* random variables satisfying

$$P\{v_t = 1\} = \begin{cases} 0.60 & \text{if } g_t^* = 1 \quad (\text{coin type } 1) \\ 0.45 & \text{if } g_t^* = 2 \quad (\text{coin type } 2). \end{cases}$$

Since $(1/3)(0.60) + (2/3)(0.45) = 0.50$, all these are rational beliefs for almost all g^*. In the RB literature the ratio 1/3 is referred to as the "frequency" of optimism. When the frequency of bull and bear states is not the same we have *market asymmetry* between these two states. The probability 0.60 is the "intensity" of optimism when optimistic. Note that in defining an RB these characteristics are selected separately: for each frequency there is a range of feasible intensities which are rational.

Now let the gamblers decide, at some date t-1, on how they want to bet. They can gamble $1 on $v_t = 1$ or on $v_t = 0$: they win $1 if they are right and they lose $1 if they are wrong. Being a small bet they will all bet. Those who put money on 1 expect to win with a probability 0.60 and those who put their money on 0 expect to win with probability 0.55. *They are all overconfident and rational.*

The constancy of (0.60 , 0.45) is not essential since we can, instead, put in any sequences of parameters which converge to 0.50 from above and from below and the result is the same.

Observe that here deviations from empirical frequencies lead to optimal behavior which exhibits overconfidence. When subjective probability is above the empirical frequency a long position is optimal and it is overconfidence. When probability is below empirical frequency a short position is optimally held with overconfidence. *Hence, all agents are then optimally overconfident at all time.*

Generalizing the example is natural. If x is a quantitative measure of success in an activity then by (22d), $E_t^i(x_{t+1}) - E_t^m(x_{t+1}) = \lambda_x^g g_t^{xi}$ where g_t^{xi} measures the belief of i about future x. Beliefs are all about deviations from empirical frequencies on the basis of which economic decisions are made. Optimistic agents engage in taking the risk of success in that activity and pessimistic agents engage in gambles against the activity. If they cannot gamble against it (e.g. short positions are not allowed) they refrain from participation. This type of behavior is then natural to the Rational Belief paradigm. Moreover, this behavior is natural to any complex environment in which aggregation of subjective probability beliefs of agents may not be equal to the empirical frequencies. But then all creative work and all innovative decisions are based on beliefs which exhibit "overconfidence." Indeed, one can hardly think of entrepreneurship, inventive activity and any speculative behavior without beliefs which exhibit rational overconfidence.

4.2   Anatomy of Market Volatility and Trade

4.2.1   On the Significance of Simulation Work: The Main Results

Although much of our discussion is analytical, significant results about excess volatility are deduced in all models we review here from simulation models. Simulations require specification of functions, parameters and beliefs, and aim to show a model replicates the statistics of the economy. Since most models reviewed have only two agent types, the beliefs selected are representative of only two classes of agents. Since one might question the validity of such approach as too narrow, it is useful to explain the features of *all simulation models reviewed* which have valid implications to the way markets function. Our view on this issue is simple: a simulation model is a very good tool to explore the impact on market volatility of the *qualitative features* of feasible belief structure.

The best way to explain the above view is to highlight the central conclusions of the work we

review below and the summation of Kurz, Jin and Motolese (2005a) is useful. This paper starts from the view that in all non REE asset market equilibria there are basically two *natural individual states* one can have: optimistic (i.e. bull states), and pessimistic (i.e. bear states). The authors then reason that given these two basic states there are three central characteristics of individual beliefs which fully account for all characteristics of market volatility and risk premia observed in real markets. Two are:

(A) *large size (i.e. high intensity) of fat tails in the belief densities of agents;*

(B) *asymmetry in the proportion of bull and bear states in the market over time.*

Large size or high intensity means the densities of the agents' beliefs have fat tails, the sizes of which vary with belief states. Intensity measures deviations from stationary probabilities as in the review of overconfidence. The *asymmetry in the time frequency of belief states* needed to reproduce the results is a subtle characteristic which says that on average, agents are in bear states at more than 50% of the dates. Equivalently, on average, at more than half of the time agents do not expect to make above normal excess return on their investments. Therefore, it follows from the RB principle that when agents are in bull states and expect *above* normal returns, their expected excess returns must be very high. We shall see later that this asymmetry is empirically supported by the fact that major abnormal rises in stock prices occur over a relatively small fraction of time. Hence, when agents believe a bull market is ahead, they expect to make excess return in relatively short periods. The third feature is:

(C) belief states are correlated resulting in regular joint dynamics of belief distributions

The correlation of beliefs is a market externality which regulates the probability of agreement or disagreement of beliefs in the market and the transitions among states of agreement and disagreement. This is important because it is the distribution of beliefs which determines prices and returns, and the dynamics of the belief distribution is crucially affected by the correlation. It turns out that asymmetry in the transition is also important since it regulates the dynamics of bear vs bull markets.

Kurz, Jin and Motolese (2005a) then make two important observations. First, exactly the same simulation model used to study market volatility is used to study all other aspects of market dynamics. In that model stock prices and returns exhibit a structure of forecastability observed in the real data. Also, the model implies that market returns exhibit stochastic volatility generated by the dynamics of the market beliefs. Second, examination of alternative configurations of belief shows that no other configuration of qualitative features as the three specified above yields predictions *which*

simultaneously reproduces the empirical record. Many feasible model parameters generate volatility of prices and returns. However, as we move away from the three features, the model fails to generate some essential components of the empirical record, frequently the riskless rate and the premium. Thus, the main reason why the models are able to explain the empirical record is that they have the needed configuration of qualitative factors. These imply a unique structure of parameters which would explain the empirical record. But the specific structure of belief in the model is not essential.

Papers on excess volatility simulate computed equilibria with finite or infinite belief states. Those with finite belief states are OLG models while those with infinite belief states are infinite horizon models. This division guides our review. Papers which fall into the first category (i.e. OLG) include Nielsen (1996),(2003),(2005),(2006), Kurz (1997b), Black (1997),(2005), Kurz and Beltratti (1997), Kurz and Motolese (2001), Kurz and Schneider (1996), Motolese (2003), Nakata (2007), Wu and Guo (2003). Papers using infinite horizon models with infinite belief states include Kurz, Jin and Motolese (2005a), (2005b), Kurz (2007), Kurz and Motolese (2006),and Guo and Wu (2007). We turn to market volatility in finite state model and use Kurz and Motolese (2001) as a prototype[10].

4.2.2 Understanding the Parametrized Structure of Beliefs

We start by discussing OLG models *with finite belief states.* All models have two assets: a stock and a riskless bond. The stock pays dividends with a two state growth rate. There are two types of agents, each living two periods with a power utility function over consumption

$$u^k(c_t^{1k}, c_{t+1}^{2k}) = [1/1 - \gamma_k](c_t^{1k})^{1-\gamma_k} + [\beta_k/1 - \gamma_k](c_{t+1}^{2k})^{1-\gamma_k} \ , \ \gamma_k > 0 \, , \, 0 < \beta_k < 1.$$

Hence, there are eight states and belief structure as described in (13)-(15). Rational Beliefs are the two pairs of matrices (F_1, F_2) of agent 1 and (G_1, G_2) of agent 2, satisfying (15). Since the exogenous variable is the growth rate of dividend, deviations from the long term mean growth (10b) could either be *above* it or *below* it. Hence, at each date an agent must be either a bull or a bear about future growth of dividend. Since the first four states are the states of high dividend growth, a bull must have

---

[10] Here again we present the belief structure and rationality conditions for finite state and continuos state models. The empirical results reported later are all deduced from the continuous state model hence it may be useful for the reader to skip, upon first reading, material which pertains to finite state models. This material would be especially relevant if the reader wants to replicate any of these results by visiting the web pages provided to download the programs with which to compute the solutions. It will be found that the finite state models are much easier to handle.

an increased probabilities of transition to states 1-4 relative to the Γ probabilities in (14). Similarly, a bear must have increased probabilities into states 5-8 relative to the Γ probabilities in (14). This literature typically uses two parameters λ and μ to specify (F_1, F_2) and (G_1, G_2) while satisfying the rationality conditions (15). To explain how this is done we denote the row vectors of A and B by:

$$A^j = (a_j, \ \alpha_1 - a_j, \ \alpha_2 - a_j, \ 1 + a_j - (\alpha_1 + \alpha_2)) \qquad j = 1, 2, 3, 4$$
$$B^j = (b_j, \ \alpha_1 - b_j, \ \alpha_2 - b_j, \ 1 + b_j - (\alpha_1 + \alpha_2)) \qquad j = 1, 2, 3, 4.$$

With this notation we define the 4 matrix functions of a real number z as follows:

$$(28) \quad A_1(z) = \begin{bmatrix} z\,A^1 \\ z\,A^2 \\ z\,A^3 \\ z\,A^4 \end{bmatrix}, \ A_2(z) = \begin{bmatrix} (1-\varphi z)A^1 \\ (1-\varphi z)A^2 \\ (1-\varphi z)A^3 \\ (1-\varphi z)A^4 \end{bmatrix}, \ B_1(z) = \begin{bmatrix} z\,B^1 \\ z\,B^2 \\ z\,B^3 \\ z\,B^4 \end{bmatrix}, \ B_2(z) = \begin{bmatrix} (1-(1-\varphi)z)B^1 \\ (1-(1-\varphi)z)B^2 \\ (1-(1-\varphi)z)B^3 \\ (1-(1-\varphi)z)B^4 \end{bmatrix}.$$

Finally we define

$$(29) \qquad F_1 = \begin{bmatrix} \varphi A_1(\lambda) \ , & A_2(\lambda) \\ (1-\varphi)B_1(\lambda), & B_2(\lambda) \end{bmatrix} \qquad G_1 = \begin{bmatrix} \varphi A_1(\mu) \ , & A_2(\mu) \\ (1-\varphi)B_1(\mu), & B_2(\mu) \end{bmatrix}.$$

By the rationality conditions (15), $F_2 = \dfrac{1}{1-\alpha_1}(\Gamma - \alpha_1 F_1)$, $G_2 = \dfrac{1}{1-\alpha_2}(\Gamma - \alpha_2 G_1)$. In (28)-(29) one can assign different values (μ_s, λ_s) to different rows. This is avoided in favor of symmetry where $\mu_s = \mu$, $\lambda_s = \lambda$ for all s. Hence, the central belief parameters (λ, μ) *measure the intensity of optimism when in an optimistic state while* (α_1, α_2) *measure the frequency of being optimistic.*

To see the implications note that λ and μ are proportional revisions of the conditional probabilities of states (1, 2, 3, 4) and (5, 6, 7, 8) relative to Γ. $\lambda > 1$ and $\mu > 1$ imply increased probabilities of (1, 2, 3, 4) in matrices F_1 and G_1 where the first four prices occur at $d_t = d^H$ states. But these are actually states of high prices as well, hence $\lambda > 1$ implies agent 1 is optimistic about high prices at $t + 1$. Similarly for $\mu > 1$. In all simulations $\lambda \geq 1$ and $\mu \geq 1$ hence we can interpret g_t^k so that $g_t^k = 1$ means agent k is optimistic (*relative to Γ*) at t about *high prices* at $t + 1$.

In the transition Γ, the matrices A and B regulate the correlation between belief states and the effect of dividends on that correlation. One must think of A and B as a *correlation externality* given to agents. We encountered this externality in correlation among the ρ_t^{ig} across i in (11), a correlation which gives rise to the dynamics of the aggregate Z_t in (20). The correlation is crucial but it turns out that it does not need to be complex. The case $\lambda = \mu = 1$, $\alpha_1 = 0.50, \alpha_2 = 0.50$ *and* $a_i = b_i = 0.25$ is the case of REE. In Kurz and Motolese (2001) the matrices A and B can be reduced

to two parameters (c_1, c_2) so that it is sufficient to have A = B with $a_1 = a_2 = a_3 = c_1$, $a_4 = c_2$. Kurz and Motolese (2001) require a = b = (.50 , .14, .14, .14) which implies the dynamics of prices have the feature that bull and bear markets are asymmetric. For the market to transit from the crash state of the lowest price to the states of the highest prices it must take several steps: it cannot go *directly* from the low to the high prices. The opposite is possible since at the bull market states there is a positive probability of reaching the crash states in one step. *This implies that a bull markets which reaches the high prices must evolve in several steps but a crash can occur in one step.*

To sum up, there are three classes of parameters and simulation work explores regions of the parameter space which are compatible with rationality. We report here a representative set of parameters under which *the model replicates the empirical record with great accuracy.* These are:

Utility function: $\qquad\qquad 2.00 \le \gamma_1 = \gamma_2 = \gamma \le 3.00$ – the common risk aversion coefficient

$\qquad\qquad\qquad\qquad 0,90 \le \beta_1 = \beta_2 = \beta \le 0.95$ – the common discount rate

Correlation parameters: $\qquad a_1 = a_2 = a_3 = 0.50$, $a_4 = 0.14$

Belief Parameters: $\qquad\qquad \lambda = \mu = 1.7542$ – the maximal intensity permitted by rationality

$\qquad\qquad\qquad\qquad \alpha_1 = \alpha_2 = \alpha = 0.57$ – frequency an agent is in an optimistic state

The model replicates well the empirical record and the results are similar to those of infinite horizon models. Hence there is no point repeating the same results twice and we report the precise numerical results only for the infinite horizon models[10]. We thus turn to models with infinite belief states.

Models with infinite horizon and infinite belief states typically have two assets: a stock paying dividends and a zero net supply bond. The model has a large number of identical agents of two types with the same utility and endowment. Across types they differ in their beliefs. For consistency we use a single model to illustrate the structure and report results of simulation models, and for that aim we

---

[10] For details of the finite belief state results see Kurz and Motolese (2001) pages 530-533. For computational procedures to reproduce these results go to http://www.stanford.edu/~mordecai/ and click on "computable models with heterogenous beliefs." Keep in mind that an OLG model has a unique market oriented feature not shared by an infinite horizon models which requires an agent to sell his position when old, regardless of his beliefs. This feature is important for a model of market volatility since agents who aim to preserve capital by holding a portfolio of a riskless asset must sell the asset into the market regardless of their beliefs. This fact tends to generate additional volatility that would not be present in an infinite horizon model. This feature has two results which are not shared by the infinite horizon model. First, the riskless rate has a much larger standard deviation in simulated OLG equilibria than in the infinite horizon models. Second, in order to generate a low average riskless rate, a results needed to replicating the 6%-7% equity premium, it is necessary to assume an asymmetry where the majority of agents are optimists about earning abnormal excess returns and the frequency of optimism is greater than 50%. In the infinite horizon model it is necessary to have the pessimists in the majority with a frequency of pessimism being more than 50%. We shall comment on this issue later again when we discuss the Equity Premium Puzzle.

use the one developed by Kurz, Jin and Motolese (2005a), referred to as KJM (2005a). This offers the advantage that all simulation results reported *were derived from a single model* in KJM (2005a). These authors assume $D_{t+1} = D_t e^{d_{t+1}}$ with an empirical distribution of the growth rate defined by

$$d_{t+1} = (1 - \lambda_d)d^* + \lambda_d d_t + \rho_{t+1}^d \qquad \rho_{t+1}^d \sim N(0, \sigma_d^2) \text{ i.i.d.}$$

Given his probability belief, agent i maximizes an infinite horizon expected utility with date t utility of $\beta^t[1/1 - \gamma](c_t^i)^{1-\gamma}$. To explain the perception models of the agents we could have postulated $d_{t+1}^i = (1 - \lambda_d)d^* + \lambda_d d_t + \lambda_d^g g_t^i + \rho_{t+1}^{id}$ as in (12a). Such a model is sufficient for the conceptual needs of the illustrative model but it would not be consistent with the finite belief state simulation model reviewed above. The reason is that the belief state g_t^i is a symmetric variable and that would not meet the three principles advocated earlier. Limited by space we cannot review the technical aspects of the very complex structure in KJM (2005a)[11]. We can only say that to attain asymmetry and fat tails in the belief densities they transform the belief index g_t^i into an index $\eta^i(g_t^i)$ which depends upon g_t^i and an asymmetry parameter a. When $g_t^i > a$ an agent has a positive outlook about future dividend growth and places heavier probability on the positive side of the distribution. When $g_t^i < a$ heavier weight is placed on the negative side of the distribution. For a given g, this creates an asymmetrical density of the random index $\eta^i(g_t^i)$ with a discontinuity at 0. Ignoring technical details which are not essential, we exhibit in Figure 1 the asymmetric density of $\eta^i(g_t^i)$

FIGURE 3 PLACED HERE

The belief of agent i about future dividend growth rate is then formulated as

(30) $\qquad d_{t+1}^i = (1 - \lambda_d)d^* + \lambda_d d_t + \lambda_g^d \eta^i(g_t^i) + \tilde{\rho}_{t+1}^{d\,i} \qquad \tilde{\rho}_{t+1}^d \sim N(0, \tilde{\sigma}_d^2) \text{ i.i.d.}$

Fat tails enter the conditional distribution d_{t+1}^i via the random function $\Psi^j(g_t^j) = \lambda_g^d \eta^i(g_t^i) + \tilde{\rho}_{t+1}^{d\,i}$ whose distribution is deduced via a convolution and has the shape illustrated in Figure 2

FIGURE4 PLACED HERE

The assumption of a power utility $\beta^t[1/1 - \gamma](c_t^i)^{1-\gamma}$ imply income effects matter and beliefs do not aggregate. Hence, the state variable in the simulation model is the actual distribution of beliefs.

---

[11] Indeed, the main deficiency of the KJM (2005a) model is its complexity which, in our view today, could have been avoided. Both the model itself as well as the computational procedures could have been drastically simplified since the basic ideas are rather simple, as explained in Section 4.2.1.

Since there is a large number of identical agents of *two types* this distribution is a vector (z_t^1, z_t^2). The fact that we denote it by (z_t^1, z_t^2) and not (g_t^1, g_t^2) is an important technical issue arising due to the assumption of anonymity. Agent i knows his own belief as g_t^i which he uses to forecast all state variables in the economy whereas (z_t^1, z_t^2) are observed state variables and the agent uses g_t^i to forecast d_{t+1} as in (30) and (z_{t+1}^1, z_{t+1}^2) with a fully developed perception model which is analogous to (22a)-(22c) in the illustrative model which we have developed here. For technical details of the perception model and the implied RB principle restrictions, see KJM (2005a) pages 12-19. We now turn to a detailed examination of the simulation results of the finite and infinite belief state model.

4.2.3 Explaining the Volatility Moments

We now report simulation results of models with infinite belief states. Please note that *all simulation results reported in this paper are derived from a single model by KJM (2005a).* In these simulations KJM (2005a) compute various measures of volatility by using 20,000 observations. Raw moment calculations were carried out by KJM (2005a) for the following list of volatility measures:

q^s -   long term price/dividend ratio;

σ_{q^s} -   standard deviation of the price/dividend ratio q^s;

R^S -   average risky return on equity;

σ_R -   standard deviation of R^S;

r -   the riskless rate of interest;

σ_r -   the standard deviation of r;

e_p -   the equity premium;

$\rho(d, R^S)$ - correlation coefficient between stock returns and growth rate of dividends. Since consumption and divided are assumed to grow at the same rate, $\rho(d, R^S)$ is also the correlation coefficient between consumption growth and the risky rate;

(Shrp) - the Sharp ratio .

Table 2 reports the results. The model clearly matches simultaneously the moments and, as we see later, it also matches most other features of market dynamics. S*imultaneous* explanation of diverse phenomena *by a single model* is a key property any good theory of market volatility must have. The added observation made by KJM (2005a) is that the results in Table 2 are not due to the particular beliefs used or their parameter values. They are due to fat tails in asset returns, to

asymmetry and to correlation of beliefs. Are these three key characteristics supported by the data?

Table 2: Simulation Results of KJM (2005a)[12]
(all moments are annualized)

Moment	Simulation Results	Empirical Record[13] (1889-1998)
q^s	25.54	25.00
σ_{q^s}	5.46	7.10
R^S	7.57%	7.00%
σ_R	18.81%	18.00%
r	1.08%	1.00%
σ_r	5.44%	5.70%
e_p	6.49%	6.00%
$\rho(d,R^S)$	0.21	0.10
Shrp	0.34	0.33

The fact that the distribution of asset returns exhibits fat tails is well documented (e.g. see Fama (1965) and Shiller (1981)). It is natural to ask where these tails come from. The theory at hand says they come from fat tails in the probability models of agents' beliefs. Correlation of beliefs across agents is documented in the BLU and other data on forecast distributions. The evidence in support of the hypothesis that the frequency of bear states is, on average, higher than 50% is more complicated.

The hypothesis is supported by the empirical fact that, on average, most above normal stock returns are realized over *relatively small proportion of time* when asset prices rally (see Shilling (1992)). It is thus reasonable that, on average, the proportion of time when agents expect to make above normal returns is less than 50%[14]. Additional indirect support comes from the psychological literature which suggests agents place heavier weight on losses than on gains. In the treatment here agents fear losses at majority of dates since on those dates they place higher probabilities of abnormally lower returns. By the RB principle, a higher frequency of bear states implies that when in bull states, an agent's intensity of optimism is higher than the intensity of pessimism. This means the average size of the positive tail in the belief densities is bigger than the average size of the negative

---

[12] For computational procedures to reproduce the simulation results of KJM (2005a) click on "computable models with heterogenous beliefs" at the address http://www.stanford.edu/~mordecai/

[13] The main data source for the empirical record is Shiller at http://www.econ.yale.edu/~shiller/data.htm. It was updated by KJM (2005a) to 1998. Since the discussion here does not aim to evaluate the precision of the estimates, the numbers in the Table 2 were rounded off to indicate orders of magnitude.

[14] Shilling (1992) shows that during the 552 months from January 1946 through December of 1991 the mean real annual total return on the Dow Jones Industrials was 6.7%. However, if an investor missed the 50 strongest months the real mean annual return over the other 502 months was -0.8%. Hence the financial motivation to time the market is very strong, as is the case with the agents in the model.

tail. That is, the asymmetry hypothesis implies optimistic agents tend to be intense. Together with correlation of beliefs across agents, this hypothesis also implies that we should observe periods of high optimism across a majority of agents. High level of optimism leads to agents' desire to borrow and finance present and future consumption. At such dates the only way for markets to clear is by exhibiting sharp rises in stock prices together with high borrowing rates. Hence, this theory predicts we should expect to observe rapidly rising stock prices induced by bursts of correlated optimism correlated with high realized growth rates of dividends. The structure of correlation also implies that we should also expect to see crashes induced by correlated pessimistic agents together with low realized growth rates of dividends.

We make one comment with respect to the low riskless rate. Matching many volatility moment *except the riskless rate* depend mostly upon intensity and correlation. These moments exhibit relatively low sensitivity to asymmetry. Hence, apart from the riskless rate, many long term volatility measures are explained by a broad configuration of the intensity parameters and correlation across agents' beliefs. The low riskless and a few others require asymmetry.

4.2.4   Why Does the Model of Diverse Beliefs Resolve the Equity Premium Puzzle?

Risk premia are compensations for risk perceived by risk averse agents. In single agent models the market portfolio is identified with a security whose payoff is aggregate consumption, mostly taken to be exogenous. The Equity Premium Puzzle is thus an observation that the small volatility of aggregate consumption growth cannot justify a 6% equity premium given the degree of risk aversion. The theory of diverse beliefs offers a clear resolution of the Puzzle by studying optimal behavior and consumption growth rate volatility on the individual, not the aggregate level.

Any theory of diverse beliefs implies that at each date the risk premium perceived by an agent is subjective. The risk premium required by an investor with a bullish outlook is smaller than the risk premium required by a bear. Hence to resolve the Equity Premium Puzzle a theory must explain why some agents are willing to hold a riskless asset paying a real return of only 1% when the average return on the risky stock is 7%. The 7% return on the stock is entirely explainable by fundamental factors of growth and productivity, together with the added high volatility of returns induced by factors of intensity and correlation of beliefs which generate Endogenous Uncertainty. The problem is the low riskless rate. Pessimistic agents who aim to preserve capital are willing to earn low return on

their investment and with enough of them around the riskless rate would indeed fall. But can a desire to preserve capital by those avoiding the risky stock be compatible with fat tails in returns? This is the role of *asymmetry*. Symmetry between bulls and bears generates fat tails only due to intensity and correlation. Agents are intense when they are bulls and correlation causes the majority sentiment to fluctuate. Fat tails then reflect fluctuations of the majority between bull or bear averages. After all when a majority of agents try to sell or buy the stock, the price fluctuates. But to push the riskless rate down we need the asymmetric persistence of the bear view by those who expect the stock to deliver low excess returns. Expecting low excess returns they would rather avoid the risky stock and hold the bond at lower return. The fact that bears are in the majority of investors at the majority of dates constitute the extra factor which lowers the riskless rate as well.

We turn to the low volatility of aggregate consumption growth. Diverse beliefs cause diverse individual consumption growth rates even if aggregate consumption is exogenous, which is the case in the models here. This is true not only because of idiosyncratic factors but also because under diverse beliefs markets are inherently incomplete and the representative agent model does not capture the conditions of individual consumers. Hence, volatility of individual consumption growth rates are higher than the volatility of the aggregate rate, an empirical fact supported by household survey data. Since agents' perceived volatility of their own consumption growth is different from the aggregate rate, they do not seek to own a portfolio whose payoff is aggregate consumption. Consequently, one must not focus on the relation between asset returns and aggregate consumption growth but instead, on the relation between perceived asset returns and *perceived volatility of individual consumption growth*. The key question is then, how volatile do individual consumption growth rates need to be in order to generate an equity premium of 6% and a riskless rate of 1%? The answer is: *not very much*. Relative to their equilibrium, KJM (2005a) report that although the volatility of the aggregate consumption growth rate is $\sigma_d = 0.03256$, the standard deviation of individual consumption growth rates supporting the premium in the simulations is only 0.039 (i.e. 3.9%) and the required correlation between individual consumption growth rate and the growth rate of dividends is only 0.83 (compared to 1.00 in a representative household model). Both figures are compatible with survey data.

4.2.5   *Predictability of stock returns*

The problem of predictability of risky returns generated a large literature in empirical finance

(e.g. Fama and French (1988a,1998b), Poterba and Summers (1988), Campbell and Shiller (1988), Paye and Timmermann (2003)). This debate is contrasted with the simple theoretical observation that under risk aversion asset prices and returns are not martingales and contain a predictable component. We focus only on the empirical record. KJM (2005a) use the basic model of Table 2 to generate simulated data with which they examine the following: (i) variance ratio statistic; (ii) autocorrelation of returns and of price/dividend ratios; (iii) predictive power of the dividend yield. We first introduce some notation. Let $\varrho_t = \log[\frac{(q_t^s + 1)e^{x_t}}{q_{t-1}^s}]$ be the log of gross one year stock return, $\varrho_t^k = \sum_{i=0}^{k-1} \varrho_{t-i}$ be the cumulative log-return of length k from t-k+1 to t, and $\varrho_{t+k}^k = \sum_{j=1}^{k} \varrho_{t+j}$ be the cumulative log-return over a k-year horizon from t+1 to t+k.

4.2.5a Variance Ratio Test

Let the variance-ratio be $VR(k) = \frac{var(\varrho_t^k)}{(k \, var(\varrho_t))}$. As k rises it converges to one if returns are uncorrelated. If returns are negatively autocorrelated at some lags, the ratio is less than one. KJM (2005a) show there exists a significant higher order autocorrelation in simulated stock returns hence a long run predictability which is consistent with U.S. data on stock returns, as in Poterba and Summers (1988). In Table 3 KJM (2005a) report the computed values of the ratios for k = 1, 2, ..., 10 and compare them with the ratios in the empirical record reported by Poterba and Summers ((1988) ,Table 2, line 3) for k = 1, 2,...,8. The model's prediction is close to the U.S. empirical record

Table 3: Variance Ratios for NYSE 1926 - 1985

k	1	2	3	4	5	6	7	8	9	10
VR(k)	1.00	0.85	0.73	0.64	0.57	0.51	0.46	0.41	0.38	0.34
U.S.	1.00	0.96	0.84	0.75	0.64	0.52	0.40	0.35	---	---

4.2.5b The Autocorrelation of Log-Returns and Price-Dividend ratios

In Tables 4 we report the KJM (2005a) autocorrelation function of log annual returns. The model predicts negatively autocorrelated returns at all lags. This implies a long horizon mean reversion of the kind documented by Poterba and Summers (1988), Fama and French (1998a) and Campbell and Shiller (1988). Thus, apart from the very short returns which exhibit positive autocorrelation, the model reproduces the empirical record.

Table 4: Autocorrelation of Log-Returns

corr(ϱ_t, ϱ_{t-i})	Model	Empirical Record
i = 1	-0.154	0.070
i = 2	-0.094	-0.170
i = 3	-0.069	-0.050
i = 4	-0.035	-0.110
i = 5	-0.040	-0.040

In Table 5 we report the autocorrelation function of the price-dividend ratio reported by KJM (2005a). The table shows the model generates a highly autocorrelated price/dividend ratio which matches reasonably well the behavior of U.S. stock market data. The empirical record in Tables 4 and 5 is for NYSE data for 1926-1995 as reported in Barberis et al. (2001).

Table 5: Autocorrelation of Price-Dividend Ratio

corr(q_t^s, q_{t-i}^s)	Model	Empirical Record
i = 1	0.695	0.700
i = 2	0.485	0.500
i = 3	0.336	0.450
i = 4	0.232	0.430
i = 5	0.149	0.400

4.2.5c Dividend Yield as a Predictor of Future Stock Returns

The papers cited above show that the price/dividend ratio is the best explanatory variable of long returns. To test this fact KJM (2005a) consider the following regression model

(31)
$$\varrho_{t+k}^k = \zeta_k + \eta_k (D_t / \tilde{q}_{t-1}^s) + \vartheta_{t,k}.$$

Fama and French (1988b) report that the ability of the dividend yield to forecast stock returns, measured by regression coefficient R^2 of (31), increases with the return horizon. KJM (2005a) find the model captures the main features of the empirical evidence as reported in Table 6.

Table 6: The behavior of the regression slopes in (31)

Time horizon	Model		Empirical Record	
k	η_k	R^2	η_k	R^2
1	5.03	0.08	5.32	0.07
2	8.66	0.14	9.08	0.11
3	11.16	0.18	11.73	0.15
4	13.10	0.21	13.44	0.17

To conclude the discussion of predictability, observe that the empirical evidence reported by

Fama and French (1998a, 1998b), Campbell and Shiller (1988), Poterba and Summers (1998) and others is consistent with asset price theories in which time-varying expected returns generate predictable, mean-reverting components of prices (see Summers (1986)). The important question left unresolved by these papers is what drives the predictability of returns implied by such mean-reverting components of prices? Part of the answer is the persistence of the dividend growth rate. KJM (2005a) offers a second and stronger persistent mechanism. It shows these results are primarily driven by the dynamics of market state of beliefs which exhibit correlation across agents and persistence over time. Agents go through bull and bear states causing their perception of risk to change and expected returns to vary over time. Equilibrium asset prices depend upon states of belief which then exhibit memory and mean reversion. Hence returns exhibit these same properties.

4.2.6 GARCH Behavior of the Price-dividend Ratio and of the Risky Returns

Stochastic volatility in asset prices and returns is well documented (e.g. Bollerslev, Engle and Nelson (1994), Brock and LeBaron (1996)). In partial equilibrium finance it is virtually standard to model asset prices by stochastic differential equations, *assuming* an exogenously driven stochastic volatility. But where does stochastic volatility come from? Dividends certainly do not exhibit stochastic volatility. One of the most important implication of theory of diverse beliefs is that *it explains why asset prices and returns exhibit stochastic volatility.* To motivate, KJM (2005a) present simulated price/dividend ratios and the associated risky rates of return (see their Figures 5 and 6 page 29). These reveal time variability of the variance of prices and returns with clustering of volatility which is familiar from market data . However, GARCH behavior is more subtle than just volatility clustering; it requires volatility to be persistent and this requires a formal test.

To formally test the GARCH property of the price/dividend ratio and of the risky returns KJM (2005a) use the 20,000 simulated observations noted in section 4.2.3. With that data they estimate the following econometric model of the dynamics of the log of the price/dividend ratio

$$\log(q_t^s) = \kappa^q + \mu^q \log(q_{t-1}^s) + \varsigma^{q_t}$$

(32a)
$$\varsigma_t^q \sim N(0, h_t^q)$$

$$h_t^q = \xi_0^q + \xi_1^q (\varsigma_{t-1}^q)^2 + v_1^q h_{t-1}^q.$$

Since the price dividend ratio is postulated to be an AR(1) process, the process in (32) is GARCH(1,1). Similarly, for the risky rates of return they postulated the model

$$\varrho_t^s = \kappa^\varrho + \mu^\varrho \log(q_t^s) + \varsigma^{\varrho_t}$$

(32b)
$$\varsigma_t^\varrho \sim N(0, h_t^\varrho)$$

$$h_t^\varrho = \xi_0^\varrho + \xi_1^\varrho (\varsigma_{t-1}^\varrho)^2 + v_1^\varrho h_{t-1}^\varrho.$$

For a specification of (32a) and (32b) they also tested ARCH(1) and GARCH(2,1) but have concluded that the proposed GARCH(1,1) as in (32a)-(32b) , describes best the behavior of the data. Due to the large sample they ignore standard errors and report that the estimated model for the log of the price-dividend ratio satisfies the GARCH(1,1) specification

$$\log(q_t^s) = 0.99001 + 0.69384 \log(q_{t-1}^s) + \varsigma^{q_t}$$

$$\varsigma_t^q \sim N(0, h_t^q)$$

$$h_t^q = 0.00592 + 0.02370(\varsigma_{t-1}^q)^2 + 0.73920 h_{t-1}^q, \quad R^2 = 0.481.$$

For the risky rates of return the estimated model satisfies the GARCH(1,1) specification

$$\varrho_t = 1.13561 - 0.33355 \log(q_t^s) + \varsigma_t^\varrho$$

$$\varsigma_t^\varrho \sim N(0, h_t^\varrho)$$

$$h_t^\varrho = 0.00505 + 0.01714(\varsigma_{t-1}^\varrho)^2 + 0.77596 h_{t-1}^\varrho, \quad R^2 = 0.180.$$

To explain we observe that stochastic volatility is a direct consequence of the dynamics of beliefs, defined by (z_t^1, z_t^2) in KJM (2005a). Persistence of beliefs and correlation across agents introduce these patterns into prices and returns. When agents disagree (i.e. $z_t^1 z_t^2 < 0$) they offset the demands of each other and as that pattern persists, prices do not need to change by very much for markets to clear. During such periods prices exhibit low volatility: persistence of belief states induce persistence of low volatility. When agents agree (i.e. $z_t^1 z_t^2 > 0$) they compete for the same assets and prices are determined by difference in belief intensities. Changes in the levels of bull or bear states generate high volatility in asset prices and returns. Persistence of beliefs cause such high volatility regimes to exhibit persistence. Market volatility is then time dependent and has a predictable component as in (32a)-(32b).

The virtue of the above argument is that it explains stochastic volatility *as an endogenous consequence of equilibrium dynamics*. Some "fundamental" shocks (i.e. an oil shock) surely cause market volatility, but it has been empirically established that market volatility cannot be explained *consistently* by "fundamental" exogenous shocks (e.g. Schwert (1989), Pesaran and Timmermann (1995), Beltratti and Morana (2006)). The KJM (2005a) explanation of stochastic volatility is thus consistent with the empirical evidence

4.3   The Endogenous Uncertainty Risk Premium

We now return to an analytical exploration of the risk premium under heterogenous beliefs in the illustrative model of Section 3.4. Using to the notation of that section we follow the ideas in Kurz and Motolese (2006). The risk premium on a long position, *as a random variable*, is defined by

(33)
$$\pi_{t+1} = \frac{p_{t+1} + d_{t+1} + \mu - Rp_t}{p_t}.$$

We want a measure which is *a known expected quantity* recognized by market participants but we have a problem since with diverse beliefs the premium is subjective. Kurz and Motolese (2006) compute three equilibrium measures to consider. One is the subjective expected excess returns by i,

(33a)
$$E_t^i(\pi_{t+1}) = \frac{1}{p_t}[(a_d+1)(\lambda_d d_t + \lambda_d^g g_t{}^i) + a_z(\lambda_z Z_t + \lambda_Z^g g_t{}^i) + \mu + P_0 - Rp_t].$$

Aggregating over i they define the *market premium* as the average market expected excess returns. It reflects what the market expects, not necessarily what the market gets:

(33b)
$$\bar{E}_t(\pi_{t+1}) = \frac{1}{p_t}[(a_d+1)(\lambda_d d_t + \lambda_d^g Z_t) + a_z(\lambda_z Z_t + \lambda_Z^g Z_t) + \mu + P_0 - Rp_t]$$

(33a) and (33b) are not necessarily "correct" and we focus on a third premium which is an objective measure, common to all agents. Econometricians who study the long term time variability of the premium would measure it by the empirical distribution of (33). Kurz and Motolese (2006) show it is

(33c)
$$E_t^m[\pi_{t+1}] = = \frac{1}{p_t}[(a_d+1)(\lambda_d d_t) + a_z(\lambda_z)Z_t) + \mu + P_0 - Rp_t]$$

We stress that (33c) is the common way all the 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

(34a)
$$E_t^i[\pi_{t+1}] - \bar{E}_t[\pi_{t+1}] = \frac{1}{p_t}[(a_d+1)\lambda_d^g + a_z\lambda_Z^g](g_t{}^i - Z_t).$$

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

(34b)
$$E_t^m[\pi_{t+1}] - \bar{E}_t[\pi_{t+1}] = -\frac{1}{p_t}[(a_d+1)\lambda_d^g + a_z\lambda_Z^g]Z_t.$$

The risk premium is different from the market perceived premium when $Z \neq 0$. But the important conclusion of Kurz and Motolese (2006) is the analytical expression of the risk premium:

(35a)
$$E_t^m[\pi_{t+1}] = \frac{1}{p_t}[(R\frac{\hat{\sigma}_Q^2}{\tau} - u_1 d_t) - a_z(R - \lambda_z)Z_t]$$

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

(35b)   *The Risk Premium* $\mathrm{E}_t^m[\pi_{t+1}]$ *is decreasing in the mean market belief* Z_t.

Conclusions (35a) -(35b) are important since they exhibit what Kurz and Motolese (2006) 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 <u>permanent mean premium</u> $R\dfrac{\hat{\sigma}_Q^2}{\tau}$. It is shown by Kurz and Motolese (2006) that there are weights $(\omega_1, \omega_{12}, \omega_2)$ such that

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

hence, the market beliefs increase the variance of ρ_{t+1}^{iZ} hence the risk premium.

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

To understand the second result 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*. From an REE perspective this sign is somewhat surprising since when $Z_t > 0$ the market expects *above normal* future dividend but instead, the risk premium on the stock *declines*. When the market holds bearish belief about dividends ($Z_t < 0$) the risk premium *rises*. Kurz and Motolese (2006) explore this result both analytically and empirically. Before proceeding let us discuss some ramifications of this result.

4.3.1 The Market Belief Risk Premium is General

    The main result (35b) was derived for the exponential utility function. Kurz and Motolese (2006) show this result is more general and depends only on the positive coefficient a_z of Z_t in the price map. To see this, 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 only on the numerator in (33) we have $E_t^m[p_{t+1} + d_{t+1} + \mu - Rp_t]$. 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 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 above note that

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

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

4.3.2   Interpretation of the Market Belief Risk Premium

Why is the effect of Z_t on the risk premium *negative*? Since this result is applicable to any asset with risky payoffs, Kurz and Motolese (2006) offer a general interpretation. The 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 the risk premium on the long position of the asset falls. Fluctuating market belief implies time variability of risk premia but fluctuations in risk premia are inversely related to the degree of market optimism about future prospects of asset payoffs.

To further explore the result, it is important to explain *what it does not say*. One may interpret it 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 the 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}. Given the available information an optimizing agent is already placed on his demand function defined relative to his own belief, hence it not optimal for him to just abandon his demand and adopt a contrarian strategy. This 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 but you dislike the sharp expected declines 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 true optimal position is to be long. Indeed, this argument explains why *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. The crucial observation to 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 used to change the estimated risk premium on investments in the bond and

stock markets. Finally, we note that Kurz and Motolese (2006) use data compiled by the Blue Chip survey of forecasts in order to test the theory proposed in (35b). They report that the data support the theoretical results.

4.4   Beauty Contests and Speculation

Although market practitioners have an intuitive idea of what "speculation" is, there is no scientific consensus on how to define this concept. Keynes (1936) viewed asset markets as a "Beauty Contest" and many writers have interpreted this to be a form of speculation. A different perspective was proposed by Kaldor (1939) who define speculation as *"the purchase (or sale) of goods with a view to resale (repurchase) at a later date."* It is clear that for such asset trades to make sense, prices of assets must regularly deviate from their fundamental values and agents must believe prices are, or will not equal their fundamental values. It is also clear that in a perfect REE world with homogenous beliefs and complete information a Kaldor speculation is not possible (e.g. see Tirole (1982), Milgrom and Stokey (1982)). Here we explore the perspective of a Rational Belief Equilibrium with respect to Beauty Contests and Kaldor (1939) speculation.

4.4.1   On Beauty Contests

The Keynes Beauty Contest metaphor has been extensively discussed. Some have associated it with asset pricing equilibrium where the price is expressed as iterated expectations of average market belief of the future *fundamental value* of the asset. In (8) we presented the Allen Morris and Shin (2006) example of such pricing with PI. But this interpretation should be questioned. An examination of Keynes' view (see Keynes (1936), page 156) shows the crux of Keynes's conception is that there is little merit in using fundamental values as a yardstick for market valuation. Hence what matters for asset demand of an agent is the perception of what the market believes the future price of that asset will be rather than what the intrinsic value is. Keynes insists future price depends upon future market belief and that *may be right or wrong* without a necessary relation to a fundamental value. The Beauty Contest parable is thus simple: a price does not depend upon an intrinsic value but rather, upon what the market believes future payoffs and valuations will be. Keynes' Beauty Contest is thus a statement that in order to forecast the price in the future an individual must forecast the future market state of belief, *when such forecasts may be "right" or "wrong*." We now show that a Rational Belief

Equilibrium (RBE) captures the essence of the Keynes Beauty Contest.

To explain we make two observations. From (27a) equilibrium price is $p_t = a_d d_t + a_z Z_t + P_0$ and this is clearly in accord with the above: *in any model of the "Beauty Contest" equilibrium price should not depend upon a true intrinsic value but rather, it should depend upon market belief.* It follows from the rationality conditions that price\earning ratios exhibit fluctuations with reversion to the long run stationary mean but such long term value is not, in general, an intrinsic fundamental value. Indeed, in a model with diverse belief there is no such thing as fundamental intrinsic value since all prices depend upon market beliefs. Second, to forecast future prices an agent in an RBE forecasts Z_{t+1} which is the market state of belief tomorrow. From (22b) we have that $Z_{t+1}^i = \lambda_Z Z_t + \lambda_Z^g g_t^i + \rho_{t+1}^{iZ}$ which means that an agent forecasts the future market belief with his own idiosyncratic model. The Beauty Contest parable explains that market volatility does not necessarily result from changes in some intrinsic values but rather, from fluctuations in market beliefs. In sum, an essential requirement of a Beauty Contest is that an "intrinsic fundamental value" is not agreed upon and hence the price map depends upon market beliefs.

4.4.2   On Speculation and RBE

Following the definition of Kaldor (1939), Harrison and Kreps (1978) study the consequences of risk-neutral investors having different beliefs about the dividend process $\{d_t\}$ of a risky asset. Investor i can expect a payment $E_t^i(\beta^{t+k} p_{t+k} + \sum_{s=0}^{k-1} \beta^{t+s} d_{t+s})$ if he chooses to resell k periods later, where $\{p_t\}$ and β denote the stock price process and the discount rate. The equilibrium market price, called a *consistent price scheme*, is the supremum over all stopping times k and across all investors. That is, this price is

$$p_t = \max_i \sup_k E_t^i(\beta^{t+k} p_{t+k} + \sum_{s=0}^{k-1} \beta^{t+s} d_{t+s}).$$

Agents hold diverse beliefs and are assumed to have infinite wealth for each class of investor type. A speculative premium is then defined to be the difference between the consistent price scheme and the value, $\max_i E_t^i(\sum_{s=0}^{\infty} \beta^{t+s} d_{t+s})$, expected when all investors are obliged to hold the asset forever. Harrison and Kreps (1978) show that under the assumptions made there exists a positive speculative premium, or a price bubble, when short sales are not allowed.

Morris (1996) further examines asset pricing during initial public offerings when investors have different prior distributions, but the difference of beliefs disappears as investors learn from

observations. A major weakness of both the works of Harrison and Kreps (1978) and Morris (1996) is their assumption of the unmodeled, arbitrary, heterogeneity of beliefs.

Wu and Guo (2003) provide the justification for the continued presence of diverse beliefs by adopting the RB theory due to Kurz (1994). Investors have diverse beliefs that are rational in the sense defined in Section 3.3.4, and investors can learn only the stationary measure of observed data. Suppose that each type i of investors adopts either Q^{i1} or Q^{i2}, depending on a state of belief, with equal probabilities. The rationality condition is then ½ Q^{i1} + ½ Q^{i2} = Y, where Y is the stationary measure of observable data. Wu and Guo (2003) show that a stationary minimal consistent price scheme in the implied Rational Belief Equilibrium (RBE) is

$$\vec{p} \ = \ \max_{(i,j \,=\, 1,2)} E^{\,Q^{ij}} (\beta \cdot \vec{p} \ + \ \beta \cdot \vec{d}),$$

where prices and dividends are vectors. Wu and Guo (2003) offer a finite algorithm for finding the unique market price $\vec{p}$ and "representative belief" Q^{*f} which satisfy the equilibrium price equation: $\vec{p} \ = \ (1 \ - \ \beta Q^{*f})^{-1} \beta Q^{*f} \cdot \vec{d}$. In general, speculative bubbles exist and Endogenous Uncertainty as defined in Section 3.2 above also emerges. They further characterize how the speculative premium increases with the degree of heterogeneity.

To explore the phenomenon of *simultaneous* increase in asset prices and trading volume, Wu and Guo (2004) study a model of heterogeneous rational beliefs held by a continuum of agents on the unit interval, as in Miller (1977). In contrast with Harrison and Kreps (1978) and Morris (1996), Wu and Guo (2004) permit limited short sales and impose a wealth constrain. Dividends follow an i.i.d. process with two realizations d_L and d_H, with $d_L < d_H$. The empirical frequency of d_H is denoted by m. These authors arrange investors in the order of their optimism along the unit interval, denoted by $\{B(i)\}_{i \in [0,1]}$, with $B'(i) > 0$. They then derive a steady state rational belief equilibrium price

$$p_S = \beta \, [\, (1 \ - \ B(i_s))(p_L + d_L) + B(i_s)(p_H + d_H) \,] \quad , \quad s = L, H$$

where i_L and i_H are "representative investors" whose willingness to pay equals the market price. In equilibrium all optimistic investors $i \in [i_s, 1]$, $s = L, H$ purchase and hold the entire supply. With correlation absent, rationality requires the average belief to be consistent with the stationary measure m hence we must have $\int_0^1 B(i) \, di \ = \ m$.

In this framework, Wu and Guo (2004) demonstrate the emergence of price amplification effect (i.e. Endogenous Uncertainty) and characterize the conditions for a positive or negative speculative premium in an RBE. They show that the positive speculative premium increases with the

level of investment fund and degree of optimism, and decreases with the amount of short sale constraint. Furthermore, the model generates a positive relationship between trading volume and the directions of price changes and a positive relationship between trading volume and price level. These results are consistent with the empirical evidence (e.g. Karpoff (1987) and Basci et al. (1996)).

4.5   Volatility of foreign exchange rates and the forward discount bias

The relevance of foreign exchange markets to our discussion in this chapter is motivated by the following problem. Estimate a regression of the form

$$(36) \qquad \frac{ex_{t+1} - ex_t}{ex_t} = c + \zeta(r_t^D - r_t^F) + \varepsilon_{t+1}$$

where $(ex_{t+1} - ex_t)$ is the change of the exchange rate between t and t + 1 and $(r_t^D - r_t^F)$ is the difference between the short term nominal interest rates in the domestic and the foreign economies. Under rational expectations $(r_t^D - r_t^F)$ is an unbiased predictor of $(ex_{t+1} - ex_t)$. It is motivated by a standard arbitrage argument: if there is a differential in nominal rates agents can borrow in one country and invest in the other and gain from the difference if the exchange rate does not move against them by the amount of the differential. In a no arbitrage REE a rationally expected change in the exchange rate must then be equal to the interest rate differential. This means that apart from a technical correction for risk aversion, the parameter ζ should be close to 1. In 75 empirical studies ζ was estimated to be significantly less than 1 and in many studies it was estimated to be negative (see Froot and Frankel (1989), Frankel and Rose (1995), and Engel (1996) for an extensive survey).

The failure of ζ to exhibit estimated values close to 1 is known as the "Forward Discount Bias" in foreign exchange markets. The empirical fact is that exchange rates are far more volatile than can be explained by differentials in nominal interest rates or inflation rates between countries. But changes in foreign exchange rates are not predictable and interest rate differentials account only for a small fraction of the movements in foreign exchange rates. However, it is not surprising that this lack of predictability decreases with the length of time involved. That is, long run differentials in nominal interest rates do exhibit better predictive power of long run movements in foreign exchange rates since long run differentials in nominal rates reflect differentials in inflation rates. Since the problem at hand is the nature of market expectations and exchange rate volatility, it is a natural for us to consider

it here and the model of diverse belief is an obvious candidate to be used to solve the problem.

Applying the RBE theory to this problem, Kurz (1997b) and Black (1997), (2005) developed a model which is similar to the Kurz and Motolese (2001) model except for treating the second agent as a second country and adding two short term nominal debt instruments. A similar model was also reformulated by Kurz and Motolese (2001). Limitation of space makes it impossible to review all technical details of these models here. Instead, we outline the key points of the model construction and note the results. Hence, the central model construction elements in these papers are as follows

- consider the first agent as the "domestic U.S." which is the home country and the second agent as a "foreign economy;"

- introduce a second shock which is associated with productivity in the foreign economy and is different from the first shock defined for productivity in the domestic economy;

- introduce a monetary system for both countries and a second currency;

- introduce two nominal interest rates and two different monetary policies;

- there are the two standard financial assets: (i) ownership shares of a domestic firm with stochastic dividend whose stock trades freely in both currencies across the countries, and (ii) a zero net supply riskless bond which pays a unit of consumption and which trades in both currencies across the countries;

- introduce a simple production structure for the foreign economy;

Note that the models above do not aim to simulate the U.S. or world economies. They merely aim to explain via simulations why a model with diverse beliefs imply $\zeta < 1$. And indeed, all models produce estimated parameters ζ which are significantly less than 1: in the RBE of Kurz (1997b) the estimated ζ is around 0.25, in Black (1997), (2005) it is around 0.15 and in Kurz and Motolese (2001) it is around 0.45. More realistic results could be obtained by formulating more realistic models but the key idea of these papers is that the result $\zeta < 1$ is virtually independent of the model formulation. We now provide an explanation for this strong conclusion.

Why does the RBE predict that ζ is less than 1? If $\zeta < 1$ in an REE agents can make an *expectational* arbitrage: they can borrow today in one currency, invest in the other and *expect* that the net return on their investment *next period* will be larger than the depreciation of the currency. In such an equilibrium all agents hold the same self-fulfilling expectations, the *expectational* arbitrage becomes a *real* arbitrage and consequently this implies that $\zeta < 1$ cannot hold in equilibrium.

In a world with diverse beliefs equilibrium exchange rate depends upon the distribution of beliefs and hence exchange rates exhibit excess volatility, reflecting the variability of investors' beliefs. Indeed, volatility of foreign exchange rates is dominated by Endogenous Uncertainty. The implication is that *regardless of the information today*, to forecast future exchange rates agents must forecast future market states of beliefs rendering exchange rate virtually unpredictable . Hence, if a condition of differential nominal interest rates across countries arises, it can never be the only factor that will determines the exchange rate next period. With risk averse agents who are unable to predict the exchange rate a condition of differential interest rates will not generate the beliefs of traders that the exchange rate will, in fact, adjust. Failing to expect the exchange rate to adjust, they will not undertake such arbitrage and the exchange rate will, in fact, not adjust. This mechanism ensures that a differential of nominal interest rates between the two countries is not an unbiased estimate of the rate of depreciation of the exchange rate one period later hence $\zeta < 1$. This reasoning does not hold in the long run since a long term differential of nominal interest rates will persuade the markets that the exchange rate must adjust in the long run and this will persuade them to engage in such arbitrage.

4.6 Macroeconomic Applications

Although there is a wide range of potential applications in Macroeconomics, so far only limited questions have been studied with the model of diverse beliefs. Motolese (2001), (2003) shows that in an economy with diverse beliefs money is not neutral. To see why it is important to observe that before Rational Expectations the case for money neutrality was based on the quantity theory of money. The main contribution of Lucas (1972) was to show that money neutrality can be proved only by an exploration of the structure of expectations. In a model with heterogenous beliefs agents hold diverse beliefs about the relative effects of productivity growth and money shocks hence they hold diverse beliefs about future inflation. With diverse expectations money cannot be neutral.

Kurz, Jin and Motolese (2005b) is a comprehensive study of the efficacy of monetary policy in an economy with diverse beliefs. The authors show that diverse beliefs constitute an important propagation mechanism of fluctuations, money non neutrality and efficacy of monetary policy. Since expectations affect demand, the theory shows that economic fluctuations are driven mostly by varying demand not supply shocks. Using a competitive model with flexible prices in which agents hold Rational Beliefs the authors arrive at six conclusions:

(i) the model economy replicates well the empirical record of fluctuations in the U.S.

(ii) Under monetary rules without discretion, monetary policy has a strong stabilization effect and an aggressive anti-inflationary policy can reduce inflation volatility to zero.

(iii) The statistical Phillips Curve changes substantially with policy instruments and activist policy rules render it vertical.

(iv) Although prices are flexible, money shocks result in less than proportional changes in inflation hence the aggregate price level is "sticky" with respect to money shocks.

(v) Discretion in monetary policy adds a random element to policy and increases volatility. The impact of discretion on the efficacy of policy depends upon the structure of market beliefs about future discretionary decisions. The paper studies two rationalizable beliefs. In one case, market beliefs weaken the effect of policy and in the second, beliefs bolster policy outcomes hence, in this case, discretion is a desirable attribute of the policy rule. That is, social gains from discretion arise only under special structures of belief of the private sector about future bank discretionary acts and such requirement complicates the bank's problem. Hence, the weight of the argument leads Kurz, Jin and Motolese (2005b) to conclude that bank's policy *should be transparent and abandon discretion except for rare and unusual circumstances.* This analysis is in contrast with the recent literature initiated by Morris and Shin (2005) and others who suggest that due to asymmetric private information central bank transparency has inherent cost of failing to retrieve useful private information by the bank. We have rejected the applicability of the private information model for the study of economic aggregates such as interest rates, inflation rate or GDP growth. Hence, the Morris and Shin (2005) model does not address the real problem associated with the objective of Central Bank transparency, which is the coordination of expectations.

(vi) One implication of the model suggests that the present day policy is only mildly activist and aims mostly to target inflation.

5.   Conclusions and Open Problems

Rational Expectations and Irrational Behavior are extreme hypotheses: with REE one cannot explain the observed data on market dynamics and with irrational behavior one can prove anything. We highlight here the merit of an intermediate concept of belief rationality that emerges from the fact

that the economy is a non-stationary system with time varying structure. This prevents agents from ever learning true structural relations and probability laws. All they learn are the empirical frequencies from which emerges a common knowledge of a stationary probability reflecting the long term dynamics. Belief rationality requires agents to hold only beliefs which are not contradicted by the empirical evidence. But since it is irrational to believe in a fixed deviation from the stationary probability, such belief rationality implies belief dynamics: individual beliefs must be time varying and correlation across agents generates a new aggregate force in market dynamics which is the dynamics of market beliefs. The main observation made in this chapter is that the dynamics of market belief is a central market force which is as important to asset pricing and allocation as the dynamics of productivity or public policy. Indeed, the dynamics of market belief explains well the 4 recessions which Samuelson noted the market predicted but which did not happen. It shows that a rational market makes forecasting mistakes and rational investors are not infallible. They may use wrong forecasting models. Once we recognize that being rational and being wrong are not incompatible and no psychological impulses are needed for this proposition, we are open to a new paradigm of market dynamics. It provides a coherent explanation to most dynamical phenomena of interest as outlined in this chapter. We thus sum up our six central conclusions:

(i)   Diverse beliefs without any private information is an empirical fact and such diversity provide a strong motive to trade assets and hedge subjectively perceived risks.

(ii)   Financial markets are the great arena for agents to trade differences in beliefs.

(iii)   The dynamics of market beliefs is a central component of asset price volatility and this component of risk has been named "Endogenous Uncertainty."

(iv)   Asset markets exhibit large excess volatility of prices, returns and high volume of trade due to the dynamics of belief.

(v)   Risk premia reflect the added market risk due to the dynamics of beliefs and in some markets the component of risk premia due to the dynamics of market belief is very significant

(vi)   Distributions on market beliefs are observable since agents are anonymous and do not consider belief as a source of private information. However, in practice, beliefs cannot be precisely observed since data on distribution of beliefs can only be deduced from samples drawn from the population, and these are subject to sampling errors,

We note that important problems, which we have not discussed, are still open. some of which

are being under investigation. Four examples are as follows:

Pareto Optimality. The concept of ex-ante Pareto Optimum is not a satisfactory concept for market with diverse beliefs . To attain any Pareto improvment all agents must believe it is an improvement and that is not likely. Hence, most stabilization policies would not be Pareto improving. Following the idea of ex-post Pareto Optimality (e.g. Starr (1973) and Hammond (1981)) progress on this issue was made by Nielsen (2003) and (2006) who argued that a currency union is superior to multiple currencies since a union would eliminate endogenous uncertainty inherent in foreign exchange rates.

Stabilization Policy. When the problem of Pareto Optimality is resolved, the door will be open to a study of the desirability of stabilization public policies. Some start has been made by Kurz, Jin and Motolese (2005b) regarding stabilizing monetary policy. But the question is broader. Should the Fed target the stock market? Should countries cooperate to avoid an international financial crisis? What is the role of an international convention regarding bank reserve requirements? Under REE these type of questions are set aside since it is often argued that the market solution is best and no cooperative policy is needed. In a world of diverse beliefs this is not true and the question is open.

Continuous time reformulation. A continuous time reformulation of the RB theory would open the door to a study of the decomposition of risk into fundamental and endogenous components. With such formulation available one can formulate the decomposition of the values of derivative securities, using Black Scholes, into the fundamental and endogenous components. Such a decomposition is likely to provide an explanation to the Smile Curves in derivative pricing.

Destabilizing speculation of futures markets. Could the opening of a future's market increase the volatility of a spot market? This is an old question which has not been fully clarified. Our conjecture is that a proper formulation of the problem will show that if margin requirements and leverage conditions are sufficiently relaxed in a futures market, its opening could give rise to endogenous uncertainty which cannot arise in the spot market if storage cost are high enough. This could increase the volatility of a spot market.

Appendix: Deriving (11) From A Bayesian Inference

A.1   *Qualitative Information and Subjective Interpretation of Public Information*

Bayesian inference is only possible with *quantitative* data. But quantitative data like d_t are always accompanied with *qualitative* information. For example, data on profits of the S&P500 are

interpreted in light of reports on abnormal productivity growth, unusual public opinion survey, new government policies, etc If d_t are profits it is just one number extracted from a financial report of the firm which contains other information. The firm may announce a new research project that could alter the way it will do business, or it may announce an organizational change. Qualitative information has indirect quantitative measures reflecting the intensity of the activity contemplated. For example, a new research project could be large or small, a new organizational structure may be major and cover the entire firm or restricted to a small portion of it. Qualitative data cannot, in general, be compared over time. When a firm starts research into something that does not exist, no past data is available. 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 statements which are the focus of diverse investors' opinions.

There is little modeling of deduction from qualitative information. Saari (2006) uses qualitative information in a competitive model of market shares. Toukan (2006) is a second example. Here we provide a simple formalization (see Kurz (2007) and Kurz and Motolese (2006)) of the use of qualitative information. Thus, qualitative information consist of *statements about the future*. Let date t statements be $C_t = (C_{t1}, C_{t2}, ..., C_{tK_t})$, each with quantitative measures in some units. The list may change with t and K_t varies with time. The activity in a statement may turn out to impact $(d_{t+1} - \lambda_d d_t)$ or not. The effects may be desirable or not. A realization at t+1 is a vector $\varphi_{t+1} = (\varphi_{t+1,1}, \varphi_{t+1,2}, ..., \varphi_{t+1,K_t})$ of numbers which are 0 or 1. A 0 means the activity turns out to have no effect and 1 means it has an effect. These can be interpreted as "success" or "failure." There are 2^{K_t} possible vectors of outcomes $\varphi_{t+1}(k)$, $k = 1, 2, ..., 2^{K_t}$. Next, agent i has a subjective map from φ_{t+1} to an expected value $\Phi^i(\varphi_{t+1})$ of $(d_{t+1} - \lambda_d d_t)$. This is an independent estimate by agent i on how different he expects $(d_{t+1} - \lambda_d d_t)$ to be from the stationary forecast conditional upon the statements C_t. Finally, conditional on C_t, agent i attaches probabilities $(a_1^i, a_2^i, ..., a_{2^{K_t}}^i)$ to the vectors $\varphi_{t+1}(k)$. This results in agent i making an alternate subjective estimate of $(d_{t+1} - \lambda_d d_t)$ based only on public C_t data:

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

Since by (12a) the long term average of $(d_{t+1} - \lambda_d d_t)$ is zero, rationality requires the Ψ_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.

A.2   *A Bayesian Inference: Beliefs are Markov State Variables*

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

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

Agents do not know true values b_t but β is known. At first decision date t (say, t = 1) an agent has two pieces of information. He knows d_t and observes qualitative information $(C_{(t)1}, C_{(t)2}, ..., C_{(t)K_t})$ with which to assess Ψ_t^i. Without Ψ_t^i the prior subjective mean 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 $E_{t+1}^i(b_t|d_{t+1}, \Psi_t^i)$ [15] in a standard Bayesian procedure. But he needs an *estimate of b_{t+1}*, not of b_t. Hence, his problem is how to go from $E_{t+1}^i(b_t|d_{t+1}, \Psi_t^i)$ to a prior of b_{t+1}? Without new information his belief about b_{t+1} is unchanged and he would use $E_{t+1}^i(b_t|d_{t+1}, \Psi_t^i)$ as the new prior of b_{t+1}. This is surely true if the b's change very slowly or when $b_{t+1} = b_t$. Hence, if agent i believes $b_{t+1} \neq b_t$ a new prior is needed. To that end he uses the qualitative information $(C_{(t+1)1}, C_{(t+1)2}, ..., C_{(t+1)K_{t+1}})$ released publicly before trading at t+1. These lead to an alternate subjective estimate Ψ_{t+1}^i of b_{t+1}. Now the agent has two independent sources for belief about b_{t+1}: the last posterior $E_{t+1}^i(b_t|d_{t+1}, \Psi_t^i)$ is to be used if $b_{t+1} = b_t$ and Ψ_{t+1}^i if $b_{t+1} \neq b_t$. With a Bayesian approach we assume:

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

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

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

(AP.2) $\qquad b_1 \sim N(\mu b + (1-\mu)\Psi_1^i, \frac{1}{\alpha})$ for some α.

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, $E_{t+1}^i(b_{t+1}|d_{t+1}, \Psi_{t+1}^i)$, before d_{t+2} is observed. The following result is then available:

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

---

[15] Recall that we use the notation $E_t^i(b_t|d_t, \Psi_t^i)$ for the *prior* belief at date t about the unknown parameter b_t used to forecast d_{t+1}. We then use the notation $E_{t+1}^i(b_t|d_{t+1}, \Psi_t^i)$ for the posterior belief about b_t given the observation of d_{t+1}. Assumption (A) will use this posterior belief as a building block in the formation of the new prior $E_{t+1}^i(b_{t+1}|d_{t+1}, \Psi_{t+1}^i)$.

$g_t^i = E_t^i(b_t | d_t, \Psi_t^i)$ and $\mu = \lambda_Z$ then the dynamics (11) holds: (AP.1) implies (11).

Theorem 1 shows that as the d_t data set increases, there is nothing new to learn. The posterior fluctuates forever but its dynamic *law of motion* converges to a time invariant Markov transition (11). This provides a foundations for individual and market beliefs. New data d_t and subjective estimates Ψ_t^i alter the conditional probability of agents, but do not change the law of motion of g_t^i.

References

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

Amato, J. and Shin, H.S. (2003): "Imperfect Common Knowledge and the Information Value of Prices." Working Paper, London School of Economics and the IBS.

Angeletos, G.M., and Pavan, A. (2006): "Efficient Use of Information and Social Value of Information." Working Paper, MIT.

Angeletos, G.M., and Werning, I. (2006): "Crises and Prices: Information Aggregation, Multiplicity and Volatility." American Economic Review, **96**, 1721 - 1737.

Bacchetta, P. and van Wincoop, E. (2005): "Higher Order Expectations in Asset Pricing" Working paper, Study Center Gerzensee.

Bacchetta, P. and van Wincoop, E. (2006): "Can Information Heterogeneity Explain the Exchange Rate Determination Puzzle?" *American Economic Review* **96***,* 552 - 576.

Basci, E., Ozyildirim, S. and Aydogan, K. (1996): "A Note On Price-Volume Dynamics in An Emerging Stock Market." *Journal of Banking and Finance*, **20**, 389-400.

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

Beltratti, A. and Morana, C. (2006): "Breaks and Persistency: Macroeconomic Causes of Stock Market Volatility." *Journal of Econometrics*, 131, 151 - 177.

Black, S. (1997): "The forward discount puzzle in a rational beliefs framework." Working paper, Department of Economics, Stanford University, May

Black, S. (2005): "The Role of Endogenous Uncertainty in Foreign Exchange Markets." Ph.D. Dissertation Submitted to the Department of Economics, Stanford University.

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

Bollerslev, T., Engle, R., Nelson, D. (1994): "ARCH Models." In Engle, R., McFadden, G., (eds.) *Handbook Of Econometrics*, **Vol. IV**, Amsterdam: North Holland.

Brock, W.A. and LeBaron, B. (1996): "A dynamic structural model for stock return volatility and trading volume." *The Review of Economics and Statistics* **78**, 94 - 110

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

Cabrales, A., Hoshi, T. (1996): Heterogeneous Beliefs, Wealth Accumulation, and Asset Price Dynamics, Journal of Economic Dynamics and Control, **20**, 1073 -1100.

Camerer, C., Lovallo, D., (1999): Overconfidence and Excess Entry: An Experimental Approach. *American Economic Review*, **89**, 306 - 318.

Campbell, J.Y., Shiller, R.J. (1988): "Stock prices, earnings, and expected dividends." *Journal of Finance* **43**, 661-676

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

Diamond, D. and Verrecchia, R. (1981): "Information Aggregation in a Noisy Rational Expectations Equilibrium." *Journal of Financial Economics*, **9**, 221 - 235

Engel, C.M. (1996): "The Forward Discount Anomaly and the Risk Premium: A Survey of Recent Evidence." *Journal of Empirical Finance* **3**, 123 - 192

Fama, E.F., French, K.R. (1988a): "Permanent and Temporary Components of Stock Prices." *Journal of Political Economy,* **96**, 246-273

Fama, E.F., French, K.R. (1988b): "Dividend Yields and Expected Stock Returns." *Journal of Financial Economics,* **22**, 3-25

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

Frankel, J.A., Rose, A.K. (1995): "A Survey of Empirical Research on Nominal Exchange Rates." In Grossman, G.M., Rogoff, K. (ed.).: *Handbook of international economics* **III**, Chapter 33, 1689-1729, Amsterdam: North Holland

Froot, K.A., Frankel, J.A. (1989): "Forward Discount Bias: Is it an Exchange Risk Premium?" *Quarterly Journal of Economics* 104, 139-161

Garmaise, M.J. (2001): "Rational Beliefs and Security Design." *Review of Financial Studies*, 14, 1183 - 1214.

Grossman, S.J. (1981): An Introduction to the Theory of Rational Expectations Under Asymmetric Information. *Review of Economic Studies* **154**, 541-559.

Grossman, S.J., Stiglitz, J. (1980): "On the impossibility of informationally efficient markets." *American Economic Review* **70**, 393-408.

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

Guo, W.C., Wu, H.M. (2007): "Pyramiding/Depyramiding and Margin Trades." Working Paper, July.

Hammond, P. (1981):" Ex-Ante and Ex-Post Welfare Optimality Under Uncertainty." *Economica*, **46**, 137-147.

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

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

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

Hellwig, C. (2002) : "Public Announcements, Adjustment Delays and the Business Cycle." Working Paper, Department of Economics, UCLA.

Hellwig, C. (2005) : "Heterogenous Information and Welfare Effects of Public Information Disclosures." Working Paper, Department of Economics, UCLA.

Hellwig, C., Mukherji, A. and Tsyvinski, A. (2006) : "Self-Fulfilling Currency Crises: The Role of Interest Rates." *American Economic Review*, 96, 1769-1787.

Judd, K. L., Bernardo, A.E.(1996) :"Volume and Price Formation in an Asset Trading Model with Asymmetric Information,", Working Paper, Stanford University.

Judd, K. L., Bernardo, A.E. (2000) : "Asset Market Equilibrium with General Tastes, Returns, and

Informational Asymmetries", *Journal of Financial Markets* **1**, 17- 43.

Kaldor, N. (1939): "Speculation and Economic Stability." *Review of Economic Studies*, **7**, 1-27

Kandel, E., Pearson, N.D. (1995): "Differential Interpretation of Public Signals and Trade in Speculative Markets." *Journal of Political Economy* 4,   831-872.

Karpoff, J. (1987): "The Relation Between Price Changes and Trading Volume: A Survey." *Journal of Financial and Quantitative Analysis*, **22**, 109 -126.

Keynes, J. M. (1936): *The General Theory of Employment, Interest and Money*. Macmillan: London.

Kurz, M. (1974): "The Kesten-Stigum Model and the Treatment of Uncertainty in Equilibrium Theory." In Balch, M.S., McFadden, D.L., and 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.) (1997a))

Kurz, M. (ed.) (1996): Symposium: Rational Beliefs and Endogenous Uncertainty. *Economic Theory*, 8, 383 - 553.

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): "Asset prices with rational beliefs." In Kurz, M. (ed.).: (1997a), Chapter 9, 211 - 250.

Kurz, M. (1997c): "On the Volatility of Foreign Exchange Rates." In Kurz, M. (ed.).: (1997a), Chapter 12, 317 - 352.

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

Kurz, M. (2007): " Private Information and Diverse Beliefs: How Different?" *Journal of Mathematical Economics*, (forthcoming).

Kurz, M., Beltratti, A. (1997) : "The Equity Premium is No Puzzle." In Kurz, M. (ed.).: (1997a), Chapter 11, 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., Motolese, M. (2006): "Risk Premia, Diverse Beliefs and Beauty Contests." Working paper, Stanford University, August, 2006.

Kurz, M., Schneider, M.(1996): "Coordination and Correlation in Markov Rational Belief Equilibria." *Economic Theory* **8**, 489 - 520.

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)).

Lindahl, E. (1939): *Studies in the theory of money and capital*. London: George Allen and Unwin Ltd.

Lucas, R. E. (1972): "Expectations and the Neutrality of Money." *Journal of Economic Theory*, **4**, 103 - 124.

Mehra, R., Prescott, E.C. (1985): " The equity premium: a puzzle." *Journal of Monetary*

Economics **15**, 145-162.

Milgrom, P., Stokey, N. (1982): "Information, trade and common knowledge." *Journal of Economic Theory* **26**, 17-27.

Miller, E. M. (1977): "Risk, Uncertainty and Divergence of Opinions." *Journal of Finance* **32**, 1151-1168.

Morris, S. (1996) : "Speculative Investor Behavior and Learning." *Quarterly Journal of Economics*, **111,** 1113 - 1133.

Morris, S., Shin, H.S. (2002) : "Social Value of Public Information." *American Economic Review*, **92,** 1521 - 1534.

Morris, S., Shin, H.S. (2005) : "Central Bank Transparency and the Signal Value of Prices ." *Brookings Papers on Economic Activity*, **2,** 1 - 66.

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.

Myrdal, G.: *Monetary equilibrium* (1939). Reprints of economic classics, New York: Augustus M. Kelley 1962.

Nakata, H. (2007): "A Model of Financial Markets with Endogenously Correlated Rational Beliefs." *Economic Theory*, (forthcoming).

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**, 293 - 315.

Nielsen, K. C. (2005): "On Rationally Confident Beliefs and Rational Overconfidence." Working paper, Catholic University of Milan.

Nielsen, K. C. (2006): "Optimal Economic Institutions Under Rational Overconfidence." Quaderni dell'Instituto di Politica Economica, No. 49, Universita Cattolica del Sacro Cuore, Milano.

Paye, B., Timmerman, A. (2003): "Instability of Return Prediction Models." Working Paper, UCSD.

Pesaran, H., Timmerman, A. (1995): "Predictability of Asset Returns: Robustness and Economic Significance." *Journal of Finance*, **50**, 1201-1228.

Phelps, E.(1970): " Introduction : The New Microeconomics in Employment and Inflation Theory." In *Microeconomic Foundations of Employment and Inflation Theory*, New York: Norton, 1970.

Pigou, A.C. (1941): *Employment and equilibrium.* London: Macmillan and Co.

Poterba, J.M., Summers, L.H. (1988): "Mean Reversion in Stock Prices: Evidence and Implications." *Journal of Financial Economics,* **22**, 27-29.

Russo, J.E., Schoemaker, P.J.H. (1992): "Managing Overconfidence." Sloan Management Review, Winter.

Schwert, G.W. (1989):"Why Does Stock Market Volatility Changes Over Time?" *Journal of Finance*, 64, 1115-1163.

Shiller, R.J. (1981): "Do stock prices move too much to be justified by subsequent changes in dividends?" *American Economics Review*, **71**, 421 - 436.

Shilling, G. A. (1992): "Market Timing: Better Than Buy-and-Hold Strategy." *Financial Analysis Journal*, **48**, 46 - 50.

Singleton, K. (1987): "Asset Prices in a Time-Series Model with Disparately Informed, Competitive

Traders." In, Barnet, W. and Singleton, K. (ed.) *New Approaches to Monetary Economics, Proceedings of the Second International Symposium in Economic Theory and Econometrics.* Cambridge: Cambridge University Press.

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

Starr, R. (1973): "Optimal Production and Allocation Under Uncertainty." *Quarterly Journal of Economics*, **87**, 81-95.

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.

Summers, L.H.(1986): "Does the Stock Market Rationally Reflect Fundamental Values?" *Journal of Finance,* **41**, 591-601.

Svenson, O. (1981): "Are We All Less Risky and More Skillful Than Our Fellow Drivers?" *Acta Psychologica,* **47**, 143 - 148.

Thornton, H. (1802): *An inquiry into the nature and effects of the paper credit of Great Britain.* Edited with an introduction by F.A. v. Hayek, New York: Augustus K. Kelley (1962).

Tirole, J. (1982): "On the Possibility of Speculations Under Rational Expectations." *Econometrica*, **50**, 1163 - 1181.

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

Townsend, R. (1978): "Market Anticipations, Rational Expectations and Bayesian Analysis." *International Economic Review*, **19**, 481 - 494.

Townsend, R. (1983): "Forecasting the Forecasts of Others." *Journal of Political Economy*, **91**, 546 - 588.

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.(1993): "A model of intertemporal asset prices under asymmetric information." *Review of Economic Studies* **60**, 249 - 282.

Wang, J. (1994): "A Model of Competitive Stock Trading Volume." *Journal of Political Economy*, 102, 127 - 167.

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**, 795 - 829.

Woodford, M. (2003) : "Imperfect Common Knowledge and the Effect of Monetary Policy." Chapter 1 in Aghion, P., Frydman, R., Stiglitz, J., Woodford, M. (ed.) *Knowledge, Information and Expectations in Modern Macroeconomics, in Honor of Edmund S. Phelps* Princeton: Princeton University Press.
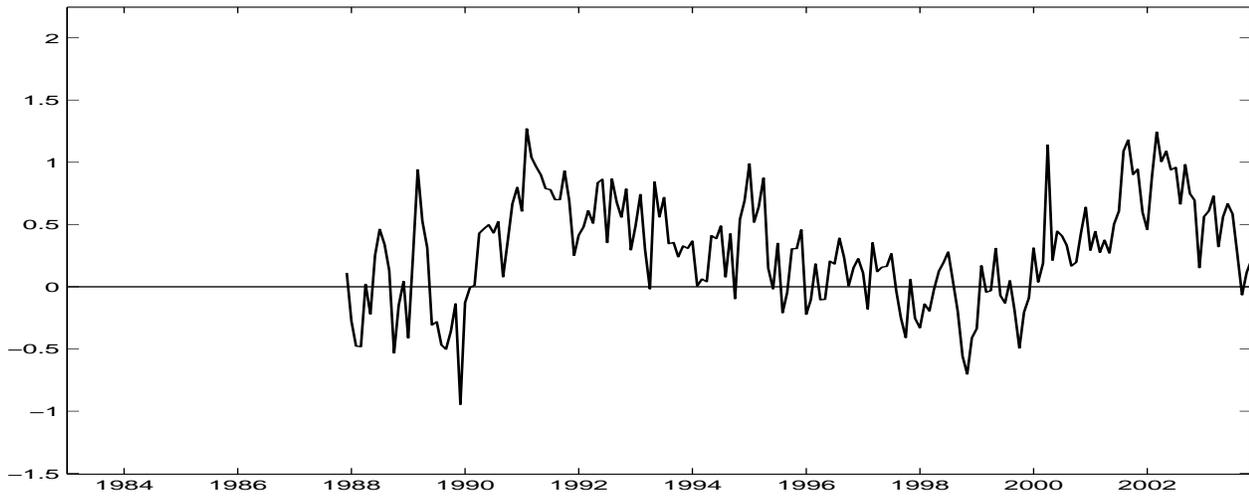
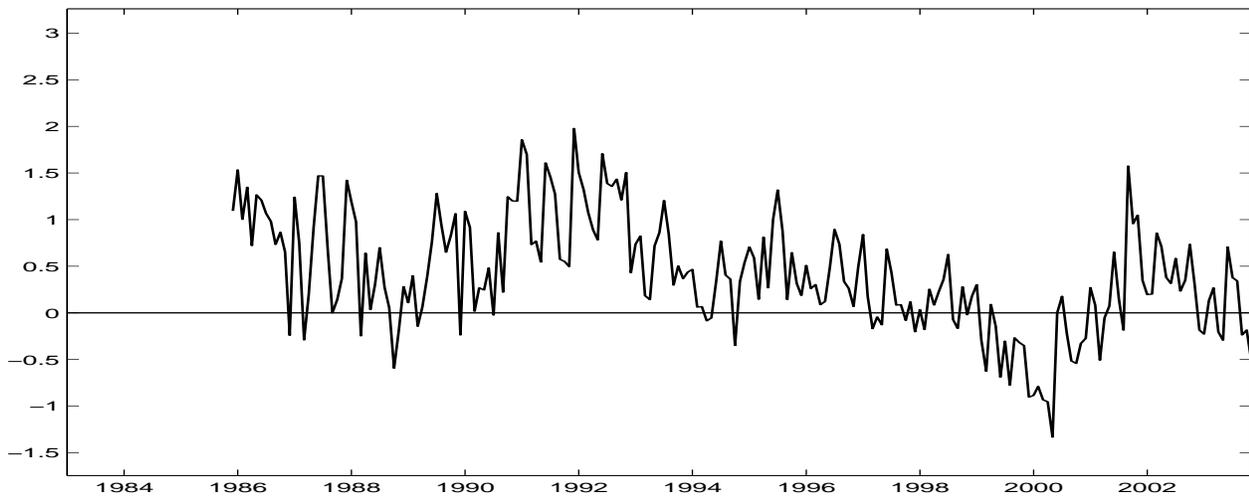Figure 1: 6-month Treasury Bill rate: 6 months ahead Market Belief.

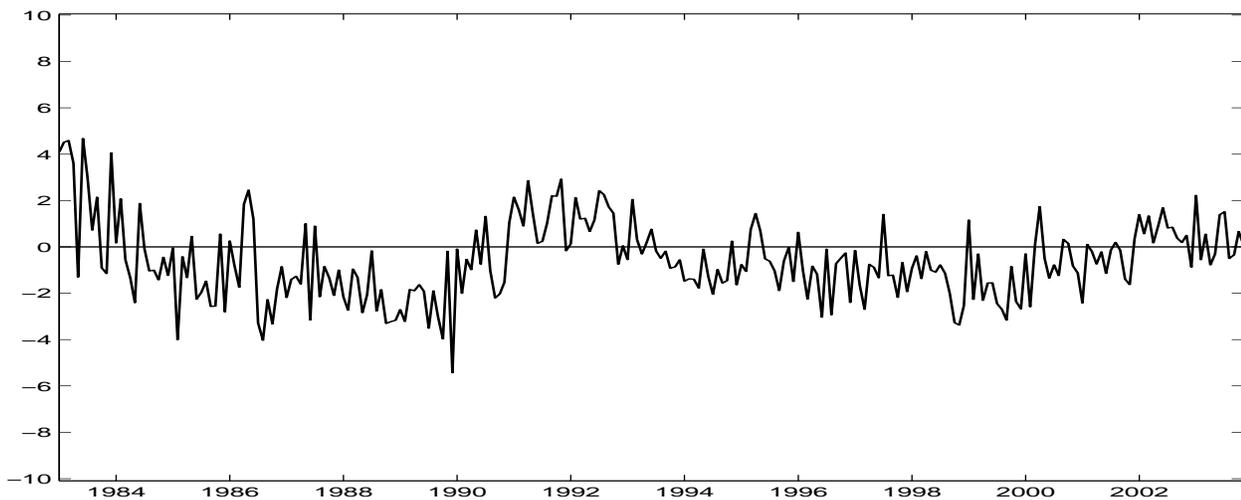Figure 2: GDP deflator for inflation rate: 2 quarters ahead Market Belief.

Figure 3: Month-over-Month, annualized growth rate of Industrial Production: 6 months ahead Market Belief.
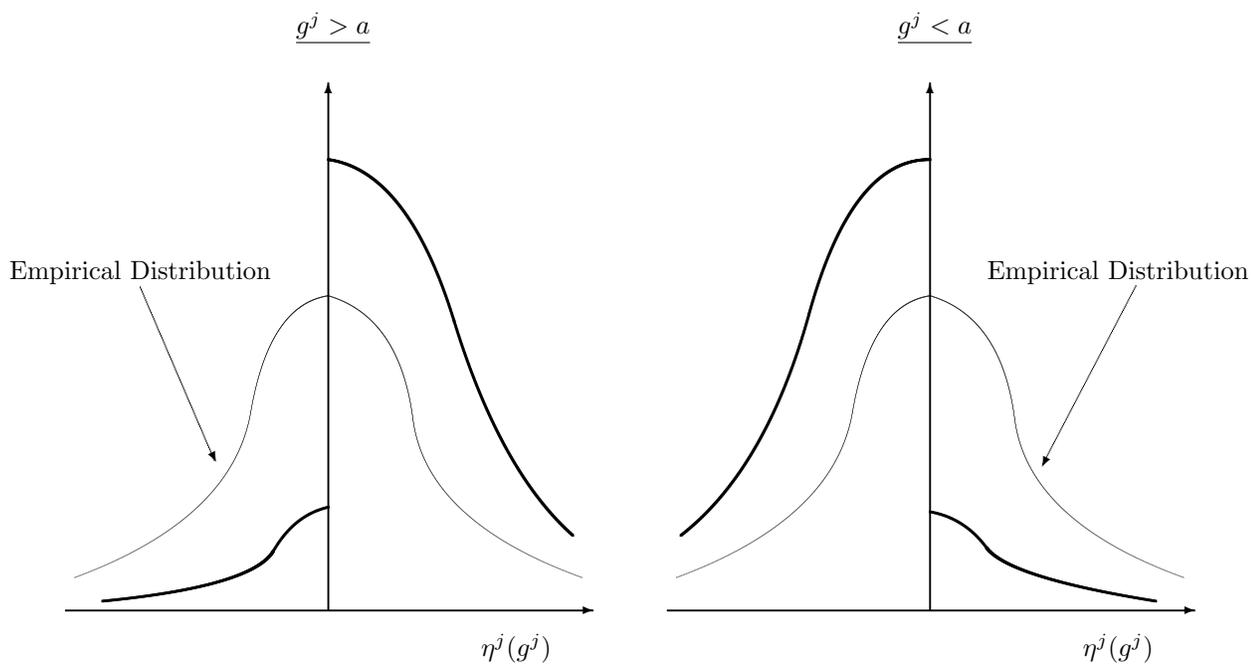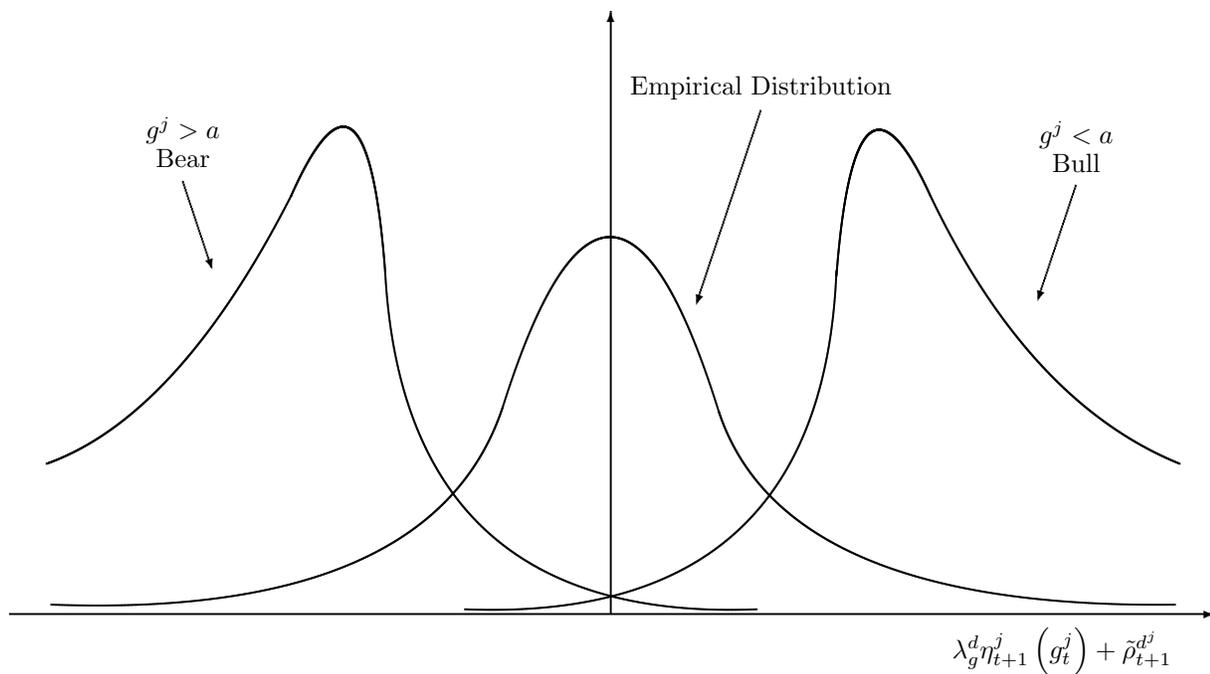
Figure 4: Non-normal belief densities.

Figure 5: Density $\Psi(g_t^j)$ with fat tails.