

# Heterogeneous Beliefs, Speculation and Trading in Financial Markets

José Scheinkman\*      Wei Xiong†

June 22, 2004

**Summary.** We survey recent developments in finance that analyze how heterogeneous beliefs among investors generate speculation and trading. We describe the joint effects of heterogeneous beliefs and short-sales constraints on asset prices, using both static and dynamic models, discuss the no-trade theorem in the rational expectations framework, and present investor overconfidence as a potential source of heterogeneous beliefs. We review recent results of Scheinkman and Xiong (2003) modeling the resale option that is embedded in share prices in the presence of short-sale constraints and heterogeneous beliefs, highlighting the implied correlation between stock prices and trading volume. Finally, we discuss the survival of investors with incorrect beliefs.

## 1 Introduction

Standard asset pricing theories have difficulty explaining episodes of asset price bubbles such as the one that seems to have occurred in the market for US internet stocks during the period of 1998-2000. In addition to asset prices that are difficult to justify by fundamentals such as expected future dividends, one typically observes inordinate increases in trading volume. For instance, Ofek and Richardson (2003) document that during the internet bubble of the late 90's, Internet stocks represented six percent of the market capitalization but accounted for 20% of the publicly traded volume of the U.S. stock market. Cochrane (2002) provides additional support for the correlation between bubbles and trading volume for US stocks during the late

---

\*Princeton University E-mail: joses@princeton.edu; Phone: (609) 258-4020.

†Princeton University. E-mail: wxiong@princeton.edu; Phone: (609) 258-0282.

90's. This evidence indicates that a satisfactory theory of bubbles should be able to explain simultaneously the level of prices and trading volume.

Several papers have been written over the last couple of decades that emphasize the role of heterogeneous beliefs in generating higher levels of asset prices and trading volume. In this chapter we present a selective survey of this literature.

We start by expounding a simple point made by Miller (1977), who argued that, if agents have heterogeneous beliefs about an asset's fundamentals and short sales are not allowed, equilibrium prices would, if opinions diverge enough, reflect the opinion of the more optimistic investor.

The Miller model is static and cannot be used to discuss the dynamics of trading. Harrison and Kreps (1978) exploit the *dynamic* consequences of heterogeneous beliefs. Since an investor knows that, in the future, there may be other investors that value the asset more than he does, the investor is willing to pay more for an asset than he would pay if he was forced to hold the asset forever. The difference between the investor's willingness to pay, and his own discounts expected dividends reflects a speculative motive, the willingness to pay more than the intrinsic value of an asset because the ownership of the asset gives the owner the right to sell it in the future. To make this right valuable, short sales must be costly - in the Harrison and Kreps model it is simply assumed that short sales are not possible.

In the Harrison-Kreps model there is a single unit of an asset and several classes of risk neutral traders that disagree about the probability distribution of future dividends. The reservation price of a buyer is the supremum, over all stopping times, of the discounted cumulative dividends until the stopping time plus the discounted (ex-dividend) price that the owner can obtain by selling the asset at the stopping time. At each time, the agent in the group with highest reservation price would buy the asset. The equilibrium price process satisfies a simple recursive relationship. Furthermore, if there is a positive probability that at some future time a group that is currently not holding the asset would have a higher reservation price than the current owner, then the current price has to strictly exceed the maximum of the discounted expected future dividends among all groups - that is the current price exceeds the maximum fundamental valuation of any agent in the economy. This difference between the current price and the maximal valuation can be identified as a bubble. Section 3 contains a summary of the Harrison-Kreps theory.

Harrison and Kreps do not discuss the source of heterogeneous beliefs. Although private information seems to be a natural source of disagreement and of trading, a series of results, known as "no-trade theorems" appear in the economics literature showing that if all agents are rational and share identical prior beliefs, heterogeneity of information cannot generate trading or cause a price bubble. These results are discussed in Section 4. Several possibilities exist to avoid these no-trade results, including the presence of noise-traders, who trade for liquidity reasons, or heterogeneous priors. Another option is to assume that agents have behavioral biases that preclude "full rationality." Several behavioral biases are suggested by the psychology literature. See Hirshleier (2001) and Barberis and Thaler (2003) for detailed reviews of these biases.

Overconfidence, the tendency to overestimate the precision of own's opinion is a well documented behavioral bias. Scheinkman and Xiong (2003) use overconfidence as a convenient way to generate a parameterized model of heterogeneous beliefs. They adopt a continuous time framework describing a market for a single risky-asset with a limited supply, and many risk-neutral agents who can borrow and lend at a fixed rate of interest  $r$ . The current dividend of the asset is a noisy observation of a fundamental variable that will determine future dividends. More precisely:

$$dD_t = f_t dt + \sigma_D dZ_t^D,$$

where  $Z^D$  is a standard Brownian motion and  $f$  is not observable. However, it satisfies:

$$df_t = -\lambda(f_t - \bar{f})dt + \sigma_f dZ_t^f,$$

In addition to the dividends, there are two other sets of information available at each instant. These signals again satisfy the linear SDEs:

$$\begin{aligned} ds_t^A &= f_t dt + \sigma_s dZ_t^A \\ ds_t^B &= f_t dt + \sigma_s dZ_t^B. \end{aligned}$$

The information is available to all agents. However, agents are divided in two groups  $A$  and  $B$ , and group  $A$  ( $B$ ) has more confidence in signal  $s^A$  (*resp.*  $s^B$ .) As a consequence, when forecasting future dividends, each group of agents place different weights in the three sets of information, resulting in different forecasts. In the parametric structure that Scheinkman and Xiong (2003) consider, linear filtering applies and the conditional beliefs are normally

distributed with a common variance and means  $f_t^A$  and  $f_t^B$ . Although agents in the model know exactly the amount by which their forecast of the fundamental variables exceed that of agents in the other group, behavioral limitations lead them to continue to disagree. As information flows, the mean forecasts by agents of the two groups fluctuate, and the group of agents that is at one instant relatively more optimistic, may become in a future date less optimistic than the agents in the other group. These changes in relative opinion generate trades.

Each agent in the model understands that the agents in the other group are placing different weights on the different sources of information. When deciding the value of the asset, agents consider their own view of the fundamentals as well as the fact that the owner of the asset has an option to sell the asset in the future to the agents in the other group. This option can be exercised at any time by the current owner, and the new owner gets in turn another option to sell the asset in the future. In the parametric example discussed by Scheinkman and Xiong (2003) it is natural to look for an equilibrium where the value of this option for the current owner, is a function of differences in opinions, that is if he belongs to group  $A$  ( $B$ ), this value equals  $q = q(f_t^B - f_t^A)$  (*resp.*  $q = q(f_t^B - f_t^A)$ .) This option is “American,” and hence the value of the option is the value function of an optimal stopping problem. Since the buyer’s willingness to pay is a function of the value of the option that he acquires, the payoff from stopping is, in turn, related to the value of the option. This gives rise to a fixed point problem that the option value must satisfy. Scheinkman and Xiong (2003) show that the function  $q$  must, in the absence of trading costs, satisfy:

$$q(x) = \sup_{\tau \geq 0} \mathbb{E}^o \left[ \left( \frac{x_\tau}{r + \lambda} + q(x_\tau) \right) e^{-r\tau} \right],$$

where  $\mathbb{E}^o$  represents the expected value using the beliefs of the current owner. Scheinkman and Xiong (2003) write down an “explicit” solution to this equation that involves Kummer functions.

In equilibrium an asset owner will sell the asset to agents in the other group, whenever his view of the fundamental is surpassed by the view of agents in the other group by a critical amount. When there are no trading costs, this critical amount is zero - it is optimal to sell the asset immediately after the valuation of the fundamentals of the asset owner is “crossed” by the valuation of agents in the other group. The agents’ beliefs satisfy simple stochastic differential

equations and it is a consequence of properties of Brownian motion, that once the beliefs of agents cross, they will cross infinitely many times in any finite period of time right afterwards. This results in a trading frenzy. Although agents' profit from exercising the resale option is infinitesimal, the net value of the option is large because of the high frequency of trades. When trading costs are positive the duration between trades increases but in a continuous manner. In this way the model predicts large trading volume in markets with small transaction costs.

When a trade occurs the buyer has the highest fundamental valuation among all agents, and because of the re-sale option the price he pays exceeds his fundamental valuation. Agents pay prices that exceed their own valuation of future dividends, because they believe that in the future they will find a buyer willing to pay even more. This difference between the transaction price and the highest fundamental valuation can be reasonably called a bubble. Sections 5 and 6 contain an exposition of the model in Scheinkman and Xiong (2004).

The bubble in the Scheinkman and Xiong model, based on the expectations of traders to take advantage of future differences in opinions is quite different from the more traditional "rational bubbles" that are discussed, for instance, in Blanchard and Watson (1982) or Santos and Woodford (1997). In the typical rational bubble model agents have identical rational expectations, but prices include an extra bubble component that is always expected to grow at a rate equal to the risk free rate. Models of rational bubbles are incapable of explaining the increase in trading volume that is typically observed in the historic bubble episodes, and have non-stationarity properties that are at odds with empirical observations. Nonetheless, because these models constituted the first attempt to develop equilibrium models of price bubbles we briefly discuss them in Section 7.1.

While rational bubble models study asset prices while not considering questions on trading volume, a complementary literature uses heterogeneous beliefs to study trading, without dealing with the impact of heterogeneous beliefs on asset prices. Section 7.2 discusses these models using a paper by Harris and Raviv (1993) as an example.

We also discuss two other related questions that have played an important role in the economics literature on asset pricing. The first one considers what economists have dubbed the equity premium puzzle. This puzzle is the observation that stock returns over the last 50 years have been too high to be justified by standard asset pricing models with reasonable risk

aversion parameters for investors. Section 7.3 briefly describes this puzzle and reviews several models that use heterogeneous beliefs to explain it.

The second question that dates back at least to Friedman (1953), is whether traders with irrational beliefs will lose money trading with rational traders and eventually disappear from the market. In Section 8, we discuss a model by Kogan, Ross, Wang and Westerfield (2004) who analyze this issue using a continuous-time equilibrium setup. In their model, in addition to a risk-free bond, there is a stock that pays a dividend  $D_T$  at the consumption date  $T$ . This terminal dividend satisfies:

$$dD_t = D_t(\mu dt + \sigma dZ_t).$$

There are two groups of traders of equal size. One group the “rational” traders, knows the true probability  $P$  determining the distribution of the final dividend. The second group believes incorrectly on a probability  $Q$  that is absolutely continuous with respect to  $P$ . As a result, the two groups disagree on the drift of the process determining  $D_T$ . Irrational traders believe that

$$dD_t = D_t[(\mu + \sigma^2\eta)dt + \sigma dZ_t^Q],$$

where  $Z_t^Q$  is a Brownian motion under  $Q$ . Irrational traders are optimistic (pessimistic) if  $\eta > 0$  (*resp.*  $\eta < 0$ .) All agents maximize a constant relative risk aversion utility function that depends on final consumption, that is they choose a trading strategy to maximize:

$$\mathbb{E}_0^\Pi \left[ \frac{1}{1-\gamma} C_T^{1-\gamma} \right]$$

where  $C_T$  is the consumption at  $T$ , and the measure  $\Pi = P$  for rational agents and equals  $Q$  for irrational agents. Kogan *et al.* use the fact that, since there is only one source of uncertainty and two traded assets, it is natural to expect that the markets are complete and as a consequence that the equilibrium is Pareto efficient. They find the set of all Pareto optimal allocations of final consumption across types, by maximizing the weighted sum of utilities and derive the corresponding support prices. Using these support prices and the budget constraint Kogan *et al.* characterize the particular Pareto efficient allocation that corresponds to a competitive equilibrium, and show that the corresponding support prices are such that markets are complete. They then combine the expression for the equilibrium allocation with the strong law of large numbers to show that, if traders are sufficiently risk averse and irrational traders are moderately

optimistic, irrational traders may survive in the long run. The reason is that they may earn higher returns albeit by bearing excessive risk.

The papers we discuss in this chapter were chosen mainly on the basis of our views concerning their contribution to the understanding of speculation and trading. Some such as Miller (1977) involve very little mathematics. Nonetheless we believe that the area is ripe for a more rigorous mathematical treatment. In fact, in section 9 we discuss some open problems related to the speculative behavior of investors and the resale option component in asset prices. Typically these problems involve optimal stopping problems with nonlinear filtering and multiple state variables.

## 2 A static model with heterogeneous beliefs and short-sales constraints

Miller (1977) argued that short-sales constraints can cause stocks to be overpriced when investors have heterogeneous beliefs about stock fundamentals. In the presence of short-sale constraints, stock prices reflect the views of the more optimistic participants. If some of the pessimistic investors that would like to short are not allowed to do it, prices will in general be higher than the price that would prevail in the absence of short-sale constraints.

We illustrate Miller's argument using a version of Lintner (1969) model, where we add short-sales constraints. Related models can also be found in Jarrow (1980), Varian (1989), Chen, Hong and Stein (2002), and Gallmeyer and Hollifield (2004). The model has one period with two dates:  $t = 0, 1$ . There is one risky asset which will be liquidated at  $t = 1$ . The final liquidation value is

$$\tilde{f} = \mu + \epsilon,$$

where  $\mu$  is a constant unknown to investors and  $\epsilon$  is normally distributed with mean zero. Investors have diverse opinions concerning the distribution of liquidation values. Investor  $i$  believes that  $\tilde{f}$  has a normal distribution with mean  $\mu_i$  and variance  $\sigma^2$ . Since all investors share the same views concerning the variance, we index investors by their mean beliefs  $\mu_i$ , and assume that  $\mu_i$  is uniformly distributed around  $\mu$  in an interval  $[\mu - \kappa, \mu + \kappa]$ . The parameter  $\kappa$  measures the heterogeneity of beliefs. In addition, we assume that all investors can borrow

or lend at a risk-free interest rate of zero, short-sales of the risky asset are prohibited, and the total supply of the asset is  $Q$ .

At  $t = 0$ , each investor chooses his asset demand to maximize his expected utility at  $t = 1$ :

$$\max_{x_i} \mathbb{E} \left[ -e^{-\gamma(W_0 + x_i(\tilde{f} - p_0))} \right],$$

where  $\gamma$  is the investor's risk aversion,  $W_0$  is the initial wealth,  $p_0$  is the market price of the asset, and  $x_i$  is the investor's asset demand, subject to  $x_i \geq 0$ . It is immediate that

$$x_i = \max \left\{ \frac{\mu_i - p_0}{\gamma\sigma^2}, 0 \right\}.$$

The investor's demand, in the absence of the short-sale constraint would be  $(\mu_i - p_0)/(\gamma\sigma^2)$ . When these constraints are present, investors with mean beliefs  $\mu_i$  below the market price stay out of the market. The market clearing condition,  $\int_i x_i = Q$  thus implies that

$$\int_{\max\{p_0, \mu - \kappa\}}^{\mu + \kappa} \frac{\mu_i - p_0}{\gamma\sigma^2} \frac{d\mu_i}{2\kappa} = Q,$$

and the equilibrium price

$$p_0 = \begin{cases} \mu - \gamma\sigma^2 Q & \text{if } \kappa < \gamma\sigma^2 Q \\ \mu + \kappa - 2\sqrt{\kappa\gamma\sigma^2 Q} & \text{if } \kappa \geq \gamma\sigma^2 Q \end{cases}$$

In the absence of short-sales constraints the equilibrium price would be  $\mu - \gamma\sigma^2 Q$ . Thus, the short-sales constraints cause the asset price to become higher when the heterogeneity of investors' beliefs  $\kappa$  is greater than  $\gamma\sigma^2 Q$ . If heterogeneity of beliefs is small enough than the no-short-sale constraint is not binding for any investor, and the equilibrium price is not affected by the presence of the constraint.

This simple model shows that short-sales constraints combined with heterogeneous beliefs can cause asset prices to become higher than they would be in the absence of the short-sales constraints. When beliefs are sufficiently heterogeneous, short-sale constraints insure that asset prices reflect the opinion of the more optimistic investors. However, because of its static nature, the model has no prediction concerning the dynamics of trading. In the following sections, we discuss the effects of heterogeneous beliefs and short-sales constraints on asset prices and share turnovers in dynamic models.

### 3 A dynamic model in discrete time with short-sales constraints

Harrison and Kreps (1978) say that investors exhibit *speculative behavior* if the right to resell an asset makes them willing to pay more for it than they would pay if obliged to hold it forever. This definition is particularly compelling when agents are risk-neutral since in this case no risk-sharing benefits arise from trading. Harrison and Kreps constructed a model where, because risk-neutral agents have heterogeneous expectations and face short-sales constraints, speculative behavior arises.

In the model, there exists one unit of an asset that pays a non-negative random dividend  $d_t$  at each time  $t$ . All agents are risk-neutral, and discount future revenues at a constant rate  $\gamma < 1$  or equivalently can borrow and lend at a rate  $r = \frac{1-\gamma}{\gamma}$ . Agents are divided into groups that differ on their views on the distribution of the stochastic process  $d_t$ . Harrison and Kreps allow for an arbitrary number of groups of agents, but the analysis they provide is well illustrated by treating the case of two groups. Thus, we consider two groups  $\{A, B\}$ , each with an infinite number of agents. For simplicity we assume that each group  $C \in \{A, B\}$  views  $d_t$  as a stochastic process defined on a probability space  $\{\Omega, \mathcal{F}, P^C, \}$  and that  $P^A \sim P^B$ . We write  $\mathbb{E}^C$  for the expected value with respect to the probability distribution shared by all agents in group  $C \in \{A, B\}$ .

Write  $\mathcal{F}_t$ ,  $t \geq 0$  for the  $\sigma$ -algebra generated by the realizations of  $d^t \equiv (d_1, \dots, d_t)$ . A price process is an  $\mathcal{F}_t$  adapted non-negative process.

The owner of an asset at  $t$  must decide on a strategy to sell all or part of his holdings in the future. Since agents are risk neutral it suffices to consider the strategy of selling one unit of the asset at a (possibly infinite) stopping time. For this reason we define a feasible selling strategy from time  $t$  as a (possibly infinite) integer valued stopping time  $T > t$ .

Because each group has an infinite number of agents and there is a single unit of the asset, competition among buyers will lead to a price that equals the reservation price of the buyers. Since agents are risk-neutral and can choose any feasible selling strategy, the value of the asset at  $t$  for an agent in group  $C \in \{A, B\}$  is given by

$$\sup_{T>t} \mathbb{E}^C \left[ \sum_{k=t+1}^T \gamma^{k-t} d_k + \gamma^{T-t} p_T | \mathcal{F}_t \right],$$

where  $\sum_{k=t+1}^T \gamma^{k-t} d_k$  represents the value of the discounted dividend stream received up to the moment of sale at  $T$ , and  $\gamma^{T-t} p_T$  represents the discounted value from selling the asset at the prevailing market price at  $T$ . The buyers will belong to the group that places the highest valuation on the asset. Hence an equilibrium price process has to satisfy:

$$p_t = \max_{C \in \{A, B\}} \sup_{T > t} \mathbb{E}^C \left[ \sum_{k=t+1}^T \gamma^{k-t} d_k + \gamma^{T-t} p_T | \mathcal{F}_t \right]. \quad (1)$$

Since  $T = \infty$  is a feasible strategy, it follows that

$$p_t \geq \max_{C \in \{A, B\}} \mathbb{E}^C \left[ \sum_{k=t+1}^{\infty} \gamma^{k-t} d_k | \mathcal{F}_t \right]. \quad (2)$$

Since the right hand side of equation (2) represents the maximal value that any agent is willing to pay for the asset if resale is impossible, speculative behavior is equivalent to a strict inequality in equation (2). Suppose  $F \in \mathcal{F}_t$  is such that A realizes the maximum in the right-hand side of (1) for each  $\omega \in F$ . Suppose further that for some  $t' > t$  and event  $F' \in \mathcal{F}_{t'}$ ,  $F' \subset F$  with  $P^A(F') > 0$  (and hence  $P^B(F') > 0$ )

$$\mathbb{E}^B \left[ \sum_{k=t'+1}^{\infty} \gamma^{k-t'} d_k | \mathcal{F}_{t'} \right] (\omega) > \mathbb{E}^A \left[ \sum_{k=t'+1}^{\infty} \gamma^{k-t'} d_k | \mathcal{F}_{t'} \right] (\omega),$$

for each  $\omega \in F'$ . Then a strict inequality must hold in equation (2), for  $\omega \in F$ .

In fact,

$$\begin{aligned} & \mathbb{E}^A \left[ \sum_{k=t+1}^{\infty} \gamma^{k-t} d_k | \mathcal{F}_t \right] \\ = & \mathbb{E}^A \left[ \sum_{k=t+1}^{t'} \gamma^{k-t} d_k | \mathcal{F}_t \right] + \mathbb{E}^A \left[ \mathbb{E}^A \left[ \sum_{k=t'+1}^{\infty} \gamma^{k-t} d_k | \mathcal{F}_{t'} \right] | \mathcal{F}_t \right] \\ < & \mathbb{E}^A \left[ \sum_{k=t+1}^{t'} \gamma^{k-t} d_k | \mathcal{F}_t \right] + \mathbb{E}^A \left[ \max_{C \in \{A, B\}} \mathbb{E}^C \left[ \sum_{k=t'+1}^{\infty} \gamma^{k-t} d_k | \mathcal{F}_{t'} \right] | \mathcal{F}_t \right] \\ \leq & \mathbb{E}^A \left[ \sum_{k=t+1}^{t'} \gamma^{k-t} d_k | \mathcal{F}_t \right] + \mathbb{E}^A \left[ \gamma^{t'-t} p_{t'} | \mathcal{F}_t \right] \leq p_t. \end{aligned}$$

Speculative behavior arises, because the owner of the asset retains in addition to the flow of future dividends an option to resell the asset to other investors. This option will become

*in-the-money* when there are investors that have a relatively more optimistic view of future dividends than the current owner.

The following proposition allows one to characterize all pricing processes that satisfy equation (1) by a two period condition.

**Proposition 1** *A price process satisfies equation (1) if and only if, for each  $t$ ,*

$$p_t = \max_{C \in \{A, B\}} \mathbb{E}^C [\gamma d_{t+1} + \gamma p_{t+1} | \mathcal{F}_t]. \quad (3)$$

Proof: Suppose (3) holds. Then for each  $C$ ,

$$p_t \geq \mathbb{E}^C [\gamma d_{t+1} + \gamma p_{t+1} | \mathcal{F}_t].$$

Hence the process  $y_t = \sum_{s=1}^t \gamma^s d_s + \gamma^t p_t$  is a non-negative supermartingale and hence  $\lim_{t \rightarrow \infty} y_t$  exists. Doob's optional stopping theorem implies that  $p_t \geq \mathbb{E}^C \left[ \sum_{k=t+1}^T \gamma^{k-t} d_k + \gamma^{T-t} p_T | \mathcal{F}_t \right]$ , what implies that (1) must hold.

Conversely, suppose (1) holds, but that:

$$p_t > \max_{C \in \{A, B\}} \mathbb{E}^C [\gamma d_{t+1} + \gamma p_{t+1} | \mathcal{F}_t].$$

The law of iterated expectations and equation (1) applied at  $t + 1$  implies that:

$$p_t > \max_{C \in \{A, B\}} \sup_T \mathbb{E}^C \left[ \sum_{k=t+1}^T \gamma^{k-t} d_k + \gamma^{T-t} p_t | \mathcal{F}_t \right],$$

a contradiction.  $\square$

Suppose  $d_t$  is a time-homogeneous Markov process, that is

$$P^C[d_{t+s} | \mathcal{F}_t] = P^C[d_{t+s} | d_t] = P^C[d_{s+1} | d_1],$$

for each  $C \in \{A, B\}$ . Then it is natural to search for equilibrium prices that are of the form  $p_t = p(d_t)$ . If we write:

$$\mathcal{T}p(d) = \max_{C \in \{A, B\}} \mathbb{E}^C [\gamma d_{t+1} + \gamma p(d_{t+1}) | d_t = d]. \quad (4)$$

Then equation (3) can be rewritten as:

$$\mathcal{T}p = p. \quad (5)$$

The operator  $\mathcal{T}$  has a monotonicity property. If  $p \geq q$  then  $\mathcal{T}(p) \geq \mathcal{T}(q)$ . The existence and uniqueness of (continuous) solutions to equation (5) are guaranteed if, for instance, the process  $d$  stays in a bounded set.

Harrison and Kreps do not explicitly address the source of heterogeneous beliefs among investors. In what follows we will examine specific mechanisms to generate beliefs' heterogeneity. In Section 4 we describe some results concerning the difficulty of generating heterogeneous beliefs from the presence of private information. We then discuss a model that utilizes *overconfidence*, a behavioral limitation that is suggested by psychological studies, to parameterize heterogeneous beliefs. The model will be then used to link Harrison and Kreps' speculative behavior to a resale option value and to explain some empirical regularities concerning trading volume and prices during asset "bubbles."

## 4 No-trade theorem under rational expectations

A possible source of heterogeneous beliefs is private information. The presence of private information suggests that investors could use their information to trade and realize a profit. However, Tirole (1982) and Milgrom and Stokey (1982) prove that this cannot happen when all agents are rational and share identical prior beliefs, the conditions that are imposed in the standard rational expectations models. Thus, private information cannot be the source of speculative trading. Results of this kind are called "no-trade theorems".

We use a static setup from Tirole (1982) to illustrate the main ideas. Consider a market with  $I$  risk neutral traders:  $i = 1, \dots, I$ . The traders exchange claims for an asset with random payoff  $\tilde{p} \in E$ , which will be realized after the trading. The set  $E \subset \mathbb{R}$  is the set of all possible payoffs. Claims are traded at an equilibrium market price  $p$ . If a trader buys  $x$  units of the asset and  $\tilde{p}$  obtains, the realized profit is

$$G = (\tilde{p} - p)x.$$

Each trader  $i$  receives a private signal  $s^i$  belonging to a possible set of signals  $S^i$ . Let  $s = (s^1, \dots, s^I) \in S = \prod_{i=1}^I S^i$ , and  $\Omega = E \times S$ . Assume that all traders have a common prior  $\nu$  on  $\Omega$ , and that each trader can only take a bounded position.

Trader  $i = 1, \dots, I$  chooses an amount  $x^i$  to maximize his conditional expected value of  $G$ . In a *Rational Expectations Equilibrium*, (REE) each trader uses all information at his disposal, including the observed market price  $p$ . In spite of its name, a REE does not involve only an equilibrium price, but a *forecast function* that maps each vector of all signals  $s \in S$  into a price that establishes equilibrium in the market if the state  $s$  obtains. This forecast function is typically not one-to-one. By observing the price  $p$  as well as his own signal  $s^i$ , trader  $i$  is not able to identify the full vector of signals  $s$ . However, a forecast function  $\Phi : S \rightarrow \mathbb{R}$ , an observed  $p$ , and a signal  $s^i$  induce a conditional distribution on  $E \times S$ ,  $\Gamma_{\Phi, p, s^i}^i$ .

**Definition 1:** A rational expectations equilibrium (REE) is a forecast function  $\Phi : S \rightarrow R$ , and a set of trades  $x^i(\Phi, p, s^i)$  for each trader  $i$  such that

1.  $x^i(\Phi, p, s^i)$  maximizes  $\mathbb{E}_{\Phi}(G^i|p, s^i) \equiv \int G d\Gamma_{\Phi, p, s^i}^i$
2. The market clears for each  $s \in S : \sum_i x^i(\Phi, p, s^i) = 0$ .

The next proposition is a no-trade theorem.

**Proposition 2** *In a REE,  $\mathbb{E}_{\Phi}(G^i|p, s^i) = 0$ . As a consequence given an REE, there exists another REE with the same forecast function and  $x^i \equiv 0$ .*

Proof: Since  $x^i = 0$  is always a possible choice,

$$\mathbb{E}_{\Phi}(G^i|p, s^i) \geq 0. \tag{6}$$

The Law of Iterated expectations thus implies that

$$\mathbb{E}_{\Phi}(G^i|p) \geq 0. \tag{7}$$

The market clearing condition implies that for each realization of  $\tilde{p}$  aggregate gains are null. Hence

$$\sum_i \mathbb{E}_{\Phi}(G^i|p) = 0,$$

and equality must hold in equations (6) and (7).  $\square$

Proposition 2 rules out the possibility that investors that share the same prior can expect to profit from speculating against each other based on differences in information. As a consequence,

they cannot do any better than by choosing not to trade. Although Proposition 2 only deals with the risk neutral case, it is intuitive that risk aversion would further reduce the net gain among investors from trading.

Tirole (1982) also analyzes a dynamic model with rational expectations and demonstrates that the no-trade theorem holds in dynamic setup. He further shows that the resale options suggested by Harrison and Kreps cannot arise in asset prices in such an environment even if short-sales constraints are imposed. Diamond and Verrecchia (1987) also study the effects of short-sales constraints on asset prices in a rational expectations model with asymmetric information. They show that short-sales constraints reduce the adjustment speed of prices to private information, especially to bad news, since agents with negative information are prohibited from shorting the asset. However Diamond and Verrecchia also confirm that short-sales constraints do not lead to an upward bias in prices since agents, when forming their own beliefs, could rationally take into account the fact that negative information may be not reflected in trading prices.

There are at least two ways to weaken the assumptions in Proposition 2 and avoid the no-trade result. First one might consider some agents who trade for non-speculative reasons such as diversification or liquidity. The presence of such traders would make the trading among speculators a positive-sum game. This is the approach that has been adopted in several models of market microstructure such as Grossman and Stiglitz (1980), Kyle (1985), and Wang (1993).

Another possibility is to relax the assumption that agents share the same prior beliefs. This approach is pursued by Morris (1996), Biais and Bossaerts (1998), and Brav and Heaton (2002). Finally one may assume that agents display behavioral biases. In this chapter, we discuss in detail overconfidence as a way to parameterize the dynamics of heterogeneous beliefs among agents. However many other behavioral biases may generate heterogeneity of beliefs. For instance, heterogeneous beliefs can arise if agents gain utility from adopting certain beliefs as discussed in Brunnermeier and Parker (2003).

## **5 Overconfidence as source of heterogeneous beliefs**

Overconfidence, the tendency of people to overestimate the precision of their knowledge, provides a convenient way to generate heterogeneous beliefs. Psychology studies suggest that

people are overconfident. Alpert and Raiffa (1982), and Brenner et al. (1996) and other calibration studies find that people overestimate the precision of their knowledge. Camerer (1995) argues that even experts can display overconfidence. Hirshleifer (2001) and Barberis and Thaler (2003) contain extensive reviews of the literature.

In finance, researchers have developed theoretical models to analyze the implications of overconfidence on financial markets. Kyle and Wang (1997) show that overconfidence can be used as a commitment device over competitors to improve one's welfare. Daniel, Hirshleifer and Subrahmanyam (1998) use overconfidence to explain the predictable returns of financial assets. Odean (1998) demonstrates that overconfidence can cause excessive trading. Bernardo and Welch (2001) discuss the benefits of overconfidence to entrepreneurs through the reduced tendency to herd. In all these studies, overconfidence is modelled as overestimation of the precision of one's information.

In this section we exposit the model in Scheinkman and Xiong (2003), that exploits the consequences of this overestimation in a dynamic model of pricing and trading. Since overconfident investors believe more strongly in their own assessments of an asset's value than in the assessment of others, heterogeneous beliefs arise.

Consider a single risky asset with a dividend process that is the sum of two components. The first component is a fundamental variable that determines future dividends. The second is "noise". The cumulative dividend process  $D$  satisfies:

$$dD_t = f_t dt + \sigma_D dZ_t^D, \quad (8)$$

where  $Z^D$  is a standard Brownian motion and  $\sigma_D$  is the volatility parameter. The stochastic process of fundamentals  $f$  is not observable. However, it satisfies:

$$df_t = -\lambda(f_t - \bar{f})dt + \sigma_f dZ_t^f, \quad (9)$$

where  $\lambda \geq 0$  is the mean reversion parameter,  $\bar{f}$  is the long-run mean of  $f$ ,  $\sigma_f > 0$  is a volatility parameter and  $Z^f$  is a standard Brownian motion, uncorrelated to  $Z^D$ . The presence of dividend noise makes it impossible to infer  $f$  perfectly from observations of the cumulative dividend process.

There are two sets of risk-neutral agents, who use the observations of  $D$  and any other signals that are correlated with  $f$  to infer current  $f$  and to value the asset. In addition to the

cumulative dividend process, all agents observe a vector of signals  $s^A$  and  $s^B$  that satisfy:

$$ds_t^A = f_t dt + \sigma_s dZ_t^A \quad (10)$$

$$ds_t^B = f_t dt + \sigma_s dZ_t^B, \quad (11)$$

where  $Z^A$  and  $Z^B$  are standard Brownian motions, and  $\sigma_s > 0$  is the common volatility of the signals. We assume that all four processes  $Z^D$ ,  $Z^f$ ,  $Z^A$  and  $Z^B$  are mutually independent.

Agents in group  $A$  ( $B$ ) think of  $s^A$  ( $s^B$ ) as their own signal although they can also observe  $s^B$  ( $s^A$ ). Heterogeneous beliefs arise because each agent believes that the informativeness of his own signal is larger than its true informativeness. Agents of group  $A$  ( $B$ ) believe that innovations  $dZ^A$  ( $dZ^B$ ) in the signal  $s^A$  ( $s^B$ ) are correlated with the innovations  $dZ^f$  in the fundamental process, with  $\phi$  ( $0 < \phi < 1$ ) as the correlation parameter. Specifically, agents in group  $A$  believe that the process  $s^A$  satisfies

$$ds_t^A = f_t dt + \sigma_s \phi dZ_t^f + \sigma_s \sqrt{1 - \phi^2} dZ_t^A.$$

Although agents in group  $A$  perceive the correct unconditional volatility of the signal  $s^A$ , the correlation that they attribute to innovations causes them to over-react to signal  $s^A$ . Similarly, agents in group  $B$  believe the process  $s^B$  satisfies

$$ds_t^B = f_t dt + \sigma_s \phi dZ_t^f + \sigma_s \sqrt{1 - \phi^2} dZ_t^B.$$

On the other hand, agents in group  $A$  ( $B$ ) believe (correctly) that innovations to  $s^B$  ( $s^A$ ) are uncorrelated with innovations to  $Z^B$  ( $Z^A$ ). We assume that the joint dynamics of the processes  $D$ ,  $f$ ,  $s^A$  and  $s^B$  in the mind of agents of each group is public information.

Since all variables are Gaussian, the filtering problem of the agents is standard. With Gaussian initial conditions, the conditional beliefs of agents in group  $C \in \{A, B\}$  is Gaussian. Standard arguments, *e.g.* section VI.9 in Rogers and Williams (1987) and Theorem 12.7 in Liptser and Shiriyayev (1977), can be used to compute the variance of the stationary solution and the evolution of the conditional mean of beliefs. The variance of this stationary solution is the same for both groups of agents and equals

$$\gamma \equiv \frac{\sqrt{(\lambda + \phi\sigma_f/\sigma_s)^2 + (1 - \phi^2)(2\sigma_f^2/\sigma_s^2 + \sigma_f^2/\sigma_D^2)} - (\lambda + \phi\sigma_f/\sigma_s)}{\frac{1}{\sigma_D^2} + \frac{2}{\sigma_s^2}}.$$

One can directly verify that the stationary variance  $\gamma$  decreases with  $\phi$ . When  $\phi > 0$ , agents have an exaggerated view of the precision of their estimates of  $f$ . A larger  $\phi$  leads to more overstatement of this precision. For this reason we refer to  $\phi$  as the “overconfidence” parameter.

The conditional mean of the beliefs of agents in group  $A$  satisfies:

$$\begin{aligned} d\hat{f}_t^A &= -\lambda(\hat{f}_t^A - \bar{f})dt + \frac{\phi\sigma_s\sigma_f + \gamma}{\sigma_s^2}(ds_t^A - \hat{f}_t^A dt) \\ &\quad + \frac{\gamma}{\sigma_s^2}(ds_t^B - \hat{f}_t^A dt) + \frac{\gamma}{\sigma_D^2}(dD_t - \hat{f}_t^A dt). \end{aligned} \quad (12)$$

Since  $f$  mean-reverts, the conditional beliefs also mean-revert. The other three terms represent the effects of “surprises.” These surprises can be represented as standard mutually independent Brownian motions for agents in group  $A$ :

$$dW_t^{A,A} = \frac{1}{\sigma_s}(ds_t^A - \hat{f}_t^A dt), \quad (13)$$

$$dW_t^{A,B} = \frac{1}{\sigma_s}(ds_t^B - \hat{f}_t^A dt), \quad (14)$$

$$dW_t^{A,D} = \frac{1}{\sigma_D}(dD_t - \hat{f}_t^A dt). \quad (15)$$

Note that these processes are only Wiener processes in the mind of group  $A$  agents. Due to overconfidence ( $\phi > 0$ ), agents in group  $A$  over-react to surprises in  $s^A$ .

Similarly, the conditional mean of the beliefs of agents in group  $B$  satisfies:

$$\begin{aligned} d\hat{f}_t^B &= -\lambda(\hat{f}_t^B - \bar{f})dt + \frac{\gamma}{\sigma_s^2}(ds_t^A - \hat{f}_t^B dt) \\ &\quad + \frac{\phi\sigma_s\sigma_f + \gamma}{\sigma_s^2}(ds_t^B - \hat{f}_t^B dt) + \frac{\gamma}{\sigma_D^2}(dD_t - \hat{f}_t^B dt), \end{aligned} \quad (16)$$

and the surprise terms can be represented as mutually independent Wiener processes:  $dW^{B,A} = \frac{1}{\sigma_s}(ds_t^A - \hat{f}_t^B dt)$ ,  $dW^{B,B} = \frac{1}{\sigma_s}(ds_t^B - \hat{f}_t^B dt)$ , and  $dW^{B,D} = \frac{1}{\sigma_D}(dD_t - \hat{f}_t^B dt)$ . These processes form a standard 3-d Brownian only for agents in group  $B$ .

Since, in the stationary solution the beliefs of all agents have constant variance, one may refer to the conditional mean of the beliefs as the agents *beliefs*. Let  $g_A$  and  $g_B$  denote the differences in beliefs:

$$g^A = \hat{f}^B - \hat{f}^A, \quad g^B = \hat{f}^A - \hat{f}^B.$$

The next proposition describes the evolution of these differences in beliefs:

**Proposition 3**

$$dg_t^A = -\rho g_t^A dt + \sigma_g dW_t^{A,g}, \quad (17)$$

where

$$\begin{aligned} \rho &= \sqrt{\left(\lambda + \phi \frac{\sigma_f}{\sigma_s}\right)^2 + (1 - \phi^2) \sigma_f^2 \left(\frac{2}{\sigma_s^2} + \frac{1}{\sigma_D^2}\right)}, \\ \sigma_g &= \sqrt{2} \phi \sigma_f, \end{aligned} \quad (18)$$

and  $W^{A,g}$  is a standard Wiener process for agents in group A.

Proof: from equations (12) and (16):

$$dg_t^A = d\hat{f}_t^B - d\hat{f}_t^A = - \left[ \lambda + \frac{2\gamma + \phi \sigma_s \sigma_f}{\sigma_s^2} + \frac{\gamma}{\sigma_D^2} \right] g_t^A dt + \frac{\phi \sigma_f}{\sigma_s} (ds_t^B - ds_t^A).$$

Using the formula for  $\gamma$ , we may write the mean-reversion parameter as in equation (18). Using equations (13) and (14),

$$dg_t^A = -\rho g_t^A dt + \frac{\phi \sigma_f}{\sigma_s} (\sigma_s dW^{A,B} - \sigma_s dW^{A,A}).$$

The result follows by writing

$$W^{A,g} = \frac{1}{\sqrt{2}} (W^{A,B} - W^{A,A}).$$

It is easy to verify that innovations to  $W^{A,g}$  are orthogonal to innovations to  $\hat{f}^A$  in the mind of agents in group A.  $\square$

Proposition 3 implies that the difference in beliefs  $g^A$  follows a simple mean reverting diffusion process in the mind of group A agents. In particular, the volatility of the difference in beliefs is zero in the absence of overconfidence. A larger  $\phi$  leads to greater volatility. In addition,  $-\rho/(2\sigma_g^2)$  measures the pull towards the origin. A simple calculation shows that this mean-reversion decreases with  $\phi$ . A higher  $\phi$  causes an increase in fluctuations of opinions and a slower mean-reversion.

In an analogous fashion, for agents in group B,  $g^B$  satisfies:

$$dg_t^B = -\rho g_t^B dt + \sigma_g dW_t^{B,g}, \quad (19)$$

where  $W^{B,g}$  is a standard Wiener process.

Notice that although we started with a Markovian structure on dividends and signals, the beliefs depend on the history of dividends and signals - only the vector involving dividends, all signals and beliefs is Markovian. This is a consequence of the inference problem faced by investors. In contrast, in this model, the difference in beliefs is a Markov diffusion, what greatly facilitates the analysis that follows.

## 6 Trading and equilibrium price in continuous time

In the previous section we specified a particular model of heterogeneous beliefs, generated by overconfidence. Equations (17) and (19) state that, in each group's mind, the difference of opinions follows a mean-reverting diffusion process. The coefficients of this process are linked to the parameters describing the original uncertainty and the degree of overconfidence. In this section we derive implications of this particular model of heterogeneity for the equilibrium prices and trading behavior. We also summarize some results from Scheinkman and Xiong (2003) concerning the effect of the parameters that determine the original uncertainty and overconfidence on the prices and trading volume that obtain in equilibrium.

As in the Harrison and Kreps model described in Section 3, assume that each group of investors is large and there is no short selling of the risky asset. To value future cash flows, assume that every agent can borrow and lend at the same rate of interest  $r$ . These assumptions facilitate the calculation of equilibrium prices.

At each  $t$ , agents in group  $C = \{A, B\}$  are willing to pay  $p_t^C$  for a unit of the asset. As in the Harrison and Kreps model described in Section 3, the presence of the short-sale constraint, a finite supply of the asset, and an infinite number of prospective buyers, guarantee that any successful bidder will pay his reservation price. The amount that an agent is willing to pay reflects the agent's fundamental valuation and the fact that he may be able to sell his holdings at a later date at the demand price of agents in the other group for a profit. If  $o \in \{A, B\}$  denotes the group of the current owner,  $\bar{o}$  the other group, and  $\mathbb{E}_t^o$  the expectation of members of group  $o$ , conditional on the information they have at  $t$ , then:

$$p_t^o = \sup_{\tau \geq 0} \mathbb{E}_t^o \left[ \int_t^{t+\tau} e^{-r(s-t)} dD_s + e^{-r\tau} (p_{t+\tau}^{\bar{o}} - c) \right], \quad (20)$$

where  $\tau$  is a stopping time,  $c$  is a transaction cost charged to the seller, and  $p_{t+\tau}^{\bar{o}}$  is the reservation value of the buyer at the time of transaction  $t + \tau$ .

Using the equations for the evolution of dividends and for the conditional mean of beliefs (equations (8), (12) and (16) above), one obtains:

$$\int_t^{t+\tau} e^{-r(s-t)} dD_s = \int_t^{t+\tau} e^{-r(s-t)} [\bar{f} + e^{-\lambda(s-t)}(\hat{f}_s^o - \bar{f})] ds + M_{t+\tau},$$

where  $\mathbb{E}_t^o M_{t+\tau} = 0$ . Hence, we may rewrite equation (20) as:

$$p_t^o = \max_{\tau \geq 0} \mathbb{E}_t^o \left\{ \int_t^{t+\tau} e^{-r(s-t)} [\bar{f} + e^{-\lambda(s-t)}(\hat{f}_s^o - \bar{f})] ds + e^{-r\tau} (p_{t+\tau}^{\bar{o}} - c) \right\}. \quad (21)$$

Scheinkman and Xiong (2003) start by postulating a particular form for the equilibrium price function, equation (22) below. Proceeding in a heuristic fashion, they derive properties that our candidate equilibrium price function should satisfy. They then construct a function that satisfies these properties, and verify that they have produced an equilibrium.

Since all the relevant stochastic processes are Markovian and time-homogeneous, and traders are risk-neutral, it is natural to look for an equilibrium in which the demand price of the current owner satisfies

$$p_t^o = p^o(\hat{f}_t^o, g_t^o) = \frac{\bar{f}}{r} + \frac{\hat{f}_t^o - \bar{f}}{r + \lambda} + q(g_t^o). \quad (22)$$

with  $q > 0$  and  $q' > 0$ . This equation states that prices are the sum of two components. The first part,  $\frac{\bar{f}}{r} + \frac{\hat{f}_t^o - \bar{f}}{r + \lambda}$ , is the expected present value of future dividends from the viewpoint of the current owner. The second is the value of the resale option,  $q(g_t^o)$ , which depends on the current difference between the beliefs of the other group's agents and the beliefs of the current owner. We call the first quantity the owner's fundamental valuation and the second the value of the resale option. Using (22) in equation (21) and collecting terms:

$$p_t^o = p^o(\hat{f}_t^o, g_t^o) = \frac{\bar{f}}{r} + \frac{\hat{f}_t^o - \bar{f}}{r + \lambda} + \sup_{\tau \geq 0} \mathbb{E}_t^o \left[ \left( \frac{g_{t+\tau}^o}{r + \lambda} + q(g_{t+\tau}^{\bar{o}}) - c \right) e^{-r\tau} \right].$$

Equivalently, the resale option value satisfies

$$q(g_t^o) = \sup_{\tau \geq 0} \mathbb{E}_t^o \left[ \left( \frac{g_{t+\tau}^o}{r + \lambda} + q(g_{t+\tau}^{\bar{o}}) - c \right) e^{-r\tau} \right]. \quad (23)$$

Hence to show that an equilibrium of the form (22) exists, it is necessary and sufficient to construct an option value function  $q$  that satisfies equation (23). This equation is similar to a Bellman equation. The current asset owner chooses an optimal stopping time to exercise his re-sale option. Upon the exercise of the option, the owner gets the “strike price”  $\frac{g_{t+\tau}^o}{r+\lambda} + q(g_{t+\tau}^o)$ , the amount of excess optimism that the buyer has about the asset’s fundamental value and the value of the resale option to the buyer, minus the cost  $c$ . In contrast to the optimal exercise problem of American options, the “strike price” in this problem depends on the re-sale option value function itself.

The region where the value of the option equals that of an immediate sale is the stopping region. The complement is the continuation region. In the mind of the risk neutral asset holder, the discounted value of the option  $e^{-rt}q(g_t^o)$  should be a martingale in the continuation region, and a supermartingale in the stopping region. Using Ito’s lemma and the evolution equation for  $g^o$ , these conditions can be stated as:

$$q(x) \geq \frac{x}{r+\lambda} + q(-x) - c \quad (24)$$

$$\frac{1}{2}\sigma_g^2 q'' - \rho x q' - r q \leq 0, \text{ with equality if (24) holds strictly.} \quad (25)$$

In addition, the function  $q$  should be continuously differentiable (smooth pasting). As usual, one first shows that there exists a smooth function  $q$  that satisfies equations (24) and (25) and then uses these properties and a growth condition on  $q$  to show that in fact the function  $q$  solves (23).

To construct the function  $q$ , guess that the continuation region will be an interval  $(-\infty, k^*)$ , with  $k^* \geq 0$ .  $k^*$  is the minimum amount of difference in opinions that generates a trade. The second order ordinary differential equation that  $q$  must satisfy, albeit only in the continuation region, is:

$$\frac{1}{2}\sigma_g^2 u'' - \rho x u' - r u = 0 \quad (26)$$

The following proposition describes all solutions of equation (26).

**Proposition 4** *Let*

$$h(x) = \begin{cases} U\left(\frac{r}{2\rho}, \frac{1}{2}, \frac{\rho}{\sigma_g^2}x^2\right) & \text{if } x \leq 0 \\ \frac{2\pi}{\Gamma\left(\frac{1}{2} + \frac{r}{2\rho}\right)\Gamma\left(\frac{1}{2}\right)} M\left(\frac{r}{2\rho}, \frac{1}{2}, \frac{\rho}{\sigma_g^2}x^2\right) - U\left(\frac{r}{2\rho}, \frac{1}{2}, \frac{\rho}{\sigma_g^2}x^2\right) & \text{if } x > 0 \end{cases} \quad (27)$$

where  $\Gamma(\cdot)$  is the Gamma function, and  $M : \mathbb{R}^3 \rightarrow \mathbb{R}$  and  $U : \mathbb{R}^3 \rightarrow \mathbb{R}$  are two Kummer functions described in chapter 13 of Abramowitz and Stegun (1964).  $h(x)$  is positive and increasing in  $(-\infty, 0)$ . In addition  $h$  solves equation (26) with

$$h(0) = \frac{\pi}{\Gamma\left(\frac{1}{2} + \frac{r}{2\rho}\right)\Gamma\left(\frac{1}{2}\right)}.$$

Any solution  $u(x)$  to equation (26) that is strictly positive and increasing in  $(-\infty, 0)$  must satisfy:  $u(x) = \beta_1 h(x)$  for some  $\beta_1 > 0$ .

Proof: Let  $v(y)$  be a solution to the differential equation

$$yv''(y) + (1/2 - y)v'(y) - \frac{r}{2\rho}v(y) = 0. \quad (28)$$

It is straightforward to verify that  $u(x) = v\left(\frac{\rho}{\sigma_g^2}x^2\right)$  satisfies equation (26). The general solution of equation (28) is

$$v(y) = \alpha M\left(\frac{r}{2\rho}, \frac{1}{2}, y\right) + \beta U\left(\frac{r}{2\rho}, \frac{1}{2}, y\right).$$

Given a solution  $u$  to equation (26) one can construct two solutions  $v$  to equation (28), by using the values of the function for  $x < 0$  and for  $x > 0$ . Write the corresponding linear combinations of  $M$  and  $U$  as  $\alpha_1 M + \beta_1 U$  and  $\alpha_2 M + \beta_2 U$ . If these combinations are constructed from the same  $u$  their values and first derivatives must have the same limit as  $x \rightarrow 0$ . To guarantee that  $u(x)$  is positive and increasing for  $x < 0$ ,  $\alpha_1$  must be zero. Therefore,

$$u(x) = \beta_1 U\left(\frac{r}{2\rho}, \frac{1}{2}, \frac{\rho}{\sigma_g^2}x^2\right) \quad \text{if } x \leq 0.$$

From the definition of the two Kummer functions, one can show that

$$\begin{aligned} x \rightarrow 0-, \quad u(x) &\rightarrow \frac{\beta_1 \pi}{\Gamma\left(\frac{1}{2} + \frac{r}{2\rho}\right)\Gamma\left(\frac{1}{2}\right)}, & u'(x) &\rightarrow \frac{\beta_1 \pi \sqrt{\rho}}{\sigma_g \Gamma\left(\frac{r}{2\rho}\right)\Gamma\left(\frac{3}{2}\right)} \\ x \rightarrow 0+, \quad u(x) &\rightarrow \alpha_2 + \frac{\beta_2 \pi}{\Gamma\left(\frac{1}{2} + \frac{r}{2\rho}\right)\Gamma\left(\frac{1}{2}\right)}, & u'(x) &\rightarrow -\frac{\beta_2 \pi \sqrt{\rho}}{\sigma_g \Gamma\left(\frac{r}{2\rho}\right)\Gamma\left(\frac{3}{2}\right)} \end{aligned}$$

By matching the values and first order derivatives of  $u(x)$  from the two sides of  $x = 0$ , we have

$$\beta_2 = -\beta_1, \quad \alpha_2 = \frac{2\beta_1\pi}{\Gamma\left(\frac{1}{2} + \frac{r}{2\rho}\right)\Gamma\left(\frac{1}{2}\right)}.$$

The function  $h$  is a solution to equation (26) that satisfies

$$h(0) = \frac{\pi}{\Gamma\left(\frac{1}{2} + \frac{r}{2\rho}\right)\Gamma\left(\frac{1}{2}\right)} > 0,$$

and  $h(-\infty) = 0$ . Equation (26) guarantees that at any critical point where  $h < 0$ ,  $h$  has a maximum, and at any critical point where  $h > 0$  it has a minimum. Hence  $h$  is strictly positive and increasing in  $(-\infty, 0)$ .  $\square$

Further properties of the function  $h$  are summarized in the following Lemma.

**Lemma 1** *For each  $x \in R$ ,  $h(x) > 0$ ,  $h'(x) > 0$ ,  $h''(x) > 0$ ,  $h'''(x) > 0$ ,  $\lim_{x \rightarrow -\infty} h(x) = 0$ , and  $\lim_{x \rightarrow -\infty} h'(x) = 0$ .*

Since  $q$  must be positive and increasing in  $(-\infty, k^*)$ , it follows from Proposition 4 that

$$q(x) = \begin{cases} \beta_1 h(x), & \text{for } x < k^* \\ \frac{x}{r+\lambda} + \beta_1 h(-x) - c, & \text{for } x \geq k^*. \end{cases} \quad (29)$$

Since  $q$  is continuous and continuously differentiable at  $k^*$ ,

$$\begin{aligned} \beta_1 h(k^*) - \frac{k^*}{r+\lambda} - \beta_1 h(-k^*) + c &= 0, \\ \beta_1 h'(k^*) + \beta_1 h'(-k^*) - \frac{1}{r+\lambda} &= 0. \end{aligned}$$

These equations imply that

$$\beta_1 = \frac{1}{(h'(k^*) + h'(-k^*))(r+\lambda)}, \quad (30)$$

and  $k^*$  satisfies

$$[k^* - c(r+\lambda)](h'(k^*) + h'(-k^*)) - h(k^*) + h(-k^*) = 0. \quad (31)$$

The next proposition shows that for each  $c$ , there exists a unique pair  $(k^*, \beta_1)$  that solves equations (30) and (31). The smooth pasting conditions are sufficient to determine the function  $q$  and the “trading point”  $k^*$ .

**Proposition 5** For each trading cost  $c \geq 0$ , there exists a unique  $k^*$  that solves (31). If  $c = 0$  then  $k^* = 0$ . If  $c > 0$ ,  $k^* > c(r + \lambda)$ .

Proof: Let  $l(k) = [k - c(r + \lambda)](h'(k) + h'(-k)) - h(k) + h(-k)$ .

If  $c = 0$ ,  $l(0) = 0$ , and  $l'(k) = k[h''(k) - h''(-k)] > 0$ , for all  $k \neq 0$ . Therefore  $k^* = 0$  is the only root of  $l(k) = 0$ .

If  $c > 0$ , then  $l(k) < 0$ , for all  $k \in [0, c(r + \lambda)]$ . Since  $h''$  and  $h'''$  are increasing (Lemma 1), for all  $k > c(r + \lambda)$

$$\begin{aligned} l'(k) &= [k - c(r + \lambda)][h''(k) - h''(-k)] > 0, \\ l''(k) &= h''(k) - h''(-k) + [k - c(r + \lambda)][h'''(k) - h'''(-k)] > 0. \end{aligned}$$

Therefore  $l(k) = 0$  has a unique solution  $k^* > c(r + \lambda)$ .  $\square$

The next proposition establishes that the function  $q$  described by equation (29), with  $\beta_1$  and  $k^*$  given by (30) and (31), solves (23). The proof consists of two parts. First, it is established that (24) and (25) hold and that  $q'$  is bounded. A standard argument, *e.g.* Kobila (1993) or Scheinkman and Zariphopoulou (2001), is then used to show that in fact  $q$  solves equation (23).

**Proposition 6** The function  $q$  constructed above is an equilibrium option value function. The optimal policy consists of exercising immediately if  $g^o > k^*$ , otherwise wait until the first time in which  $g^o \geq k^*$ .

Proof: Let

$$b \equiv q(-k^*) = \frac{1}{(r + \lambda)} \frac{h(-k^*)}{(h'(k^*) + h'(-k^*))}, \quad (32)$$

then equation (29) implies

$$q(-x) = \begin{cases} \frac{b}{h(-k^*)} h(-x) & \text{for } x > -k^* \\ \frac{-x}{r+\lambda} + \frac{b}{h(-k^*)} h(x) - c & \text{for } x \leq -k^*. \end{cases}$$

Equation (24) may be rewritten as  $U(x) = q(x) - \frac{x}{r+\lambda} - q(-x) + c \geq 0$ . A simple calculation shows that

$$U(x) = \begin{cases} 2c & \text{for } x < -k^* \\ \frac{-x}{r+\lambda} + \frac{b}{h(-k^*)} [h(x) - h(-x)] + c & \text{for } -k^* \leq x \leq k^* \\ 0 & \text{for } x > k^* \end{cases}$$

Thus,  $U''(x) = \frac{b}{h(-k^*)}[h''(x) - h''(-x)]$ ,  $-k^* \leq x \leq k^*$ . Lemma 1 guarantees  $U''(x) > 0$  for  $0 < x < k^*$ , and  $U''(x) < 0$  for  $-k^* < x < 0$ . Since  $U'(k^*) = 0$ ,  $U'(x) < 0$  for  $0 < x < k^*$ . On the other hand,  $U'(-k^*) = 0$ , so  $U'(x) < 0$  for  $-k^* < x < 0$ . Therefore  $U(x)$  is monotonically decreasing for  $-k^* < x < k^*$ . Since  $U(-k^*) = 2c > 0$  and  $U(k^*) = 0$ ,  $U(x) > 0$  for  $-k^* < x < k^*$ . Hence equation (24) holds in  $(-\infty, k^*)$ .

Equation (25) holds by construction in the region  $x \leq k^*$ . Therefore it is sufficient to show that equation (25) is valid for  $x \geq k^*$ . In this region,  $q(x) = \frac{x}{r+\lambda} + \frac{b}{h(-k^*)}h(-x) - c$ , thus  $q'(x) = \frac{1}{r+\lambda} - \frac{b}{h(-k^*)}h'(-x)$  and  $q''(x) = \frac{b}{h(-k^*)}h''(-x)$ . Hence,

$$\begin{aligned} & \frac{1}{2}\sigma_g^2 q''(x) - \rho x q'(x) - r q(x) \\ &= \frac{b}{h(-k^*)} \left[ \frac{1}{2}\sigma_g^2 h''(-x) + \rho x h'(-x) - r h(-x) \right] - \frac{r+\rho}{r+\lambda} x + r c \\ &= -\frac{r+\rho}{r+\lambda} x + r c \leq -(r+\rho)c + r c = -\rho c < 0 \end{aligned}$$

where the inequality comes from the fact that  $x \geq k^* > (r+\lambda)c$  (see Proposition 5.)

Also  $q$  has an increasing derivative in  $(-\infty, k^*)$  and has a derivative bounded in absolute value by  $\frac{1}{r+\lambda}$  in  $(k^*, \infty)$ . Hence  $q'$  is bounded.

If  $\tau$  is any stopping time, the version of Ito's lemma for twice differentiable functions with absolutely continuous first derivatives (*e.g.* Revuz and Yor (1999), Chapter VI) implies that

$$e^{-r\tau} q(g_\tau^o) = q(g_0^o) + \int_0^\tau \left[ \frac{1}{2}\sigma_g^2 q''(g_s^o) - \rho g_s^o q'(g_s^o) - r q(g_s^o) \right] ds + \int_0^\tau \sigma_g q'(g_s^o) dW_s$$

Equation (25) states that the first integral is non positive, while the bound on  $q'$  guarantees that the second integral is a martingale. Using equation (24) we obtain,

$$\mathbb{E}^o \left\{ e^{-r\tau} \left[ \frac{g_\tau^o}{r+\lambda} + q(-g_\tau^o) - c \right] \right\} \leq \mathbb{E}^o [e^{-r\tau} q(g_\tau^o)] \leq q(g_0^o).$$

This shows that no policy can yield more than  $q(x)$ .

Now consider the stopping time  $\tau = \inf\{t : g_t^o \geq k^*\}$ . Such  $\tau$  is finite with probability one, and  $g_s^o$  is in the continuation region for  $s < \tau$ . Using exactly the same reasoning as above, but recalling that in the continuation region (25) holds with equality we obtain

$$q(g^o) = \mathbb{E}^o \left\{ e^{-r\tau} \left[ \frac{g_\tau^o}{r+\lambda} + q(-g_\tau^o) - c \right] \right\}.$$

□

It is a consequence of Proposition 6 that the process  $g^o$  will have values in  $(-\infty, k^*)$ . The value  $k^*$  acts as a barrier, and when  $g^o$  reaches  $k^*$ , a trade occurs, the owner's group switches and the process is restarted at  $-k^*$ .  $q(g^o)$  is the difference between the current owner's demand price and his fundamental valuation and can be legitimately called a "bubble".

The model determines the magnitude of the bubble and the duration between trades. The magnitude of the bubble can be measured by  $b$ , as in equation (32), the value of the resale option when a trade occurs.

If we write  $h$  in equation (27) as a function of both  $x$  and  $r$ , Scheinkman and Xiong (2003) show that the expected duration between two trades is given by

$$\mathbb{E}[\tau] = -\frac{\partial}{\partial r} \left[ \frac{h(-k^*, r)}{h(k^*, r)} \right] \Big|_{r=0}.$$

When the trading cost is zero ( $c = 0$ ), the trading barrier  $k^* = 0$ , which implies that the expected duration between trades is also zero. This is due to the fact that once a Brownian motion hits a point, it will hit the same point for infinite many times for any given period immediately afterward.

Various comparative statics results are described in Scheinkman and Xiong (2003). As investors become more overconfident ( $\phi$  increases), the volatility parameter of the difference in beliefs ( $\sigma_g$ ) increases, resulting in more trades (shorter duration between trades) and a larger price bubble. As the signals become more informative, the mean reverting speed of the difference in beliefs ( $\rho$ ) becomes larger, resulting in shorter duration between trades and a larger price bubble. As the trading cost  $c$  increases from zero, the duration between trades increases and the magnitude of the bubble is reduced, in a continuous manner. In particular the model predicts large trading volume in markets with sufficiently small transaction costs. The effect from an increase in trading cost is most dramatic for the duration between trades but the effects on the bubble are modest. This suggests that while a tax on transactions (Tobin Tax) would have some effect on trading volume, it would have a small effect on the size of a bubble.

In a risk-neutral world, one may consider several assets and analyze the equilibrium in each market independently. In this way the comparative statics properties described in the previous paragraph can be translated into results about correlations among equilibrium variables in the

different markets. Thus this model is potentially capable of explaining the observed cross-sectional correlation between market/book ratio and turnover for U.S. stocks in the period of 1996-2000 as documented by Cochrane (2002). It is also able to account for the analogous cross-sectional correlation that has been found by Mei, Scheinkman, and Xiong (2003) between the price ratio of Chinas A shares to B shares and turnover.

## 7 Other related models

In this section, we discuss other models that have been proposed in the finance literature to study price bubbles and effects of heterogeneous beliefs on trading and asset prices.

### 7.1 Models on rational bubbles

There has been a large literature studying rational bubbles including Blanchard and Watson (1982), Santos and Woodford (1997) and others. In these papers, all agents have identical rational expectations, and the asset prices can be decomposed into two parts, a fundamental component and a bubble component which is expected to grow at a rate equal to the risk free rate. In fact, such a rational bubble component can also be built into the models discussed in Sections 3 and 6. Given a price process  $p_t$  that satisfies equation (1) and  $m_t$ , a non-negative  $\mathcal{F}_t$ -martingale,  $\hat{p}_t = p_t + \gamma^{-t}m_t$  also satisfies the equation (1). A corresponding remark holds for equation (20) that describes the equilibrium of the model based on overconfidence. We have ruled out such rational bubbles in our previous discussion.

Campbell, Lo, and MacKinlay (1997, pages 258-260) provide a detailed discussion on the properties of rational bubbles. To make this type of bubbles sustainable, the asset must have a potentially infinite maturity. Another property of rational bubbles is that the asset price grows on average without bounds. In addition, the models of rational bubbles provide no explanation for the increase in trading that is often observed during historical bubble episodes.

As we discussed in the previous sections, the resale option provides an alternative way to analyze asset price bubbles, since its value is determined by heterogeneous beliefs among investors which is a variable orthogonal to the fundamental value of the asset. In contrast to rational bubbles, the resale option component does not need to be explosive although its magnitude could be very significant due to its recursive structure. Consequently, in the model

exposed in Section 6, variables such as the asset price in equation (22) and its change have stationary distributions. In addition, the size of the bubble generated from the resale option is positively correlated with trading volume, a property that is apparent in several actual episodes of price bubbles.

Finally, the resale option still exists for an asset with a fixed finite maturity, which is not possible for rational bubbles, that depend on a potentially infinite life. It should be apparent from the analysis in Section 6 that one can, in principle, treat an asset with a fixed terminal date. Equations (20) to (21) would apply with the obvious changes to account for the finite horizon. However, the option value  $q$  will now depend on the remaining life of the asset, introducing another dimension to the optimal stopping problem. The infinite horizon problem is stationary, greatly reducing the mathematical difficulty.

## 7.2 Trading caused by heterogeneous beliefs

Several other models have been proposed to analyze asset trading based on heterogeneous beliefs, such as Varian (1989), Harris and Raviv (1993), Kandel and Pearson (1995), Kyle and Lin (2003), and Cao and Ou-Yang (2004).

Harris and Raviv (1993) analyze a model with two groups of risk neutral speculators who trade a risky asset at dates  $t = 1, 2, \dots, T - 1$ . The final liquidation value of the asset at  $T$  is random and can be either high (H) or low (L). At each date a public noisy signal is revealed to the two groups of speculators, who assign different probability distributions for the signal and therefore would hold heterogeneous expectations of the final liquidation value of the asset. This mechanism for generating heterogeneous beliefs is similar to what we discuss in Section 5 with investor overconfidence, but with a different random process and noise distribution. The beliefs of the two groups are denoted by  $\pi_H^j(t)$ , the posterior probability that the final liquidation value will be high ( $j = 1, 2$ ).

To analyze trading between the two groups of speculators resulting from difference in their beliefs, the model also imposes a short-sales constraint, and that one group has sufficient market power to offer a price on a take-it-or-leave-it basis to the other group. As a result, the trading price of the asset will always equal the reservation price of the other group (the “price-taking” group). The existence of such a price-taking group, is not a natural assumption and rules out

the presence of a bubble in the observed trading prices, but greatly simplifies the recursive structure in the determination of equilibrium prices that arises in the models of Harrison and Kreps (1978) and Scheinkman and Xiong (2003). Harris and Raviv demonstrate in their model that trade will only occur when the two groups switch side ( $\pi_H^1$  and  $\pi_H^2$  flip order), and that there is a positive correlation between trading volume and absolute price changes (but not necessarily the level of prices.)

### 7.3 Effects of heterogeneous beliefs on risk premia

Lucas (1978) provides a simple and elegant equilibrium model to analyze the relation between equity premium and aggregate consumption. Consider an economy with a representative agent and an infinite time horizon ( $t = 1, 2, 3, \dots$ ). The agent maximizes his lifetime utility from consumption:

$$\mathbb{E} \left[ \sum_{t=0}^{\infty} \beta^t u(c_t) \right]$$

where  $\beta$ , ( $0 < \beta < 1$ ), is the agent's subjective discount factor,  $c_t$  is the consumption in period  $t$ , and the agent's utility from consumption  $u(c)$  is often assumed to have a power form:  $u(c) = \frac{1}{1-\gamma} c^{1-\gamma}$ . There are two assets – one is a risk-free asset and the other is a claim to aggregate endowment in the economy  $c_t$  ( $t=1, 2, \dots$ ). The risk-free asset is in zero net supply. The agent's marginal rate of consumption provides his pricing kernel for future cashflow:

$$m_{t+1} = \beta \frac{u'(c_{t+1})}{u'(c_t)}.$$

More specifically, a random cashflow of  $x_{t+1}$  at  $t + 1$  is worth

$$p_t = \mathbb{E}_t(m_{t+1}x_{t+1})$$

at period  $t$ . Thus, the riskfree rate is given by

$$R^f = 1/\mathbb{E}(m_t),$$

and the return on a risky asset  $R^i$  should satisfy

$$\mathbb{E}(m_t R_t^i) = 1.$$

By decomposing the last equation, we obtain  $\mathbb{E}(m_t)\mathbb{E}(R_t^i) + cov(m_t, R_t^i) = 1$ . Therefore,

$$\mathbb{E}(R_t^i) - R^f = -\frac{cov(m_t, R_t^i)}{\mathbb{E}(m_t)}.$$

This relation implies an upper bound on any risky asset's Sharpe ratio:

$$\left| \frac{\mathbb{E}(R_t^i) - R^f}{\sigma(R_t^i)} \right| = \left| -\frac{cov(m_t, R_t^i)}{\mathbb{E}(m_t)\sigma(R_t^i)} \right| \leq \frac{\sigma(m_t)}{\mathbb{E}(m_t)}. \quad (33)$$

The return from the stock market portfolio provides a way to calibrate the relationship between the equity risk premium and the pricing kernel implied by aggregate consumption. In the case of the market portfolio  $R^{mv}$ , the relation in formula (33) holds exactly, and a power utility function would imply

$$\left| \frac{\mathbb{E}(R^{mv}) - R^f}{\sigma(R^{mv})} \right| = \frac{\sigma(m_{t+1})}{\mathbb{E}(m_{t+1})} = \frac{\sigma[(c_{t+1}/c_t)^{-\gamma}]}{\mathbb{E}[(c_{t+1}/c_t)^{-\gamma}]} \approx \gamma\sigma(\Delta \ln c),$$

where  $R^{mv}$  is the return on the market portfolio and  $R^f$  the risk-free rate. As summarized by Cochrane (2001, page 23), “over the last 50 years in the United States, real stock returns have averaged 9% with a standard deviation of about 16%, while the real return on treasury bills has been about 1%. Thus, the historical annual market Sharpe ratio has been about 0.5. Aggregate nondurable and services consumption growth had a mean and standard deviation of about 1%. We can only reconcile these facts if investors have a risk-aversion coefficient of 50,” which is much higher than what economists have usually assumed. This is called an “equity premium puzzle” by Mehra and Prescott (1985).

The equity premium puzzle has motivated a large number of papers since the mid eighties. The objective of these papers is to present modifications of the standard model that justify a Sharpe ratio that is considerably higher than the one implied by the standard model. Part of this literature has used heterogenous beliefs. Williams (1977), Abel (1990) Detemple and Murthy (1994), Zapatero (1998), and Basak (2000), among others, analyze the effects of heterogenous beliefs on the equilibrium risk premium and interest rates.

Detemple and Murthy (1994) consider a continuous time production economy of Cox, Ingersoll and Ross (1985) with a Brownian uncertainty structure. They assume a risky production technology with a return that is invariant to scale and that has an unobservable mean. In addition, there are two groups of risk-averse agents who have heterogeneous prior beliefs on

the mean return of the production technology. There are two assets - a claim to the aggregate output and a risk-less asset in zero net supply. Heterogeneous beliefs motivate agents in one group to borrow from agents in the other group using the risk-less asset. In equilibrium, the interest rate and risk premium on risky securities are determined by the wealth distribution across the two groups and each groups' estimates of the production growth rate. Zapatero (1998) considers a similar model in a pure exchange economy of Lucas (1978). He shows that heterogeneous beliefs can lead to a reduced risk-free interest rate in equilibrium from an otherwise identical economy with homogeneous beliefs. This lower interest rate induces a higher excess return. Basak (2000), using a model similar to Detemple and Murthy (1994), further shows that heterogeneous beliefs could add risk to investors' financial investment and therefore may lead to a greater equity premium than an economy with homogeneous beliefs.

## 8 Survival of traders with incorrect beliefs

In the earlier sections, we discussed various effects that can arise when traders with heterogeneous beliefs interact with each other in an asset market. Some traders may have incorrect beliefs which are generated from incorrect prior beliefs or from incorrect information processing rules, while some others may be smarter and have beliefs that are closer to the objective ones. In such an environment, an important question is whether traders with incorrect beliefs will lose money trading with smarter traders and eventually disappear from the market.

There has been a long debate on this fundamental issue. Friedman (1953) argues that irrational traders who use wrong beliefs cannot survive in a competitive market, since they will eventually lose their wealth to rational traders in the long run. More recently, De Long, Shleifer, Summers and Waldman (1991) suggest that traders with wrong beliefs may survive in the long-run since they may hold a portfolio with excessive risk but also higher expected return and therefore their wealth can eventually outgrow that of rational traders. Several recent studies have been devoted to analyze this issue, *e.g.* Sandroni (2000), Blume and Easley (2001), and Kogan, Ross, Wang and Westerfield (2004). However, the answer is still inconclusive. Depending on the model assumptions, different results have been found in these studies. Here, we discuss a model from Kogan *et al.* (2004) as an example.

Kogan *et al.* consider an economy that has a finite horizon and evolves in continuous time.

Uncertainty is described by a one-dimensional, standard Brownian motion  $Z_t$  for  $0 \leq t \leq T$ , which is defined on a complete probability space  $(\Omega, F, P)$ . There is a single share of a risky asset in the economy, the stock, which pays a terminal dividend payment  $D_T$  at time  $T$ , determined by process

$$dD_t = D_t(\mu dt + \sigma dZ_t)$$

where  $D_0 = 1$  and  $\sigma > 0$ . There is also a zero coupon bond available in zero net supply. Each unit of the bond makes a sure payment of one at time  $T$ . We use the risk-free bond as the numeraire and denote the price of the stock at time  $t$  by  $S_t$ .

There are two types of competitive traders in the economy. Each set corresponds to half of the total number of traders. At time zero, all traders have an equal endowment of the stock (that we normalize as  $1/2$ ) and have zero units of the bond. Traders differ in their beliefs of the drift parameters of the dividend process. One set of traders, the rational traders, knows the true probability measure  $P$ , while the other set of traders, the irrational traders, believes incorrectly that the probability measure is  $Q$ , under which

$$Z_t^Q = Z_t - \int_0^t (\sigma \eta) ds$$

is a standard Brownian motion. Hence, for an irrational trader

$$dD_t = D_t[(\mu + \sigma^2 \eta)dt + \sigma dZ_t^Q],$$

where  $Z_t^Q$  is a Brownian motion. The constant  $\eta$  parameterizes the irrationality of these traders. When  $\eta$  is positive, the irrational traders are optimistic about the prospects of the evolution of the dividend and overestimates the rate of growth of the dividend. Conversely, a negative  $\eta$  corresponds to a pessimistic view. According to this specification, the two types of traders do not update their beliefs of the drift rate, and one group will always stay as the optimistic one. This structure is different from the one introduced in Section 5, where agents' beliefs fluctuate with the flow of information.

Rational traders and irrational traders consume only at time  $T$ , and they share the same constant relative risk aversion. They can trade continuously before time  $T$  and would aim to maximize their utility based on their own probability measure. Thus, a rational trader's

objective is to maximize

$$\mathbb{E}_0^P \left[ \frac{1}{1-\gamma} C_{r,T}^{1-\gamma} \right]$$

where  $C_{r,T}$  is the consumption of the typical rational trader at time  $T$ . Correspondingly an irrational trader's objective is to maximize

$$\mathbb{E}_0^Q \left[ \frac{1}{1-\gamma} C_{n,T}^{1-\gamma} \right]$$

where  $C_{n,T}$  is the consumption of an irrational trader at time  $T$ . In addition, the model assumes that there is no trading cost and short-sales of shares are allowed.

The probability measure used by irrational traders  $Q$  is absolutely continuous with respect to the objective measure  $P$ . Thus, the expectation using probability measure  $Q$  can be transformed into an expectation using probability measure  $P$  through the density (Radon-Nikodym derivative) of the probability measure  $Q$  with respect to  $P$ :

$$\xi_t \equiv \left. \frac{dQ}{dP} \right|_t = e^{-\frac{1}{2}\eta^2\sigma^2 t + \eta\sigma Z_t}. \quad (34)$$

Hence the irrational trader maximizes

$$\mathbb{E}_0^Q \left[ \frac{1}{1-\gamma} C_{n,T}^{1-\gamma} \right] = \mathbb{E}_0^P \left[ \xi_t \frac{1}{1-\gamma} C_{n,T}^{1-\gamma} \right].$$

Effectively, an irrational trader is maximizing a state dependent utility function,  $\xi_t \frac{1}{1-\gamma} C_{n,T}^{1-\gamma}$ , under the true probability measure  $P$ .

The competitive equilibrium of the economy described above can be solved analytically. Since there is only one source of uncertainty in the economy and there are no trading cost or short-sales constraints, it is expected that continuous trading on the stock and the bond is sufficient to dynamically complete the markets. Since complete markets yield Pareto efficient allocations, it is natural to first examine the set of Pareto efficient allocations and then show that with the corresponding support prices the two assets actually yield complete markets. This is the route chosen by Kogan *et al.* First they examine the set of allocations that solve:

$$\begin{aligned} \max \quad & \frac{1}{1-\gamma} C_{r,T}^{1-\gamma} + b\xi_t \frac{1}{1-\gamma} C_{n,T}^{1-\gamma} \\ \text{s.t.} \quad & C_{r,T} + C_{n,T} = D_T \end{aligned}$$

where  $b$  is the ratio of the utility weights for the two types of traders. The optimal allocations are

$$C_{r,T} = \frac{1}{1 + (b\xi_t)^{1/\gamma}} D_T, \quad C_{n,T} = \frac{(b\xi_t)^{1/\gamma}}{1 + (b\xi_t)^{1/\gamma}} D_T. \quad (35)$$

Based on the traders' marginal utilities, one can derive the supporting state price density given the information at time  $t$  as

$$\frac{(1 + (b\xi_t)^{1/\gamma})^\gamma D_T^{-\gamma}}{\mathbb{E}_t[(1 + (b\xi_t)^{1/\gamma})^\gamma D_T^{-\gamma}]}.$$

This state price density allows us to price any contingent claim in the economy, such as the dividend from the stock that is paid at  $T$  and the traders' consumption allocations. Kogan *et al.* show that the equilibrium stock price is given by

$$P_t = \frac{\mathbb{E}_t[(1 + (b\xi_t)^{1/\gamma})^\gamma D_T^{-\gamma} Z_T]}{\mathbb{E}_t[(1 + (b\xi_t)^{1/\gamma})^\gamma D_T^{-\gamma}]}.$$

The utility weights between the two groups of traders,  $b$ , is determined so that the budget constraints for the two traders are satisfied. Since the traders start with the same endowments at time  $t = 0$ , the values of their consumption allocations at time  $t = T$  should have the same value at  $t = 0$ . This gives an equation to identify  $b = e^{(\gamma-1)\eta\sigma^2 T}$ . With these calculations, Kogan *et al.* further show that in fact the stock and the bond dynamically complete markets.

Given the equilibrium consumptions for the two types, Kogan *et al.* use the asymptotic properties of the two consumptions to discuss the survival of traders. More specifically, they say that irrational traders experience relative extinction in the long-run if

$$\lim_{T \rightarrow \infty} \frac{C_{n,T}}{C_{r,T}} = 0 \quad \text{a.s.}$$

The relative extinction of rational traders is defined symmetrically. Kogan *et al.* use the expression for consumption allocations in equation (35) above, the formula for the Radon-Nykodim derivative of  $Q$  with respect to  $P$  (equation (34)), and the strong law of large numbers for Brownian motion to establish:

**Proposition 7** *Suppose  $\eta \neq 0$ . Let  $\eta^* = 2(\eta - 1)$ . For  $\gamma > 1$  and  $\eta \neq \eta^*$ , typically only one type of traders survives. Furthermore:*

*Case 1:  $\eta < 0$  (pessimistic irrational traders), the rational traders survive.*

*Case 2:  $0 < \eta < \eta^*$  (modestly optimistic irrational traders), the irrational traders survive.*

*Case 3:  $\eta > \eta^*$  (strongly optimistic irrational traders), the rational trader survive.*

Interestingly, the irrational trader could survive in the long-run, as in case 2 with a modestly optimistic belief. In such a case, the irrational trader takes more risk and therefore able to outgrow the rational trader. This effect has been pointed out by De Long, Shleifer, Summers and Waldman (1991) using a partial equilibrium model in which irrational traders' trading has no impact on the price. Once their price impact is taken into account as in this model, Kogan *et al.* show that the survival of irrational traders becomes less likely, although still possible. Kogan *et al.* also discuss the price impact of the irrational traders in the long run. They also show that even if the irrational traders do not survive in the long run, irrational traders can still have a persistent impact on the stock price since they are willing to bet strongly on some small probability events when their probability assessment of these events differ greatly from that of rational traders.

## 9 Some Remaining Problems

Many interesting problems remain in modelling the effects of heterogeneous beliefs on financial markets. Rather than present a long list of unsolved questions we select a few problems that are particularly related to the models discussed above.

The model in Section 5 specifies a normal process for the unobservable fundamental variable. This assumption allows one to use the standard linear filtering technique to analyze the agents' learning process. It also generates a particularly tractable form for the process determining the difference in beliefs. However, a lognormal process would help capture the limited liability feature of many assets such as stocks and bonds. A model of this kind would have to contemplate the difficulties involved in nonlinear filtering and optimal stopping with multiple state variables. Panageas (2003) provides an attempt along this line to analyze the effect of stock market bubbles on firm investment. Multi-dimensional stopping time problems will also result from the introduction of multiple classes of agents.

As discussed above, heterogeneous beliefs should be able to support a bubble even if an asset

has a finite life in the context of the model in Section 6. However, in this case, the optimal stopping problem that defines the equilibrium would involve an extra dimension. Tackling this additional complication would allow one to analyze the exact impact of horizon on speculative trading and price bubbles, in addition to establishing rigorously that a bubble generated by heterogeneous beliefs can prevail in assets with a finite life.

The models in Sections 3 and 6 ignore risk aversion of agents by assuming risk neutrality. When agents are risk averse, subtle effects arise in equilibrium. As shown by Hong, Scheinkman and Xiong (2003) in a model with finite periods and myopic risk averse agents, the payoff from the resale option appears to be similar to a standard call option with a strike price determined by the asset float (number of tradable shares) and the risk bearing capacity of investors. The model suggests that asset float could have an important effect on price bubble and trading volume. A more elaborate model remains to be developed to analyze the effects of risk aversion and asset float when agents have heterogeneous beliefs and short-sales of assets are constrained.

In the model in Section 6 the level of overconfidence is a constant. It is perhaps more realistic to assume that the level of investor's overconfidence fluctuates as investors learn of their own ability to forecast. Gervais and Odean (2001) analyze a model in which a trader infers his ability to forecast from his past successes and failures. They show that overconfidence can be generated in this learning process if the trader attributes success to himself and failure to external forces, an attribution bias that has been documented in psychological studies of human behavior. It remains an open problem to analyze the dynamics of heterogeneous beliefs in the presence of endogeneously generated overconfidence and the market equilibrium that would obtain.

## References:

Abel, Andrew (1990), Asset prices under heterogeneous beliefs: implications for the equity premium, Working paper, Wharton School, University of Pennsylvania.

Abramowitz, Milton and Irene Stegun (1964), *Handbook of Mathematical Functions*, Dover Publications, New York.

Alpert, Marc and Howard Raiffa (1982), A progress report on the training of probability assessors, in Daniel Kahneman, Paul Slovic, and Amos Tversky, ed.: *Judgement under Uncertainty: Heuristics and Biases*, Cambridge University Press, Cambridge.

Barberis, Nicholas and Richard Thaler (2003), A survey of behavioral finance, in George Constantinides, Milton Harris and Rene Stulz (ed.), *Handbook of the Economics of Finance*, North-Holland.

Basak, Suleyman (2000), A model of dynamic equilibrium asset pricing with heterogeneous beliefs and extraneous beliefs, *Journal of Economic Dynamics and Control* 24, 63-95.

Bernardo, Antonio, and Ivo Welch (2001), On the evolution of overconfidence and entrepreneurs, *Journal of Economics and Management Strategy* 10, 301-330.

Biais, Bruno and Peter Bossaerts (1998), Asset prices and trading volume in a beauty contest, *Review of Economic Studies* 65, 307-340.

Blanchard, Olivier and Mark Watson (1982), Bubbles, rational expectations and financial markets, in P. Wachtel (ed.) *Crisis in the Economic and Financial Structure: Bubbles, Bursts, and Shocks*, Lexington Press, Lexington, MA.

Blume, Lawrence and David Easley (2001), If you're so smart, why aren't you rich? Belief selection in complete and incomplete markets, Cowles Foundation Discussion Papers, 1319.

Brav, Alon and J.B. Heaton (2002), Competing Theories of Financial Anomalies, *Review of Financial Studies* 15, 575-606.

Brenner, Lyle, Derek Koehler, Varda Liberman, and Amos Tversky (1996), Overconfidence in probability and frequency judgements, *Organizational Behavioral and Human Decision Processes* 65 (3), 212-219.

Brunnermeier, Markus and Jonathan Parker (2003), Optimal expectations, Working paper, Princeton University.

Camerer, C. (1995), Individual decision making, in John Kagel and Alvin Roth (ed.) *The*

*Handbook of Experimental Economics*, Princeton University Press, Princeton, NJ.

Campbell, John, Andrew Lo, and Craig MacKinlay (1997), *The Econometrics of Financial Markets*, Princeton University Press, Princeton, NJ.

Cao, Henry and Hui Ou-Yang (2004), Differences of opinion and options, Working paper, University of North Carolina and Duke University.

Chen, Joseph, Harrison Hong and Jeremy Stein (2002), Breadth of ownership and stock returns, *Journal of Financial Economics* 66, 171-205.

Cochrane, John (2001), *Asset Pricing*, Princeton University Press, Princeton, NJ.

Cochrane, John (2002), Stocks as money: convenience yield and the tech-stock bubble, NBER Working Paper 8987.

Cox, John, Jonathan Ingersoll and Stephen Ross (1985), An intertemporal general equilibrium model of asset prices, *Econometrica* 53, 363-384.

Daniel, Kent, David Hirshleifer, and A. Subrahmanyam (1998), Investor psychology and security market under- and overreaction, *Journal of Finance* 53, 1839-1885.

De Long, Bradford, Andrei Shleifer, Lawrence Summers, and Robert Waldman (1991), The survival of noise traders in financial markets, *Journal of Business* 64, 1-19.

Detemple, Jerome and Shashidhar Murthy (1994), Intertemporal asset pricing with heterogeneous beliefs, *Journal of Economic Theory* 62, 294-320.

Diamond, Douglas and Robert Verrecchia (1987), Constraints on short-selling and asset price adjustment to private information, *Journal of Financial Economics* 18, 277-311.

Friedman, Milton (1953), The case for flexible exchange rates, *Essays in Positive Economics*, University of Chicago Press.

Gallmeyer, Michael and Burton Hollifield (2004), An examination of heterogeneous beliefs with a short sale constraint, Working paper, Carnegie Mellon University.

Gervais, Simon and Terrance Odean (2001), Learning to be overconfident, *Review of Financial Studies* 14, 1-27.

Grossman, Sanford and Joseph Stiglitz (1980), On the impossibility of informationally efficient markets, *American Economic Review* 70, 393-408.

Harris, Milton and Artur Raviv (1993), Differences of opinion make a horse race, *Review of Financial Studies* 6, 473-506.

Harrison, Michael and David Kreps (1978), Speculative investor behavior in a stock market with heterogeneous expectations, *Quarterly Journal of Economics* 92, 323-336.

Hirshleifer, David (2001), Investor psychology and asset pricing, *Journal of Finance* 56, 1533-1597.

Hong, Harrison, José Scheinkman, and Wei Xiong (2003), Asset float and speculative bubbles, Working paper, Princeton University.

Jarrow, Robert (1980), Heterogeneous expectations, restrictions on short-sales, and equilibrium asset prices, *Journal of Finance* 35, 1105-1113.

Kandel, Eugene and Neil Pearson (1995), Differential interpretation of public signals and trade in speculative markets, *Journal of Political Economy* 103, 831-872.

Kogan, Leonid, Stephen Ross, Jiang Wang, and Mark Westerfield (2004), The price impact and survival of irrational traders, Working paper, MIT.

Kyle, Albert (1985), Continuous auctions and insider trading, *Econometrica* 53, 1315-1336.

Kyle, Albert and Tao Lin (2003), Continuous trading with heterogeneous beliefs and no noise trading, Working paper, Duke University.

Kyle, Albert and Albert Wang (1997), Speculation duopoly with agreement to disagree: Can overconfidence survive the market test? *Journal of Finance* 52, 2073-2090.

Lintner, John (1969), The aggregation of investor's diverse judgements and preferences in purely competitive security markets, *Journal of Financial and Quantitative Analysis* 4, 347-400.

Liptser, R. S. and A. N. Shiriyayev (1977), *Statistics of Random Processes*, Springer-Verlag, New York.

Lucas, Robert (1978), Asset prices in an exchange economy, *Econometrica* 46, 1429-1446.

Mehra, Rajnish and Edward Prescott (1985), The equity premium puzzle, *Journal of Monetary Economics* 15, 145-161.

Mei, Jianping, Jose Scheinkman and Wei Xiong (2003), Speculative trading and stock prices: An analysis of Chinese A-B share premia, Working paper, Princeton University.

Milgrom, Paul and Nancy Stokey (1982), Information, trade and common knowledge, *Journal of Economic Theory* 12, 112-128.

Miller, Edward (1977), Risk, uncertainty and divergence of opinion, *Journal of Finance* 32, 1151-1168.

Morris, Stephen (1996), Speculative investor behavior and learning, *Quarterly Journal of Economics* 110, 1111-1133.

Odean, Terrance (1998), Volume, volatility, price, and profit when all traders are above average, *Journal of Finance* 53, 1887-1934.

Ofek, Eli and Matthew Richardson (2003), Dotcom mania: The rise and fall of internet stock prices, *Journal of Finance* 58, 1113-1137.

Panageas, Stavros (2003), Speculation, overpricing and investment – theory and empirical evidence, Job market paper, MIT.

Revuz, Daniel and Marc Yor (1999), *Continuous Martingales and Brownian Motion*, Springer, New York.

Rogers, L. C. G. and David Williams (1987), *Diffusions, Markov Processes, and Martingales, Volume 2: Ito Calculus*, John Wiley & Sons, New York.

Sandroni, Alvaro (2000), Do markets favor agents able to make accurate predictions? *Econometrica* 68, 1303-1341.

Santos, Manuel, and Michael Woodford (1997), Rational asset pricing bubbles, *Econometrica* 65, 19-57.

Scheinkman, Jose and Thaleia Zariphopoulou (2001), Optimal environmental management in the presence of irreversibilities, *Journal of Economic Theory* 96, 180-207.

Scheinkman, Jose and Wei Xiong (2003), Overconfidence and Speculative Bubbles, *Journal of Political Economy* 111, 1183-1219.

Tirole, Jean (1982), On the possibility of speculation under rational expectations, *Econometrica* 50, 1163-1181.

Varian, Hal (1989), Differences of opinion in financial markets, in Courtenay Stone (ed.) *Financial Risk: Theory, Evidence and Implications*, Kluwer Academic Publishers, Boston.

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

Williams, Joseph (1977), Capital asset prices with heterogeneous beliefs, *Journal of Financial Economics* 5, 219-239.

Zapatero, Fernando (1998), Effects of financial innovations on market volatility when beliefs are heterogeneous, *Journal of Economic Dynamics and Control* 22, 597-626.