AP
7/13/06
2:15 pm

Is Noise Trading Cancelled Out by Aggregation?

Hongjun Yan*

Yale School of Management

June 6, 2006

*I am grateful to Nicholas Barberis, James Choi, Lauren Cohen, Jon Ingersoll, Owen Lamont, Rajnish Mehra, Antti Petajisto, Nagpurnanand Prabhala, Steve Ross, Robert Shiller, Matt Spiegel, and seminar participants at Yale for helpful discussions. Dong Lou and Steve Malliaris provided excellent research assistance. I thank the Whitebox Advisors Grant for its generous support of this research. Please direct all correspondence to Hongjun Yan, Yale School of Management, 135 Prospect Street, P.O. Box 208200, New Haven, CT 06520-8200. Email: hongjun.yan@yale.edu, Web: http://www.som.yale.edu/faculty/hy92/.

Is Noise Trading Cancelled Out by Aggregation?

Abstract

Conventional wisdom suggests that investors' independent biases would cancel out each other and have little impact on the equilibrium at the aggregate level. In contrast to this intuitive argument, this paper analyzes models with biased investors and finds that biases often have a significant impact on the equilibrium even if the biases are independent across investors. First, biases decrease the expected stock return if investors' demand function for the stock is convex in the biases, and increase the expected stock return if the demand function is concave. This also provides various novel predictions on the relation between differences of opinion and cross-sectional stock returns. Second, biases may have a significant impact on the equilibrium due to the fluctuation of wealth distribution. In particular, this implies mean reversion in stock returns. An initial run-up of stock price makes optimistic investors richer, which then pushes the stock price up and leads to lower future returns. Similarly, an initial drop of stock price leads to higher future returns.

JEL Classification Numbers: B40, D90, G12.

Keywords: Aggregation, bias, noise trading.

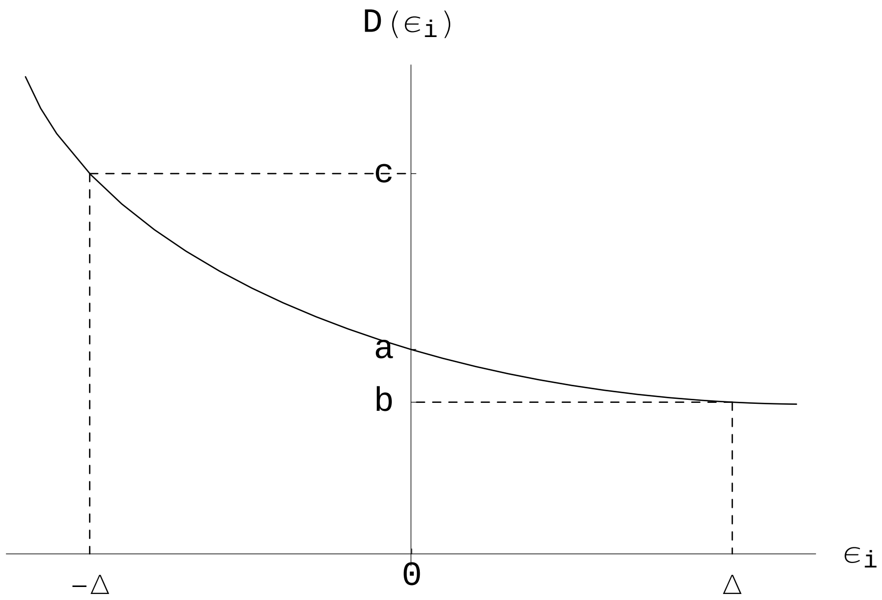
1 Introduction

While it is almost beyond debate that individuals have biases, there is less consensus on whether the biases have a significant impact on the equilibrium at the aggregate level. One conventional argument suggests that if biases are independent across investors, they generally should not have a large impact on the equilibrium since they would cancel out each other. While the existing behavioral finance literature has emphasized that individual biases tend to be correlated and so cannot be cancelled out by aggregation (see, e.g., Shleifer (2000)), this paper directly examines this aggregation argument by analyzing whether *independent* biases have a large impact on the equilibrium.

In contrast to this conventional aggregation argument, the findings suggest that individual biases often have a significant impact on the equilibrium even if they are independent across investors. To understand the intuition behind this result, let's first recapitulate the conventional aggregation argument. Suppose an unbiased investor's demand for a stock is D, which presumably is derived from utility maximization and depends on the price of the stock and other parameters. There are N biased investors, and investor i's demand is $D_i = D + \epsilon_i$, where ϵ_i is a random draw from $\tilde{\epsilon}$, which is a random variable with a mean of 0. If ϵ_1, ϵ_2,...,ϵ_N are independent and N is large, then the aggregate demand, $\sum_{i=1}^{N} D_i$ is approximately $N \times D$, which is the aggregate demand in the case without biases. As a result, the biases have little impact on the equilibrium stock price.

There are two important cases where this argument fails. First, the argument implicitly assumes that the bias affects one's demand in a linear way, i.e. D_i is a linear function of ϵ_i. If, for example, investors' demand is a convex function of their biases, then the biases increase the aggregate demand and so increase the stock price, even if the biases are independent across investors. Similarly, biases decrease the stock price if the demand function is concave in the bias. This is analogous to Jensen's inequality: if x is a random variable and f is a convex function, then $E[f(x)] > f[E(x)]$; if f is a concave function, then $E[f(x)] < f[E(x)]$. The economic intuition is illustrated in the following example. Typically, an investor's demand for the stock is a decreasing and convex function of stock return volatility σ: $D(\sigma)$ (see, e.g., Liu (2005)). Suppose investor i is biased about σ, that is, investor i's perceived volatility is $\sigma + \epsilon_i$. Then his demand for the stock is $D(\sigma + \epsilon_i)$. Intuitively, an investor invests less in the stock market if he overestimates the risk, i.e., $\epsilon_i > 0$. Similarly, an investor invests more in the stock market if he underestimates the risk, i.e., $\epsilon_i < 0$. Note that D is a convex function of ϵ_i. This implies, as illustrated in Figure 1, that the increase of stock holding induced by underestimation outweighs the decrease of stock holding induced by the same amount of overestimation. As a result, biases increase the aggregate demand and hence increase the stock price even though the biases are independent across investors and the population is unbiased.

Figure 1: The impact of biases on demand. This figure plots an investor's demand D as a function of his bias ϵ_i. If an investor is unbiased, i.e., $\epsilon_i = 0$, his demand is a. The investor's demand is b if he has a bias of Δ, and c if he has a bias of $-\Delta$. The convexity in the demand function implies that $c - a > a - b$, that is, the increase of demand induced by $-\Delta$ outweighs the decrease of demand induced by Δ.



In order to elaborate further on the above intuition and to evaluate its implications quantitatively, I also analyze three examples based on a typical demand function. The analysis shows that the impact of biases critically depends on the form of the bias. Example 1 shows that, if investors are only biased about the expected stock return, the biases are cancelled out by aggregation and have no impact on the equilibrium. Example 2, however, shows that if investors are only biased about volatility, this bias can substantially increase the stock price and so decrease the expected return. Finally, investors have two biases in Example 3: one bias is about the expected return and the other is about the volatility. This leads to the following three main predictions. 1) The biases decrease the expected return, if the correlation between these two biases is negative, that is, investors who overestimate the expected return tend to underestimate the volatility and vice versa. 2) Holding everything else constant, the expected return increases with respect to the correlation. 3) The biases can increase the expected return if the correlation is high enough. The economic intuition can be illustrated in the following example. Suppose the correlation between these two biases is negative. Then, the investors who are optimistic about the expected return tend to underestimate the volatility. As a result, they have a high demand for the stock. On the other hand, the investors who are very pessimistic about the expected return would like to short the stock. However, their short position is limited since they also tend to overestimate the volatility and so feel that the short position is risky. Therefore, the biases increase the aggregate demand and lead to lower future returns. Similar intuition leads to the results in 2) and 3).

The second case in which the aggregation argument fails is that, even if the biases affect demand in a linear way, they may still have a significant impact on the equilibrium due to the fluctuation of wealth distribution. The reason can be illustrated in the following simple example. Suppose there are two investors, A and B, and both have the same amount of initial wealth. A is optimistic about a stock and B pessimistic. At the initial date, relative to an investor without bias, A wants to hold more stock, and B less. If the demand is a linear function of the bias, the biases do not affect the total demand from A and B. So, the equilibrium stock price is not affected by the biases. This is essentially the traditional aggregation argument. The drawback of this argument arises in a dynamic setting. Suppose, after one period, the stock price goes up. Then the optimistic investor A has a larger wealth share relative to B since A chose to hold more stock in the previous period. Similarly, the pessimistic investor B has a larger wealth share if the stock goes down. This implies mean reversion in stock returns. An initial run-up of stock price makes optimistic investors richer, which then pushes the stock price up and leads to lower future returns. Similarly, an initial drop of stock price makes pessimistic investors richer, which then presses the stock price down and leads to higher future returns. Note that the previous argument does not rely on the assumption that A and B's biases don't change over time, and similar results arise as long as investors' biases are persistent over time.

The impact from the fluctuation of wealth share can be significant after large positive or negative stock returns, if investors' beliefs are widely dispersed. In one example, it is shown that the wealth fluctuation may lead to stock price "overshooting": a large positive stock return makes the optimistic investors relatively so rich that they may push the stock price up so much so that the future expected return becomes negative.

This paper is related to the literature on short sales constraint and differences of opinion. Due to short sales constraint, pessimistic investors stay out of the market and so the stock is overvalued and tends to have lower future returns (Miller (1977), Harrison and Kreps (1978), Scheinkman and Xiong (2003)). A large body of empirical evidence seems to be consistent with this argument (Chen, Hong and Stein (2002), Diether, Malloy and Scherbina (2002), Jones and Lamont (2002), Lamont and Thaler (2003), Hong, Scheinkman and Xiong (2005), Mei, Scheinkman and Xiong (2005), and Nagel (2005)). This literature shows that even if the population is unbiased and individual biases are independent across the population, the biases still affect the equilibrium because the short sales constraint partially restrains the impact from pessimistic investors. My paper complements this literature by showing that the difference in opinions may still have a large impact on the equilibrium even without short sales constraint. In addition to various novel predictions on the relation between differences of opinion and stock returns, this paper also sheds some light on the potential impacts on stock price behavior when the short sales constraint is alleviated. Analyzing a similar mechanism, Xiong and Yan (2006) focus on the impact of heterogenous beliefs on bond yields.

This paper is also related to the literature on the impact of the fluctuation of wealth distribution. For example, Dumas (1989), Wang (1996) and Chan and Kogan (2002) study the impact of wealth share fluctuation induced by the heterogeneity in risk aversion. The current paper analyzes whether independent biases are cancelled out by aggregation, and points out that one of the causes for the failure of the traditional aggregation argument is the impact of wealth share fluctuation induced by biases. One salient distinction is that, due to biased beliefs, the expected stock return in my model may become negative, while this does not happen in models with only heterogeneity in risk aversion. Shefrin and Statman (1994), Shefrin (2005) and Yan (2005) analyze dynamic models with biased investors having log or power utility functions, and find that, due to wealth share fluctuation, biases almost always affect the equilibrium. This paper considers general demand functions and derives testable predictions from wealth share fluctuation. More importantly, this paper demonstrates that the form of the biases plays a crucial role in determining the equilibrium and this has not been studied in the literature.

The rest of the paper is organized as follows. Section 2 presents a one-period model to show that the aggregation argument may fail if biases affect investors' demand in a nonlinear way. Section 3 presents a dynamic model to illustrate that, even if biases affect the demand in a linear way, the aggregation argument may still fail because of the impact from wealth share fluctuation. Section 4 concludes. All proofs are provided in the Appendix.

2 A Static Model

Let's consider a one-period (two dates) model with $t = 0, 1$. There are two assets in the economy: a riskless bond and a stock. The riskless interest rate is r_f. The stock, which is normalized to one share, is a claim to a positive dividend v_1 at $t = 1$. There are N investors and each is endowed with $1/N$ share of the stock. At $t = 0$, investors make portfolio decisions and the stock price P_0 is determined in the equilibrium by equating the aggregate demand for the stock to the supply.

Investor i is assumed to allocate a fraction θ_i of his wealth to the stock market:

$$\theta_i = \theta\left(P_0, W_i, \Psi, \epsilon_i\right), \tag{1}$$

where W_i is investor i's wealth; Ψ includes parameters such as the investor's preference and the distribution of v_1, ϵ_i denotes investor i's bias and $\epsilon_i = 0$ corresponds to the case in which investor i is unbiased. The unbiased investor's decision rule $\theta\left(P_0, W_i, \Psi, 0\right)$ can be interpreted as the one that maximizes his expected utility. $\theta\left(P_0, W_i, \Psi, \epsilon_i\right)$ can be interpreted as the "optimal" decision from the perspective of an investor with a bias ϵ_i.

The demand function θ is assumed to satisfy the following conditions

$$\frac{\partial \theta}{\partial P_0} < 0, \tag{2}$$

$$\lim_{P_0 \to \infty} \theta = -\infty, \tag{3}$$

$$\lim_{P_0 \to 0} \theta = \infty. \tag{4}$$

The above condition (2) implies that investors demand less stock if the stock price is higher. For simplicity, technical conditions (3) and (4) are made to ensure the existence and uniqueness of the equilibrium.

Investors' biases, ϵ_i for $i = 1, ..., N$, are independent draws from $\tilde{\epsilon}$, which is a random variable with

$$E[\tilde{\epsilon}] = 0, \tag{5}$$

$$E[\theta(P_0, W_i, \Psi, \tilde{\epsilon})] < \infty. \tag{6}$$

Intuitively, this assumes that the biases are independent across investors and equation (5) implies that the population is unbiased. Condition (6) is imposed for technical reasons. In contrast to the conventional aggregation argument, the following proposition shows that independent biases are not always cancelled out by aggregation.

Proposition 1 *In the above described economy, the impact of the biases can be summarized as follows:*

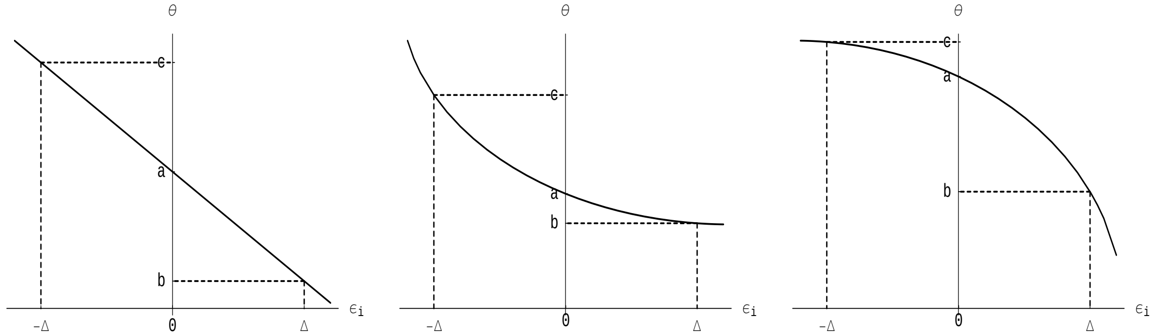
i) if θ is a linear function of ϵ_i then the biases have no impact on the stock price;

ii) if θ is a convex function of ϵ_i then the biases increase the stock price;

iii) if θ is a concave function of ϵ_i then the biases decrease the stock price.

This proposition reveals that the conventional aggregation argument holds when the demand for stock is a linear function of the bias, but not when the demand function is convex or concave. The intuition can be illustrated by Figure 2, where θ is assumed to be a decreasing function of ϵ_i. If an investor is unbiased, $\theta = a$. The bias reduces θ from a to b if $\epsilon_i = \Delta$, and increases θ from a to c if $\epsilon_i = -\Delta$. In Panel A, θ is a linear function of ϵ_i, which implies $a - b = c - a$. That is, the decrease of demand induced by a bias Δ is equal to the increase of demand induced by a bias $-\Delta$. Therefore, if investor 1 has a bias Δ and investor 2 has a bias $-\Delta$, the total demand from these two investors is the same as the total demand from two unbiased investors. The same argument can be applied to the case with a large number of investors. Therefore, the biases have no impact on the aggregate demand and so do not affect the stock price. In Panel B, however, θ is a convex function of ϵ_i, and so the decrease of demand induced by a bias Δ is outweighed by the increase of demand induced by a bias $-\Delta$. Therefore, the biases increase the aggregate demand and so increase the stock price. Similar arguments can be applied to the case where θ is a concave function of ϵ_i, as illustrated in Panel C. To further elaborate the above intuition and also evaluate the impact quantitatively, I consider the following three examples.

Figure 2: The form of the biases matters. This figure plots the demand θ as a function of the bias ϵ_i. If an investor is unbiased, i.e., $\epsilon_i = 0$, his demand is a. The investor's demand is b if he has a bias of Δ, and c if he has a bias of $-\Delta$. In Panel A, the demand function is linear, which implies that $c - a = a - b$. The convexity of the demand function in Panel B implies that $c - a > a - b$, and the concavity of the demand function in Panel C implies that $c - a < a - b$.



2.1 Example 1

Suppose v_1 is lognormally distributed: $\ln v_1 \sim \mathcal{N}(\bar{v}, \sigma^2)$ and the realized stock return is $r_1 \equiv \ln \frac{v_1}{P_0}$. Assume an investor without biases allocates a fraction θ^* of his wealth to the stock market:

$$\theta^* = \frac{E[r_1] - r_f}{\sigma^2} + \frac{1}{2}, \tag{7}$$

where $E[r_1]$ denotes the expected stock return. The decision rule (7) is approximately optimal for an investor with a logarithm preference (see Campbell and Viceira (1999)).[1] If all investors follow the decision rule (7), one can easily verify from the market clearing condition that the stock price at $t = 0$ is given by

$$P_0^* = \exp\left(\bar{v} - r_f - \frac{1}{2}\sigma^2\right), \tag{8}$$

and the risk premium of the stock is

$$E[r_1^*] - r_f + \frac{1}{2}\sigma^2 = \sigma^2, \tag{9}$$

where the third term on the left hand side, $\frac{1}{2}\sigma^2$, is the standard adjustment for convexity.

Suppose investors are biased about the expected stock return

$$E^i[r_1] = E[r_1] + \phi\epsilon_i,$$

[1] Alternatively, one can specify a preference and derive the demand function endogenously. This approach leads to similar insights, while the analysis becomes more cumbersome.

that is, investor i thinks the expected return is $E[r_1] + \phi\epsilon_i$, where $\phi \geq 0$, and ϵ_i, for $i = 1, ..., N$, are independent draws from $\tilde{\epsilon}$, which is uniformly distributed between $[-1, 1]$. Hence, investors have independent biases and the most optimistic investor overestimates the expected return by ϕ while the most pessimistic investor underestimates the expected return by ϕ. As a result, investor i allocates a fraction θ_i of his wealth to the stock market:

$$\theta_i = \frac{E^i[r_1] - r_f}{\sigma^2} + \frac{1}{2}. \tag{10}$$

The following corollary summarizes the impact of the biases on the equilibrium.

Corollary 1 *If investors follow the decision rule (10), the equilibrium stock price and risk premium are not affected and investors' wealth distribution at $t = 1$ is given by*

$$W_i = \frac{v_1}{N} + \frac{\phi P_0^*}{N\sigma^2}\left(e^{r_1} - e^{r_f}\right)\epsilon_i. \tag{11}$$

Since the bias in (10) affects the demand in a linear way, the aggregate demand is not affected and so the equilibrium stock price is given by (8). Although the biases have no impact on the equilibrium at the aggregate level, they may affect each investor's wealth, as shown in (11). If an investor is unbiased, i.e., $\epsilon_i = 0$, his wealth at $t = 1$ is v_1/N. A biased investor's wealth is generally different unless the stock return happens to be the same as the bond return, i.e., $r_1 = r_f$. In particular, if the stock outperforms the bond, the optimistic investors are richer relative to pessimistic investors. Similarly, the pessimistic investors become relatively richer if the bond outperforms the stock.

2.2 Example 2

Let's now consider the case where the unbiased investors still have a decision rule (7). Investors are assumed to be biased about the standard deviation of the stock return:

$$\sigma_i = \sigma + \phi\epsilon_i,$$

where $0 \leq \phi < \sigma$, ϵ_i for $i = 1, ..., N$, are independent draws from $\tilde{\epsilon}$, which is uniformly distributed between $[-1, 1]$. That is, investor i thinks the standard deviation of the stock return is σ_i. The biases are independent across investors and the most optimistic investor underestimates the standard deviation by ϕ while the most pessimistic investor overestimates the standard deviation by ϕ. As a result, investor i allocates a fraction θ_i of his wealth to the stock market:

$$\theta_i = \frac{E[r_1] - r_f}{\sigma_i^2} + \frac{1}{2}. \tag{12}$$

The following corollary characterizes the equilibrium.

Corollary 2 *If investors follow the decision rule (12), the risk premium is given by*

$$E[r_1] - r_f + \frac{1}{2}\sigma^2 = \sigma^2 - \frac{1}{2}\phi^2. \tag{13}$$

Since the demand is a convex function of the biases (see (12)), biases increase the aggregate demand and so increase the equilibrium stock price and decrease the risk premium. Equations (9) and (13) show that biases reduce the risk premium from σ^2 to $\sigma^2 - \frac{1}{2}\phi^2$. Suppose $\sigma = 0.25$ and $\phi = 0.2$, that is, the true volatility is 25% while investors' beliefs range from 5% to 45%. Then the biases reduce the risk premium from 6.25% to 4.25%.

Merton (1980) points out that with high frequency data one can estimate volatility accurately if the true volatility is slow-moving. This argument does not nullify Example 2, which assumes investors have dispersed opinions on the standard deviation of the stock return. Rather, it helps to identify cases where Corollary 2 is relatively more important. For example, in new industries where there is not much data, it is more likely that people end up having different opinions about volatility. Moreover, people tend to have different opinions about the volatility of an industry if the uncertainty of this industry tends to change dramatically. Therefore, Corollary 2 implies that the risk premium for these industries tends to be lower than for those with a long and stable history.

2.3 Example 3

Let's now consider the case where the unbiased investors still have a decision rule (7). But investors are biased about both the expected return and the standard deviation of the stock:

$$E^i[r_1] = E[r_1] + \phi_1\left((1-\rho)\epsilon_{1i} + \rho\epsilon_{2i}\right), \tag{14}$$

$$\sigma_i = \sigma + \phi_2\epsilon_{2i}, \tag{15}$$

where $0 \leq \phi_1$, $0 \leq \phi_2 < \sigma$. For $i = 1, ..., N$, ϵ_{1i} are independent draws from $\tilde{\epsilon}_1$, and ϵ_{2i} are independent draws from $\tilde{\epsilon}_2$. $\tilde{\epsilon}_1$ and $\tilde{\epsilon}_2$ are independent and uniformly distributed between

$[-1, 1]$. ρ captures the correlation between an investor's bias on the expected return and his own bias on the volatility. That is, investor i thinks the expected stock return is $E^i[r_1]$ and the standard deviation is σ_i. These biases are independent *across investors.* But for each investor, his two biases might be correlated. In the case of $\rho > 0$, for example, if an investor overestimates the volatility, then he also tends to overestimate the expected return. The biases in (14)–(15) imply that investor i allocates a fraction θ_i of his wealth to the stock market:

$$\theta_i = \frac{E^i[r_1] - r_f}{\sigma_i^2} + \frac{1}{2}. \tag{16}$$

The following corollary characterizes the impact of the biases on the risk premium.

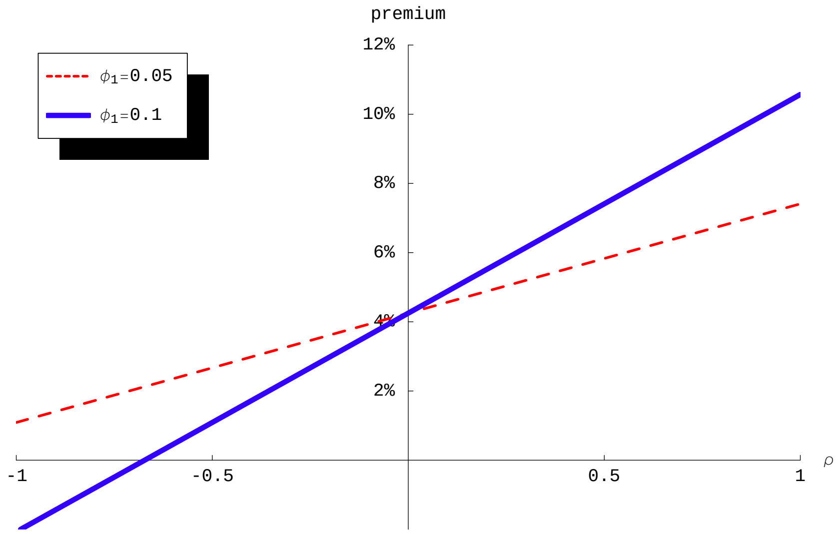
Corollary 3 *If investors follow the decision rule (16), the risk premium is given by*

$$E[r_1] - r_f + \frac{1}{2}\sigma^2 = \sigma^2 - \frac{1}{2}\phi_2^2 + \rho\phi_1 \left(\frac{\sigma}{\phi_2} + \frac{\sigma^2 - \phi_2^2}{2\phi_2^2} \log \frac{\sigma - \phi_2}{\sigma + \phi_2} \right). \tag{17}$$

The above result includes Corollaries 1–2 as two special cases: one can obtain the result in Corollary 1 by letting ϕ_2 go to 0 and obtain the result in Corollary 2 by setting $\phi_1 = 0$. If investors' biases about the expected return are independent from their biases about the volatility, i.e., $\rho = 0$, only the biases about the volatility affect the stock price and the risk premium is the same as in Corollary 2.

More interesting results arise when the biases about the expected return and the biases about the volatility are correlated. Suppose $\sigma = 0.25$, $\phi_2 = 0.2$. That is, investors' beliefs about the volatility range from 5% to 45% when the true volatility is 25%. Figure 3 plots the premium against the correlation ρ for different values of ϕ_1. It shows that the risk premium of the stock increases substantially with respect to ρ. Suppose $\phi_1 = 5\%$, that is, investors' biases about the expected return range from an overestimation of 5% to an underestimation of 5%. The stock risk premium increases from 1.1% to 7.4% when ρ increases from -1 to 1. The impact of the correlation ρ is more significant when the biases about the expected return are bigger. For example, in the case of $\phi_1 = 10\%$, when ρ increases from -1 to 1, the stock risk premium increases from -2% to 10.5%. It is interesting to note that the risk premium is negative when $\rho = -1$, and this happens without short sales constraint. In the case of $\rho = 0$, the risk premium is 4.25%, which is the same as the risk premium in the case of $\phi_1 = 0$, $\phi_2 = 0.2$. Therefore, only the biases about the volatility affect the equilibrium when these two biases are independent.

Figure 3: The risk premium and the correlation. This figure plots the risk premium of the stock on ρ, the correlation between the biases on the expected stock return and the biases on the volatility. For example, $\rho > 0$ implies that investors who overestimate the expected stock return also tend to overestimate the stock volatility. Parameter values: $\sigma = 0.25$, $\phi_2 = 0.2$.



The underlying driving force here is similar to that in Proposition 1. If $\rho < 0$, θ_i is a convex function of ϵ_{2i} and the convexity decreases the risk premium. In addition, the convexity decreases with respect to ρ and hence the risk premium increases with respect to ρ. Finally, θ_i may become concave in ϵ_{2i} when ρ is high enough and this explains why the biases increase the risk premium when ρ is high enough. For example, in the case of $\phi_1 = 5\%$, the risk premium is higher than 6.25%, the risk premium that would prevail when all investors are unbiased, if $\rho > 0.63$.

The economic intuition is also straightforward. In the case of $\rho < 0$, for example, if an investor overestimates the expected stock return he tends to underestimate the volatility. Hence this investor has a high demand for the stock. On the other hand, if an investor underestimates the expected return, he would like to short the stock. However, his short position is limited since he also tends to overestimate the volatility and so feels the short position is risky. Therefore, the biases increase the aggregate demand and lead to a lower, or even negative, risk premium. For a similar argument, if ρ is large enough the biases decrease the aggregate demand and lead to a higher expected return. Note that the negative risk premium happens here in the equilibrium without short sales constraint. The pessimistic investors choose to limit their short position because they feel it is risky rather than that it is prohibited.

2.4 Discussions of the static model

Examples 1–3 demonstrate that whether biases affect the equilibrium critically depends on the form of the biases, and that biases can have a large impact on the equilibrium even if they are independent across investors and the population is unbiased. This suggests that the traditional aggregation argument might have understated the importance of individual biases for asset pricing.

These examples also have implications on cross-sectional expected returns. If investors are only biased about the volatility, a stock's expected return tends to be lower if investors' beliefs about the volatility are more dispersed. Moreover, if investors are biased about both the expected return and volatility, the correlation between these two biases plays a key role in determining the expected return. Holding everything else constant, the higher the correlation between these two biases, the higher the expected stock return.

It is also interesting to compare these implications with the literature on short sales constraint and differences of opinion, which shows that short sales constraint makes it more difficult for pessimistic investors to express their view on the stock market, and so higher dispersion in beliefs leads to lower future returns (e.g., Miller (1977), Harrison and Kreps (1978), Scheinkman and Xiong (2003)) A large body of empirical evidence also seems to support this theory (Chen, Hong and Stein (2002), Diether, Malloy and Scherbina (2002), Jones and Lamont (2002), Lamont and Thaler (2003), Hong, Scheinkman and Xiong (2005), Mei, Scheinkman and Xiong (2005), and Nagel (2005)). The analysis in Examples 1–3 complements this theory by showing that the dispersion in opinions alone may have a large impact on the expected stock return even when short sales constraint is not stringent. Example 3 also has a new prediction that if investors' opinions are dispersed about both the expected return and the volatility, then the expected stock return is positively related with the correlation between these two biases.

In the above examples, independent biases are cancelled out by aggregation only when the biases affect the demand in a linear way, as illustrated in Example 1. The next section shows, however, that even the biases in Example 1 can significantly affect the stock price in a dynamic setting.

3 A Dynamic Model

There are two periods (three dates) $t = 0, 1, 2$. The riskless interest rate is r_f for both periods. In the first period, the stock is a claim to a positive dividend v_1 at $t = 1$. There are N investors and each is endowed with $1/N$ share of the stock. The stock price at $t = 0$, P_0, is determined in the equilibrium by equating the aggregate demand and supply. At $t = 1$, after v_1 is realized, investors can trade another stock, which is a claim to a positive dividend v_2 at $t = 2$. The stock price P_1 is also determined by equating the aggregate demand and supply. Investors' decision rule is as follows. At time t ($t = 0, 1$), investor i allocates a fraction θ_{it} of his wealth to the stock market:

$$\theta_{it} = a\left(P_t, \Psi_t\right) + \epsilon_i, \tag{18}$$

where Ψ_t ($t = 0, 1$) includes parameters such as the distribution of v_{t+1}; ϵ_i for $i = 1, ..., N$, are independent draws from $\tilde{\epsilon}$ with $E[\tilde{\epsilon}] = 0$ and $Var[\tilde{\epsilon}] = \sigma^2$. In addition, $a\left(P_t, \Psi_t\right)$ satisfies the following conditions:

$$\begin{aligned} \frac{\partial a\left(P_t, \Psi_t\right)}{\partial P_t} &< 0, & (19) \\ \lim_{P_t \to \infty} a\left(P_t, \Psi_t\right) &= -\infty, & (20) \\ \lim_{P_t \to 0} a\left(P_t, \Psi_t\right) &= \infty. & (21) \end{aligned}$$

The above condition (19) assumes that investors demand less stock if the stock price is higher. For simplicity, technical conditions (20) and (21) are imposed to ensure the existence and uniqueness of the equilibrium. I now characterize the equilibrium and postpone further discussions of the model to Section 3.2.

Let's first compute the stock prices when all investors are unbiased, i.e., $\epsilon_i = 0$ for $i = 1, ..., N$. The market clearing conditions imply that the stock price at $t = 0$, P_0^*, solves

$$a\left(P_0^*, \Psi_0\right) = 1, \tag{22}$$

and the stock price at $t = 1$, P_1^*, solves

$$v_1 a\left(P_1^*, \Psi_1\right) = P_1^*. \tag{23}$$

Let's now characterize the equilibrium when investors have biases described in (18). Note from equation (18) that investors' demand is a linear function of the bias. Hence, Proposition 1 implies that the stock price at $t = 0$ is not affected by the biases and is determined by (22). Then one can easily verify that investor i's wealth at $t = 1$ is

$$W_{i1} = \frac{v_1}{N} + \frac{1}{N} P_0^* \left(e^{r_1} - e^{r_f}\right) \epsilon_i, \tag{24}$$

where $r_1 \equiv \ln \frac{v_1}{P_0}$ is the stock return for the first period. Note that if investor i has no bias, that is $\epsilon_i = 0$, then his wealth at $t = 1$ is v_1/N. Therefore, although the biases don't affect the stock price at $t = 0$, they affect investors' wealth distribution at $t = 1$, except in the special case of $r_1 = r_f$, that is, the *ex post* stock return happens to be the same as the riskless bond return.

Equation (24) also reveals that a biased investor may go bankrupt, that is, his wealth may become negative. Intuitively, if an investor is optimistic about the stock, he might borrow in the bond market to invest in the stock market. This makes it possible that his wealth could become negative. Similarly, a pessimistic investor may short the stock and so his wealth may become negative. While the bankruptcy issue is interesting on its own right, this paper only focuses on the case where no investor goes bankrupt, that is, for $i = 1, ..., N$,

$$\frac{v_1}{N} + \frac{1}{N} P_0^* \left(e^{r_1} - e^{r_f}\right) \epsilon_i > 0. \tag{25}$$

The following proposition characterizes the stock price at $t = 1$.

Proposition 2 *Suppose investors follow the decision rule (18). Under the condition (25), the equilibrium stock price at* $t = 1$*,* P_1 *solves the following equation*

$$v_1 a\left(P_1, \Psi_1\right) + P_0^* \left(e^{r_1} - e^{r_f}\right) \sigma^2 = P_1. \tag{26}$$

Therefore, biases affect P_1 *if and only if* $r_1 \neq r_f$*:*

$$P_1 > P_1^* \quad \text{if } r_1 > r_f, \tag{27}$$

$$P_1 = P_1^* \quad \text{if } r_1 = r_f, \tag{28}$$

$$P_1 < P_1^* \quad \text{if } r_1 < r_f. \tag{29}$$

This proposition shows that even if biases affect demand in a linear way, they still affect the equilibrium if the stock return in the first period is different from the bond return. The intuition can be illustrated as follows. Suppose there are two investors A and B, and both have the same initial wealth. A is optimistic about the stock and B pessimistic. Hence A demands more stock and B less. If the biases affect the demand in a linear way as assumed in (18), they do not affect the aggregate demand and so have no impact on the stock price at $t = 0$. Suppose the first period stock return is higher than the bond return ($r_1 > r_f$). Then, A becomes richer relative to B since A chose to hold more stock at $t = 0$. As a result, the optimistic investor A has a larger wealth share and this pushes the stock price up (equation (27)). Similarly, if the first period stock return is lower than the bond return ($r_1 < r_f$), the pessimistic investor B has a larger wealth share and this leads to a lower stock price (equation (29)). In the special case where the stock return happens to be the same as the bond return ($r_1 = r_f$), the wealth share is not affected and so the stock price is the same as in the benchmark case (equation (28)). The following example further elaborates the above intuition and also evaluates the impact quantitatively.

3.1 Example 4

Let's now consider an example of the above model. Suppose v_1 and v_2 are lognormally distributed: $\ln v_1 \sim \mathcal{N}(\bar{v}, \sigma^2)$, and $\ln v_2 \sim \mathcal{N}(\ln v_1, \sigma^2)$. Investors' decision rule is as follows. At time t ($t = 0, 1$), if an investor is unbiased, he allocates a fraction θ_t^* of his wealth to the stock market:

$$\theta_t^* = \frac{E_t[r_{t+1}] - r_f}{\sigma^2} + \frac{1}{2}.$$

Investors are biased about the expected return, that is, investor i thinks the expected return is

$$E_t^i[r_{t+1}] = E_t[r_{t+1}] + \phi \epsilon_i,$$

where $\phi \geq 0$ and ϵ_i, for $i = 1, ..., N$, are independent draws from $\tilde{\epsilon}$, which is uniformly distributed between $[-1, 1]$. Therefore, investor i allocates a fraction θ_t^i ($t = 0, 1$) of his wealth to the stock market:

$$\theta_t^i = \frac{E_t^i[r_{t+1}] - r_f}{\sigma^2} + \frac{1}{2}.$$

As illustrated in Example 1, the biases have no impact on the stock price at $t = 0$, which is given by (8), and the wealth distribution at $t = 1$ is given by (11). The following discussion will focus on the case where no investor goes bankrupt at $t = 1$, that is, for $i = 1, ..., N$,

$$\frac{v_1}{N} + \frac{\phi P_0^*}{N\sigma^2} \left(e^{r_1} - e^{r_f}\right) \epsilon_i > 0. \tag{30}$$

Corollary 4 *Under condition (30), the expected stock return at* $t = 1$, $E_1[r_2]$, *is given by*

$$E_1[r_2] - r_f + \frac{1}{2}\sigma^2 - \gamma\sigma^2 e^{-E_1[r_2]} = \frac{\phi^2}{3\gamma\sigma^2}\left(e^{r_f - r_1} - 1\right), \tag{31}$$

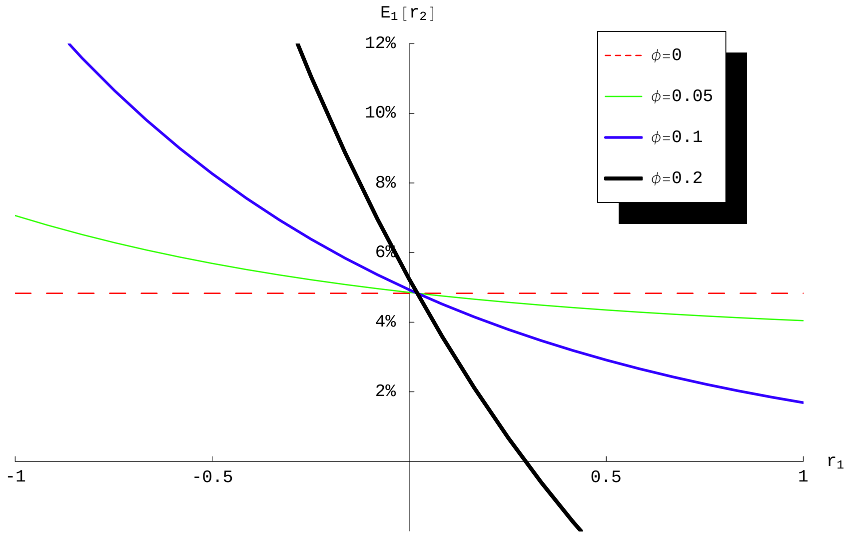
hence,

$$\frac{\partial E_1[r_2]}{\partial r_1} \begin{cases} < 0 & \text{if } \phi > 0 \\ = 0 & \text{if } \phi = 0 \end{cases}. \tag{32}$$

This corollary shows that the expected stock return depends on "past stock performance." Figure 4 plots the expected stock return in the second period $E_1[r_2]$ on the realized stock return in the first period r_1. As suggested by (32), the plot for the case of $\phi = 0$ is flat, implying that the expected stock return at $t = 1$ does not depend on the realized stock return for the previous period. For the cases with $\phi > 0$, the plots are downward-sloping, suggesting that, in the presence of biases, stock returns have a "mean reversion property": a higher stock return in the first period implies a lower expected stock return in the second period. Intuitively, the run-up in stock price in the first period means that optimistic investors have a larger wealth share $t = 1$. This pushes up the stock price and leads to a lower expected stock return. Similarly, the drop in stock price in the first period makes pessimistic investors relatively richer. This leads a lower stock price and hence a higher expected return for the second period.

Figure 4 also reveals that the biases have a large impact on the expected stock return only when investors' beliefs about the expected return are highly dispersed and the stock return in the previous period is large. This is because the underlying drive force is the fluctuation of investors' wealth share, and the fluctuation is significant only when investors have widely dispersed opinions and the stock price has a big rise or drop. In fact, Figure 4 also shows that the stock price may "overshoot" in the sense that the expected stock return in the second period may become negative. Suppose $\phi = 0.2$, that is, the most optimistic investor overestimates the expected return by 20% and the most pessimistic investor underestimates the expected return by 20%. Figure 4 shows that if the first period stock return is 30%, then

Figure 4: The expected future return and the past return. This figure plots the expected stock return in the second period $E_1[r_2]$ on the realized stock return in the first period r_1. Parameter values: $r_f = 2\%$, $\sigma = 0.25$.



expected stock return for the second period is -0.4%. Note that, in the case without biased investors, the expected return for the second period is 4.8% irrespective of the stock return in the first period.

This example provides testable cross-sectional and time series predictions on stock returns. Cross-sectionally, if investors' beliefs are highly dispersed, then stocks with high past returns tend to have low future returns. Over time series, the model predicts a negative autocorrelation for stocks with dispersed opinions. It is important to note that since these impacts arise due to the fluctuation in wealth distribution, they tend to be small except for the following two cases. First, the impact can be large at long horizons since the wealth share fluctuation can be significant at long horizons. One caution for this implication is that it also relies on the assumption that investors' biases are persistent over long horizons. Second, the impact can be significant at short horizons after a large rise or drop in stock price when investors' beliefs are widely dispersed. This case might be more relevant for new industries or technologies, on which people tend to have different assessments.

3.2 Discussions of the dynamic model

The above assumptions on the stock are non-standard: the stock at $t = 0$ is the claim to v_1, while the stock at $t = 1$ is the claim to v_2. This structure, together with the asset allocation rule (18), shuts down investors' intertemporal consideration. This allows me to illustrate the interaction between biases and the fluctuation of wealth distribution in a transparent way; the intertemporal consideration, while interesting on its own right, obscures this interaction. One consequence of these assumptions is that the stock in the first period and the stock in the second period are different stocks since they are claims to different dividends. Hence one interpretation of the previously described "mean reversion" property is that a higher past stock return implies a lower expected future return for other newly issued "similar" stocks. For example, large past returns in one industry implies that if a private firm in that industry goes public, its valuation tends to be high and the expected return from the IPO tends to be low. An alternative and more standard way to model the stock is to assume the stock is a claim to two dividends: one at $t = 1$ and one at $t = 2$. Although this assumption introduces additional intertemporal consideration, the main driving force remains the same: past performance changes investors' wealth distribution, which in turn affects the equilibrium stock price.

Another important feature of the model is that it assumes that an investor's bias at $t = 0$ is the same as his bias at $t = 1$. This assumption can be relaxed. In fact, as long as investors' biases are persistent over time, they will have an impact on the stock price. To illustrate this, let's modify the decision rule (18) as follows. Investor i allocates a fraction θ_{it} $(t = 0, 1)$ of his wealth to the stock market:

$$\theta_{it} = a\left(P_t, \Psi_t\right) + \epsilon_{it}, \tag{33}$$

where, for $i = 1, ..., N$, ϵ_{i0} are independent draws from $\tilde{\epsilon}_0$, ϵ_{i1} are independent draws from $\tilde{\epsilon}_1$ and

$$\begin{aligned} E\begin{pmatrix} \tilde{\epsilon}_0 \\ \tilde{\epsilon}_1 \end{pmatrix} &= \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \\ Var\begin{pmatrix} \tilde{\epsilon}_0 \\ \tilde{\epsilon}_1 \end{pmatrix} &= \sigma^2 \begin{pmatrix} 1 & c \\ c & 1 \end{pmatrix}. \end{aligned}$$

That is, at each period, the biases are independent across investors. For each investor, however, his biases might be correlated over time. If, for example, $c > 0$, then investors' biases are positively correlated over time. That is, if investor i is optimistic at $t = 0$, i.e., $\epsilon_{i0} > 0$, then he also tends to be optimistic at $t = 1$, i.e., ϵ_{i1} tends to be positive. The decision rule (33) includes the one in (18) as a special case of $c = 1$. Similar arguments lead to the result that the biases affect P_1 if and only if $c \neq 0$.

It is also interesting to note that even in the case of $c = 0$, biases may still affect the stock price if the decision rule in (33) is generalized to

$$\theta_{it} = a\left(P_t, W_{it}, \Psi_t\right) + \epsilon_{it}. \tag{34}$$

Note that in (33), the fraction of the wealth that an investor allocates to the stock market does not depend on his wealth level. While this property generally arises from utility maximization for standard preferences such as power utility, it certainly is violated empirically (see, e.g., Heaton and Lucas (2000)). Note that, at $t = 1$, an investor's demand for the stock is $W_{i1}a\left(P_1, W_{i1}, \Psi_1\right)$. Suppose $W_{i1}a\left(P_1, W_{i1}, \Psi_1\right)$ is a convex function of W_{i1}. This implies that the increase of demand induced by a positive shock to wealth is larger than the decrease of demand induced by a negative shock. As a result, the biases increase the aggregate demand for the stock and so increase the stock price at $t = 1$. Similarly, the biases decrease the stock price if $W_{i1}a\left(P_1, W_{i1}, \Psi_1\right)$ is a concave function of W_{i1}.

The impact of wealth share fluctuation has been studied in different contexts. For example, Dumas (1989), Wang (1996) and Chan and Kogan (2002) study the impact of wealth share fluctuation induced by the heterogeneity in risk aversion. The model in this section analyzes the impact of independent biases. The most salient distinction between my model and the previous models is that due to biased beliefs, the expected stock return may become negative, but this does not happen in models with only heterogeneity in risk aversion.

4 Conclusion

One conventional argument suggestions that if biases are independent across investors, they should not have a large impact on the equilibrium at the aggregate level since they would cancel out each other. This paper formally analyzes this argument and shows that it fails for the following two main reasons. First, if biases affect investors' demand in a non-linear way, they may have a significant impact on the equilibrium even if the biases are independent across investors and the population is unbiased. Second, even if the biases affect investors' demand linearly, the aggregation argument may still fail due to the fluctuation of wealth share in a dynamic setting.

This paper also provides various novel testable cross-sectional and time-series predictions on stock returns. Moveover, since all the analysis in this paper is conducted without the assumption of short sales constraint, it also sheds light on the potential impacts on stock returns when short sale constraint is alleviated.

References

Campbell, John, and Luis Viceira, 1999, Consumption and Portfolio Decisions When Expected Returns are Time Varying, *Quarterly Journal of Economics* 114, 433–495.

Chan, Yeung and Leonid Kogan, 2002, Catching Up with the Joneses: Heterogeneous Preferences and the Dynamics of Asset Prices, *Journal of Political Economy* 109, 1255–1285.

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

Diether, Karl, Christopher Malloy, and Anna Scherbina (2002), Differences of opinion and the cross section of stock returns, *Journal of Finance* 57, 2113-2141.

Dumas, Bernard, 1989, Two-person dynamic equilibrium in the capital market, *Review of Financial Studies* 2, 157–88.

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

Heaton, John and Deborah Lucas, 2000, Portfolio Choice and Asset Prices: The Importance of Entrepreneurial Risk, *Journal of Finance* 55, 1163–1198.

Hong, Harrison, Jose Scheinkman and Wei Xiong, 2006, Asset Float and Speculative Bubbles, *Journal of Finance* 61, 2006, 1073-1117.

Jones, Charles and Owen Lamont, 2002, Short-sale constraints and stock returns, *Journal of Financial Economics* 66, 207-239.

Lamont, Owen and Richard Thaler, 2003, Can the market add and subtract? Mispricing in tech stock carve-outs, *Journal of Political Economy*, 111, 227–268.

Liu, Jun, 2005, Portfolio Selection in Stochastic Environments, *Review of Financial Studies*, forthcoming.

Mei, Jianping, Jose Scheinkman and Wei Xiong, 2005, Speculative Trading and Stock Prices: Evidence from Chinese A-B Share Premia, working paper.

Merton, Robert, 1980, On estimating the expected return on the market: An exploratory investigation. *Journal of Financial Economics* 8, 323–362.

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

Nagel, Stefan, 2005, Short Sales, Institutional Investors, and the Cross-Section of Stock Returns, *Journal of Financial Economics*, forthcoming.

Scheinkman, Jose and Wei Xiong (2003), Overconfidence and speculative bubbles, *Journal of Political Economy* 111, 1183–1219.

Shefrin, Hersh, 2005, *A Behavioral Approach to Asset Pricing Theory*, 2005, Academic Press.

Shefrin, Hersh and Meir Statman, 1994, Behavioral Capital Asset Pricing Theory, *Journal of Financial and Quantitative Analysis*, 29, 323–349.

Shleifer, Andrei, 2000, *Inefficient Markets: An Introduction to Behavioral Finance*, Oxford University Press.

Wang, Jiang, 1996, The term structure of interest rates in a pure exchange economy with heterogeneous investors, *Journal of Financial Economics* 41, 75–110.

Xiong, Wei, and Hongjun Yan, 2006, Heterogeneous Beliefs and Bond Yields, working paper.

Yan, Hongjun, 2005, On the Market Selection Hypothesis, working paper.

Appendix

Proof of Proposition 1

Equating the aggregate demand and supply for the stock, we obtain

$$\sum_{i=1}^{N} \frac{P_0}{N}\theta_i = P_0.$$

This implies

$$\frac{1}{N}\sum_{i=1}^{N} \theta_i = 1.$$

When N is large, the above expression becomes

$$E\left[\theta\left(P_0, W_i, \Psi, \tilde{\epsilon}\right)\right] = 1. \tag{35}$$

If all investors are unbiased, the stock price solves

$$\theta\left(P_0, W_i, \Psi, 0\right) = 1. \tag{36}$$

Jensen's inequality implies that

$$E\left[\theta\left(P_0, W_i, \Psi, \tilde{\epsilon}\right)\right] = \theta\left(P_0, W_i, \Psi, 0\right) \text{ if } \theta \text{ is linear in } \tilde{\epsilon}, \tag{37}$$

$$E\left[\theta\left(P_0, W_i, \Psi, \tilde{\epsilon}\right)\right] > \theta\left(P_0, W_i, \Psi, 0\right) \text{ if } \theta \text{ is convex in } \tilde{\epsilon}, \tag{38}$$

$$E\left[\theta\left(P_0, W_i, \Psi, \tilde{\epsilon}\right)\right] < \theta\left(P_0, W_i, \Psi, 0\right) \text{ if } \theta \text{ is concave in } \tilde{\epsilon}. \tag{39}$$

Equations (35)–(39) lead to the results in Proposition 1.

Proof of Corollary 1

Substituting (7) into (36), after some algebra, we obtain (8)–(9). At $t = 1$, the stock's liquidation value is v_1. So the allocation rule (10) leads to the result in (11).

Proof of Corollary 2

Substituting (12) into (36), we obtain

$$\int_{-1}^{1} \left(\frac{E[r_1] - r_f}{(\sigma + \phi x)^2} + \frac{1}{2} \right) \frac{1}{2} dx = 1.$$

After some algebra, we obtain (13).

Proof of Corollary 3

Substituting (16) into (36), we obtain

$$\int_{-1}^{1}\int_{-1}^{1}\left(\frac{E[r_1]+\phi_1\left((1-\rho)y+\rho x\right)-r_f}{(\sigma+\phi x)^2}+\frac{1}{2}\right)\frac{1}{4}dxdy=1.$$

After some algebra, we obtain (13).

Proof of Proposition 2

Under the condition (25), the market clearing condition at $t=1$ is

$$\sum_{i=1}^{N} W_{i1}\theta_{i1}=P_1.$$

When N is large the above equation leads to (26). Comparing (23) and (26) leads to (27)–(29).

Proof of Corollary 4

Equation (31) is a special case of (26). Differentiating (31) leads to (32).