Working Paper No. 202

Are stock markets really like beauty contests?

Empirical evidence of higher order belief's impact on asset prices

Pierre Monnin

August 2004

Are stock markets really like beauty contests?

Empirical evidence of higher order belief's impact on asset prices

Pierre Monnin[1]

Swiss National Bank and University of Zurich

August 2004

Abstract

The goal of this paper is to assess, for the first time, the empirical impact of
"Keynes' beauty contest", or "higher order beliefs", on asset price volatility. The paper shows that heterogeneous expectations induce higher order beliefs and that heterogeneous expectation asset pricing models theoretically generate more volatility than rational expectation models. The paper also explains how, with some assumptions on the distribution of public and private information, a model with higher order beliefs can be empirically estimated. The model is then applied to annual data of the American stock market. The results show that a model with higher order beliefs generates a level of volatility in line with the price volatility observed on the market.

Keywords: Asset pricing; Excess volatility; Higher order beliefs.

JEL Classification: D84, G12, G14.

1 Introduction

In the past 25 years, financial economists have spent a lot of time and attention in assessing the empirical validity of rational expectation asset pricing models. The common hypothesis of these models is that the stock prices should reflect the present value of rationally expected future payoffs. Although the rational expectation hypothesis constitutes the basis of most contemporary asset pricing models, its empirical support is rather weak. The first major critics against such models came from Shiller (1981) and LeRoy and Porter (1981), who argued that the price volatility observed on the market is too large to be justified by rational expectation models. These two papers are at the origin of numerous other articles, which have tried to explain this excess volatility puzzle in the framework of rational expectations.

After some early technical improvements (see, e.g., Campbell and Schiller 1987 and 1988), research has concentrated on asset pricing models with stochastic discount factors. The idea is to build an asset pricing model with additional economic variables, which generate a time-varying discount factor, and check if the additional variability brought by these new factors can match the excess volatility.[1] However, neither of these factors seems to explain all of the excess volatility (Shiller 2003). Furthermore, recent studies, which do not use the classical volatility test, have also found evidence against rational expectations models (Zhong, Darrat and Anderson 2003).

Confronted with the apparent empirical failure of rational expectation models, many researchers have argued that some "non-fundamental" factors may be at the origin of price movements. Behavioral finance deals specifically with such market "irregularities" and has put forward different possible explanations to the excess volatility puzzle.[2] In particular, Campbell and Cochrane (1999) propose a model with habit formation, which theoreti-

---

[1] Many factors have been proposed as possible source of excess volatility. The most common are consumption and expected inflation (see Wickens 2003 for a recent study with these two variables). Other factors are, for example, tax rate changes, production volatility changes or transaction costs changes.

[2] See Barberis and Thaler (2002) or Shiller (2003) for a survey of the answer that behavioural finance gives to excess volatility in stock markets.

cally generates excess volatility through changes in risk aversion. Barberis, Huang and Santos (2001) introduce loss aversion to explain the puzzle. After calibration, both models replicate several distinctive features of the stock markets, such has, in particular, the observed volatility.

This paper explores another potential "non-fundamental" factor, namely the impact of "Keynes' beauty contest" on asset prices. The name of this effect comes from Keynes' famous metaphor, in which he suggests that, in order to form their demand for an asset, investors not only forecast the future payoffs but also try to guess other market participants' forecasts and others' forecasts of others' forecasts, etc (Keynes 1936). In this situation, investors are said to have "higher order beliefs". Townsend (1983), in a general framework, and Basak (2000), in the context of stock markets, theoretically show that higher order beliefs induce higher price volatility than rational expectations do.[3] This additional volatility is caused by the fact that investors react to variations generated by decisions of others and to the noise in such decisions. This phenomenon is called endogenous uncertainty by Kurz (1974). Bacchetta and van Wincoop (2004) also show that higher order beliefs can induce a disconnection between the price and its fundamental value.

Even if the role of higher order beliefs in explaining the excess volatility puzzle is theoretically acknowledged, no empirical estimation of their effect has been made yet. The aim this paper is to fill this gap and to quantify the empirical impact of Keynes' beauty contest on stock price volatility. To pursue this empirical goal, we contribute to the theoretical literature by elaborating a Heterogeneous Expectation Asset Pricing Model (HEAPM) with constant relative risk aversion, time-varying discount rate and a time-varying risk premium (cf. Section 2.1).[4] We explain how the HEAPM generates additional price volatility (cf Section 2.2). We then present a

---

[3]Miller (1977), Harrison and Kreps (1978), Morris (1996) or Wu and Guo (2004) also show that higher order beliefs generate excess volatility, but they need to add the hypothesis of limited short-selling to get this feature.

[4]Bacchetta and van Wincoop (2004) have independently derived a similar model but with constant absolute risk aversion, constant discount rate and a constant risk premium. Biais, Bossaerts and Spatt (2003) use a set-up similar to Bacchetta and van Wincoop (2004) for a one-period asset pricing model.

framework based on public and private information about future dividend, which allows the empirical estimation of the HEAPM (cf. Section 3). We show, in this context, how the price can be used as a public signal by market participants. We then estimate the HEAPM with American stock market data for a period between 1871 and 2003. The main result of this paper is that *the volatility implied by the HEAPM seems to correspond to the one observed on the market* (cf. Sections 6 and 7). In that sense, higher order beliefs might be a plausible explanation to the excess volatility puzzle. We finally present some additional empirical results on how the investors incorporate public and private information into their expectations.

2   An heterogeneous expectation asset pricing model

2.1   The model

As suggested by Keynes (1936), when agents have heterogeneous expectations about asset's future payoffs, the demand for asset of each agent will not only reflect his own expectations, but also his beliefs about other agents' expectations. This can be formally shown in a simple asset pricing model. Consider an overlapping generation economy where, at time t, a new generation of investors, indexed on the unit interval $[0; 1]$, enters the market.[5] At the beginning of period t, each investor chooses a portfolio, which maximizes the expected utility of her future wealth (W_{t+1}). In the second period (period $t + 1$), she sells this portfolio to the next generation of investors.[6] Each investor can either invest in a risky asset or in a risk free bond. We

---

[5]The choice of an overlapping generation model with very short-lived agents is an important characteristic of this model. This hypothesis has been made for two reasons: firstly, it significantly simplifies the solution for the price equation, and secondly, as shown by Allen, Morris and Shin (2003), the effect of higher order beliefs on the price is bigger when investors have a short investment horizon. Thus, short-lived agents allow us highlighting the effects of higher order beliefs on the price. Note that this assumption finds its justification in Section 3 and does not have any implications for the results of this section.

[6]Two hypotheses are implicit in this model: 1) the investors live only two periods and 2) they consume only in the second period. Even though these hypotheses are clearly restrictive, they are not implausible for an economy where markets are mainly driven by traders who are regularly assessed on the basis of their portfolio wealth.

3

assume that investors have power utility $U(W_{t+1}) = (W_{t+1}^{1-\gamma_i} - 1)/(1 - \gamma_i)$, where γ_i is the constant relative risk aversion of investor i.[7] We also assume that, at the end of period t, each investor is replaced by a new investor with the same relative risk aversion. Finally, we assume that asset returns are log normal.

The solution to the individual maximization problem described above is (see e.g. Campbell and Viceira 2002, p. 29):

$$\alpha_{i,t} = \frac{E_t^i[r_{t+1}] - r_{t+1}^f + \frac{1}{2}\sigma_{i,r_{t+1}}^2}{\gamma_i \sigma_{i,r_{t+1}}^2} \tag{1}$$

where $\alpha_{i,t}$ is the fraction of wealth that investor i puts into the risky asset at time t, $E_t^i[r_{t+1}]$ is the log return of the risky asset on the next period expected by investor i at time t, r_{t+1}^f is the risk free log return on the next period and $\sigma_{i,r_{t+1}}^2$ is the asset return volatility at time $t+1$ expected by investor i at time t. Note that this model allows for heterogeneous expectations about return and volatility among the investors.

Assuming that the relative risk aversion coefficient, the expected return and the expected volatility are jointly independent for each investor, the aggregating of Equation (1) over all the agent gives:

$$\alpha_t = \frac{1}{\gamma}\left(\frac{1}{\nu_{t+1}}\bar{E}_t\left[r_{t+1} - r_{t+1}^f\right] + \frac{1}{2}\right) \tag{2}$$

where $\bar{E}_t = \int_0^1 E_t^i di$ is the *average* expectation over all the investors, α_t is the average fraction of wealth invested in the risky asset, $\gamma = \left(\int_0^1 \frac{1}{\gamma_i} di\right)^{-1}$ is the inverse of the *average of the inverse* of the relative risk aversion coefficient and $\nu_{t+1} = \left(\int_0^1 \frac{1}{\sigma_{i,r_{t+1}}^2} di\right)^{-1}$ the inverse of the *average of the inverse* of the asset return expected volatility. For simplicity, we will called γ the average relative risk aversion and ν_{t+1} the average asset return expected volatility.

From Equation (2), we can recover the asset price by using the following first order Taylor approximation of the asset log return (see e.g. Campbell and Shiller 1988):

---

[7]The advantages of using power utility rather than exponential utility or quadratic utility are discussed by Cambpell and Viceira (2002), p. 24.

$$r_{t+1} = \log \frac{P_{t+1} + D_{t+1}}{P_t} = b + \rho p_{t+1} + (1 - \rho) d_{t+1} - p_t$$

where P_t is the asset price at the beginning of period t (and p_t its logarithm), D_{t+1} is the dividend distributed at the end of period t (and d_{t+1} its logarithm), and b and ρ are parameters. Plugging this approximation into Equation (2) yields:

$$p_t = \bar{E}_t \left[\rho p_{t+1} + (1 - \rho) d_{t+1} - r_{t+1}^f \right] - \left(\gamma \alpha_t - \frac{1}{2} \right) \nu_{t+1} + b \qquad (3)$$

The first part of the right hand side of Equation (3) tells us that the price is a function of the *average* expected asset return. The second part can be understood as a time-varying risk premium, which increases with the average expected asset return volatility, the average fraction invested in the risky asset and the average relative risk aversion.

As the price and the dividend are usually non stationary variables, we can rewrite Equation (3) in terms of the price-to-dividend (P/D) ratio $\delta_t = p_t - d_t$:[8]

$$\delta_t = \bar{E}_t \left[\rho \delta_{t+1} + z_{t+1} \right] - \beta_t \nu_{t+1} + b \qquad (4)$$

where $z_{t+1} = \Delta d_{t+1} - r_{t+1}^f$ is the "adjusted dividend growth rate" and $\beta_t = \gamma \alpha_t - \frac{1}{2}$. By solving this equation forward, we get the final equation:

$$\delta_t = \sum_{k=0}^{\infty} \rho^k \bar{E}_t^{k+1} \left[z_{t+k+1} - \beta_{t+k} \nu_{t+k+1} \right] + c_1 \qquad (5)$$

where $c_1 = \frac{b}{1-\rho}$ and $\bar{E}_t^k [x_{t+k}] = \bar{E}_t \left[\bar{E}_{t+1} \left[\ldots \bar{E}_{t+k} [x_{t+k}] \right] \right]$ is the *average expectation of order k of the variable x_{t+k}*, which is the average expectation at time t of the average expectation at time $t+1$ of the average expectation at time $t+2$, etc.. of the variable x_{t+k}. Equation (5) is called the Heterogeneous Expectation Asset Price Model (HEAPM).

Equation (5) is the formal counterpart of Keynes' beauty contest metaphor. Indeed, the investor i will have the following expectation for the future P/D ratio:

---

[8]The P/D ratio is a stationary variable in the sample used for this paper (cf. Section 5.2)

$$E_t^i[\delta_{t+1}] = \sum_{k=0}^{\infty} \rho^k E_t^i \left[\bar{E}_{t+1}^{k+1} \left[z_{t+k+2} - \beta_{t+k+1} \nu_{t+k+2} \right] \right] + c_1 \qquad (6)$$

This equation shows that, in order to forecast the P/D ratio, investor i will use her expectation of the *average* of the other agents' expectations. In particular, in this model, the investor must guess the *average* expectation of the future adjusted dividend growth, of the future risk free return, of the future asset return volatility and of the future fraction invested in the risky asset. Thus, forecasting the price implies forecasting others' forecasts. This shows that, on a market with heterogeneous expectations, higher order beliefs do theoretically play a role in the explanation of the price.

2.2 HEAPM as a possible solution to the volatility puzzle

One consequence of the HEAPM is that it induces a higher price volatility than the traditional rational expectation model. To see that, let us first define the rational expectation P/D ratio as:

$$\delta_t^* = E_t \left[\rho \delta_{t+1}^* + z_{t+1} \right] - \beta_t \sigma_{r_{t+1}}^2 + b \qquad (7)$$

where E_t is the traditional rational expectation operator and $\sigma_{r_{t+1}}^2$ is the rational expectation of the asset return volatility. This equation is the equivalent of Equation (4) when the investors have an identical expectation and when this expectation is rational. We can define f_t^i and g_t^i such that:

$$E_t^i \left[\rho \delta_{t+1} + z_{t+1} \right] = E_t \left[\rho \delta_{t+1}^* + z_{t+1} \right] + f_t^i$$
$$\sigma_{i,r_{t+1}}^2 = \sigma_{r_{t+1}}^2 \left(1 + g_t^i \right)$$

where f_t^i and g_t^i represents the deviation of individual i's expectation from the rational expectation (in terms of volatility for g_t^i). Aggregating the previous equations over all individuals yields:

$$\bar{E}_t \left[\rho \delta_{t+1} + z_{t+1} \right] = E_t \left[\rho \delta_{t+1}^* + z_{t+1} \right] + \bar{f}_t$$
$$\nu_{t+1} = \sigma_{r_{t+1}}^2 \left(1 + \bar{g}_t \right)$$

where $\bar{f}_t$ and $\bar{g}_t$ correspond to the average deviation from the rational expectation. Combining Equation (4), (7) and the two previous equations gives:

$$\delta_t = \delta_t^* + \bar{f}_t - \beta_t \sigma_{r_{t+1}}^2 \bar{g}_t \tag{8}$$

If we take the unconditional variance of this expression, we have:

$$V(\delta_t) = V(\delta_t^*) + V\left(\bar{f}_t - \beta_t \sigma_{r_{t+1}}^2 \bar{g}_t\right) + 2Cov\left(\delta_t^*, \bar{f}_t - \beta_t \sigma_{r_{t+1}}^2 \bar{g}_t\right)$$

As the rational P/D ratio δ_t^* is independent from the average deviations from the rational expectation $\bar{f}_t$ and $\bar{g}_t$, the covariance term is equal to zero and we have:

$$V(\delta_t) = V(\delta_t^*) + V\left(\bar{f}_t - \beta_t \sigma_{r_{t+1}}^2 \bar{g}_t\right) \geq V(\delta_t^*)$$

Thus, if the average expectation of asset returns or its average expected volatility diverge from the rational expectation and if this difference varies in time, the HEAPM implies a volatility, which is higher than with the rational expectation model. This could constitute a theoretical explanation to the excess volatility puzzle. The question is now to determine if this explanation is empirically relevant.

3 The HEAPM with private and public information

Unfortunately, the validity of the HEAPM in Equation (5) cannot be checked directly with empirical data. One way to find a testable version of this model is to impose further assumptions about the informational structure of the economy (Biais and Bossaerts 1998). This paper adopts a structure of information inspired by Allen, Morris and Shin (2003), which mixed a private heterogeneous signal and a public common signal (cf. Section 3.1). Given this structure, it is possible to deduct agent i's expectations and his beliefs about others' expectations (cf. Section 3.2) and thus, to compute an equilibrium equation for the P/D ratio (cf. Section 3.3). As, in equilibrium,

the past P/D ratios contain some information about the future adjusted dividend growth, the investor can use them as the public signal (cf. Section 3.4). Our final equilibrium equation will therefore be a mix of heterogeneous non observable private signals about future adjusted dividend growth and a public observable signal based on the past P/D ratios.

3.1 The information structure

Let us assume that each investor has two sources of information about the future adjusted dividend growth z_{t+k}: a public signal and a private signal. The public signal z_{t+k}^* is the *best forecast of z_{t+k} given by the past P/D ratios* available at time t. Expressed in a different way: $z_{t+k}^* = E[z_{t+k}|\Omega_t]$, where $\Omega_t = \{\delta_{t-1}, \delta_{t-2}, ...\}$ is the public information set. Note that the current ratio δ_t is not included in the public information set. This corresponds to a market with the following sequence of decisions: 1) the agents form their expectations, 2) they place their orders on the market according to their expectations and 3) the price P_t is set in order to clear the market. Therefore, the agents have to form their expectation before knowing the price P_t and, thus, before knowing δ_t.

In addition to the public forecast, each investor i observes an unbiased private signal x_{t+k}^i on the future values of z_{t+k}. This signal is based on his private information set Θ_t^i only, which does not contains any public information ($\Omega_t \cap \Theta_t^i = \oslash$). Thus the private signal is $x_{t+k}^i = E[z_{t+k}|\Theta_t^i]$. We assume that the average signal over all the agent is unbiased, thus $\int_0^1 x_{t+k}^i di = z_{t+k}$. We finally assume that both signals are normally distributed and that the relative precision of the private signal to the public signal is the same for each investor and is constant in time.

To be complete, we have to specify the information about the average fraction of wealth invested in the asset (α_t) available to each investors. We make the hypothesis that this variable is equal to $\alpha_t = \alpha + \epsilon_t$, where ϵ_t are i.i.d. and α is not directly observable. Therefore, the expectation about α_t is constant in time and the same for everyone, which implies in particular that $\beta_{t+k} = \beta$.

3.2 Forecasting the others' forecasts

Equation (5) can be split in two parts: firstly, the iterated average expectation of z_{t+k+1} and, secondly, the iterated average expectation of the average variance ν_{t+k+1}. With the informational framework describes in the previous section, it is possible to compute a solution for these two terms. The next two sections present the solution for the iterated average expectation of adjusted dividend growth and for the iterated expectations of the average variance, respectively.

3.2.1 Forecasting the iterated average expectation

Allen *et al.* (2003) have shown that the traditional law of iterated expectations $E_t\left[E_{t+1}\left[\ldots E_{t+k}\left[z_{t+k}\right]\right]\right] = E_t\left[z_{t+k}\right]$ does not hold for the iterated *average* expectations with heterogeneous private information and a public signal. However, when the weight given to each signal is constant, an alternative law of iterated average expectations exists. Allen, Morris and Shin's basic idea is the following: consider investor i, who tries, at time t, to forecast the adjusted dividend growth rate at time $t+1$ using the public signal and her private signal. Her expectation will be a weighted average of the two signals:

$$E_t^i\left[z_{t+1}\right] = (1-\lambda)\, z_{t+1}^* + \lambda x_{t+1}^i \tag{9}$$

where λ is the relative weight given to the private signal. This relative weight reflects the relative precision that the agents associate to each signal.[9] No assumption is made on how the agents assess this relative precision. They can, for example, give a subjective weight to each signal or use their objective precisions. Taking the average of Equation (9) over all agents yields:

$$\bar{E}_t\left[z_{t+1}\right] = (1-\lambda)\, z_{t+1}^* + \lambda z_{t+1}$$

---

[9]The relative precision is, by assumption, identical for every investors and constant in time (cf. section 3.1). The latter assumption will be relaxed in Section 7. Note that the precision of each signal can vary in time; only their *relative* precision is assumed to be constant.

Now consider the case where the investor is still situated at time t, but wants to forecast the average expectation at time $t + 1$ of the adjusted dividend growth rate at time $t + 2$. Using the previous result for the average expectation of z_{t+2} and taking its expectation for investor i yields:

$$E_t^i \left[\bar{E}_{t+1} \left[z_{t+2} \right] \right] = (1 - \lambda) z_{t+2}^* + \lambda E_t^i \left[z_{t+2} \right]$$

Plugging equation (9) for z_{t+2} into the previous equation gives:

$$E_t^i \left[\bar{E}_{t+1} \left[z_{t+2} \right] \right] = \left(1 - \lambda^2 \right) z_{t+2}^* + \lambda^2 x_{t+2}^i$$

More generally:

$$E_t^i \left[\bar{E}_{t+1}^k \left[z_{t+k} \right] \right] = \left(1 - \lambda^k \right) z_{t+k}^* + \lambda^k x_{t+k}^i \qquad (10)$$

Thus, investor i's expectation of the average expectation is a weighted average of the public and the private signal where the weight of the public signal increases with the forecast horizon. Equation (10) is equivalent to:

$$E_t^i \left[\bar{E}_{t+1}^k \left[z_{t+k} \right] \right] = z_{t+k}^* + \lambda \left(x_{t+k}^i - z_{t+k}^* \right) - \lambda \left(1 - \lambda^{k-1} \right) \left(x_{t+k}^i - z_{t+k}^* \right) \quad (11)$$

The first term of the sum in the right hand sight of Equation (11) is the future adjusted dividend growth rate's expectation given the public information. Investor i adjusts this forecast with the second term of the sum to take into account her private information. These first two terms represent the agent's expectation about the future adjusted dividend growth rate given the public and the private information. In addition, the investor makes a last adjustment to her expectation since she tries to guess the *average expectation* of future adjusted dividend growth rate and not its *true value*. Therefore, the last term of the sum in Equation (11) reflects the "beauty contest" effect on investor i's expectations. Note that the coefficient of this last term is negative, which implies that the weight of the public signal is bigger in the final expectation than it would be if the investor had to guess the true future value of the dividend. This reflects the fact that each agent knows that the other agents also observe the public signal and that everybody uses it in their forecast. Therefore, as the public signal enters into

every individual expectations, it is a better predictor of the average opinion than the private signal. Note also that the weight of the public signal becomes bigger with the forecast horizon. This is due to the fact that with a longer horizon, the number of average expectations' layers is higher and therefore, the resemblance between the average expectation and the best forecast of the dividend decreases.

The aggregation of Equation (10) over all the agents yields:

$$\bar{E}_t^k [z_{t+k}] = \left(1 - \lambda^k\right) z_{t+k}^* + \lambda^k z_{t+k} \tag{12}$$

3.2.2 Forecasting the variance

In the particular framework of Section 3.1, the problem of the iterated average expectation of the "average variance" ($\bar{E}_t^{k+1} [\nu_{t+k+1}]$) reduces to the traditional iterated expectation solution. The intuition behind this result is the following: since we assumed that the precision of the distribution of the signal was the same for everybody, then the expected variance of z_{t+k} is also the same for each investor. Then, if everyone has the same expected variance, its average expectation is known to everybody and is equal to the traditional iterated expectation.

More formally, the proof is the following: recall first that $\sigma_{i,r_{t+j}}^2$ is the expected variance of r_{t+j} for investor i at time $t + j - 1$. By using the same first-order Taylor approximation as in Section 2.1 yields:

$$\sigma_{i,r_{t+j}}^2 = \rho^2 \sigma_{i,\delta_{t+j}}^2 + \sigma_{i,z_{t+j}}^2 \tag{13}$$

The variance of z_{t+j} can be inferred from the variances of the two signals. As, by assumption, both signals are normally distributed with the same variances for everybody, the inferred variance of z_{t+j} is the same for each agent.[10] We therefore have:

$$\sigma_{i,r_{t+j}}^2 = \rho^2 \sigma_{i,\delta_{t+j}}^2 + \sigma_{z_{t+j}}^2$$

The last step is to compute the volatility of the P/D ratio expected by each investor ($\sigma_{i,\delta_{t+j}}^2$). For this, we assume that each investor believes

---

[10]See e.g. Hogg and Craig (1995) p.149.

that the P/D ratio volatility is a function of the future adjusted dividend growth volatility. Under this assumption, the P/D ratio volatility is indeed a function of the future adjusted dividend growth volatility,[11] which makes the investor's ex-ante beliefs consistent with their ex-post observation. Finally, if the P/D ratio volatility is a function of the adjusted dividend growth volatility, then as the latest is the same for everybody, the former is also identical for every investor. Thus:

$$\nu_{t+k+j} = \rho^2 \sigma^2_{\delta_{t+k+j}} + \sigma^2_{z_{t+k+j}} \tag{14}$$

and

$$\bar{E}_t^{k+1}\left[\nu_{t+k+j}\right] = E_t\left[\nu_{t+k+j}\right] = \nu^*_{t+k+j} \tag{15}$$

3.3   HEAPM with public and private information

Plugging the result about the iterated average expectation of the adjust dividend growth rate in Equation (12) and of the iterated expected return volatility in Equation (15) into Equation (5) yields the P/D ratio equation:

$$\delta_t = \sum_{k=0}^{\infty} \rho^k \left(\left(1 - \lambda^{k+1}\right) z^*_{t+k+1} + \lambda^{k+1} z_{t+k+1} - \beta \nu^*_{t+k+1} \right) + c_1 \tag{16}$$

Note that if the agents give all the weight to the private signal ($\lambda = 1$), the P/D ratio is a function of the discounted sum of future dividends. This is equivalent to a model with perfect foresight. This result is due to the fact that, if everybody follows her private signal only, the individual errors will be cancelled out by the aggregation among investors.

Furthermore, as explained in Section 3.2.1, if the investors were trying to guess the true value of the adjusted dividend growth and not the average expectations, or in other words, if they were not taking into account the beauty contest effect, the price would be equal to:

$$\delta_t = \sum_{k=0}^{\infty} \rho^k \left(\left(1 - \lambda\right) z^*_{t+k+1} + \lambda z_{t+k+1} - \beta \nu^*_{t+k+1} \right) + c_1 \tag{17}$$

---

[11]See proof in Appendix A.

The "pure" beauty contest effect on the price can be isolated by subtracting Equation (17) from Equation (16).

$$BCE_t = \lambda \sum_{k=0}^{\infty} \rho^k \left(1 - \lambda^k\right) \left(z_{t+k+1}^* - z_{t+k+1}\right) \tag{18}$$

Note that the beauty contest effect is equal to zero when $\lambda = 1$ and when $\lambda = 0$. This result is not surprising since in the first case, each investor relies on her private signal only and does not try to guess the average expectation by using the public signal. In the second case, everybody uses the public signal only, thus, everybody has the same expectation and we end up in the traditional case of homogeneous rational expectation.

3.4   How to extract the public signal from the price

If the P/D ratio is driven by Equation (16), then it value partly reflects the true value of future adjusted dividend growth (z_{t+k}). Therefore, its value contains some public information about the future adjusted dividend growth. The next two sections show how to extract this public information from the P/D ratio.

3.4.1   Public signal on the adjusted dividend growth

Let us first define the new variable y_t as:

$$y_t = \delta_t - \sum_{k=0}^{\infty} \rho^k \left(\left(1 - \lambda^{k+1}\right) z_{t+k+1}^* - \beta \nu_{t+k+1}^* \right) + c_1$$

Note that at time t, the variable y_t is known since it is constituted of the current P/D ratio and the best forecasts of the adjusted dividend growth rate and of the volatility of the asset return given the public information. Equation (16) can be rewritten with this new variable:

$$y_t = \lambda \sum_{k=0}^{\infty} (\rho\lambda)^k z_{t+k} \tag{19}$$

From this last equation, it can be deduced that:

$$z_{t+k} = \frac{1}{\lambda} y_{t+k-1} - \rho y_{t+k} \tag{20}$$

13

At time t, none of the variables in the last equation are know, but if is possible to forecast them using the public information. This yields:

$$E_t\left[z_{t+k}\left|\Omega_t\right.\right] = \frac{1}{\lambda}E_t\left[y_{t+k-1}\left|\Omega_t\right.\right] - \rho E_t\left[y_{t+k}\left|\Omega_t\right.\right] \tag{21}$$

The left hand side of this equation is the expected adjusted dividend growth rate given the public information, which precisely corresponds to the definition of the variable z_{t+k}^*.[12] After replacing y_t by its definition, it is possible to solve the equation for the variable z_{t+k}^*. This yields:

$$z_{t+k}^* = \delta_{t+k-1}^* - \rho\delta_{t+k}^* + \beta\nu_{t+k}^* + b \tag{22}$$

where $\delta_{t+k}^* = E_t\left[\delta_{t+k}\left|\Omega_t\right.\right]$ is the best forecast of the P/D ratio at time $t+k$, given the public information available at time t and ν_{t+k}^* is the best forecast of the asset volatility at time $t+k$, given the public information available at time t.

3.4.2 Public signal on the asset return volatility

Similarly to the public signal on the adjusted dividend growth rate, it is possible to compute the best forecast of the asset return volatility given the public information (ν_{t+k}^*). To do so, we combine Equation (14) and (15):

$$\nu_{t+k}^* = E_t\left[\nu_{t+k}\right] = \rho^2 E_t\left[\sigma_{\delta_{t+k}}^2\right] + E_t\left[\sigma_{z_{t+k}}^2\right] \tag{23}$$

We have that $\sigma_{z_{t+k}}^2 = Var_{t+k-1}\left[z_{t+k}\right]$. Using Equation (20) again yields:

$$Var_{t+k-1}\left[z_{t+k}\right] = \rho^2 Var_{t+k-1}\left[y_{t+k}\right]$$

Combining the previous equation, Equation (19) and Equation (26) gives:

$$Var_{t+k-1}\left[z_{t+k}\right] = \rho^2\sigma_{\delta_{t+k}}^2$$

Taking the expectation at time t reduces it to:

$$E_t\left[\sigma_{z_{t+k}}^2\right] = Var_t\left[z_{t+k}\right] = \rho^2 Var_t\left[\delta_{t+k}\right]$$

---

[12] See Section 3.1.

14

Finally, by plugging this last equation into Equation (23), we get:

$$\nu^*_{t+k} = 2\rho^2 Var_t\left[\delta_{t+k}\right] \qquad (24)$$

4 How to test the model

One way to test the validity of the HEAPM is to compute the theoretical P/D ratio given by Equation (16) and then to check if its dynamic has the same empirical characteristics as those observed on the market (cf. Section 4.2). But before doing that, we have to estimate the empirical value of the public signals z^*_{t+j} and ν^*_{t+j}, which are necessary to compute the theoretical P/D ratio.

4.1 Estimation of the public signals

As Equation (24) shows, estimating the public signal ν^*_{t+k} is equivalent to forecasting the *future volatility of the P/D ratio* given the public information available at time t (which is constituted of the past P/D ratios). For the public signal z^*_{t+k}, Equation (22) tell us that we have to forecasts the *future value of P/D ratio* given the information available at time t. It is possible forecasts the P/D ratio and volatility given its past values by using an autoregressive equation of order p with conditional heteroskedasticity of order q (AR(p)-ARCH(q) model). Then, using this estimated model, it is possible to compute the best public forecast for δ_{t+j} and ν_{t+j}. Concretely, the AR(p)-ARCH(q) model takes the following form:

$$\begin{aligned}
\delta_t &= \vartheta + \phi_1\delta_{t-1} + \ldots + \phi_p\delta_{t-p} + u_t \\
u_t^2 &= \zeta + \alpha_1 u_{t-1}^2 + \ldots + \alpha_q u_{t-q}^2 + w_t
\end{aligned}$$

where w_t is a white noise and $h_t = Var_t\left[\delta_t\right]$. This system can be rewritten in vectors and matrix terms to facilitate the forecasts.[13] It takes the following form:

---

[13]See Hamilton (1994), p. 7, to see how to rewrite a nth-order difference equation in vector and matrix terms.

$$\boldsymbol{\delta}_t = \boldsymbol{\vartheta} + \boldsymbol{\phi}\boldsymbol{\delta}_{t-1} + \mathbf{u}_t$$
$$\mathbf{u}_t^2 = \boldsymbol{\zeta} + \boldsymbol{\alpha}\mathbf{u}_{t-1}^2 + \mathbf{w}_t$$

Given this process, we can compute the best forecasts, which are equal to:

$$\boldsymbol{\delta}_{t+i}^* = \boldsymbol{\phi}^{i+1}\boldsymbol{\delta}_{t-1} + \left(1 - \boldsymbol{\phi}^{i+1}\right)\left(1 - \boldsymbol{\phi}\right)^{-1}\boldsymbol{\vartheta}$$
$$\mathbf{u}_{t+i}^{2*} = \boldsymbol{\alpha}^{i+1}\mathbf{u}_{t-1}^2 + \left(1 - \boldsymbol{\alpha}^{i+1}\right)\left(1 - \boldsymbol{\alpha}\right)^{-1}\boldsymbol{\zeta}$$

Thus, once the parameters of the AR(p)-ARCH(q) model estimated, it is possible to compute the z_{t+k}^*'s and the ν_{t+k}^*'s in Equation (16) given the past information by using the best forecasts given above and Equation (22) and (24). The theoretical P/D ratio is then equal to:

$$\delta_t = \boldsymbol{\Psi}\boldsymbol{\delta}_{t-1} - \boldsymbol{\Gamma}\mathbf{u}_{t-1}^2 + \lambda\sum_{k=0}^{\infty}(\rho\lambda)^k z_{t+k+1} + c_3 \tag{25}$$

where:

$$\boldsymbol{\Psi} = \mathbf{g}_p\boldsymbol{\phi}\left(\mathbf{I}_p - \lambda\left(\mathbf{I}_p - \rho\boldsymbol{\phi}\right)\left(\mathbf{I}_p - \rho\lambda\boldsymbol{\phi}\right)^{-1}\right)$$
$$\boldsymbol{\Gamma} = 2\lambda\beta\rho^2\mathbf{g}_q\boldsymbol{\alpha}^2\left(\mathbf{I}_q - \rho\lambda\boldsymbol{\alpha}\right)^{-1}$$

where $\mathbf{I}_{p,q}$ are identity matrices of dimension $(p \times p)$ and $(q \times q)$, respectively, and $\mathbf{g}_{p,q}$ are p and q raw vectors, respectively, with all element equal to zero except for the first one, which is equal to one. These vectors select the first row or the first element of the next matrix or vector, respectively.

4.2 Indirect tests of the model

Unfortunately, Equation (25) cannot be directly tested since the future adjusted dividend growth rate is not known. However, it is possible to get an approximation of the theoretical P/D ratio by using the adjusted dividend

growth rates observed ex-post.[14] Then, one can get an idea of the validity of the model by examining the importance of the differences between the theoretical P/D ratio, which corresponds to the price-to-dividend ratio that would prevail if the model given by Equation (25) was true, and the observed P/D ratio. One way to do that concretely is to test if the variance of the theoretical P/D ratio is equal to the variance of the observed P/D ratio.

5 Data and parameters

5.1 Dataset

The HEAPM is estimated with annual data on American stocks' prices, dividends and interest rates for the period between 1871 and 2003. The prices and dividends are taken from the Standard & Poors 500 Composite Stock Price Index, extended back to 1871 by using the data in Cowles (1939).[15] The interest rate is the 6 -month prime commercial paper rate.

5.2 Preliminary verifications

Before going estimating the model, we checked that the variables are stationary. Table 1 displays the *t*-statistics of the Augmented Dickey-Fuller test for the null hypothesis of unit root.

The test shows that the hypothesis of unit root is rejected at a 1% confidence level for the adjusted dividend growth rate on the entire period. For the P/D ratio, the hypothesis of unit root is rejected at a 1% confidence level for the period 1871-1995. If we introduce the next eight years, this hypothesis is rejected at a confidence level slightly higher than 5%. Formally,

---

[14]We made the assumption that the adjusted dividend growth rates after 2003 are equal to their historical mean. The effect of this assumption should be marginal for the major part of the sample since the mean of dividend difference is close to zero and is discounted. However, it might affect more significantly the last observations.

[15]The index reflects the total market value of all 500 component stocks at a given date. The market value of a company is determined by multiplying the stock by the number of common shares outstanding. The dividends can be recovered from an index, which is based on the sum of the total monthly dividend for the same 500 stocks. This data set is kindly provided by Rober J. Schiller on his website (http://www.econ.yale.edu/~shiller/data.htm).

Table 1: Unit root test		
	1871-1995	1871-2003
δ_t	-4.2466***	-2.8340*
z_t	-10.0796***	-10.3925***

the P/D ratio can be considered as stationary at a 6% confidence level, but this first test suggests that the behavior of the market might have change during the period 1996-2003.

5.3   Estimation of the parameters

Finally, before estimating the theoretical P/D ratio given by Equation (25), we must estimate the different parameters of the model. The parameters ρ and b are taken from the estimation of equation: $\log\left(P_{t+1} + D_{t+1}\right) = b + \rho p_{t+1} + (1 - \rho) d_{t+1} + \mu_t$ where is a μ_t white noise. The vectors of parameters ϕ and α come from the estimation of the AR(p)-ARCH(q) model presented in Section **??**. Finally, the parameter β can be estimated by using Equation (22). To do so, we use the hypothesis made in Section (3.1), which states that the public signal is an unbiased signal of the future adjusted dividend growth rate. Therefore, the parameter β can be estimated by the following regression:

$$z_t = \delta_{t-1}^* - \rho\delta_t^* + \beta\nu_t^* + b + \varepsilon_t$$

where $E\left[\varepsilon_t\right] = 0$. The best forecasts δ_{t-1}^*, δ_t^* and ν_t^* can be replaced by their value derived from the AR(p)-ARCH(q) model as described in Section **??** to get the final regression equation for the parameter β:

$$z_t = \mathbf{g}_p\left(\mathbf{I}_p - \rho\phi\right)\delta_{t-1} + 2\beta\rho^2\mathbf{g}_q\alpha\mathbf{u}_{t-1}^2 + b + \varepsilon_t$$

For example, Table 2 gives the estimated parameters for an AR(3)-ARCH(1) model estimated on the entire sample.

Table 2: Estimated parameters

Parameter	Test value
b	0.1664
ρ	0.9613
ϑ	0.3542
ϕ_1	0.8477
ϕ_2	-0.0689
ϕ_3	0.1160
ζ	0.0208
α_1	0.5652
β	4.0643

6 Model with constant relative weight: empirical results

6.1 Overall sample vs. rolling sample estimation

Once that the parameters of the model are known, it is possible to compute the theoretical P/D ratio for any given relative weight λ by using Equation (25). It is then possible to estimate the relative weight λ^* for which the theoretical price is the closest from the price observed on the market. To find this optimal relative weight, we computed the sum of the squared differences between the observed and the theoretical P/D ratio given λ. We then define the optimal λ^* as the relative weight which minimize the sum of the squared differences.

Two methods are possible to estimate the parameters of the model and the optimal weight. Firstly, one can estimate the parameters of Equation (25) by using the observations of the entire sample ("overall sample estimation") and then compute the theoretical P/D ratio for each period. This is equivalent to the situation where investors in 1920, for example, use the same parameter as investors in 2000. This can be the case if investors do not infer the model parameters from the past observations but use some constant rule to set the value of these parameters. The second method is

to estimate the parameters of the model for each period, using only the available information at this time, and then compute the theoretical P/D ratio given the parameter at this time ("rolling sample estimation"). This estimation method corresponds to the case where, at each period, investors use the new information to re-evaluate the model parameters and make their forecast with these new values. The results of both methods are compared in the next section.

6.2   Estimated relative weight and volatility tests

The optimal relative weight λ^* is equal to 0.227 with the overall sample method and to 0.187 with the rolling sample method. These two results are similar. They mean that, between 1871 and 2003, the investors seem to have given a weight to the public information about four times bigger than the one given to the private information. Figure 1 and 2 show the observed and the theoretical P/D ratio with the overall sample method and the rolling sample method respectively.[16]   With both methods, the theoretical price follows relatively closely the price observed on the market. This is particularly true for the overall sample method. In both case, the picture is very different from the traditional graph given by a rational expectation model where the theoretical and the observed price diverge significantly (see, e.g., Shiller 1981 or 2003).
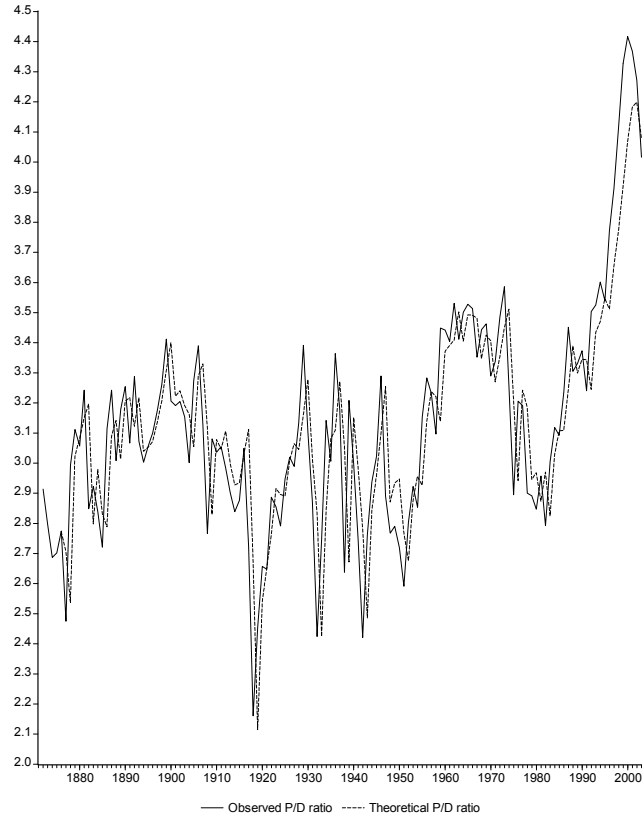
Table 3: Volatility test

Method	Overall sample		Rolling sample	
	Test value	p-value	Test value	p-value
F-test	1.374090	0.0745	1.623560	0.0146
Siegel-Tukey	1.477986	0.1394	2.224521	0.0261
Bartlett	3.180563	0.0745	5.960385	0.0146
Levene	2.051350	0.1533	4.681654	0.0316
Brown-Forsythe	1.938616	0.1650	3.889712	0.0499

The volatility tests confirm these conclusions. Table 3 displays the test

---

[16]With the rolling sample method, the theoretical P/D ratio is computed from 1900. The observation before this date are used to give the first estimation of the parameters.

Figure 1: Theoretical P/D ratio with constant relative weight (overall sample method)

statistics and the p-values of the volatility test for the optimal value with both methods. The null hypothesis is the equality between the two variances. For the overall sample method, the tests indicate that it is not possible to reject the hypothesis of equal variances at a 5% confidence level. For the rolling sample method, the equal variance hypothesis is rejected at a 5% confidence level for all tests but is accepted at a 1% confidence level. Note that, with the rolling sample method, the volatility of the model ($\sigma^2 = 0.2749$) is bigger than the observed one ($\sigma^2 = 0.1693$). The evidence in favour of the model with higher order beliefs are not unquestionable, but formally, both models pass the volatility test at a 1% confidence level. Therefore, a model

Figure 2: Theoretical P/D ratio with constant relative weight (rolling sample method)
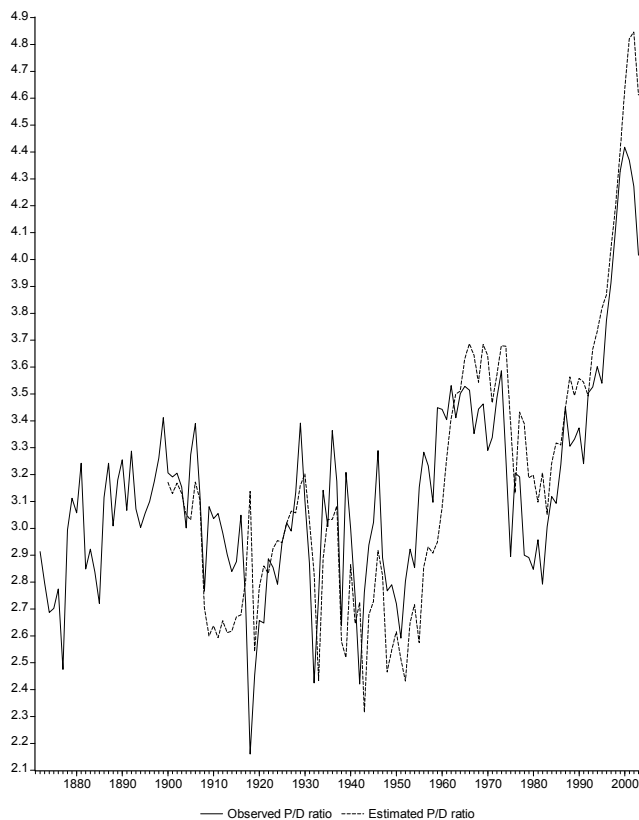

with higher order beliefs seems to give a volatility which is more in line with the one observed on the market, what traditional rational expectation models fail to do.

6.3 Stability of the relative weight

In our HEAPM, the relative weight λ is assumed to be constant in time. This is a rather strong hypothesis. To check if this hypothesis corresponds to the reality, we computed the optimal relative weight over different periods of 20 years. The result for the entire sample and the rolling sample method are displayed in Figure 3. In these two figures, it is clear that the relative

Figure 3: Stability of relative weight (L: overall. R: Rolling)
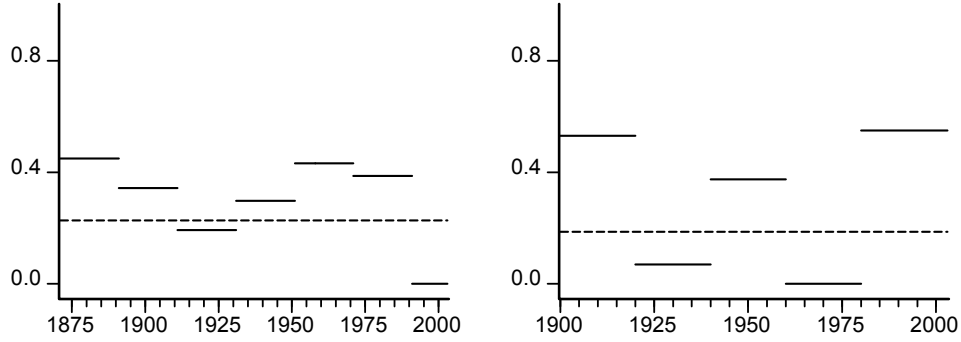
weight seems to vary with time. The next section deals with this problem by slightly modifying the original HEAPM and introducing a time-varying relative weight.

# 7 Model with time-varying relative weight: empirical results

7.1 Modification of the original model

With a time-varying relative weight, the original HEAPM of Equation (16) can be re-written as:

$$\delta_t = \sum_{k=0}^{\infty} \rho^k \left(\left(1 - \lambda_t^{k+1} \right) z_{t+k+1}^* + \lambda_t^{k+1} z_{t+k+1} - \beta \nu_{t+k+1}^* \right) + c_1$$

This model implies that, at each time t, investors choose a different relative weight between the private and the public signal. This relative weight varies for each period, but at time t, it remains constant for each forecast horizon $t + k$. It is still assumed that the relative weight is identical for each investor. Similarly to the constant relative weight, the optimal λ_t^*, which is different for each period, is defined as the relative weight that

23

Figure 4: Estimated relative weight (L: overall. R: rolling)
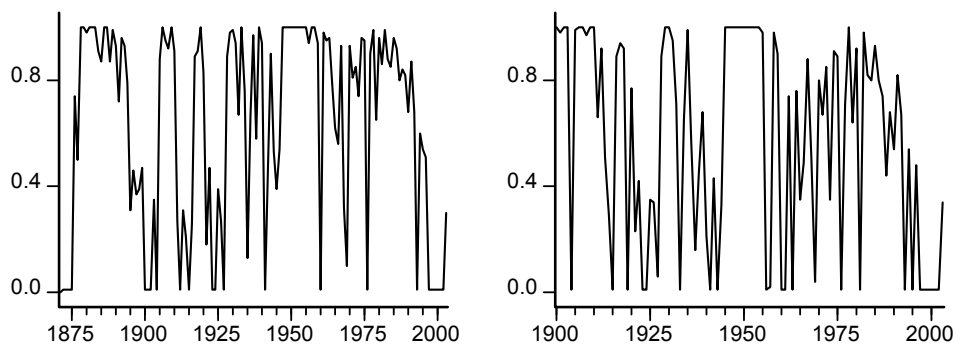
minimizes the squared difference between the theoretical and the observed δ_t at time t.

7.2 Estimated relative weight and volatility tests

The estimated relative weights obtained with the overall sample and the rolling sample method are presented in Figure 4. As suggested by in the previous section, the optimal relative weight seems to vary significantly in time. Note that with both methods, the periods with a relative weight equal to zero, which corresponds to a situation where investors give no weight to their private signal, are relatively rare and short. The end of the 90s is an exception since both methods indicates that investors have used only the public signal for several years in a row. This could explain the irregularities in the stationary of the P/D ratio mentioned in Section 5.2.

Figures 5 and 6 (in Appendix B) show the observed and the theoretical P/D ratios with a time-varying relative weight. Once again, the P/D ratios given by the model are relatively close to the observed ones. This conclusion is confirmed by the p-values of the volatility tests presented in Table 4. With a time-varying relative weight, the hypothesis of equal variance between the observed and the theoretical P/D ratio cannot be rejected at a 5% confidence level.[17] This is a significant empirical evidence in favour of the HEAPM

[17]The hypothesis cannot even be rejected at a 10% confidence level, if we do not take

24

developed in this paper. It seems that a model with higher order beliefs is able to generate a volatility, which is similar to the one observed on the market, whereas rational expectation models fail to produce this feature.

Table 4: Volatility test

	p-value	
Method	Overall sample	Rolling sample
F-test	0.9988	0.1205
Siegel-Tukey	0.6959	0.0503
Bartlett	0.9988	0.1205
Levene	0.7879	0.1385
Brown-Forsythe	0.8269	0.1497

8 Conclusion

This paper proposes an asset pricing model, which takes into account investors' higher order beliefs. The particularity of this model is that it can be estimated and thus help to determine if the effect of higher order beliefs on the stock price is empirically significant. The model is estimated with American data on stock prices, dividends, and interest rates for the period between 1871 and 2003. The main conclusion is that higher order beliefs seem to have a significant impact on asset prices. In particular, the price volatility induced by the model does not significantly differ from the volatility observed on the market. In this sense, higher order beliefs appear to be a plausible explanation of the excess volatility puzzle.

In addition to the main conclusion, the paper sheds light on a few other points. First, it shows that heterogeneous expectations induce the beauty contest phenomenon described by Keynes (1936). In the asset price equation, the beauty contest effect takes the form of an *iterated average expectation*. This iterated average expectation replaces the iterated expectation used in traditional asset price models. Second, after making some further assumptions about the information available to each agent, we give a testable

into account the Siegel-Tukey test.

25

asset price equation. This equation is a useful tool to understand how the agents combine their private and public information to take into account the beauty contest effect. In particular, it shows that, in order to guess others' expectations, the agent put more weight on the public signal than they would do if they were trying to guess the future dividends. They do so because, as the public signal influences everyone's expectation, it constitutes a better predictor of the average opinion that the private signal.[18] Finally, in our model, the price still contain some information about the future payoffs. This paper shows how this information can be extracted from the past prices. The direct consequence of this is that the past prices can be used as the pubic signal described above.

In conclusion, our empirical results indicate that higher order beliefs might play a significant role in the stock markets. A significant part of the volatility observed in the price seems to be explained by this phenomenon, rather than by the movements of the fundamentals. This conclusion suggests that adding higher order beliefs to traditional present value model could improve their empirical performance. The model used in this paper is based on some restrictive hypotheses, but its simplicity and its value added in terms of empirical performances might constitute a promising basis for further developments.

References

[1] Allen, F., Morris, S. and Shin, H. S. (2003). "Beauty contests, bubbles and iterated expectations in asset markets", *Cowles Foundation Discussion Paper* No 1406.

[2] Bacchetta, P. and van Wincoop (2004). "Higher order expectations in asset pricing", *Studienzentrum Gerzensee Working Papers*, no. 04.03.

[3] Barberis, N., Huang, M. and Santos, T. (2001). "Prospect theory and asset prices", *Quarterly Journal of Economics*, 116(1), p. 1-53.

---

[18]This result has been highlighted by Morris and Shin (2002).

[4] Barberis, N. and Thaler, R. (2003). "A survey of behavioral finance" in Constantinides, G. M., Harris, M and Stulz, R. M. (eds), *Handbook of the Economics of Finance: Financial Market and Asset Pricing*, North-Holland, Amsterdam.

[5] Basak, S. (2000). "A model of dynamic equilibrium asset pricing with heterogeneous beliefs and extraneous risk", *Journal of Economic Dynamics and Control*, 24, p. 63-95.

[6] Biais, B. and Bossaerts, P. (1998). "Asset prices and trading volume in a beauty contest", *Review of Economic Studies*, 65(2), p. 307-340.

[7] Biais, B., Bossaerts, P. and Spatt, C. (2003). "Equilibrium asset pricing under heterogeneous information", *IDEI Working Paper*.

[8] Campbell, J. Y. and Cochrane, J. H. (1999). "By force of habit: a consumption-based explanation of aggregate stock market behavior", *Journal of Political Economy*, 107(2), p. 205-251.

[9] Campbell, J. Y. and Schiller, R. J. (1987). "Cointegration and tests of present value models", *Journal of Political Economy*, 95(5), p. 1062-1088.

[10] Campbell, J. Y. and Shiller, R. J. (1988). "The dividend-price ratio and expectations of future dividends and discount factors", *Review of Financial Studies*, 1(3), p. 195-228.

[11] Campbell, J. Y. and Viceira, L. M. (2002). *Strategic Asset Allocation.* Oxford; Oxford University Press.

[12] Hamilton, J. D. (1994). *Time Series Analysis.* Princeton: Princeton University Press.

[13] Harrison, M. and Kreps, D. (1978). "Speculative investor behavior in a stock market with heterogeneous expectation", *Quarterly Journal of Economics*, 92, p. 323-336.

[14] Hogg, R. V. and Craig, A. T. (1995). *Introduction to Mathematical Statistics*, 5th ed., Englewood Cliffs NJ: Prentice Hall International.

[15] Keynes, J. M. (1936). *The General Theory of Employment, Interest and Money.* London: MacMillan.

[16] Kurz, M. (1974). "The Kersten-Stigum model and the treatment of uncertainty in equilibrium theory" in: Balch, M. S., McFadden, D. L. and Wu, S. Y., *Essays on Economic Behavior Under Uncertainty*, North Holland, Amsterdam.

[17] LeRoy, S. F. and Porter, R. (1981). "The present-value relation: tests based on implied variance bounds", *Econometrica*, 49(3), p. 555-574.

[18] Miller, E. (1977). "Risk, uncertainty, and divergence of opinion", *Journal of Finance*, 32, p. 1151-1168.

[19] Morris, S. (1996). "Speculative investor behavior and learning", *Quarterly Journal of Economics*, 111, p. 1111-1133.

[20] Morris, S. and Shin, H. S. (2002). "Social value of public information", *American Economic Review*, 92(5), p. 1521-1534.

[21] Shiller, R.J. (1981). "Do stock price move too much to be justified by subsequent changes in dividends?", *American Economic Review*, 71, p. 421-436.

[22] Shiller, R. J. (2003). "From efficient markets theory to behavioral finance", *Journal of Economic Perspectives*, 17(1), p. 83-104.

[23] Townsend, R. M. (1983). "Forecsting the forecats of the others", *Journal of Political Economy*, 91(4), p. 546-588.

[24] Wickens, M. R. (2003). "Microeconomic sources of equity risk", *Center for Economic Policy Research Discussion Paper Series*, no. 4070.

[25] Wu, H.-M. and Guo, W.-C. (2004). "Asset price volatility and trading volume with rational beliefs", *Economic Theory*, 23, p. 795-829.

[26] Zhong, M., Darrat, A. F. and Anderson, D. C. (2003). "Do US stock prices deviate from their fundamental values? Some new evidence", *Journal of Banking and Finance*, 27, p. 673-697.

A   Volatility of the P/D ratio

From Equation (16), we can compute the volatility of δ_{t+j} expected by investor i at time $t+j-1$. This yields:

$$Var_{t+j-1}^i\left(\delta_{t+j}\right) = Var_{t+j-1}^i\left(\sum_{k=0}^{\infty}\rho^k\left(\left(1-\lambda^{k+1}\right)z_{t+j+k+1}^* + \lambda^{k+1}z_{t+j+k+1} - \beta\nu_{t+j+k+1}^*\right)\right)$$

As the expected adjusted dividend growth z_{t+j+k}^* and the expected volatility ν_{t+j+k}^* are known at time $t+j-1$, their conditional variance is null and the previous equation is equivalent to:

$$Var_{t+j-1}^i\left(\delta_{t+j}\right) = Var_{t+j-1}^i\left(\sum_{k=0}^{\infty}\rho^k\lambda^{k+1}z_{t+j+k+1}\right)$$

Thus, the variance of the P/D ratio is a function of the variance and autocovariance of the future adjusted dividend growth. As we have seen in Section 3.2.2, these variances and autocovariances are identical for each investor and therefore we have:

$$Var_{t+j-1}^i\left(\delta_{t+j}\right) = \lambda^2 Var_{t+j-1}\left(\sum_{k=0}^{\infty}\left(\rho\lambda\right)^k z_{t+j+k+1}\right) = \sigma_{\delta_{t+j}}^2 \qquad (26)$$

B Theoretical and observed P/D ratios with time varying relative weight

Figure 5: Theoretical P/D ratio with time-varying relative weight (overall sample method)
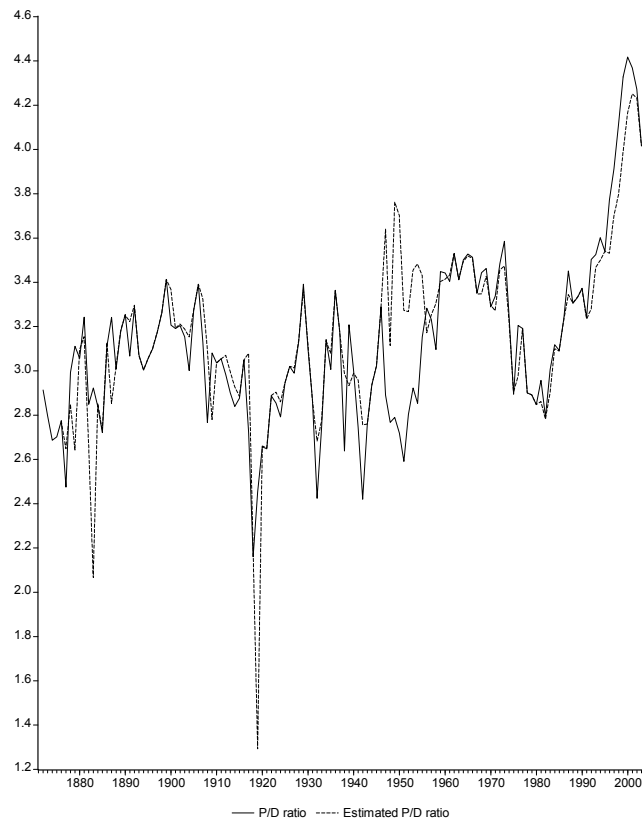
Figure 6: Theoretical P/D ratio with time-varying relative weight (rolling sample method)