QUANTITATIVE FINANCIAL ECONOMICS

(2ND Edition 2004)

CHAPTER 15
BEHAVIOURAL MODELS

AIMS:

- Show how rational traders and noise traders interact to give equilibrium prices which diverge from fundamental value and how rational traders may be destabilizing.
- Examine how short-termism could lead to mispricing.
- Analyze how noise traders (e.g. momentum traders, style investors) and fundamentals traders interact to product underreaction followed by overreaction in stock prices and cross-correlations between different classes of stocks
- Demonstrate how a non-standard utility function incorporating both consumption and 'other variables' in an intertemporal optimizing model where investors suffer from loss-aversion, can explain the 'stylised facts' of stock returns such as excess volatility, predictability and the equity premium puzzle.

So far we have been discussing the implications of noise traders in fairly general terms and now it is time to examine more formal models. One set of models assumes noise traders and rational traders interact to give equilibrium prices which clear the market. A second broad approach assumes all agents are identical but have non-standard preferences – here utility depends on the change in wealth as well as the level of consumption and agents may fear losses more than gains. These models often assume intertemporal utility maximization as in the 'standard approach'.

1. A SIMPLE MODEL

Shiller (1989) provides a simple piece of analysis in which stylized empirical results can be explained by the presence of noise traders. The proportionate demand for shares by the smart money is Q and is based (loosely) on the mean-variance model:

[1] $$Q_t = (E_t R_{t+1} - \rho)/\theta$$

If $E_t R_{t+1} = \rho$ demand by the smart money equals zero. If $Q_t = 1,$ the smart money holds all the outstanding stock and this requires an expected return $E_t R_{t+1} = \rho + \theta$. Hence θ is a kind of risk premium to induce the smart money to hold all the stock.

We now let (Y / P_t) equal the *proportion* of stock held by noise traders. In equilibrium, the proportions held by the smart money and the noise traders must sum to unity:

[2] $$Q_t + (Y / P_t) = 1$$

substituting [1] in [2]

[3] $$E_t R_{t+1} = \theta[1 - (Y / P_t)] + \rho$$

Hence the expected return as perceived by the smart money depends on what they think is the current and future demand by noise traders: the higher is noise trader demand, the higher are current prices and the lower is the expected return perceived by the smart money. Using [3] and the definition

[4] $$E_t R_{t+1} = E_t[(P_{t+1} + D_{t+1}) / P_t - 1]$$

we obtain

[5] $$P_t = \delta E_t (P_{t+1} + D_{t+1} + \theta Y_t)$$

where $\delta = 1/(1 + \rho + \theta)$. Hence by repeated forward substitution:

[6] $$P_t = \sum_0^{\infty} \delta^t (E_t D_{t+1} + \theta E_t Y_{t+1})$$

If the smart money is *rational* and recognizes demand by noise traders, then the smart money will calculate that the market clearing price is a weighted average of fundamentals (i.e. $E_t D_{t+1}$) and future noise trader demand $E_t Y_{t+1}$. The weakness of this 'illustrative model' is that noise trader demand is completely exogenous. However, as we see below, we can still draw some useful insights.

If E_tY_{t+1} and hence aggregate noise trader demand is random around zero then the moving average of E_tY_{t+1} in [6] will have little influence on P_t which will be governed primarily by fundamentals. Price will deviate from fundamentals but only randomly. On the other hand, if demand by noise traders is expected to be persistent (i.e. 'large' values of Y_t are expected to be followed by further large values) then small changes in current noise trader demand can have a powerful effect on current price, which might deviate substantially from fundamentals over a considerable period of time.

Shiller (1989) uses the above model to illustrate how tests of market efficiency based on regressions of *returns* on information variables known at time t, have low power to reject the EMH when it is false. Suppose dividends (and the discount rate) are constant and hence the EMH (without noise traders) predicts that the stock price is constant. Now suppose that the market is actually driven *entirely* by noise traders. Let noise trader demand be characterized by

[7] $$Y_t = u_{t+1} + u_{t+2} + u_{t+3} + \dots \qquad + u_{t-n}$$

where u_t is white noise. Equation [7] has the property that a unit increases in u_t at time t generates changes in Y in future periods that follow a 'square hump' which dies away after n-periods[(1)]. Using [6] price *changes* ($P_{t+1} - P_t$) only arise because of *revisions* to expectations about future noise trader demand which are weighted by $\delta, \delta^2, \delta^3$ etc. Because $0 < \delta < 1$, price changes are heavily dominated by u_t (rather than by past u_{t-j}). However as u_t is random, price changes in this model, which by construction is dominated by noise traders, are nevertheless largely unforecastable.

Shiller generates a ΔP_{t+1} series using [6] for various values of the persistence in Y_t (given by the lag length *n*) and for alternative values of ρ and θ. The generated data for ΔP_{t+1} is regressed on the information set consisting only of P_t. Under the EMH we expect the R-squared of this regression to be zero. For $\rho = 0, \theta = 0.2$ and $n = 20$ he finds $R^2 = 0.015$. The low R-squared supports the constant returns EMH, but it results from a model where price changes are *wholly determined* by noise traders. In addition, the price *level* can deviate substantially from fundamentals even though price *changes* are hardly forecastable. He also calculates that if the generated data includes a constant dividend price ratio of 4% then the 'theoretical R-squared' of a regression of the return R_{t+1} on the dividend price ratio $(D/P)_t$ is 0.079. Hence empirical evidence that returns are only weakly related to information at time is not necessarily inconsistent with prices being determined by noise traders (rather than by fundamentals).

Overall, Shiller makes an important point about empirical evidence. The evidence using real world data is not that stock returns are unpredictable (as suggested by the EMH) but that stock returns are not *very* predictable. However, the latter evidence is also not inconsistent with possible models in which noise traders play a part.

If the behaviour of Y_t is exogenous (i.e. independent of dividends) but is stationary and mean reverting then we might expect *returns* to be predictable. An above average level of Y will eventually be followed by a fall in Y (to its mean long-run level). Hence prices are mean reverting and current returns are predictable from previous periods returns.

In addition this sample noise trader model can explain the positive association between the dividend-price ratio and next periods return on stocks. If dividends vary very little over time, a price rise caused by an increase in EY_{t+1} will produce a fall in the dividend price ratio. If Y_t is mean reverting than prices will fall *in the future*, so returns R_{t+1} also fall. Hence $(D/P)_t$ is positively related to returns R_{t+1} as found in empirical studies. Shiller also notes that if noise trader demand Y_{t+1} is influenced either by post returns (i.e. bandwagon effect) or past dividends then the share price might overreact to current dividends compared to that given by the first term in [6], that is the fundamentals part of the price response.

2. OPTIMISING MODEL OF NOISE TRADER BEHAVIOUR

In the 'neat' model of De Long, Shleifer, Summers and Waldmann (DSSW, 1990), both smart money and noise traders are risk averse and maximize utility of terminal wealth. There is a finite horizon, so that arbitrage is risky. The (basic) model is constructed so that there is no fundamental risk (i.e. dividends are known with certainty) but only noise trader risk. The noise traders create risk for themselves and the smart money by generating fads in demand for the risky asset. The smart money forms optimal forecasts of the future price based on the correct distribution of price changes but noise traders have biased forecasts. The degree of *price misperception* of noise traders ρ_t represents the *difference* between the noise trader forecasts and optimal forecasts:

[8] $$\rho_t \sim N(\rho^*, \sigma^2)$$

If $\rho^* = 0$, noise traders forecasts agree with those of the smart money (on average). If noise traders are on average pessimistic (e.g. bear market) then $\rho^* < 0$, and the stock price will be below fundamental value and vice versa. As well as having this long-run view ρ^* of the divergence of their forecasts from the optimal forecasts, 'news' also arises so there can be *abnormal but temporary* variations in optimism and pessimism, given by a term, $(\rho_t - \rho^*)$. The specification of ρ_t is ad-hoc but does have an intuitive appeal based on introspection and evidence from behavioural/group experiments.

In the DSSW model the *fundamental value* of the stock is a constant and is arbitrarily set at unity. The market consists of two types of asset: a risky asset and a safe asset. Both noise traders and smart money are risk averse and have mean-variance preferences, so their demand for the risky asset depends positively on expected return and inversely on the noise trader risk (see Appendix I). The noise trader demand also depends on whether they feel bullish or bearish about stock prices (i.e. the variable ρ_t). The risky asset is in fixed supply (set equal to unity) and the market clears to give an equilibrium price P_t. The equation which determines P_t looks rather complicated but we can break it down into its component parts and give some intuitive feel for what's going on:

[9] $$P_t = 1 + \frac{\mu}{r}\rho^* + \frac{\mu}{(1+r)}[\rho_t - \rho^*] - \frac{2\gamma\mu^2\sigma^2}{r(1+r)^2}$$

where μ = the proportion of investors who are noise traders, r = the riskless real rate of interest, γ = the degree of (absolute) risk aversion, σ^2 = variance of noise trader misperceptions. If there are no noise traders μ = *0* and [9] predicts that market price equals fundamental value (of unity). Now let us suppose that at a particular point in time, noise traders have the same long run view of the stock price as does the smart money (i.e. ρ^* = *0*) and that there are no 'surprises', (i.e. no abnormal bullishness or bearishness), so that, $(\rho_t - \rho^*) = 0$.

We now have a position where the noise traders have the same view about future prices as does the smart money. However, the equilibrium market price still *does not solely reflect fundamentals* and the market price is less than the fundamental price - given by the last term on the RHS of equation [9]. This is because of the presence noise trader risk, since their potential actions may influence future prices. The price is below fundamental value so the smart money (and noise traders) may obtain a positive expected return (i.e. capital gain) as compensation for the noise trader risk. This mispricing is probably the key result of the model and involves a *permanent* deviation of price from fundamentals. We refer to the effect of the third term in [9] as the amount of 'basic mispricing'.

Turning now to the second term in equation [9] we see for example that the noise traders will push the price above fundamental value if they take a *long-term view* that the market is bullish $(\rho^* > 0)$. The third term reflects *abnormal* short-term bullishness or bearishness. These terms imply that at particular time periods price may be above or below fundamentals.

If only ρ_t varies so that $\rho_t - \rho*$ is random around zero then the actual price would deviate randomly around its 'basic mispricing' level. In this case stock prices would be 'excessively volatile' (relative to fundamentals) where volatility is zero. From [9] the variance of prices is :

[10] $$E_t(P_t - E_t P_t)^2 = \frac{\mu^2 \operatorname{var}(\rho_t - \rho*}{(1+r)^2} = \frac{\mu^2 \sigma^2}{(1+r)^2}$$

Hence excess volatility is more severe, the greater is the variability in the misperceptions of noise traders σ^2, the more noise traders there are in the market μ and the lower is the cost of borrowing funds r. To enable the model to reproduce *persistence* in price movements and hence the broad bull and bear movements in stock prices that we observe, we need to introduce 'fads' and 'fashions'. Broadly speaking this implies, for example, that periods of bullishness are followed by further periods of bullishness. Which can be represented by a random walk in ρ_t^* :

[11] $$\rho_t^* = \rho_{t-1}^* + \omega_t$$

where $\omega_t \sim N(0, \sigma_w^2)$. (Note that σ_ω^2 is different from σ^2, above). At any point in time the investor's optimal forecast of $\rho*$ is its current value. However as 'news' ω_t arrives, noise traders alter their views about ρ_t^* and this 'change in perceptions' persists over future periods. It should be clear from the second term in [9] that the random walk in ρ_t^* implies P_t will move in long swings and hence there will be 'bull and bear' patterns in P_t.

Fortune (1991) assumes for illustrative purposes, ω_t and $(\rho_t - \rho^*)$ are *niid* and uses representative values for $r, \mu, \gamma,$ in [9]. He then generates a time series for P_t shown in (figure 1). The graph indicates that on this *one* simulation, price falls to 85% of fundamental value (which itself may be rising) with some dramatic rises and falls in the short-run.

Figure 1 – Fortune-here

An additional source of *persistence* in prices could be introduced into the model by assuming that σ^2 is also autoregressive (e.g. ARCH and GARCH processes). It is also not unreasonable to assume that the 'conversion rate' from being a smart money trader to being a noise trader may well take time and move in cycles. This will make μ (i.e. the proportion of noise traders) exhibit persistence and hence so might P_t. It follows that in this model, price may differ from fundamentals for substantial periods of time because arbitrage is incomplete. Also, persistence in $\rho_t - \rho*$ could be mean reverting. This would imply that prices are mean reverting and that *returns* on the stock market are partly predictable from past returns or from variables such as the dividend-price ratio.

CAN NOISE TRADERS SURVIVE?

De Long *et al* show that where the proportion of noise traders is fixed in each period (i.e. μ is constant) it is possible (although not guaranteed) that noise traders do survive, even though they tend to buy high and sell low (and vice versa). This is because they are over-optimistic and under-estimate the true riskiness of their portfolio. As a consequence, they tend to 'hold more' of the assets subject to bullish sentiment. In addition, if noise trader risk σ^2 is large, the smart money will not step in with great vigour, to buy underpriced assets because of the risk involved.

The idea of imitation can be included in the model by assuming that the conversion rate from smart money (s) to noise traders (n) depends on the excess returns earned by noise traders over the smart money $(R^n - R^s)$:

[12] $$\mu_{t+1} = \mu_t + \psi(R^n - R^s)_t$$

where μ is bounded between 0 and 1. De Long *et al* also introduce fundamental risk into the model. The per period return on the risky asset becomes a random variable, $r + \varepsilon_t$ where $\varepsilon_t \sim N(0, \sigma_\varepsilon^2)$. In this version of the model, the probability of noise trader survival is always greater than zero. This is because of what they call the 'create space' effect, whereby risk is increased to such an extent that it further inhibits risk averse smart money from arbitraging any potential gains. (The latter result requires ψ to be 'small', since otherwise the newly converted noise traders may influence price and this would need to be predicted by the 'old' noise traders who retire).

CLOSED END FUNDS

We have noted that closed-end funds often tend to sell at a discount and this discount varies over time, usually across *all* funds. Sometimes such funds sell at a premium. Using our noise trader model we can get a handle on reasons for these ‘anomalies’. Let the risky asset be the closed end fund itself and the safe asset the actual underlying stocks. The smart money will try and arbitrage between the fund and the underlying stocks (e.g. buy the fund and sell the stocks short, if the fund is at a discount). However, even if $\rho_t = \rho^* = 0$, the fund (risky asset) will sell at a discount because of inherent noise trader risk (see [9]). Changes in noise trader sentiment (i.e. in ρ^* and $\rho_t - \rho^*$) will cause the discounts to vary over time and as noise trader risk is systematic we expect discounts on most funds to move together.

In the noise trader model a number of closed end funds should also tend to be started at the same time, namely when noise trader sentiment for closed end funds is high (i.e. $\rho^* > 0, \rho_t > 0)$. When existing closed end funds are at a premium it pays the smart money to purchase shares (at a relatively low price) bundle them together into a closed end fund and sell them at a premium to optimistic noise traders.

CHANGES IN BOND PRICES

Empirically, when the long-short spread $(R - r)$ on bonds is positive, then long rates tend to fall, and the *prices* of long bonds rise. This is the opposite of the pure expectations hypothesis of the term structure. The stylised facts of this anomaly are consistent with our noise trader model with the long bond being the risky asset (and the short bond the safe asset). When $R_t > r_t$ then the *price* of long bonds as viewed by noise traders may be viewed as abnormally low. If noise trader fads are mean reverting, they will expect bond prices to rise in the future and hence long rates R to fall. This is what we observe in the empirical work. Of course, even though the noise trader model explains the stylised facts this will leaves us a long way from a formal test of the noise trader model in the bond market.

Overall, the key feature of the De Long *et al* model is to demonstrate the possibility of underpricing in *equilibrium.* The other effects mentioned above depend on one's adherence to the possibility of *changes* in noise trader sentiment, which are persistent. However, 'persistence' is not the outcome of an optimising process in the model, although it is an intuitively appealing one.

SHORT TERMISM

In a world of only smart money, the fact that some of these investors take a 'short-term' view of returns should not lead to a deviation of price from fundamentals. The argument is based on the implicit forward recursion of the rational valuation formula. If you buy today at time t, in order to sell tomorrow, your return depends (in part) on the expected capital gain and hence on the price you can get tomorrow. But the latter depends on what the person you sell to at t+1 thinks the price will be at t+2 etc. Hence a linked chain of short term 'rational fundamental' investors performs the same calculation as an investor with an infinite horizon.

With a finite investment horizon and the presence of noise traders the above argument doesn't hold. True, the longer the horizon of the smart money the more willing she may be to undertake risky arbitrage based on divergences between price and fundamental value. The reason being that in the meantime she receives the insurance of dividend payments each period and she has a number of periods over which she can wait for the price to rise to fundamental value. However, even with a 'long' but finite horizon there is some price resale risk. The share in the total return from dividend payments over a 'long' holding period is large but there is still substantial risk present from uncertainty about price in the 'final period'.

We note from the noise trader model that if a firm can make its equity appear less subject to noise trader sentiment (i.e. to reduce σ^2) then its underpricing will be less severe and its price will rise. This reduction in uncertainty might be accomplished by

a) raising current dividends (rather than investing profits in an uncertain long term investment project for example, R&D expenditures)
b) swapping debt for equity
c) share buybacks

Empirical work by Jensen (1986) and many others has shown that items (a)-(c) do tend to lead to an increase in the firm's share price and this is consistent with our interpretation of the influence of noise traders described above. It follows that in the presence of noise traders one might expect changes in capital structure to affect the value of the firm (contrary to the Modigliani-Miller hypothesis).

DESTABILISING RATIONAL TRADERS

In the DeLong et al (1990a) model above, the rational traders always move prices back towards fundamentals but not sufficiently to eliminate the mispricing of the ‘momentum’ noise traders. Suppose we accept the presence of momentum traders, who buy (sell) after a price rise (fall). The anecdotal evidence for this is quite strong. Some Chartists are known to chase ‘trends’, while stop-loss orders lead to selling after a price fall as does forced liquidation of your short position if you face increased margin calls that you cannot meet. Portfolio insurance can also lead to selling (buying) after a price fall (rise). For example, if you have sold (written) a call option (to a customer) then you hedge the position by buying stocks (i.e. delta hedging). If stock prices subsequently fall, the value of the written call also falls and your new hedge position requires that you sell some of your stocks. (Strictly portfolio insurance applies to replication of a ‘stock plus put’ portfolio with a ‘stock plus futures’ portfolio but our delta hedge example gives a similar result). So, portfolio insurance *logically* implies that you sell stocks after a price fall (and buy stocks after a price rise) – see Cuthbertson and Nitzsche (2001b). There is also experimental evidence (Andreassen & Kraus 1988) where economics students when faced with actual stock price data, tend to sell after a *small* price rise (and vice-versa) but buy after a *run* of price rises (i.e. momentum).

DeLong et al (1990b) use the analytic framework developed above but now allow the rational traders, (who are aware of momentum buying by noise traders) to *anticipate* momentum behaviour. Hence rational traders also buy after a price rise. The rational traders hope to 'ride the wave' caused by the momentum traders, but to sell before price begins to fall back to its fundamental value. The short-run behaviour of arbitrageurs is therefore *destabilising* and creates even larger *short-run* positive autocorrelation in returns (after the arrival of new fundamental's news), but returns are mean-reverting over long-horizons (i.e. negative autocorrelation).

Space constraints dictate that we cannot develop the model fully here, but merely sketch out the salient features. The model has 4 time periods (0, 1, 2, 3) and three types of trader. Rational traders maximise end-of-period wealth (consumption) and have mean-variance asset demands, (i.e. proportional to next period's expected return and inversely related to the variance of returns). Momentum traders' demand depends on the previous period's price *change* while the 'passive investors' demand for the risky asset depends on last period's price (i.e. a mechanical buy low – sell high strategy). The market clears in each period and there is a noisy signal about fundamentals (i.e. dividends) which influences the demand of the rational traders and hence their view of future actions by momentum traders. Good public news at time-t = 1 about (end-of-period-3) dividends leads to a rise in prices due to the fundamentals traders. In turn this causes an increase in demand by momentum traders at t=2 and finally a fall in prices in period-3, back to their fundamental level. There can be overshooting even if rational traders do not anticipate future increased momentum traders' demand, but the overshooting is exacerbated if rational traders 'jump on the bandwagon'. So an increase in the *number* of forward-looking rational traders entering the model can increase the 'overshooting'.

The model is consistent with anecdotal evidence about the 'players' in the market. Investment banks have 'insider' information about customer order flow and may use this to anticipate future demands by momentum traders (see chapter ?? on market microstructure). A stronger variant of this is 'front running', where marketmakers purchase (sell) on *their own account* before executing the buy-sell orders given to them by their customers. (This attracted the attention of the New York Attorney General, Eliot Spitzer in 2003). In the US, 'Investment Pools' often generate interest in 'hot-stocks', so as to attract momentum investors. George Soros' (1987) investment strategy in the 1960s and 1970s could also be viewed as betting on future crowd behaviour when he took (successful) long horizon bets on stocks involved in conglomerate mergers throughout the 1960 and in Real Estate Investment Trust stocks throughout the 1970s. In September 1992 he also successfully implemented a short-horizon strategy of selling the pound sterling from which he is reputed to have made $1bn over a few weeks (although whether this was 'chasing trends' or a fundamental misalignment is debateable).

Evidence from survey date on FX forecasting services (Frankel & Froot 1988) indicates that during the mid 1980s the 'inexorable' rise of the USD (with unchanged interest differentials) led forecasters to predict both a rise in the US dollar over one month and a depreciation by the end of the year. Their recommendation to investors was to buy today *even though they thought the USD was overpriced relative to fundamentals*. The DeLong (1990b) model demonstrates that this is a perfectly rational statement.

A weakness of the model is that the momentum traders are really dumb and should lose money and hence be forced out of the market. This criticism can be answered by assuming additional momentum traders (e.g. using new techniques such as neutral networks and genetic algorithms) enter the market or, existing momentum traders return with new backers. (If you read the *Financial Times* or *Wall Street Journal* you will have noted this occurs quite frequently, although with the recent tougher environment on Wall Street this may occur less often in the future). Also, if lots of momentum traders lose money over the same period, they can claim 'everybody did badly' and they may retain their investment mandates. There is evidence to suggest that pension fund mandates are not altered because of absolute losses but because of *worse* losses than your competitors – so some momentum traders may remain in the market. Also, such mandates are decided by a wide variety of factors other than 'return' (e.g. management costs, provision of analysts' research and investment 'style') The earlier model of DeLong et al (1990a) shows that noise traders may carry more market risk than the rational traders, so even if they make judgement errors, they can earn positive returns.

3. SHLEIFER-VISHNY MODEL: SHORT-TERMISM

The underpricing of an *individual* firm's stock is not a direct result of the formal noise-trader model of De Long *et al* since this formal model requires noise trader behaviour to be systematic across all stocks. However, the impact of high borrowing costs on the *degree* of mispricing in individual shares has been examined in a formal model by Shleifer and Vishny (1990). They find that current mispricing is most severe for those stocks where mispricing is *revealed* at a date in the distant future (rather than next period, say). Suppose physical investment projects with uncertain long horizon payoffs are financed with shares whose true value is only revealed to the market after some time. In the Shleifer-Vishny model these shares will be severely underpriced. It follows that the firm might be less willing to undertake such long horizon, yet profitable projects and short-termism on the part of the firm's managers might ensue. That is, they choose less profitable short-term investment projects rather than long term projects since this involves less current undervalution of the share price and less risk of them loosing their jobs from a hostile takeover or management reorganisation by the Board of Directors. This is a misallocation of real resources. We begin our description of this model by considering the infinite horizon case where the smart money is indifferent as to *when* the actual price moves to its fundamental value.

TIMING OF ARBITRAGE PROFITS IN A PERFECT CAPITAL MARKET

If the smart money can borrow and lend unlimited amounts then she does not care how long it takes a mispriced security to move to its fundamental value. In table 1 we consider a simple case of underpricing where the cost of borrowing r, and the fundamentals return on the security (i.e. dividend return q) are identical at 10%. If the mispriced security moves from $5 to its fundamental value of $6 after only one period, the price including the dividend payout is $6 (1+q) = $6.6 in period-1. At the end of period-1 the arbitrageur has to pay back the loan plus interest, that is $5 (1+r) = $5.5. If the price only achieves its fundamental value in period-2 the arbitrageur receives $6 $(1+q)^2$ = $7.26 at t+2 but has to pay out additional interest charges between t+1 and t+2. However in *present value terms* the arbitrageur has an equal gain of $1 regardless of when the mispricing is irradicated. Also, with a perfect capital market she can take advantage of any further arbitrage possibilities that arise since she can always borrow more money at any time.

Table 1 – Arbitrage Returns – here

In the case of a *finite horizon*, fundamentals and noise trader risk can lead to losses from arbitrage. If suppliers of funds (e.g. banks) find it difficult to assess the ability of arbitrageurs to pick genuinely underpriced stocks, they may limit the amount of funds to the arbitrageur. Also, they may charge a higher interest rate to the arbitrageur because they have less information on her true performance than she herself does (i.e. the interest charge under asymmetric information is higher than that which would occur under symmetric information).

If r = 12% in the above example, while the fundamentals return on the stock remains at 10% then the arbitrageur gains more if the mispricing is eliminated *sooner* rather than *later*. If a strict credit limit is imposed then there is an additional cost to the arbitrageur, namely that if money is tied up in a long-horizon arbitrage position then she cannot take advantage of other potentially profitable arbitrage opportunities.

An arbitrageur earns more potential $ profits the more she borrows and takes a position in undervalued stocks. She is therefore likely to try and convince (signal to) the suppliers of funds that she really is 'smart', by engaging in repeated short-term arbitrage opportunities since long horizon positions are expensive and risky. Hence, smart money may have an incentive to invest over short horizons rather than eliminating long horizon arbitrage possibilities.

The formal model of Shleifer and Vishny (1990) has both noise traders and smart money (see appendix 2). Both ‘short’ and ‘long’ assets have a pay-out at the *same time* in the future but the *true value* of the short asset is *revealed* earlier than that for the 'long' asset. In equilibrium, arbitrageurs' rational behaviour results in greater *current* mispricing of 'long assets', where the mispricing is revealed at long horizons. The terms 'long' and 'short' therefore refer to the date at which the mispricing is revealed (and not to the actual cash payout of the two assets). Both types of asset are mispriced but the long term asset suffers from *greater mispricing* than the short asset.

In essence the model relies on the cost of funds to the arbitrageur being greater than the fundamentals return on the mispriced securities. Hence, the longer the arbitrageur has to *wait* before she can liquidate her position (i.e. sell the underpriced security) the more it costs. The sooner she can realise her capital gain and pay off 'expensive' debts the better. Hence, it is the 'carrying cost' or per period costs of borrowed funds that is important in the model. The demand for the long-term mispriced asset is lower than that for the short-term mispriced asset and hence the *current price* of the long-term mispriced asset is lower than that for the short-term asset.

Investment *projects* which have uncertain payoffs (profits) which accrue in the distant future, may be funded with assets whose true fundamental value will not be revealed until the distant future (e.g. the Channel Tunnel between England and France, where passenger revenues were to accrue many years after the finance for the project had been raised). In this model these assets will be (relatively) strongly undervalued.

The second element of the Schleifer and Vishny (1990) argument which yields adverse outcomes from short-termism, concerns the behaviour of the managers of the firm. They conjecture that managers of a firm have an asymmetric weighting of mispricing. *Underpricing* is perceived as being relatively worse than an equal amount of overpricing. This is because underpricing either encourages the Board of Directors to change its managers or managers could be removed after a hostile takeover based on the underpricing. *Overpricing* on the other hand gives little benefit to managers who usually don't hold large amounts of stock or whose earnings are not strongly linked to the stock price. Hence incumbent managers might under-invest in long-term physical investment projects.

A hostile acquirer could abandon the long-term investment project, hence improve short-term cash flow and current dividends, all of which reduce uncertainty and the likely duration of mispricing. She could then sell the acquired firm at a higher price, since the degree of underpricing is reduced when she cancels the long term project. The above scenario implies that some profitable (in DPV terms) long-term investment projects are sacrificed because of (the rational) short termism of arbitrageurs, who face 'high' borrowing costs or outright borrowing constraints. This is contrary to the view that hostile takeovers involve the replacement of inefficient (i.e. non-value maximizing) managers by more efficient acquirers. Thus, if smart money cannot wait for long-term arbitrage possibilities to unfold they will support hostile takeovers which reduce the mispricing and allow them to close-out their arbitrage position more quickly.

4. CONTAGION

Kirman's (1993) ‘cute’ model is very different to that of De Long *et al* in that it explicitly deals with the interaction between individuals, the rate at which individuals' opinions are altered by recruitment and hence the phenomenon of 'herding' and 'epidemics'. The basic phenomenon of 'herding' was noted by entomologists. It was noted that ants, when 'placed' equidistant from two identical food sources which were constantly replenished, were observed to distribute themselves between each source in an asymmetric fashion. After a time, 80 percent of the ants ate from one source and 20 percent from the other. Sometimes a 'flip' occurred which resulted in the opposite concentrations at the two food sources. The experiment was repeated with one food source and two symmetric bridges leading to the food. Again, initially, 80 percent of the ants used one bridge and only 20 percent used the other, whereas intuitively one might have expected that the ants would be split 50-50 between the bridges. One type of recruitment process in an ant colony is 'tandem recruiting' whereby the ant that finds the food, returns to the nest and recruits by contact or chemical secretion. Kirman notes that Becker (1991) documents similar herding behaviour when people are faced with very similar restaurants in terms of price, food, service etc. on either side of the road. A large majority choose one restaurant rather than the other even though they have to 'wait in line' (queue). Note that here, there may be externalities in being 'part of the crowd' which we assume does not apply to ants.

We have already noted that stock prices may deviate for long periods from fundamental value. A model that explains 'recruitment', and results in a concentration at one source for a considerable time period and then a possibility of a 'flip', clearly has relevance to the observed behaviour of speculative asset prices. Kirman makes the point that although economists (unlike entomologists) tend to prefer models based on optimising behaviour, optimisation is not necessary for survival (e.g. plants survive because they have evolved a system whereby their leaves follow the sun but they might have done much better to develop feet which would have enabled them to walk into the sunlight). Kirman's model of recruitment has the following assumptions :

i) There are 2 views of the world 'black' and 'white' and each agent holds one (and only one) of them at any one time.

ii) There are a total of N-agents and the system is defined by the number $(=k)$ of agents holding the 'black' view of the world.

iii) The evolution of the system is determined by individuals who meet at random and there is a probability $(1-\delta)$ that a person is converted (δ = probability not converted) from black to white or vice versa. There is also a small probability ε that an agent changes his 'colour' independently before meeting anyone (e.g. due to exogenous 'news' or the replacement of an existing trader by a new trader with a different view).

iv) The above probabilities evolve according to a statistical process known as a Markov chain and the probabilities of a conversion from k to $k+1, k-1$ or 'no change' is given by :

$$k \quad \rightarrow \quad \begin{cases} \text{k + 1 with probability } p_1 = p(k,k+1) \\ \text{no change with probability } = 1 - p_1 - p_2 \\ \text{k-1 with probability } p_2 = p(k,k-1) \end{cases}$$

where

$$p_1 = \left[\frac{1-k}{N}\right]\left[\varepsilon + \frac{(1+\delta)k)}{N-1}\right]$$

$$p_2 = \frac{k}{n}\left[\varepsilon + \frac{(1-\delta)(N-k)}{N-1}\right]$$

In the special case $\varepsilon = \delta = 0,$ the first person always gets recruited to the second person's viewpoint and the dynamic process is a martingale with a final position at $k = 0$ or $k = N$. Also, when the probabiliity of being converted (1-δ) is relatively low and the probability of self-conversion ε is high then a 50-50 split between the two ensues. Kirman works out what proportion *of time* the system will spend in each state (i.e. the equilibrium distribution). The result is that the smaller is the probability of spontaneous conversion ε relative to the probability of not being converted δ, the more time the system spends at the extremes that is, 100 percent of people believing the system is in one or other of the two states. (The required condition is that ε < *(1-δ)/(N-1)*, see figure 2). The *absolute* level of δ, that is how 'persuasive' individual's are, is not important here but only that ε is small *relative to* 1-δ. Although persuasiveness is independent of the number in each group, a majority once established will tend to persist. Hence individuals are more likely to be converted to the *majority* opinion of their colleagues in the market and the latter is the major force in the evolution of the system (i.e. the probability that any single meeting will result in an increase in the majority view is higher than that for the minority view).

Figure 2 - 8.6 – kirman here

Kirman (1991), uses this type of model to examine the possible behaviour of an asset price such as the exchange rate which is determined by a *weighted average* of fundamentalists and noise traders' views. The proportion of each type of trader w_t depends on the above evolutionary process of conversion via the Markov chain process. He simulates the model and finds that the asset price (exchange rate) may exhibit periods of tranquility followed by bubbles and crashes as is figure 3. In a later paper Kirman (1993) assumes the fundamentals price p_{ft} follows a random walk and hence is non-stationary, while the noise traders forecast by simple extrapolation. The change in the *market* price ρ_t is:

Figure 3 – Kirman- (8.7) – here

[13] $$\Delta p_{t+1} = w_t \Delta p_{f,t+1} + (1 - w_t)\Delta p_{n,t+1}$$

where the noise trader forecast is:

[14a] $$\Delta p_{n,t+1} = p_t - p_{t-1}$$

and the fundamentals forecast:

[14b] $$\Delta p_{f,t+1} = \upsilon(\bar{p}_t - p_t)$$

is an error correction model around the long run equilibrium $\bar{p}_t$. The weights w_t depend on the parameters governing the rate of conversion of market participants. The weights are endogenous and incorporate Keynes' beauty queen idea. Individuals meet each other and are either converted or not. They then try and assess which opinion is in the majority and base their forecasts on who they think is in the majority, fundamentalists or noise traders. Thus the agent does not base her forecast on her own beliefs but on what she perceives is the majority view. The model is then simulated and exhibits a pattern that resembles a periodically collapsing bubble. When the noise traders totally dominate prices are constant and when the fundamentalists totally dominate prices follow a random walk. Standard tests for unit roots are then applied (e.g. Dickey-Fuller 1979, Phillips-Perron 1988) and cointegration tests between p_t and $\overline{p}_t$ tend (erroneously) to suggest there are no bubbles present. A modification of the test by Hamilton (1989) which is designed to detect points at which the system switches from one process to another was only moderately successful. Thus as in the cases studied by Evans (1991), when a periodically collapsing bubble is present, it is very difficult to detect.

Of course none of the models discussed in this chapter are able to explain what is a crucial fact, as far as public policy implications are concerned. That is to say, they do not tell us how far away from the fundamental price, a portfolio of particular stocks might be. For example, if the deviation from fundamental value is only 5% for a portfolio of stocks, then even though this persists for some time it may not represent a substantial misallocation of investment funds, given other uncertainties that abound in the economy. Noise trader behaviour may provide an *a priori* case for public policy in the form of trading halts, during specific periods of turbulence or of insisting on higher margin requirements. The presence of noise traders also suggests that hostile takeovers may not always be beneficial for the predators since the actual price they pay for the stock of the target firm may be substantially above its fundamental value. However, establishing a *prima facie* argument for intervention is a long way short of saying that a specific government action in the market is beneficial.

5. MOMENTUM AND NEWSWATCHERS

An interesting model involving the interaction of heterogeneous agents is that of Hong and Stein (1999). Momentum traders base their investment decisions only on past price changes while ‘newswatchers’ only observe *private* information which diffuses gradually across the newswatcher population. Both sets of agents are boundedly rational since they do not use all information available. If there are only newswatchers then prices respond monotonically and there is underreaction as news slowly becomes assimilated, but there is no overreaction until we introduce momentum traders. After some good news at time-t, prices rise due to the increased demand of newswatchers. This leads to increased demand by some momentum traders at t + 1 which causes an acceleration in prices and further momentum purchases. Momentum traders make most of their profits early in the ‘momentum cycle’. Momentum traders who buy later (at t + i for some i>1)) lose money because prices overshoot their long run equilibrium and therefore some momentum traders buy after the peak (i.e. there is negative externality imposed on the ‘late’ momentum traders). The dynamics are the outcome of a market clearing equilibrium model but the bounded rationality assumption (e.g. momentum traders only use univariate forecasts and do not know when the ‘news’ arrives) is crucial in establishing both short-horizon positive autocorrelation and long horizon price reversals (i.e. overshooting).

If we allow the introduction of *fully informed* 'smart money' traders into the model, the above conclusions continue to hold as long as the risk tolerance of the smart money traders is finite (When the risk tolerance of the smart money is infinite then prices follow a random walk).

The model therefore yields predictions that are consistent with observed profits from momentum trading (Jegadeesh and Titman 1993, 1999 which may be due to the slow diffusion of initially *private* information. The 'events literature' demonstrates that observed *public* events (e.g. unexpectedly good earnings, new stock issues or repurchases, analysts recommendations) lead to post-event price drift (in the same direction as the initial event) over horizons of 6-12 months. The Hong-Stein model can only generate this price drift if newswatchers after observing public news require additional *private* information in order to translate public news into views about future valuation. Otherwise newswatchers would be able to immediately incorporate the public news into prices. Hence, the model is capable of predicting a different price response to private information than to public information.

The basic structure of the model is as follows. The newswatchers purchase a risky asset that pays a single dividend D_T at $t = T$ where:

$$D_T = D_O + \sum_{j=o}^{T} \varepsilon_j$$

where the $\varepsilon_j 's$ are $iid(0, \sigma^2)$. There are z groups of newswatchers so that each $\varepsilon_j = \varepsilon_j^1 + \varepsilon_j^2 + ... + \varepsilon_j^z$. At time-t, newswatcher group-1 observes ε_{t+z-1}^1, group-2 observes ε_{t+2-1}^2 etc. so each group observes a fraction 1/z of the innovation ε_j. Then at t+1, group-1 observes ε_{t+z-1}^2, group-2 observes ε_{t+z-1}^3 etc. so each subinnovation of ε_{t+z-1} has now been seen by a fraction 2/z of newswatchers. Hence ε_{t+z-1} becomes totally public by time $t + z - 1,$ but the diffusion of information is slow. On average all the newswatchers are equally well informed. Each newswatcher has CARA utility and at time-t their asset demands are based on a static optimisation (i.e. buy and hold until T) and they do not condition on past prices. Without loss, the riskless rate is assumed to be zero and the newswatchers have an infinite horizon, T.

The momentum traders have CARA utility but have finite horizons. At each time t a new group of momentum traders enters the market and holds their position until $t + j$. Momentum traders base their demand on predictions $P_{t+j} - P_t$ and make forecasts based on ΔP_{t-1} only (i.e. univariate forecast). The order flow demand F_t from momentum traders is

$$F_t = A + \phi \Delta P_{t-1}$$

There are *j* momentum traders at *t* and their demand which is absorbed by the newswatchers (who act as market makers) is:

$$Q - jA - \sum_{i=1}^{j} \phi \Delta P_{t-i}$$

where *Q* = fixed supply of assets. It can then be shown that equilibrium prices (which clear the market) are given by

$$P_t = D_t + \{(z-1)\varepsilon_{t+1} + (z-2)\varepsilon_{t+2} + ...\varepsilon_{t+z-1}\}/z - Q + jA + \sum_{i=1}^{j} \phi \Delta P_{t-i}$$

where ϕ in equilibrium can be shown to depend on the coefficient of risk tolerance $1/\gamma$, of the momentum traders. The model is solved numerically and sensitivity analysis provides the following qualitative results.

i) overshooting is greatest for the holding period horizon *j* = 12 months, where the overshooting is around 34%

ii) As risk tolerance increases, momentum traders respond more aggressively to past price changes, equilibrium ϕ increases and there is greater overshooting

iii) As the information diffusion parameter *z* increases, the newswatchers become 'slower', momentum traders are more aggressive (i.e. ϕ increases) and their profits higher, as short run continuation is more pronounced. Also, overshooting increases leading to larger negative autocorrelations in the reversal phase

The result in (i) is consistent with the findings of Jegadeash & Titman (1993, 1999) where momentum profits occur up to horizons of around 12 months. The last result (iii) provides a further test of the model providing we can isolate stocks where information diffusion is likely to be relatively slow. Hong, Lin & Stein

(2000) use ‘firm size’ and ‘analysts (residual) coverage’ (i.e. coverage after correcting for firm size) as proxies for slow information dissemination. They find that six month momentum profits decline with market cap and with increased analysts’ coverage. Also, in low analyst coverage stocks, momentum profits persist for horizons of about 2 years as opposed to less than 1 year in high coverage stocks.

The above model gets one a long way with a minimum of assumptions but there are some limitations of the model. The newswatchers are time inconsistent in that they decide their asset demands in a static framework at *t*, but they change their demands as they absorb the demands of the noise traders. The model has underreaction to private news but this does not necessarily imply underreaction to public news (e.g. earnings announcements) found in the data. However if the assimilation of this public news takes time as private agents undertake their respective calculations of its implications for prices, then the model can deliver post event drift. Of course, if momentum traders are allowed to observe this public news, they may trade more quickly in which case there may be no eventual overreaction to public news and hence no price reversals. However, it should also be noted that representative agent models cannot yield predictions of the type (i)-(iii) above (e.g. linking trading horizons with the pattern of autocorrelations, information diffusion and the size of continuation and long run reversals) as they only condition on public news which is immediately available.

6. STYLE INVESTING

In the previous chapter we noted the prevalence of ‘style investing’. Casual empiricism suggest that many mutual funds are based on styles (e.g. value versus growth, momentum, small cap, tech stocks, real estate etc.). It may be the case that investors allocate their wealth across a limited number of styles and are not particularly concerned about the allocation to *individual* stocks, within any given style category. Certainly there are economies of monitoring and transactions cost to style investing, compared with building a portfolio based on the analysis of *individual* stocks. (Bernstein 1995, Swensen 2000). We have already encountered models that rely on the interaction of momentum traders and some form of rational traders or arbitrageurs such as De Long et al (1990) and Hong & Stein (1999) where the demands for stocks by momentum (noise) traders depend on their *absolute past* performance. Barberis & Shleifer (2003) in their model of style investing assume that investors momentum demand for stocks of a particular style-X depend on past returns on X *relative to* past returns on the alternative style-Y (e.g. X = old economy stocks, Y = new economy stocks). Hence, momentum investors move into stocks in style-X and out of style-Y, if past returns on X exceed those on Y. This increases the returns on assets in style-X and decreases the returns on assets in style-Y. There is negative autocorrelation across assets in the two different styles at short lags. But Y eventually rises towards its fundamental value and hence at long lags the autocorrelation between ΔP_{xt} and $\Delta P_{y,t-k}$ (for large-k) will be positive.

There are arbitrageurs or fundamental traders in the model whose demand depends on their estimate of expected returns based on fundamentals (i.e. final dividends). The arbitrageurs also act as market makers for the 'switchers', absorbing their changing demands. The model delivers a market clearing price for all assets but because the fundamental traders are boundedly rational and do not know the time series properties of the change in demand of the switchers, the market clearing price differs from that when there are only fundamentals traders.

From the above description one can see that the broad set up of the equilibrium model has features in common with Hong & Stein (1999) but the different behavioural assumptions of the switchers and fundamental traders does lead to some different predictions. Before we examine the latter, we briefly present the main elements of the model and we derive the equilibrium market price for assets in style-X and style-Y.

There are 2n (=100 say) risky assets in fixed supply (and the risk free asset has infinitely elastic supply and a zero return). Each risky asset is a claim to a single liquidating dividend to be paid at a future time:

$$D_{i,T} = D_{i,O} + \varepsilon_{i,1} + \varepsilon_{i,2} + \dots \qquad \varepsilon_{i,T}$$

where ε_{it} are announced at time *t* and $\varepsilon_t = (\varepsilon_{1,t}, \varepsilon_{2,t} \dots \varepsilon_{2n,t})' \sim N(0, \sum_D)$ and *iid* over time. There are 2 styles X and Y with assets 1 to…n in style-X and *n* + 1 through 2n in style-Y. The return (= price change) to style-X is:

$$\Delta P_{x,t} = P_{x,t} - P_{x,t-1} \qquad \text{where} \qquad P_{x,t} = \sum_{i \in X} P_{i,t}$$

Each asset's shock ε_{it} depends on a market factor (common to both styles), one style factor (either X or Y) and an idiosyncratic cash flow shock, specific to a single asset-i. Each of these factors has a unit variance and is orthogonal to the other factors so that:

$$\sum\nolimits_D^{ij} = \text{cov}(\varepsilon_{it}, \varepsilon_{jt}) = \begin{cases} 1 & i = j \\ \psi_m^2 + \psi_s^2 & i, j \quad \text{in the same style } i \neq j \\ \psi_m^2 & i, j \quad \text{in different styles} \end{cases}$$

The demand by switchers depends on the relative past performance of the two styles. For assets-i in style-X, the number of shares demanded is:

$$N_{it}^{S} = \frac{1}{n}\sum_{k=1}^{t-1}\theta^{k-1}\left(\frac{\Delta P_{x,t-k} - \Delta P_{y,t-k}}{2}\right) = \frac{N_{x,t}^{S}}{n}$$

with $0 < \theta < 1$ giving declining weights on past relative performance. Symmetrically, for assets-*j* in style-Y:

$$N_{j,t}^{S} = \frac{1}{n}\sum_{k=1}^{t-1}\theta^{k-1}\left(\frac{\Delta P_{y,t-k} - \Delta P_{x,t-k}}{2}\right) = \frac{N_{y,t}^{S}}{n}$$

The above assumes that style investors move their demand from one style to another (e.g. value to growth stocks) within the same asset class (i.e. stocks) and do not move funds out of other asset classes (e.g. cash, bonds, FX) when they wish to switch styles. This may be largely true of institutional investors who have fairly constant 'strategic' asset allocations across alternative asset classes. Also transactions costs might imply that a favoured style is financed from sales of one (or a few) badly performing 'styles', rather than many. There may also be 'rules of thumb' that result in natural *twin styles* (e.g. value versus growth), so when one style is doing well the 'twin style' nearly always does badly.

The fundamental traders maximises expected end-of-period utility of wealth using a CARA utility function. Hence when returns are normally distributed, optimal asset demands are linear in expected returns.

$$N_t^F = (1/\gamma)(V_t^F)^{-1}\left[E_t^F(P_{t+1}) - P_t\right]$$

where $V_t^F \equiv \text{var}_t^F(P_{t+1} - P_t)$, $N^F = (N_1, N_2, ...N_{2n})'$, $P_t = (P_1, P_2 ... P_{2n})'$ and γ is the degree of risk aversion. If the fixed supply of the 2n assets is given by the vector $Q = N_t^F + N_t^S$, then substituting $N_t^F = Q - N_t^S$ in the above equation gives:

$$P_t = E_t^F(P_{t+1}) - \gamma V_t^F(Q - N_t^S)$$

Iterating forward and noting that $E_t^F P_T = E_t^F(D_T) = D_t$ then

$$P_t = D_t - \gamma V_t^F (Q - N_t^S) - E_t^F \sum_{k=1}^{T-t-1} \gamma V_{t+k}^F (Q - N_{t+k}^S)$$

where $D_t = (D_{1t}..., D_{2n,t})'$. The covariance matrix V^F is assumed to be time-invariant and have the same structure as the cash flow covariance matrix $\sum_D$. The term $E_t^F(N_{t+k}^S)$ is assumed to be constant so that fundamental traders are boundedly rational and do not calculate the time series properties of N_{t+k}^S but merely absorb some of the demands by the switchers. Dropping all the non-stochastic terms gives:

$$P_t = D_t + \gamma V N_t^S$$

with the price of asset-i in style-X:

$$P_{it} = D_{it} + \left(\frac{\psi}{n}\right) \sum_{k=1}^{t-1} \theta^{k-1} \left(\frac{\Delta P_{x,t-k} - \Delta P_{y,t-k}}{2}\right)$$

where ψ depends directly on γ and on the parameters of the covariance structure of $\sum_D = V$. The price of asset-j in style-Y is the same form as the above equation but with the sign on $(\Delta P_x - \Delta P_y)$ reversed (i.e. symmetry). With only fundamental traders $P_t = D_t$ but with switchers, price deviates from fundamentals and the deviation is persistent if θ is close to 1.

The model is calibrated with $\psi_m = 0.25, \psi_s = 0.5, \theta = 0.95$ and $\lambda = 0.093$ (so that equilibrium return volatility broadly matches US data) and in turn this gives $\phi = (\psi / n)^{-1} = 1.25$. They take *n* = 50 so there are 50 assets in each style, (X and Y) and $D_{io} = 50$ for all-i.

PREDICTIONS

First consider the co-movement of style-returns. If there is a one-time cash-flow shock to style-X (i.e. $\varepsilon_{i,1} = 1, \varepsilon_{it} = 0$ for $t > 1, \forall i \in X$) then P_x follows a long damped oscillation around its (new higher) fundamental value, with P_x initially overshooting its long run equilibrium and then slowly mean-reverting.

This positive autocorrelation at short horizons and negative autocorrelation at long lags is also predicted by other momentum models (e.g. Hong &Stein 1999, De Long et al 1990). The reason for this is straightforward. Good news about cash flows in assets of style-X lead to price rises which stimulate the demand of switchers, pushing prices above fundamental value. A new 'feature' of this style approach is that the prices of assets in style-Y initially fall as they are sold to help finance purchases of assets in style-X (i.e. symmetry effect). This makes style-Y look even worse *relative to* style-X returns, so there is increased momentum sales of style-Y assets and hence Y's prices also overshoot. Note that the price of Y moves without any cash flow news about the stocks in Y and the autocorrelation *across styles* is negative at short horizons. Eventually fundamental traders sell the overpriced stocks in style-X and price moves to its long run fundamental level. Bad news about stocks in style-X or good news about style-Y, would accelerate this process.

So, style-X imposes a negative externality on assets in style-Y, the magnitude of which depends on how investors finance purchases in X. If they sell small amounts from many other different style portfolios then the externality will be less than if they just sell from the single *twin-style* assets in Y.

It follows from the above analysis that if (say) only *one* asset in style-X experiences a one-time positive cash flow, then the price of other assets in style-X will also experience increases (unrelated to cash flows), while stocks in style-Y will again experience a fall in price (again unrelated to cash flows). In the Hong & Stein (1999) and De Long (1990) models good news about asset-i $(i \in X)$ only affects the price of asset-i and not the prices of other assets in the same style $(i \in X)$, nor assets belonging to style-Y. The predictions of these models differ in this respect.

The above results are consistent with the success of investing in small caps between 1979 (Banz 1979) to around 1983, after which returns on small caps were poor. Data snooping could also set off changes in returns unrelated to cash flow news, while the poor returns on some styles (e.g. value stocks in the US in 1998 and 1999), even though cash flows were good (Chan et al 2000) may have been due to an increased demand for stocks viewed as being in *alternative* styles (e.g. large growth stocks).

Investors move into *all* securities in a particular style category, if the past style return has been relatively good. Hence there may be positive co-movement in individual asset returns within a particular style category which is unrelated to common sources of cash flows. This is consistent with the co-movement in prices of closed end funds, even when their net asset values are only weakly correlated (Lee et al 1991) – this would not be predicted by De Long (1990) and Hong & Stein (2003) where *individual* asset returns are driven by underlying cash flows, nor in certain learning models (Veronesi 1999, Lewellyn & Shanken 2002). Of course, other models are capable of explaining some co-movement which is unrelated to cash flows. For example, in Kyle & Xiong (2001), after banks suffer trading losses they may sell stocks to restore their capital base. This is consistent with the financial crisis of 1998 where aggregate stock prices in different countries fell simultaneously (even though different countries had different economic fundamentals) but it is not a strong candidate to explain co-movement in sub-categories of stocks (e.g. small caps).

According to the fundamentals approach, prices of Royal Dutch and Shell which are claims to the same cash flow stream, should move very closely together. But Froot & Dabora (1999) show that Royal Dutch moves closely with the US market, while Shell moves mainly with the UK market. Royal Dutch is traded mostly in the US and Shell mostly in London and hence they may 'belong to' these two 'styles'. Similarly if a stock is added to the S&P500 index (i.e. 'a style') then in the future one might expect it to co-vary more with the S&P500 and its correlation with stocks outside the S&P500 to fall (Barberis et al 2001).

Finally, Barberis & Shleifer show that a momentum strategy based on style, that is buy into styles that have good recent performance, should offer Sharpe Ratios that are at least as good as momentum strategies based on the momentum performance of *individual* assets. This is consistent with the momentum strategies based on industry portfolios (Moskowitz & Grinblatt 1999) and on size-sorted and book-to-market sorted momentum portfolios (Lewellen 2002).

Even when more sophisticated arbitrageurs are introduced into the model so they understand the time-variation in the momentum traders demands N_t^S, this does not necessarily reduce the size and persistence in the mispricing. This arises because as in the model of Delong et al (1990), the arbitrageurs do not sell when price is above fundamental value – they *buy,* knowing that the increasing demand by feedback traders will raise the price even further, after which the arbitrageurs exit at a profit. In other words the arbitrageurs mimic the behaviour of the momentum traders after a price rise (or fall).

The style model we have discussed above provides an analytic framework in which boundedly rational arbitrageurs and momentum style investors inter-react and gives new predictions compared to models based purely on momentum in *individual* stocks.

7. PROSPECT THEORY

Prospect theory is a descriptive model of decisions under risk, based on the psychological experiments of Kahneman & Tversky (1979). They noted that individuals take gambles that violate the axioms of expected utility theory. Their experiments indicate that individuals are concerned about *changes* in wealth (rather than the absolute level of wealth) and they are much more sensitive to losses than to gains – known as *loss aversion*.

In the models below which incorporate the prospect theory approach there are no exogenous noise traders interacting with the smart money. Instead it is a fully optimising approach but agents have a non-standard utility function: lifetime utility depends not only on consumption but on recent gains and losses on risky assets (i.e. the change in stock market wealth). This is the 'narrow framing' assumption since gains and losses on other assets are ignored and even though investors have long horizons, they are worried by *annual* gains and losses.

In a one-period model Benartzi & Thaler (1995) show that loss aversion can produce a high equity premium but in their i*ntertemporal* model Barberis et al (2001) find that they need an additional effect, a form of integral control which they refer to as *prior losses*. (Thaler and Johnson 1990). The idea is very straightforward and intuitively plausible, namely that if you have suffered losses over *several* previous periods (i.e. cumulatively), then a loss in the current period will be relatively more painful, so that your risk aversion increases. The converse also applies, so that after a series of gains (e.g. in a casino) individuals become more willing to gamble, since they are now 'playing with the house money' – that is the Casino's money, not their 'own'. Barberis et al (2001) show that loss aversion plus 'prior losses' can explain the stylised facts of a high equity premium and high volatility of stock returns, the low level and volatility of interest rates and the predictability in stock returns (e.g. a low price-dividend ratio leads to higher future returns).

Clearly the idea of time varying risk aversion is similar to that in Campbell-Cochrane habit persistence model, but it is cumulative changes in wealth rather than cumulative levels of past exogenous consumption, that lead to changing risk aversion. As we shall see, Barberis *et al* use the calibration approach and show that with 'reasonable' parameter values, the model yields time series behaviour for stock returns and the interest rate that are consistent with the stylised facts.

Empirical evidence indicates a tendency for investors to sell winning rather than losing stocks (Shafrir & Statman 1985, Odean 1998) and this is broadly consistent with prior loss and loss-aversion, although these studies assume utility gains and losses only occur when they are *realised* via a sale of stock. In Barberis et al utility gains and losses occur even when gains and losses are not realised.

THE MODEL

There are two assets, a risk-free rate in zero net supply paying a gross rate of $R^*_{f,t}$ and a risky-asset paying R^*_{t+1} (between t and t+1). The risky asset has a total supply of 1 unit and is a claim on a dividend sequence $\{D_t\}$. There is a continuum of infinitely lived agents each endowed with one-unit of the risky asset at t=0, which they hold at all times. In 'Economy I' agents consume the dividend stream (Lucas 1978) and aggregate consumption C_t equals aggregate dividends D_t which therefore have the same stochastic process (ie. *iid* lognormal):

[15] $$c_{t+1} - c_t = g_c + \sigma_c\,\varepsilon_{t+1} \qquad\qquad \varepsilon_{t+1} \sim niid(0,1)$$

In 'Economy II' consumption and dividends are separate stochastic processes which are still *individually iid,* but with different means and standard deviation and non-zero correlation between their respective errors – see below.

Utility is a power function of consumption, with an additional function $v(.)$ reflecting the dollar gain or loss X_{t+1} experienced between t and t+1.

[16] $$U = E_t\left[\sum_{t=0}^{\infty}\theta^t \frac{C_t^{1-\gamma}}{1-\gamma} + b_t\theta^{t+1}v(X_{t+1}, z_t)\right]$$

where $\gamma>0$ is the coefficient of relative risk aversion over consumption, $\upsilon(X_{t+1}, z_t)$ is the utility from gains and losses and z_t represents prior gains or losses. If there are no prior gains or losses z_t=1 (which is further explained below). The term b_t is an exogenous scaling factor. An annual horizon is chosen. If S_t = \$100 is the reference level then

[17] $$X_{t+1} = S_t(R^*_{t+1} - R^*_{f,t})$$

so for $R^*_{t+1} = 1.20$ and $R^*_f = 1.05$ then $X_{t+1} = \$15$. The Kahneman–Tversky (1979) loss aversion utility function chosen is a piecewise linear function:

[18] $$v\left(X_{t+1}, z_t = 1\right) = \begin{cases} X_{t+1} & X_{t+1} \geq 0 \\ \lambda X_{t+1} & X_{t+1} \leq 0 \end{cases}$$

where $\lambda > 1$ to reflect loss aversion and here there are no prior gains or losses. This utility function is consistent with the experimentally observed risk aversion for small wealth bets. The usual smooth utility functions (e.g. power) calibrated to match individuals' risk aversion over small bets, lead to absurd results over larger gambles. For example, Rabin (1999) shows that an expected utility maximiser who turns down a 50:50 bet of losing $100 or gaining $110, will also turn down a 50:50 bet of losing $1,000 and gaining *any* (including an infinite) amount of money. Loss-aversion avoids this 'Rabin paradox'.

In Barberis et al, it is the expected utility of $\upsilon(X_{t+1})$ that is important and they assume equal subjective probabilities of gains and losses (rather than Kahnemann & Tversky's (1979) non-linear transformation of these probabilities). From [17] and [18] utility depends on returns

[19] $$v(X_{t+1}, z_t = 1) = v\left[S_t(R^*_{t+1} - R^*_{f,t})\right] = S_t \hat{v}(R^*_{t+1})$$

where

[20] $$\hat{v}(R^*_{t+1}, z_t = 1) = \begin{cases} R^*_{t+1} - R^*_{f,t} & \text{for} \quad R^*_{t+1} \geq R^*_{f,t} \\ \lambda(R^*_{t+1} - R^*_{f,t}) & \phantom{\text{for}} \quad R^*_{t+1} < R^*_{f,t} \end{cases}$$

The scaling term b_t in [16] is required so that as wealth S_t increases over time, it does not dominate the utility function. We can use either *aggregate* consumption or wealth as the scaling factor and Barberis *et al* choose the former (which is exogenous to the individual investor).

[21] $$b_t = b_o \overline{C}_t^{-\gamma} \qquad b_o > 0$$

For $b_o = 0$, utility in [16] reverts to the familiar power function over consumption only, while the larger is b_o the greater the weight given to utility from wealth changes (i.e. returns) rather than to consumption.

PRIOR OUTCOMES

The idea that risk aversion is lower (higher) after a sequence of gains (losses) comes from responses to survey questions (Thaler and Johnson 1990) such as:

1. You have just won \$30. Choose between:
   a) a 50% chance to gain \$9 and a 50% chance to lose \$9 [8%]
   b) no further gain or loss [1%]

2. You have just lost \$30. Choose between:
   a) a 50% chance to gain \$9 and a 50% chance to lose \$9 [3%]
   b) No further gain or loss [6%]

The percentage of respondents choosing each option are shown in parenthesis and therefore you are much less willing to gamble with the 'house money' after you have lost \$30.

Gertner (1993) also shows that this 'playing with the house money' effect also works for larger bets, where the participants in a TV game show have to place bets on whether the next card drawn at random will be higher or lower than the card currently showing. Linville & Fisher (1991) using survey evidence find that people prefer unpleasant events to occur far apart rather than close together and they also prefer a 'bad' followed by a 'pleasant' event to occur close together so one cushions the other.

To implement the descriptive notion of prior outcomes, we require a *historic benchmark level* Z_t for the risky asset and then $S_t - Z_t$ measures how much you are 'up' or 'down'. When $S_t > Z_t$ the investor becomes less risk-averse than usual. The state variable measuring prior outcomes is $z_t = Z_t / S_t$ and a value of z_t<1 represents substantial prior gains and hence less risk-aversion. Utility is determined by X_{t+1} and z_t so we now have $v(X_{t+1}, z_t)$. The functional form for $v(X_{t+1}, z_t)$ is simple but ingenious and incorporates a) utility loss depends on the size of *prior* gains or losses b) prior gains are penalised less (in utility terms) than prior losses. We set $R_f^* = 1$ for simplicity (see Barberis et al 2001 for the case of $R_f^* \neq 1$) and we split returns into two parts, relative to a benchmark level.

Suppose we begin with a benchmark level of $Z_t = \$90$ and $S_t = \$100$, so $z_t = 0.9$ and we have prior gains. If stock prices fall in t+1 to $S_t R_{t+1}^* = \$80$ we do not penalize all of the loss of \$20 by $\lambda = 2.0$ say. The loss from \$100 to the benchmark $Z_t = \$90$ is only penalized with a weight of 1 and the loss below the benchmark (i.e. from \$90 to $S_{t+1} = \$80$) is penalized at $\lambda = 2.0$. Hence the overall disutility of the \$20 loss is:

[22] *Change in Utility* = (90 – 1) (1) + (80-90) 2.0 = -30

[23] *Change in Utility* $= (Z_t - S_t)(1) + (S_t R^*_{t+1} - Z_t)\lambda$

$$= S_t(z_t - 1)(1) + S_t(R^*_{t+1} - z_t)\lambda$$

If losses are small enough $S_t R^*_{t+1} > Z_t$ or equivalently $R^*_{t+1} > z_t$, then the entire loss is only penalized at the lower rate of 1. Consider the impact on utility in two cases: prior losses and prior gains.

CASE A: PRIOR GAINS: $z_t \le 1$

[24] $$v(X_{t+1}, z_t) = \begin{cases} S_t(R^*_{t+1} - 1) & R^*_{t+1} \ge z_t \\ S_t(z_t - 1) + \lambda S_t(R^*_{t+1} - z_t) & R^*_{t+1} < z_t \end{cases}$$

Hence if $S_t = \$100, Z_t = \$90$ and $S_t R^*_{t+1} = \$90$ then $R^*_{t+1} = z_t = 0.9$ and current returns of $R^*_{t+1} - 1$ = –10% just offset the historic prior gains of 10% and receive a utility loss weight of 1. But if $S_t R^*_{t+1} = \$80$ then $R^*_{t+1} = 0.8 < z_t = 0.9$ and the current loss of 20% implies the value of $S_t R^*_{t+1} = \$80$ is below its historic reference level of Z_t =\$90 and the 'pain' in terms of utility loss is greater (i.e. λ comes into play in [24]).

CASE B: PRIOR LOSSES $z_t > 1$

A similar scenario applies to when we begin from a situation of prior losses, except we assume that any *further* loss, inflicts *even greater* pain so that here we make λ an increasing function of prior losses:

[25] $$\lambda(z_t) = \lambda + k(z_t - 1) \qquad \text{for } z_t > 1 \text{ and } k > 0$$

and

[26] $$v(X_{t+1}, z_t) = \begin{cases} X_{t+1} & X_{t+1} \ge 0 \\ \lambda(z_t) X_{t+1} & X_{t+1} < 0 \end{cases}$$

Let $\lambda = 2$ and suppose k is set at 10. Let Z_t =\$110 and $S_t = \$100$ which implies prior losses of \$10 and z_t =1.1. So when z_t increases from 1.0 to 1.1 (i.e. higher prior losses) the pain of *additional* losses are now penalized at $\lambda(z_t)$ = 2.0+10(0.1) = 3 rather than $\lambda = 2.0$. The only further 'realistic' requirement is that z_t should move sluggishly relative to S_t so that when S_t rises (falls) then Z_t rises (falls) *but by less than S_t*.

[27] $$z_{t+1} = \eta\left(z_t \frac{R}{R_{t+1}}\right) + (1-\eta) \qquad 1 > \eta > o$$

where R = constant and η measures the degree of sluggishness. Now η can be varied, so that $\eta = 0$ implies $z_{t+1} = 1$ so that Z_t tracks S_t one-for-one while $\eta = 1$ implies z_{t+1} responds sluggishly. For $\eta \neq 1$, when $R^*_{t+1} > R$ then z_{t+1} falls relative to z_t

Optimisation

The intertemporal optimisation problem is

[28] $$\max\ E_t\left\{\sum_{t=0}^{\infty}\left[\theta^t\left(\frac{C_t^{1-\gamma}}{1-\gamma}\right) + b_o\theta^{t+1}\overline{C}_t^{-\gamma} S_t\ v(X_{t+1}, z_t)\right]\right\}$$

where $X_{t+1} = S_t(R^*_{t+1} - R^*_{f,t})$ and the 'wealth' constraint (with no labour income) is:

[29] $$W_{t+1} = (W_t - C_t)R^*_{f,t} + S_t(R^*_{t+1} - R^*_{f,t})$$

and the relationships in [26] and [27]. The model is calibrated and then simulated to see if it produces the stylised facts in the real world data, with particular reference to the equity premium puzzle. Some of the parameters are chosen using historic average values, some based on behavioural studies while others are chosen to 'match' certain properties found of the data. The 'baseline' solution is obtained numerically, using consumption growth g_c = 1.84% p.a. with σ_c = 3.79%, risk aversion (over consumption) $\gamma = 1.0$ and the time preference rate θ=0.98. These parameters ensure that the equilibrium risk free rate $R^*_f - 1$ = 3.86, (see [31]) equals that found in the data and hence ensures there is no 'risk-free rate puzzle'.

A range of values for k are used, with increasing values of k indicating the increased pain of a loss when it follows earlier losses. For example k =3 allows average loss aversion to remain around 2.25. Suppose z_t is initially equal to one and the stock market falls by 10%. Then with η =1 for example, z_t moves to 1.1 and any additional losses are penalized at 2.25 + 3 (0.1) = 2.25 – only a slight increase in 'pain' (see [25]). But with the pain of a loss k = 50, the above implies that if the stock market falls by 10%, then (with η = 1) any additional losses are penalised with a weight of 2.25 + 5 = 7.25. As we see below, increasing k tends to increase the simulated equity premium and bring it closer to that observed in the data. The constant 'R' is set at a level to ensure the unconditional mean of z_t = 1. Simulation results are given for various values of b_0 , since we have no priors on the likely value for this parameter.

The loss aversion parameter λ is taken to be 2.25. The parameter η controls the persistence in z_t and hence the persistence in the price-dividend ratio: η =0.9 (i.e. sluggish response) is chosen so that the simulated price-dividend ratio has autocorrelation properties close to that found in the data.

The numerical solution procedure is complex and uses an iterative technique since the state variable z_{t+1} is a function of both the dividend-price ratio and ε_{t+1} and is of the form $z_{t+1} = h(z_t, \varepsilon_{t+1})$. Using Monte Carlo simulation, 10,000 draws of ε_{t+1} give a series for z_{t+1} which in turn can be used to simulate returns and the dividend-price ratio.

RESULTS: 'ECONOMY I'

When dividends have exactly the same stochastic process as consumption, results on the average equity premium and its volatility are not very impressive. For $b_0 = 2, k = 3$ the simulated equity premium is 0.88% p.a. (s.d. = 5.17% p.a.) while in the real data these are 6.03 (s.d. =20.02). Even when b_0 is increased to 100 (giving the prior loss part of the utility function more weight) and k to 50 (ie. higher loss aversion for any given prior losses), the simulated mean equity premium only rises to 3.28% p.a. (s.d. = 9.35).

RESULTS: 'ECONOMY II'

We now turn to 'Economy II' where dividends follow:

$$d_{t+1} - d_t = g_d + \sigma_d \varepsilon_{t+1} \qquad \textit{with} \quad g_d = 1.84\% p.a. \quad \textit{and} \quad \sigma = 12.0\% p.a.$$

and the correlation between shocks to consumption and dividend growth is taken to be 0.15. The model generated 'statistics of interest' are shown in table 2 along with their empirical counterparts.

[Table 2 – Barberis- here]

For $b_0 = 2, k = 3$ the model generates an equity premium of 2.62% p.a. (s.d. = 20.87) which is better than for 'Economy I' but still less than the empirically observed 6.03 (s.d. =20.02). If we increase the pain due to prior losses so that $b_0 = 2$ *but now* $k = 10$ the model delivers an equity premium of 5.02%p.a. (s.d. = 23.84) much closer to that in the real data, but accompanied by only a modest increase in the average level of loss aversion from 2.25 to 3.5.

The solution for the stock return can be written:

$$R_{t+1} = \frac{1 + f(z_{t+1})}{f(z_t)} e^{g_d + \sigma_d \varepsilon_{t+1}}$$

so allowing a separate process for dividends with $\sigma_d = 12\%$ (whereas $\sigma_c = 3.79\%$) provides the extra volatility in stock returns. Intuitively, the higher volatility arises because if there is a positive dividend innovation, this leads to higher stock prices and a higher return. But this increases prior gains so the investor is less risk averse, which lowers the rate at which future dividends are discounted thus leading to even higher prices and greater movement (volatility) in returns. (The reverse applies for a negative dividend innovation, with the added 'kicker' if there are prior losses). Since returns are more volatile on average, and the investor experiences more losses, then the loss aversion requires a higher equity premium. Note that without the assumption of prior outcomes (i.e. set z_t to zero) the model with b_0 = 2 generates a very small equity premium of around 2.0% (std dev = 12%) and this only rises to 2.88% (s.d. = 12%) for b_0 =100.

The model uses as one input a low correlation of innovations in dividends and consumption growth of 0.15 – close to that in the actual data of around 0.1. The model then generates a low correlation between consumption and *stock returns.*. This is because returns respond to dividend news and any change in risk aversion due to changes in returns. In the model both of the latter are largely driven by shocks to dividends which have a low correlation with consumption – hence returns and consumption are only weakly correlated in the model (and in the real world data). This low correlation between consumption growth and stock returns is not present in the Campbell-Cochrane (1999) habit persistence model where changes in risk aversion and hence returns are driven by consumption, implying a high correlation between these variables.

Long horizon predictability also arises from the slowly changing degree of risk aversion, due to the sluggish response to prior gains or losses (see table 2). A positive dividend innovation leads to rising prices, hence lower risk aversion and even higher prices. The price-dividend ratio will now be high. The investor is now less risk averse and therefore subsequent desired returns will be lower. Hence the price dividend ratio helps predict future returns, which is consistent with the empirical work of Campbell-Shiller (1988). Hence, investor's risk aversion changes over time because of prior losses or gains, so expected returns also vary over time in the model, which leads to predictability. The price-dividend ratio in the model is also highly persistent and autocorrelated (table 2).

The model also produces negatively autocorrelated returns since high prices (returns) lead to lower risk aversion and hence lower returns in the future (table 2). Negatively autocorrelated returns imply long-horizon mean reversion (e.g. Poterba and Summers 1988, Fama-French 1988, Cochrane 2001 – as noted in earlier chapters).

What about the risk-free rate? In this model the risk-free rate is decoupled from the prospect theory portion of the utility function and is given by the standard Euler equation:

[30] $$1 = \theta R_f^* E_t\left[(C_{t+1}/C_t)^{-\gamma}\right]$$

and in equilibrium:

[31] $$\ln R_f^* = -\ln\theta + \gamma g_c - (\gamma^2/2)\sigma_c^2$$

The mean of the risk-free rate and its volatility are primarily determined by (γg_c) and the volatility of consumption growth, where g_c = 1.84% and σ_c = 3.79%. Hence, the model implies a low value for the mean of the risk-free rate, with relatively low volatility. Although the Euler equation for R_f is decoupled from the prospect theory utility function, it is the latter that allows more volatility in the return on the *risky* asset. The Euler equation is:

[32] $$1 = \theta E_t\left[R_{t+1}^*(\overline{C}_{t+1}/\overline{C}_t)^{-\gamma}\right] + b_o\theta E_t\left[\hat{v}(R_{t+1}^*, z_t)\right]$$

and the average return on the risky asset depends on changes in wealth relative to prior losses, as well as the 'standard' consumption growth term.

Note in table 2, that the model explains the average price-dividend ratio but not its volatility which has a standard deviation of 2.25%-2.5% in the model but 7.1% in the data. Additional state variables (e.g. consumption relative to habit) would increase the volatility of the models' price-dividend ratio but this is left for further work.

Is it the loss aversion parameter λ or the prior loss parameter z_t which generates the key results in 'Economy II' ? The solution for returns (given $z_t = 0$) is

$$R_{t+1} = \frac{P_{t+1} + D_{t+1}}{P_t} = \frac{1 + f_{t+1}}{f_t} \, e^{g_d + \sigma_d \varepsilon_{t+1}}$$

where f_t is the dividend-price ratio, which does not now depend on z_t. The volatility in R_{t+1} now depends only on the volatility of log dividend growth σ_d of 12% , which is not sufficient to match the data. Hence prior losses are needed in the utility function, in order to change the degree of loss aversion and hence move expected returns more than cash flows by allowing changes in the discount rate. So the key factor in this model is that returns must move more than cash flows. Note however that changes in loss aversion are not the only possible reason for this result. One could also have changing perceptions of risk or 'over-reaction' to dividend news or perhaps, learning about some key parameters in the model (e.g. the mean rate or volatility of dividend growth) – these are issues for future research.

The basic conceptual ideas behind this prospect theory model are rather similar (but not identical) to those in the Campbell-Cochrane (1999) habit persistence model. Both models require utility to depend on a state variable relative to its recent past. For Campbell-Cochrane this is excess consumption $S_t = (C_t - X_t)/X_t$ and changing risk aversion is most sensitive when C_t is close to X_t (i.e. close to habit consumption). In Barberis et al it is current stock returns relative to recent prior gains or losses (i.e. the z_t variable) that is important. In both models, the variables S_t and z_t are assumed to be sluggish, since some 'sluggishness' is required to 'fit' the observed persistence in the dividend-price ratio, which gives rise to long-horizon predictability.

In both models risk aversion changes over time depending on their respective state variables. Campbell-Cochrane use only consumption (relative to habit) in the utility function while Barberis et al have utility depending on consumption and (prior) returns (i.e. changes in wealth). The Barberis *et al* model has a low value for $\gamma = 1.0$ but you need the additional (prior) loss aversion term in the maximand to fit the stylized facts found in the data. The Campbell-Cochrane model requires a high degree of risk aversion (which depends on consumption relative to habit) but the utility function is 'more conventional' and parsimonious. Broadly speaking, since these are calibration models and both broadly mimic the stylised facts, which model you favour depends in part on how realistic you find the assumptions of each model. But that takes us outside the realms of positive economics. These are both interesting models and maybe they can be usefully extended to include other elements such as learning behaviour and even the addition of some non-rational traders.

The reader should be aware that there are now a plethora of behavioural models incorporating for example: systematic forecasting errors for earnings using public information (Barberis, Shleifer and Vishny 1998), initial overconfidence about one's forecasts based on private information which is tempered by the arrival of public information (Hershleifer and Subrahmanyam 1998) and the interaction of rational 'newswatchers' and momentum traders (who only look at past prices) – Hong and Stein 1999. An excellent overview can be found in Barberis and Thaler (2001).

8. SUMMARY

- There has been an explosion in the behavioural finance in recent years resulting in a wide variety of models which attempt to explain observed anomalies (e.g. closed end fund discounts, momentum profits) and the 'stylised facts' (e.g. equity premium, predictability in stock returns). The explicit models are usually not tested by using regression techniques but by some form of calibration and simulation.

- Some behavioural models amend the standard utility function to include variables other than consumption, for example changes in wealth due to stock market fluctuations and investors may suffer from 'loss aversion'..

- Other behavioural models concentrate on the interaction between rational or smart money traders and noise traders who (often) are assumed to base investment decisions on past price movements. The timing of private and public information on future cash flows between 'newswatchers' and momentum traders is often crucial in producing short term momentum profits and long term price reversals and hence predictability. Cross correlations amongst stocks (whose cash flows are not correlated) can be rationalized in a mode of *style investing.*

- The De Long et al model has rational agents and noise traders maximizing expected end of period wealth but the noise traders can be irrationally optimistic or pessimistic about future returns. This noise trader risk implies that equilibrium price can be permanently below the fundamentals price (even when the noise traders have the same expectations as the rational traders).

- The Schleifer-Vishny model highlights the interaction between the costs of borrowing (to purchase shares) and the time information about a firm's future prospects *is revealed*. Mispricing is greater, the longer the time it takes to reveal to the market, the success of the firms investment decisions. This might encourage short-termism by managers in their choice of investment project.

- Some noise trader models such a Kirman (1993) rely on contagion and conversion of opinion to generate rapid and large price movements, which are consistent with the observed data.

- Prospect theory assumes individuals care about gains and losses, as well as the level of consumption and investors suffer from loss aversion. In Barberis et al (2001) individuals maximize lifetime utility from consumption and changes in wealth. The model is then calibrated (from observed consumption and dividend data) and simulated. The predictions of the model 'fit' a number of stylized facts, which include a low level and volatility in the risk free rate, long horizon predictability of stock returns and a relatively high equity premium (but lower than that observed in the real data).

    - There has been tremendous progress in producing a wide variety of 'behavioural models', which will continue to influence our views of the underlying causes of observed phenomena such as excess volatility, the equity risk premium, stock return predictability and the anomalies literature.

APPENDIX I
THE DE LONG ET AL MODEL OF NOISE TRADERS

The basic model of DeLong et al (1990) is a two period overlapping generations model. There are no 1st-period consumption or labour supply decisions: the resources agents have to invest are therefore exogenous. The only decision is to choose a portfolio in the 1st-period (i.e. when young) to maximise the expected utility of end of period wealth. The 'old' then sell their risky assets to the 'new young' cohort and use the receipts from the safe asset to purchase the consumption good. The safe asset *s* is in perfectly elastic supply. The supply of the uncertain/risky asset *u* is fixed and normalised at unity. Both assets pay a known real dividend r (riskless rate) so there is no fundamental risk. One unit of the safe asset buys one unit of the consumption good and hence the real price of the safe asset is unity.

The proportion of noise traders NT is μ, with $(1-\mu)$ smart money SM operators in the market. The SM correctly perceives the distribution of returns on the risky asset at t+1. NT can be 'bullish' or 'bearish' and misperceive the true price distribution. The NT *average* misperception of the expected price is denoted $\rho *$ and at any point in time the *actual* misperception ρ_t is:

[A1] $$\rho_t \sim N(\rho*, \sigma^2)$$

Each agent maximises a constant absolute risk aversion utility function in end of period wealth, W :

[A2] $$u = -\exp(-2\gamma W)$$

If returns on the risky asset are normally distributed then maximising [A2] is equivalent to maximizing

[A3] $$EW - \gamma\sigma_w^2$$

where EW = expected final wealth, γ = coefficient of absolute risk aversion. The SM therefore chooses the amount of the risky asset to hold, λ_t^s by maximizing

[A4] $$E(U) = c_o + \lambda_t^s[r + {}_tP_{t+1}^e - P_t(1+r)] - \gamma\left(\lambda_t^s\right)^2 {}_t\sigma_{pt+1}^2$$

where c_0 is a constant and ${}_t\sigma_{pt+1}^2$ is the one period ahead conditional expected variance of price:

[A5] $\sigma^2_{pt+1} = E_t(P_{t+1} - E_t P_{t+1})$

The NT have the same objective function as the SM except her expected return has an additional term $\lambda_t^n \rho_t$ (and of course λ_t^n replaces λ_t^s in [A4]). These objective functions are of the same form as those found in the simple two-asset, mean-variance model (where one asset is a safe asset). **(check KC still included?)** Setting $\partial E(U)/\partial \lambda_t = 0$ in [A4] then the objective function gives the familiar mean-variance asset demand functions for the risky asset for the SM and the NTs

[A6] $\lambda_t^s = E_t R_t \;/\; 2\gamma({}_t\sigma^2_{p+1})$

[A7] $$\lambda_t^n = \frac{E_t R_{t+1}}{2\gamma({}_t\sigma^2_{pt+1})} + \frac{\rho_t}{2\gamma({}_t\sigma^2_{pt+1})}$$

where $R^e_{t+1} = r_t + {}_tP_{t+1} - (1+r)P_t$. The demand by NTs depends in part on their abnormal view of expected returns as reflected in ρ_t. Since the 'old' sell their risky assets to the young and the fixed supply of risky assets is 1, we have :

[A8] $$(1-\mu)\lambda_t^s + \mu\lambda_t^n = 1$$

Hence using [A6] and [A7], the equilibrium pricing equation is :

[A9] $$P_t = \frac{1}{(1+r)}\left(r +_t P_{t+1} - 2\gamma_t \sigma^2_{pt+1} + \mu\rho_t\right)$$

The equilibrium in the model is a steady state where the *un*conditional distribution of P_{t+1} equals that for P_t. Hence solving [A9] recursively :

[A10] $$P_t = 1 + \frac{\mu(\rho_t - \rho^*)}{(1+r)} \quad + \frac{\mu\rho^*}{r} - \frac{2\gamma\,{}_t\sigma_{pt+1}}{r}$$

Only ρ_t is a variable in [A10] hence :

[A11] $${}_t\sigma^2_{pt+1} = \sigma^2_{pt+1} = \mu^2\sigma^2 /(1+r)^2$$

where from [A1], $\rho_t - \rho^* = N(0, \sigma^2)$. Substituting [A11] in [A10] we obtain the equation for the price level given in the text :

[A12] $$P_t = 1 + \frac{\mu(\rho_t - \rho^*)}{(1+r)} + \frac{\mu\,\rho^*}{r} - \frac{2\gamma\mu^2\sigma^2}{r(1+r)^2}$$

APPENDIX II

THE SHLEIFER-VISHNY MODEL OF SHORT TERMISM

This appendix formally sets out the Shleifer-Vishny (1990) model whereby long-term assets are subject to greater mispricing than short-term assets. As explained in the text this may lead to managers of firms to pursue investment projects with short-horizon cash flows in order to avoid severe mispricing and the risk of a takeover.

There are 3-periods 0, 1, 2 and firms can invest either in a 'short-term' investment project with a $ payout of V_s in period-2 or a 'long-term' project also with a payout only in period-2 of V_g. The key distinction between the projects is that the *true value of* the short-term project becomes *known* in period-1, but the true value of the long-term project doesn't become known until period-2. Thus arbitrageurs are concerned not with the timing of the cash flows from the project but with the *timing* of the mispricing and in particular, the point at which such mispricing is revealed. The riskless interest rate is zero and all investors are risk neutral.

There are two types of trader, noise traders NT and smart money SM (arbitrageurs). Noise traders can either be pessimistic $(S_i > 0)$ or optimistic at time t = 0 about the payoffs V_i from both types of project (i = s or g). Hence *both* projects suffer from systematic optimism or pessimism. We deal only with the pessimistic case (i.e. 'bearish' or pessimistic views by NTs). The demand for the equity of firm engaged in project i (= s or g) by noise traders is :

[A1] $$q(NT, i) = (V_i - S_i) / P_i$$

For the bullishness case q would equal $(V_i + S_i)/P_i$. Smart money (arbitrageurs) face a borrowing constraint of $\$b$ at a gross interest rate R > 1 (i.e. greater than one plus the riskless rate). The SM traders are risk neutral so they are indifferent between investing all $\$b$ in either of the assets-i. Their demand curve is:

[A2] $$q(SM, i) = n_i b / P_i$$

where n_i = number of SM traders who invest in asset-i (= s or g). There is a unit supply of each asset-i so equilibrium is given by:

[A3] $$1 = q(SM, i) + q(NT, i)$$

and hence using [A1] and [A2] the equilibrium price for each asset is given by :

[A4] $$P_i^e = V_i - S_i + n_i b$$

It is assumed that $n_i b_i < S_i$ so that both assets are mispriced at time t = 0. If the SM invests $\$b$, at t = 0, she can obtain b / P_s^e shares of the short-term asset. At t ≠ 1 the payoff per share of the short-term asset V_s is revealed. There is a total \$ pay-off in period-1 of $V_s(b / P_s^e)$. The net return NR_s in period-1 over the borrowing cost of bR is :

[A5] $$NR_s = \frac{V_s b}{P_s^e} - bR = \frac{bV_s}{(V_s - S_s + n_s b)} - bR$$

where we have used equation [A4]. Investing at t = 0 in the long-term asset, the SM purchases b / P_g^e shares. In period t = 1, she does nothing. In period t = 2, the true value V_g per share is revealed which discounted to t+1 at the rate R implies a \$ payoff of $bV_g / P_g R$. The amount owed at t = 2 is bR^2 which when discounted to t+1 is bR. Hence the net return in period-1 NR_g is :

[A6] $$NR_g = \frac{bV_g}{P_g R} - bR = \frac{}{R} \frac{bV_g}{(V_g - S_g + n_g b)} - bR$$

The only difference between [A5] and [A6] is that in [A6] the return to holding the (mispriced) long-term share is discounted back to t=1, since its true value is not revealed until t=2. In equilibrium the returns to arbitrage over one period, on the long and short assets must be equal $(NR_g = NR_s)$ and hence from [A5] and [A6] :

[A7] $$\frac{(V_g / R)}{P_g^e} = \frac{V_s}{P_s^e}$$

Since $R < 1$, then in equilibrium the long-term asset is *more* underpriced (in percentage terms) than the short-term asset (when the noise traders are pessimistic, $S_i > 0)$. The differential in the mispricing occurs because payoff uncertainty is resolved for the short-term asset in period-1 but for the long-term asset this does not occur until period-2. Price moves to fundamental value V_s for the short asset in period-1 but for the long asset not until period-2. Hence the long-term fundamental value V_g has to be discounted back to period-1 and this 'cost of borrowing' reduces the return to holding the long-asset.

TABLE 1
ARBITRAGE RETURNS : PERFECT CAPITAL MARKET

Assumptions :

Fundamental Value	=	$6
Current Price	=	$5
Interest rate, r	=	10% per period
Return on risky asset, q	=	10% per period (on fundamental value)

Smart Money borrows $5 at 10% and purchases stock at t=0.

	Selling Price (including dividends)	**Repayment of Loan**	**Net Gain**	**DPV of Gain (at r=10%)**
Period-1	6(1+q) = $6.60	5(1+r) = $5.50	$1.10	$1
Period-2	$6(1+q)^2$ = $7.26	$5(1+r)^2$ = $6.05	$1.21	$1

TABLE 2

MODEL GENERATED STATISTICS AND EMPIRICAL VALUES ECONOMY II

(BARBERIS ET AL 2001)

	Model Values		Empirical Values
	b_o=2,k=3	b_o=2,k=10	
Excess Stock Return (Equity Premium)			
Mean	2.62	5.02	6.03
Standard Deviation	20.87	23.84	20.02
Sharpe Ratio	0.13	0.21	0.30
Average Loss Aversion	2.25	3.5	-
Price-Dividend Ratio			
Mean	22.1	14.6	25.5
Std.Dev.	2.25	2.5	7.1
Return Autocorrelations			
Lag –1	-0.07	-0.12	0.07
-2	-0.03	-0.09	-0.17
-3	-0.04	-0.06	-0.05
-4	-0.04	-0.04	-0.11
-5	-0.02	-0.03	-0.04
Corr $\{(P/D)_t,(P/D)_{t-k}\}$			
k=1	0.81	0.72	0.70
k=3	0.53	0.38	0.45
k=5	0.35	0.20	0.40
Regression $R_{t,t+k} = \alpha_k+\beta_k(D/S)_t$			
β_1	4.6(2%)	4.4(6%)	4.2 (7%)
β_2	8.3(4%)	7.5(10%)	8.7 (16%)
β_3	11.6(5%)	9.7(12%)	12.3 (22%)
β_4	13.7(6%)	11.5(14%)	15.9 (30%)
(%) = % R-squared of regression			