3. The self-referential hypothesis in finance

André Orléan

According to Eugene Fama, 'a market in which prices always "fully reflect" available information is called "efficient"'. The Efficient Market Hypothesis (EMH) has long dominated financial theory, so much so, that Michael Jensen, one of the major advocates of this theory, claimed in 1978 that 'the efficient market hypothesis is the best established fact in all social sciences' (1978, p.95). It is not certain that he would still be able today to make the same claims with equal conviction. In fact, a startling number of empirical and theoretical studies has emerged which call the theory into question. These studies are so numerous that a summary, even a selective one, would exceed the scope of this article. Instead, by way of introduction, I shall highlight the most important points derived from these studies, through a schematic reading of the rich and turbulent financial history of the last few years.

From this point of view, the crisis of 19 October 1987 undeniably played a significant role in waking intellects from their ‘dogmatic slumber’[1]. The 22.6 per cent fall in the Dow Jones Index – the most significant drop ever recorded in one day, astonishingly inexplicable from the economic information available – served to reveal with searchlight clarity the limitations of the EMH. Furthermore, a few days after this crash, when Robert Schiller (1991) asked investors the reasons for their actions, contrary to what the EMH would have one believe, they did not cite the changes in the real economic environment as an explanation, but rather the 200 point fall in the Index that occurred as the markets opened on 19 October 1987! Soon afterwards, an ‘irrational exuberance’ reigned through March 2000, which precipitated a massive over-estimation of stock prices, greatly exceeding even the over-estimation of 1929. This rise was followed by an equally dramatic downturn, punctuated by short cycles of optimism and pessimism lasting a few months. The whole process served to exemplify the extent to which financial markets can be governed by speculative bubbles, extreme volatility and drastic changes in expectations. Finally, the criteria of corporate management endorsed by the market also varied dramatically. Those same regular cash-flow sector businesses that were absolute sells in 1999 were, a few years later, core portfolio holdings. Analysts rejected overtly the theories that they had defended previously.

The events described above (speculative bubbles, extreme volatility, wavering opinions, whimsical fads and investors showing greater faith in the market and its gurus than in fundamental facts and figures) are universally recognized. They create an image of the financial market that is far from the one painted by the EMH. It is no longer enough to talk about anomalies as we once did when describing the 'PER', 'size' or 'January' effects. The deviations are too frequent and too significant. In fact, the model itself needs to be revised. This is precisely what the behavioural finance theorists seek to do (Schleifer, 2000; Thaler, 1993). This strand of analysis aims to make sense of these 'anomalies'. It works on the basis that the behaviour of investors is not perfectly rational due to the cognitive biases which affect the way information is handled. It is these biases that lead them to make systematic errors of judgment when anticipating the market.

While this is an interesting perspective to take, with a good grounding in cognitive psychology, it is not the particular avenue that this article seeks to explore. According to our own analytical perspective, it is not so much this questioning of participants' rationality that is problematical, but rather the very definition of financial rationality itself. Orthodox financial theory considers stock price to be a reflection of the fundamental value of the firm, where this value is equal to the discounted stream of future expected dividends. It assumes that every participant of the market seeks to estimate this objective value. In such a perspective**,** their rationality can be said to be 'fundamentals-based' or 'fundamentalist', because it aims to understand the way in which fundamental variables such as technologies, preferences and market situations affect the objective profitability of a firm. In my opinion, this interpretation of financial rationality is erroneous. It merely gives an imperfect description of what actually takes place at the Stock Exchange. Far more important for investors than the value of the discounted stream of future expected dividends is the price of shares at any given moment, as decided by market opinion. Each investor seeks to anticipate prices, for it is the accuracy of this crucial prediction that will determine the true earnings of the market participants. It is not necessary to assume that this price is a reflection of the fundamental value, or that all participants in the market base their predictions on this notion, or even that they believe others to be doing so. These are specific hypotheses elaborated by a theoretician working outside the market. From the position of market participants, each seeks simply to be aware of future market opinion, in whatever form it may take. This 'basic' description of the market (Orléan, 2002) does not require ad hoc hypotheses pertaining to the beliefs of all market participants. It can be described as 'self-referential', that is to say, the point of reference for understanding price is an endogenous product of the market itself. The orthodox vision of the market is not a self-referential one in the sense that it considers the market price to be the product of forces external to the marketplace which decide the fundamental value. This vision is 'fundamentalist', or fundamentals-driven. The rationality of the self-referential theory is far removed from the fundamentalist rationality inasmuch as it focuses upon the opinions of other participants who all influence each other. It is this that Keynes was describing in Chapter 12 of The General Theory.

The study of what I have termed 'self-referential finance' will consist in demonstrating that those particular phenomena succinctly discussed above, namely speculative bubbles, extreme volatility and whimsical fads, are completely compatible with the rationality of market participants, so long as the environment is considered to be self-referential. In order to understand these phenomena, one does not need to assume that the participants are irrational. This is not to say that the market and all of its participants function in a perfectly rational manner at all times. As far as this matter is concerned, we are inclined to accept the notions of behavioural finance. Our problem with such an approach stems from our refusal to accept that the source of **all** the phenomena listed above is essentially psychological, resulting from cognitive biases. Even in the absence of these biases, 'anomalies' would still occur. In fact, we would have great difficulty understanding the systematic recurrence of these events otherwise. The financial behaviour which lies at the root of these phenomena must be subject to some very strong, and indeed, rational influences if their continued regular reoccurrence is to be explained, despite the lessons that history has taught us.

This article seeks to justify the self-referential hypothesis and to prove that it gives rise to some original interpretations and to a myriad of previously un-asked questions. In the first section, we shall look at the results of experimental economics in the field of finance. These results are extremely interesting as far as our project is concerned. This is because they show unequivocally that the production of speculative bubbles is engrained in the very fabric of what we have termed the 'basic financial game' (Orléan, 2002). This is independent of all social contexts likely to distort individual beliefs aberrantly. Indeed, bubbles also emerge in controlled experimental conditions, where situations are created in which actors neither speak to each other nor collectively create fads like the 'New Economy' or 'Tulipomania'. The fact that these speculative bubbles appear quite clearly and systematically confirms that they must be seen as a 'normal product' of financial interaction. In other words, they reveal conclusively the underlying rationality. Bubbles are not inexplicable peculiarities; rather they are rational. This is an essential conceptual component of the thesis that we seek to defend.

In the second section, we will see how it is not necessary to abandon the hypothesis of rationality in order to understand these experimental results. What we must question is the much stronger hypothesis that rationality is 'common knowledge'.

If we abandon this latter, altogether unrealistic hypothesis, it becomes apparent that rational behaviour, as far as the participants are concerned, consists of guessing how the other participants will act. This leads us directly back to the self-referential hypothesis, which is laid out in the third section and can be summarized as follows: <u>each participant acts according to a specific model which he or she expects will provide him or her with an accurate explanation of how market opinions are formed.</u> This is the hypothesis in its most simple and naïve form. Furthermore, in accordance with the self-referential approach, it follows logically that we must distinguish between two types of individual belief: the one pertaining to individual opinions on how much a security should be worth, and the one referring to how individuals believe the market will operate in the future. The first type will be called 'personal beliefs', and the second 'market beliefs'. This distinction will be the prime subject matter of the fourth section. It shows that no economic force is capable of pushing market beliefs to fall into line with individual ones. The intuitive <u>bottom-up</u> model which sees market beliefs as aggregates of personal ones is not valid. Indeed, market beliefs retain a certain amount of autonomy, as can be seen in situations where the two types of beliefs systematically diverge. Take for example the case of everyone believing that a security should be valued at X, but simultaneously accepting the market valuation of it at Y. In such a situation, in accordance with the self-referential hypothesis, it is the market value Y that would prevail. This type of belief configuration lies at the basis of bubbles. These bubbles can be described as rational insofar as nobody is mistaken. Indeed, it is possible that as far as fundamentalist criteria are concerned, X is in fact the correct value and that each of the participants realizes this deep down. Nonetheless, they simultaneously believe that the market retains the estimation Y. It is the latter judgment that is finally quoted as the market price, which effectively corroborates all of the market beliefs.

1. The lessons of experimental economics

To begin, we should bear in mind that experimental economics is a recent branch of economics. Smith (1988) seeks to accord the discipline the status of a true experimental science. To achieve this aim, it constructs experimental protocols which are geared to reproducing in a controlled situation the interactions that form the basis for economic theory. For example, an artificial market structure is constructed using a given number of participants in order to examine the ways in which transactions take place. This structure can then be used to test the accuracy of predictions made by economic theories relating to price, information and trade volumes. Experimental economics deals particularly with financial markets and has produced numerous results that disprove the EMH. Indeed, as Noussair and Ruffieux (2001) emphasize in their summary of the research that has been undertaken, the vast majority of financial market experiments has led to the appearance of a speculative bubble followed by a crash. For example, among the earliest results obtained, Smith, Suchanek and Williams (1988) study an asset with a lifespan covering either 15 or 30 time periods.. During each period, the holder of the asset receives a dividend. This dividend is random, and its law of probability is 'common knowledge' for all market participants.[2] When one observes the price series that were derived from this experiment, one sees that in 14 out of 22 occurrences prices rose very high, well above their fundamental value; and then, near the end, a crash brought them plummeting down to their fundamental value.[3] This same scenario can be found in all of the experimental work that ensued. As Noussair and Ruffieux (2001) note, such a result surprised many observers, who created numerous tests to confirm or disprove it. They write that 'after the work of Smith et al. (1988) revealed that speculation tended to prevent prices from tracking the fundamental value of the asset being traded, many studies have been conducted in search of conditions that eliminate the effect of speculation on prices. These attempts have so far been in vain' (p.4, italics added). The appearance of a speculative bubble followed by a crash has thus turned out to be a very robust finding.

This was confirmed again by Noussair, Robin and Ruffieux (2000) in their study of an experimental situation similar to the one devised by Smith et al., except for the fact that here fundamental value is assumed to remain constant. To ensure this, a dividend is drawn each period in a random manner: it is either worth +36, or +4, or -16, or –24 with a uniform probability of 0.25. Its average is therefore 0. At the end of the game, the asset is repurchased at a price of 360 by the experimenter. As such, during the entire experiment, the fundamental value of the asset remains equal to 360. This is an experimental situation that is even simpler than the one studied by Smith, Suchanek and Williams (1988). For Noussair et al., the fact that the fundamental value was decreasing in Smith, Suchanek and Williams' experiment might explain the actors' difficulty in coordinating themselves around this value. A priori, it seems easier for prices to converge towards a fundamental value that is constant. Yet this is not what happened, leading the authors to conclude that 'although the constant fundamental value means that prices only need to converge once during the session to a value that can serve as a focal point over the whole horizon of the experiment, we continue to observe bubbles' (Noussair et al., p.14). In fact, until now, the only experimentally tested condition that has led to the elimination of bubbles occurs when the game is repeated in an identical manner. 'If all subjects have participated previously in at least two sessions with markets with exactly the same structure, bubbles tend not to occur' (p.18). This is a highly restrictive condition, inasmuch as the repetition has to be identical in nature. More specifically, the involvement of new players is sufficient to invalidate the result.

Theoreticians should view the recurrent appearance of bubbles in these experimental configurations as an essential fact. They do not, however, tend to do this. After all, although it is appropriate to question the capacity of the experimental framework to reproduce a true image of actual financial markets, there is no doubt that experimental markets constitute a quasi-perfect transcription of the theoretical models which economists have devised as a means of representing stock market dynamics.[4] It is, therefore, all the more surprising and interesting to note that even in this purified sort of environment, efficiency is contradicted. The integration of such a result has major theoretical consequences, as will be seen below.

In fact, to explain the emergence of speculative bubbles, the following hypothesis has often been proposed[5]: events of this nature are attributable to dynamics of collective exuberance which inflame investors' imaginations and cause them to lose sight of fundamentals. Remember the Dutch Tulipomania, or even more recently, the 1998-2000 infatuation with the 'New Economy'. In other words, a bubble appears because agents make errors of calculation with respect to the fundamental value of firms. Following this hypothesis, it is the irrationality of expectations – itself a product of interpersonal influences – that should be the main cause of price inefficiency.

Experimental economics shows that this hypothesis is incorrect. The configurations studied above neglect the idea of ignorance of fundamental value, and they do not recognize any collective phenomena. Nevertheless speculative bubbles still appear. At no time can players justify higher prices by increases in the fundamental value, because each player knows perfectly well what the fundamental value is. Moreover, this evaluation is not about to be upset by any collective beliefs, since players have no contact with one another (except for the fact that each is aware of current prices). They do not even speak with each other. One might assume that certain players cannot calculate the fundamental value, and that this failure might be partially responsible for the bubble. However, this is not the case either. When players are explicitly told what the fundamental value of the share is for each period, the same findings are preserved. As can be seen, the emergence of speculative bubbles is a highly robust phenomenon that crops up even in situations where no collective belief exists to trouble investors' evaluations. This should lead theoreticians to focus on whether the pricing mechanism itself is a source of inefficiency. If so, the very concept of financial liquidity, namely the fact that stock can be bought and sold at each period of time, should be questioned.

A final comment should convince the reader of the seriousness of the findings derived from experimental economics. After reading the preceding results, readers might possibly feel that this discipline possesses an 'anti-market' bias, that by its very nature, it is hostile to competitive efficiency, and that it is this bias which is responsible for the negative findings that have been obtained. I cannot emphasize strongly enough that there is no truth whatsoever in this conjecture. Whenever experimental economics considers 'traditional' goods (that is, goods with a short life span[6]), and not multi-period assets, it generally reaffirms the efficiency of market procedures, including those situations where the hypotheses of pure and perfect competition have not been entirely fulfilled. It even arrives at this conclusion when the experimentation has placed great demands on the market's ability to compile and diffuse information.

Findings from experimental economics should therefore be taken very seriously. They tell us that market inefficiency is not only a consequence of the uncertainty of the fundamental value, or of the presence of social contexts that are likely to blind investors' individual judgements; inefficiency is also encrusted in the very structure of financial interaction. In our perspective, bubble phenomena may therefore be seen as rational, but in a way that must be specified.

2. What sort of rationality for investors?

For greater clarity, let us consider the experiment analysed by Noussair, Robin and Ruffieux (2000), during which fundamental value was kept constant. How can the appearance of bubbles followed by crashes in this experimental situation be explained? I propose the following hypothesis: even when agents are placed in an environment in which fundamental share value is known, their actions, at moment t, are not based on their estimation of this fundamental value, but on their expectations of what market opinion will be during the time interval $[t+1, t+T]$ where T is their horizon. In Noussair, Robin and Ruffieux's experiment, their own time horizon T does not exceed two or three periods. It is only when the experiment reaches its conclusion that the fundamental value becomes important. This is because they know that at the end of the game, the experimenter will be re-purchasing the share at this price, something that fully suits the preceding hypothesis without any need to assume that agents 'believe in fundamental value'.

The way in which each agent creates his or her own model of market opinion during the entire interaction depends on their past price observations, and also on their initial thoughts about finance and about the psychology of the group participating in the interaction. Overall, this involves a whole array of disparate elements, mixing up economic, sociological and psychological factors. It is known, for example, that references to well-known and pivotal episodes play a major role in financial actors' self-representation of equity phenomena. Shiller (1991) demonstrates that the reference to the 1929 Crash had a major effect on the way in which the Crash of 19 October 1987 developed. This is a partial explanation for the magnitude of the price drop that was observed at the time. Shiller writes that 'investors had expectations before the 1987 Crash that something like a 1929 Crash was a possibility, and comparisons with 1929 were an integral part of the phenomenon. It would be wrong to think that the crash could be understood without reference to the expectations engendered by this historical comparison' (p.399). Moreover, individuals change their personal models of market opinion as a result of what they have discovered and learned. The way in which agents modified their model of market opinion after the 1987 crisis provides an example of such a change. During this crisis, contrary to the events of 1929, economic actors experienced a very important stock crash without suffering any lasting consequences. Since then, it has become apparent that sporadic crashes can occur without necessarily leading to a major macroeconomic crisis. Actors have subsequently modified their expectations of market opinion after a crash. This new model had a major influence during the crisis of 1997. A large number of investors kept calm and decided not to withdraw their funds. Instead, they waited for a market recovery which they considered likely. This optimism was based on the pattern observed in 1987.

Returning to Noussair, Robin and Ruffieux's experiment, it is important to emphasize that there is nothing irrational about not using fundamental value for price forecasting. Quite the contrary, everyone can observe the lasting gaps that can exist between the two numbers. Even the most fundamentalist players have to concede that the market engenders a specific pricing dynamic that has nothing, or at best, very little to do with fundamentals. In the players' view, market opinion is a specific entity imbued with its own life form. Therefore, it is necessary to possess a reliable model of market opinion, or to remain indifferent to short-term changes in prices, if one is to operate in the marketplace.

This conjecture about how financial agents behave is something that is presented as the 'self-referential' hypothesis in the next section. It can be inferred from the important experiences which operators have had in the markets. On the one hand, they have witnessed the ability of markets to create surprises and to disprove fundamental predictions; on the other hand, however 'fantastic' the price may be, it is this price, and this price alone, that will make operators rich or poor. In both cases, market opinions appeared to be autonomous.

Reverting again to Noussair, Robin and Ruffieux's experiments, we need to note the players' marked preference for the hypothesis that prices will tend to rise before returning to their fundamental value. A player who believes in such a model faces the following alternative: either he or she does nothing; or else he or she participates in the bull-run by buying shares and by hoping to resell them at the right time. Here the financial interaction resembles, as Keynes already noted, a game of musical chairs:

> For it is, so to speak, a game of Snap, of Old Maid, of Musical Chairs - a pastime in which he is victor who says Snap neither too soon nor too late, who passes the Old Maid to his neighbour before the game is over, who secures a chair for himself when the music stops. These games can be played with zest and enjoyment, though all the players know that it is the Old Maid which is circulating, or that when the music stops some of the players will find themselves unseated (pp. 155-6).

Taking part in this type of game might seem to be a particularly reasonable thing to do when the players have a strong belief in market liquidity, that is to say when they believe that they will be able to resell the speculative shareholding at any moment in time. As soon as the majority of people believes that an increase in share price will occur, they will jump into the market. This leads to share prices rising, followed by a crash when the game approaches its close. In other words, what is witnessed is the model's self-realization. The self-validating ability of the model is certainly one of the elements that explains the significant role such dynamics play in experimental and actual markets. Note that this self-realization can be enacted in real markets without participants being aware of it. Market players are then induced to see this outcome as a straightforward confirmation of their beliefs, something that is not completely untrue. This type of phenomenon is apt to reinforce belief in the idea that the market has a life of its own, independent of whatever fundamentals may exist.

Nevertheless, the self-referential description creates a problem for theoreticians inasmuch as it does not mesh with what a rational analysis of the game would lead one to expect. To analyse the given situation, a game theorist proceeds as follows. If at the end *T* of the game the experimenter repurchases a share for 360 accounting units, at time *T*-1 an investor will also evaluate the share as being worth 360 accounting units, since by holding this share he or she will be able to obtain 360 units during the following period, and because the expected dividend at *T*-1 is equal to 0. Repeating the same reasoning by means of backwards induction, it is clear that the share price should be equal to 360 (that is, to its fundamental value) during each period. This analysis leads to a conclusion that contradicts the dynamics observed in experimental economics. However, we do know that reasoning by backwards induction is based on extremely strong hypotheses which are open to criticism. Dupuy (1989) contains a very penetrating analysis of this point, based on the work of game theorists such as Aumann (1994), Kreps and Wilson (1982) and Kreps, Milgrom, Roberts and Wilson (1982). I shall examine briefly Dupuy's conclusions, as these will be very useful in later sections.

Dupuy starts by stressing the fact that reasoning by backwards induction is based on a very strong hypothesis, namely, the fact that rationality is common knowledge. Thus, not only are all the players rational, but each knows that all the others are rational as well; and each knows that all the others know that all the others are rational as well; and so on, ad infinitum. He writes that:

> [The] implicit hypothesis underlying most games theory analyses is that rationality is common knowledge. Yet nowadays it is clear that many of the theory's "paradoxical" findings result from this hypothesis – and that their paradoxical nature stems from the fact that they differ radically from common sense solutions that laboratory experiments or daily life can come up with spontaneously (p.380).

The 'paradoxes' which are traditionally studied are Selten's 'Chain-Store Paradox' (1981); 'Prisoners' Dilemma' (repeated a finite number of times, see Kreps et al., 1982) and Rosenthal's Centipede (1981). In all of these situations, reasoning by backwards induction has led to a counter-intuitive outcome, one that has been disproved by observation. This same divergence from reality can be found in the example that interests us here. To clarify this, Dupuy offers the following explanation: 'to return to reasonableness, all that is needed is to make an infinitesimally small move away from absolute rationality,' defined as common knowledge of rationality (p.380).

Let us consider the configuration suggested by Noussair, Robin and Ruffieux (2000). A rejection of the common knowledge of rationality can manifest itself in many forms. One could assume that certain individuals are not acting rationally (H1); or that everyone is acting rationally but that certain operators believe others to be acting irrationally (H2); or even that everyone is acting rationally and believes the other operators to be doing so too, but certain operators believe that others suspect that they may be acting irrationally (H3)…. The possibilities for confusion and misinformation are endless. Let us begin with the simplest situation. What takes place in the game analysed by Noussair et al. when certain operators are not acting rationally (viz., H1)? The solution derived using reasoning by backwards induction is clearly no longer valid. When investor *i* intervenes erratically, estimating incorrectly that the share is worth 400, the quoted price will obviously no longer be equal to the fundamental value. This is in line with Dupuy's affirmation. Nevertheless, one might feel that this is an overly strong hypothesis: why should an individual *i* be irrational? The power behind the argument offered here stems from the fact that a hypothesis of this nature is not necessary, and can therefore be weakened, as in the case of H2. Individual *i* may be perfectly rational; it suffices that he or she believes that individual *j* will behave irrationally for the same outcome to be preserved. Indeed, if individual *i* believes that individual *j* will value the share at 410 irrationally at time t+1, then it is rational for individual *i* to pay 400 for it at time *t* in order to resell it to *j* at t+1. Once again, this hypothesis can be seen to be too strong: why should individual *i* be led to expect that individual *j* will behave irrationally? Again, this hypothesis can be weakened by postulating not only that individual *i* is rational, but that he or she also thinks that individual *j* is rational. It suffices then, as in the case of H3, that we assume that individual *i* believes that individual *j* believes that individual *k* (who may be *i* himself) is going to behave irrationally for a gap to open up between the share's market price and its fundamental value. If *i* believes that *j* believes that *k* will irrationally value the share at 420 at t+2, then *i* believes that *j* will be rationally induced to pay 410 for it at t+1, and as a result *i* will buy it for 400 at *t*. By so doing, other participants' presumed irrationality can be pushed further and further upwards along a chain of mutual expectations. Reasoning by backwards induction thus continues to be invalidated.[7]

A variance with a magnitude of ε, however small this may be in terms of common knowledge rationality, would then be enough to cause divergence with the fundamental value. This variance can be measured in 'two separate ways: by the probability ε that one of the players will stray from his/her own equilibrium strategy…; and/or by the level of expectation where this variance has been envisaged' (p. 387). Here the issue is the behaviour of agents when such a gap exists. For Dupuy, the intuitive or 'sociological' approach, which consists of trying to predict the future 'by means of the representations, models and theories devised by actors on the basis of past behaviours', becomes valid again. He concludes by saying that 'this involves introducing a tiny dose of imperfection into the game's information system and showing that this slight impurity is enough to rupture the paradoxical logic of backwards induction reasoning, thus returning the time arrow to its customary direction and restoring a power of causal determination to collective objects' (p. 383). This is the perspective that the self-referential hypothesis defends. However, before getting to this point, a final experiment undertaken by R. Nagel in 1995 will enable us to hone our analysis further.

The game that Nagel proposes is the following one: a large number of players are asked to choose a whole number in the interval (0,100), the winner being the one whose number is closest to the average of all the other numbers that have been chosen, multiplied by a parameter p, with $0 \leq p \leq 1$. This situation[8] is similar to the one in which we are interested, inasmuch as a unique Nash equilibrium exists, that is, 'choosing 0'. This can be seen directly, or else by the following line of reasoning. To clarify things, say that $p = 0.5$. If so, I know with certainty that all of the numbers included in the interval (50,100) have no chance of winning. Therefore, I will not play them. In addition, I can also hypothesize that other participants will not play them, that is, if they too are rational. However, if this hypothesis is verified, I then know that the numbers that are going to be chosen are necessarily included in the interval (0,50). From that point on, I can reiterate my reasoning, this time by hypothesizing that everyone has followed the same logic as I have done, that is to say, everyone believes that everyone else is rational. This procedure then leads me to exclude the numbers that are included in the (25,50) interval. By repeating this reasoning for an indefinite period of time, I find that only the number 0 can resist this process of elimination – as long as the hypothesis of the existence of a rationality which is common knowledge remains valid. As in the experiments proposed by Noussair, Robin and Ruffieux (and in the situations considered by Dupuy), Nagel's experiment shows us that this solution is not the one that will prevail.

Nagel carried out three experiments with *p* equal to 0.5 and four with *p* equal to 2/3; and with the number of players varying between 15 and 18. She obtained the following findings: no player chose 0, and a small number of players could be found who actually chose a number above 100*p*. In the case of $p = 0.5$, the average choice was equal to 27.05 and the winning choice was 13.5; in the case of $p = 2/3$, the average choice was equal to 36.73 and the winning one was 24.5. What is so interesting about this experiment is that we can measure the degree of iterated expectations involved in the choices. A person who randomly chooses a number between 0 and 100 is defined as a strategic player of degree 0. A player is strategic of degree 1 if he or she hypothesizes that the other players' strategic degree is equal to 0. He or she then opts for the number 50*p*. A strategic degree 2 player hypothesizes that the others are degree 1 players, and therefore chooses $50p^2$, and so on. The result is that the choice $50p^n$ corresponds to a strategic player of degree *n*. Nagel found that: (1) in all of the cases analysed, the winning choice corresponded to strategic degree 2; (2) that second degree choices are in fact modal choices in terms of their distribution, concentrating 30 per cent of all choices if $p = 0.5$ and 25 per cent if $p = 2/3$. She concludes: 'thus, many players are observed to be playing approximately optimally, given the behaviour of others' (p.1318).

This experiment strengthens our hypothesis. The players have no reason to emphasize the fundamental solution, that is, 0. This solution is predicated on an hypothesis of common knowledge rationality that has demonstrably been invalidated –

after all, we even find some players who are irrational enough to play for more than 100*p*. But how do people actually play? In keeping with the self-referential assumption, they form a model of the group's opinion, which they use as a basis for calculating their best response. In our example, the winner is the person who hypothesizes that the group's opinion is represented by strategic degree 1. As such, he or she plays a strategic degree 2 solution. In this case, however, a more convincing analysis can be offered. It assumes that the operators of the market are following a second-degree reasoning. First, each player assumes that all individuals draw their numbers randomly. It follows that the average of their choices is worth 50. If this is the case, every member of the group will choose 50*p*. Because every player is aware of this, they will respond optimally and choose $50p^2$. This second-degree reasoning is a possibility. Faced with a model of opinions, it is possible that operators do not choose the best response, but instead behave in accordance with the assumption that others will make their choice in this manner, thus opting for a second-order type of expectation. We know that Keynes was in favour of it[9] when he wrote about guessing the winner of a beauty contest:

> …so that each competitor has to pick, not those faces which he himself finds prettiest, but those which he thinks likeliest to catch the fancy of the other competitors, all of whom are looking at the problem from the same point of view. It is not a case of choosing those which, to the best of one's judgement, are really the prettiest, nor even those which average opinion genuinely thinks the prettiest. We have reached the third degree where we devote our intelligences to anticipating what average opinion expects the average opinion to be. And there are some, I believe, who practise the fourth, fifth and higher degrees (p. 156).

In such a configuration of second-level reasoning, the model of opinion can be very simple without being unrealistic, in that it is not meant to describe the actual behaviour of members of the group. In fact, the players do not assume that the other members of the group conform to this model. Rather, the players assume that they are calculating the best response to give in light of this model. This is how we understand Nagel's experiment. Within this hypothesis, individuals compensate for the simplicity of the group model by assuming that the other members of the group are playing a strategic level of degree 1. This degree of strategy balances out the simplicity of the initial model of group opinion.

3. The self-referential hypothesis

The section above leads me to propose a very general hypothesis as a means of understanding the way in which financial operators forecast their expectations. Each agent acts upon a specific model that is supposed to provide him or her with a relevant description of how market opinion develops. In other words, investors are not adopting an individualistic approach; they are not seeking to anticipate each others' behaviour in order to be able to infer future price levels by means of a process of aggregation. Rather, they start out with a general conception of the market and of its opinion, a view that might be called 'holistic'. They construct this on the basis of a set of elements of disparate origins and characteristics,[10] for example, a chronicle of past interactions, certain pivotal historical episodes, such as the 1929 crisis, or a priori models such as the fundamentalist model. This idiosyncratic model enables investors to develop their own personal idea of future price movements, something that will in turn be the basis for determining their present actions. In certain cases, agents operate slightly differently, because they have recourse to a higher level of speculation: they choose what they consider to be the best action by assuming that other agents have made the same calculation. This is the general form assumed by individual rationality in complex interaction configurations where the hypothesis of the common knowledge of rationality is no longer valid. In this view, emphasis is placed on agents' learning capabilities, thanks to which their model can be adapted to the events that have been observed, as well as to the performances that are being realized.

B. Arthur (1994) also offers a similar conception of rationality. As with the analysis here, he concludes that conventional economic rationality (which he describes as being 'perfect, logical and deductive') is no longer operative once interactions become complex:

> In interactive situations of complication, agents cannot rely upon the other agents they are dealing with to behave under perfect rationality, and so they are forced to guess their behaviour. This lands them in a world of subjective beliefs, and subjective beliefs about subjective beliefs…As a result, deductive rational reasoning can no longer be applied (p. 406).

Recycling Arthur's terminology, we observe a shift from a deductive to an inductive type of rationality that is consistent with the concept of constant endogenous learning.

Why should this hypothesis be called 'self-referential'? From a theoretician's point of view, the justification is that the term refers to a general vision of the stock market which sees the financial market itself as the only relevant point of reference for thinking about prices. The self-referential vision of finance is at loggerheads with 'hetero-referential' approaches which consider that a reference exists outside of the market (that is to say in the real economy). Most of the time such approaches assume the existence of objective economic forces which automatically correct any deviation between the quoted price and the assumed external norm. One example of a 'hetero-referential' approach is provided by the fundamentalist approach, which defends the idea that price expresses an objective value that can be defined outside of the marketplace. This is called the fundamental value.

Within our theoretical framework, price is entirely different in nature, for it reflects the way in which participants are all reacting to their own predictions about what the market will do. This is a perfect example of a self-referential loop inasmuch as the starting point for any understanding of the market and of prices is the way in which the people involved happen to be thinking about the market and about prices. Expressed in a different way, price is the product of opinion, or better yet, of opinion about opinions. One should reject the belief in the existence of economic forces that inexorably pull share prices back to a level that can be defined a priori on the sole basis of real data. Quite the contrary, the self-referential stance defends the idea that, potentially, prices can assume any value. For this to occur, it suffices that everyone involved believes that this value is perceived as fair by the market opinion. This notion of a price's self-referentiality is based on a strong institutional reality. It is a fact that however fantastic a price might be, it is this price alone that determines operators' wealth.

This radical non-determination of prices does not mean that a price can be anything and everything at all times. Our approach suggests that price is an expression of agents' beliefs, such as they have been forged by financial and economic history at a given moment. These beliefs have no reason to be structurally disconnected from economic realities. Our hypothesis only denies that there exists an automatic relationship between prices and fundamental data. This leads to the possibility of long interludes during which stock prices will find themselves at 'deviant' levels. This sort of situation will certainly have negative effects on economic dynamics, as bad investment decisions will be taken based on faulty signals, leading to some destruction of capital. Nevertheless, such shortcomings do not automatically undermine the legitimacy of the market itself. This is because, in the view expressed here, stock markets do not derive their social usefulness from their ability to manufacture relevant valuations, but from their ability to ensure market liquidity for financial securities.[11] This is totally unrelated to the 'fairness' of share prices. Rather, it depends on the stability of individual beliefs.

These reflections on the appropriateness of the term 'self-referential' have been developed from the viewpoint of a theoretician. However, they are also valid, partially at least, for investors themselves, as they are described in this theory. We say 'partially' because the analysis is clearly compatible with the fact that certain investors believe in fundamentalist models. However, in almost all situations, the players who survive and whose actions are significant for theoreticians, display pragmatism towards their models, including those which are fundamentalist. They allow for the possibility of a lasting gap between fundamentals and prices. As demonstrated forcefully in 'Noise

Trader Approach' models (see De Long et al., 1990a and 1990b; Shleifer and Summers, 1990), the nature of markets is such that they force people to adopt an opportunistic type of rationality, one which incorporates into the calculations the fact that the market can behave irrationally. In this approach, one moves closer to an entirely self-referential type of behaviour, in other words to considering that however it has been shaped, it is market opinion itself (and nothing else) that counts. In other words, everyone recognizes that the market possesses the ability for self-legitimation. This idea was expressed forcefully by an experienced practitioner, Pierre Balley, when he wrote:

> The quality of a reasoning is of little importance if the stock market doesn't agree, meaning if it is contradicted by the dominant collective opinion. In the same way that a politician, manager or analyst cannot be right about something if this contradicts the majority opinion, it is the market that has the final say. For this reason, and above and beyond the study of firms, it is important to take stock of the various currents of opinion that can move the stock market and which lead it, at the various phases of its existence, to make radically different assessments of one and the same business and sometimes of an entire economic situation (1987, p. 137).

This quote expresses precisely what constitutes the self-referential hypothesis from the individual actor's perspective. It describes the immediate experience of stock market participants, an experience that is both social and psychological: market opinion is the ultimate reality, an entity which is imbued with a life-form of its own, because it can refute any forecast and any model, and because it is the only legitimate source for valuing securities. Note that the points of view of theoreticians and operators are mutually supportive: (1) inasmuch as prices can potentially take on any value, actors are induced to make market opinion the central object of their concern; (2) if they behave in this manner, the self-referential hypothesis is confirmed.

Once we assume the hypothesis of self-referentiality, what is the theoretician's task? The first objective is to analyse the set of individual models (M) as well as the individual rules for learning (R). If we call one market participant *i* and the set of participants I, the theoretician studies a population of $\{M_i, R_i\}_{i \in I}$. In this view, research first involves the specification of the models M_i. Examples include N. Barberis, A. Sheifer and R. Vishny (1998); K. Daniel, D. Hirshleifer and Subrahmanyam (1998); H. Hong and J. Stein (1999); Shiller (2000); and most of what is usually called Behavioural Finance, as in Thaler (1993). Consistent with the intuitions of Dupuy, we note that psychological and sociological analyses exert a great deal of influence on this body of work. A second type of research studies price dynamics for some of the more interesting populations, for example, fundamentalist investors, chartists, trend chasers and noise traders. Examples include work by De Long, Shleifer, Summers and Waldmann (1990a and 1990b); Frenkel and Froot (1986); and more generally what is called the 'Noise Trader Approach' (Shleifer and Summers, 1990). This list does not pretend to come even close to recapitulating all of the research that has been carried out, given the multitude of issues raised by a self-referential approach. For example, are some models more robust than others? Do some models or populations possess self-validating characteristics? To put it differently, can they confirm all of the actors' initial hypotheses? How is it possible to select the information that the market will consider relevant at any given moment in time? Can we determine general forms of historical price evolutions?

4. Personal beliefs and market beliefs in rational bubbles

If we pursue the self-referential hypothesis, it appears that individuals do not act according to what they consider to be the right value, that is to say, the value that is justified in light of the real economic situation as they understand it; instead, they act according to their beliefs about what constitutes market opinion. This 'schizophrenia' is one specificity of the self-referential model: faced with new information, investors do not wonder about its real contents (in the sense that the fundamentalist model attributes to this term), but about the way in which the market is going to interpret this information. To put it differently, they do not react to news, but to how they believe others will react to it.[12]

One particularly illuminating example involved a situation in the FX market as reported in The New York Times on 12 November 1987. The dollar had risen following a statement by President Reagan that its value 'had fallen far enough'. The situation seems easy to understand: it was fundamentalist in nature. Traders took stock of the new information provided by President Reagan in his public statement, leading them to modify upwards their evaluation of the dollar's fundamental value. However, further study has shown this explanation to be erroneous. In actual fact, and after having interviewed traders, the reporter for The New York Times observed that 'no one believed the President'. Why then did people buy the dollar? Interviewees answered that even though they did not personally have any faith in President Reagan, they nevertheless thought that others could be influenced by his statement. This provides an illustrative situation in which each participant's belief is the personal opinion P that 'President Reagan's statement does not contain any information'. Simultaneously, there is the thinking that everyone else will believe that he said something important. If we use the term $C_i Z$ to signify that 'individual *i* believes proposition *Z* to be true', and if we call R the proposition that 'information is contained in President Reagan's statements', we have a situation where for all traders *i*, there is $C_i(P)$ and simultaneously $C_i C_M(R)$, which can be read that 'individual *i* believes that the market *M* believes that the proposition R is true'. At the personal level, if all traders use fundamental analysis they might feel that a certain proposition, P, is true. Moreover, at the same time, they might have a representation of the market that leads them to believe that it will end up behaving differently, as per their belief R. Now, the self-referential hypothesis tells us that speculators do not act according to what they themselves believe, but in light of what they expect from everyone else. 'Despite his scepticism about Reagan's comments, the trader bought dollars as soon as he learned what the President had said. The trader explained that he expected other traders to buy dollars on the news, driving up the currency's price, and he hoped to sell later in the day after the price had risen', (New York Times, 12 November 1987).

$C_i(P)$ types of beliefs can be qualified as a form of personal beliefs, wherein the investor places his faith in his own intuition. Theoretical analysis demonstrates that personal beliefs have no direct effect on the market. What counts for the investor are $C_i C_M(R)$ types of market belief. Yet contrary to what a reflex analysis might lead us to think, nothing forces the gap between these two different levels of belief to decrease. They are fundamentally different in that they both relate to different experiences. One relates to the actual value of the security and the other to the value awarded to it by the market.

This distinction between personal and market beliefs can be useful for thinking about fundamentalist evaluations. I will hypothesize that, by its very nature, this is a $C_i(P)$ type of private assessment. For the person involved, this corresponds to a normative stance. When the agent carries out this calculation, he or she is thinking that 'this is what the market price should be if all of the conditions of efficiency were fulfilled'. In other words, one could say that fundamentalist evaluations express the personal vision of the individual concerned with respect to the way in which things should be happening. Some investors go even further and assume that the market is actually going to behave in this manner. However, as previously mentioned, those who act according to such an assumption usually proceed with a great deal of pragmatism. The preceding analysis demonstrates that personal fundamentalist beliefs are only poorly transmitted to the marketplace, or sometimes not at all as with the example concerning Reagan.

To grasp fully the private nature of fundamentalist evaluations, it is useful to analyse what can be gleaned from the 'Noise Trader Approach' (Shleifer and Summers, 1990). The central objective of this approach is to demonstrate that actions by fundamentalist arbitrageurs are not sufficient to bring prices back down to the level of their fundamental value. According to these theoreticians, where a gap is observed between a value and a quoted price, the possibility of intervention by arbitrageurs is highly restricted by the presence of two major risks, so that in the end this gap is never totally eliminated. The first risk relates to the intrinsically random nature of fundamental data. The second is what one might call market risk. If a fundamentalist investor believes that the price of security A is overvalued today, by selling it short (to enact his or her personal belief) he or she exposes him/herself to the risk that the security will be even more over-valued the day it has to be repurchased. This can cause substantial losses. Shleifer and Summers illustrate this point by describing the situation for Japanese shares during the 1980s:

> During this period, Japanese equities have sold at price earning multiples of between 20 and 60…Expected growth rates of dividends and risk premia required to justify such multiples seem unrealistic. Nonetheless, an investor who believes that Japanese equities are overvalued and wants to sell them short must confront two types of risk. First, what if Japan actually does perform so well that prices are justified? Second, how much more out of line can prices get, and for how long, before Japanese equities return to more realistic prices? Any investor who sold Japanese stocks short in 1985, when price earning multiple was 30, would have lost his shirt as the multiples rose to 60 in 1986 (pp. 21-22).

As was the case with the self-referential hypothesis, this analysis shows that financial rationality forces fundamentalist arbitrageurs to take into account the evolution of market opinion. This seems clear when analysing the models being offered by way of illustration of such theses (see Appendix). Here the arbitrageur's rationality involves defining his/her own actions on the basis of what is being done by noise traders even if it is inappropriate. If one takes a closer look at these models, the fact that a particular speculator might be fundamentalist is at best of secondary importance. It is the outcome of a perfectly artificial hypothesis, according to which the fundamental value is announced publicly at the end of the game, serving as a basis for the securities' liquidation. If this unrealistic hypothesis is rejected, the notion of fundamental value loses all relevance. The only thing that is left is the idea that however the individual has privately evaluated what the situation should be when he or she intervenes in the market, what really counts is how market opinion takes shape. The best way to describe this reality is with terminology such as 'smart money' or 'rational speculators' rather than 'fundamentalist arbitrageur'. This is because the former expressions translate the idea that what is paramount in the situation is anticipating how the other parties are going to behave. Keynes wrote about this as follows:

> [Professional investors] are concerned not with what an investment is really worth to a man who buys it "for keeps", but with what the market will value it at, under the influence of mass psychology, three months or a year hence. Moreover, this behaviour is not the outcome of a wrong-headed propensity...For it is not sensible to pay 25 for an investment of which you believe the prospective yield to justify a value of 30 if you also believe that the market will value it at 20 three months hence (p. 167).

Thus, the only thing that the market needs to know is how each agent thinks that the market is going to develop – that which we have termed market beliefs. In Keynes's example, someone who on the basis of his or her personal evaluations should be bullish will act in the market like a person who is bearish.

There is a final reason why fundamentalist evaluations are purely personal. This is because they propose calculations relating to future quantities (profits and dividends) about which not a lot is known. Thus, if we abandon models such as those that have been proposed by Noussair, Robin and Ruffieux (2000) or by De Long et al. (1990a and 1990b), which arbitrarily assume that this value will be publicly announced by a deus ex machina, it would appear that there are as many fundamental evaluations as there are investors. A security's price, whatever this may be, can always be justified by some ad hoc fundamental scenario. In this view, the concept of fundamental value is empty of any useful dimension for people interested in the actual dynamics underlying stock prices.[13] Where an evaluation is portrayed as being fundamentalist, this should be seen as a claim for legitimacy by those agents who present their evaluations as such. This claim may, or may not, be justified and/or succeed. It will only have an impact on price dynamics to the extent that it influences the models of market opinion that operators use, and not because it conforms to fundamental values.

This distinction between personal beliefs and market beliefs helps us to overcome the overly hasty accusations of irrationality that are so often thrown at financial investors, for example when a significant disconnection is noted between a quoted price and what the community of economists considers to be its right value. As an example, let us consider a currency that is already undervalued, yet which is still being sold massively, leading to even greater under-valuation. Here traders are accused of being irrational in the sense that they are attributing an incorrect value to the currency in question. However, this hypothesis does not hold. It is, in any case, unnecessary. Traders, like everyone else, may be perfectly aware that a currency is undervalued, yet continue to sell it. This is because what counts for them when they act in the market is not what they believe to be the 'right value' for the currency (as far as they can judge), but what they expect from the rest of the market. People make profits in markets when they successfully predict the way in which group opinion is going to evolve. These are the rules of the game. Agents are not asked to be 'right'. In this view, the statement below (made by a trader as the euro dropped sharply in September 2000) reveals a great deal about the dichotomy between personal beliefs and market beliefs. It shows someone who is convinced that the euro is undervalued, but who explains that he is still forced to sell it if he does not want to end up ruined.

> I might be bullish on the euro, but that doesn't mean much when you see that almost everyone else in the FX market is positioned to sell it. So even if I feel that the euro should be higher against the dollar, I'd still hesitate before buying it. After all, if I am the only buyer of euros and I'm up against 50 sellers, it's going to be tough...I am not necessarily doing what I personally believe in, but what I think that the market is generally going to do. And in the end, the market is always right. A trader's job is to try to assess currency market sentiment as closely as possible.[14]

This is a perfect illustration of self-referential rationality. Despite his personal conviction that the euro is undervalued, this trader is behaving rationally by betting on a further fall.

The model developed by De Long et al. (1990a) offers the same paradox: the actions of perfectly rational fundamentalist speculators destabilize prices. A description of this model can be found in the Appendix. The core of this argument revolves around the fact that rational speculators act not only on the information they have regarding fundamentals,[15] but also on their knowledge of the market and, more specifically, on the fact that they know that there are some ‘positive feedback investors’ in the market. Positive feedback investors are investors who buy when they detect a bullish trend in the market. If such investors are present in the market, it is profitable for rational speculators to push prices above their fundamental value in order to try to provoke positive feedback investors into buying at a very overvalued price. This is what the aforementioned model has demonstrated. Again, a superficial vision would conclude that fundamentalist speculators are irrational, because they are buying a security during period 1 at a price above its fundamental value. This is totally wrong, however. As shown previously, rationality does not mean at all that people act in accordance with fundamentals. It requires that maximum benefits be derived from the behaviour of markets, that is, paying a high price in period 1 in order to sell it even higher in period 2. Such decisions in no way signify an ignorance of the fundamentals. Quite the opposite is the case, given the assumption that rational speculators are very aware of fundamentals even if they choose to ignore them in order to take advantage of the erroneous judgment of other participants and in the process make a profit.

These analyses are very enlightening but require noise traders, that is to say, agents who assess the fundamentals inaccurately**,** in order to generate a deviation between fundamental valuations and prices. We shall call these market participants irrational agents. For example, the euro-trader interviewed by Libération justifies his imitative behaviour by the fact that he was up against ‘50 sellers’. The same analysis applies to the irrational speculators found in the work of De Long et al., as it is the existence of positive feedback investors that justifies their actions. Such logic is not necessarily false. In a given financial environment, naïve, ill-informed or irrational investors can be found. That is a fact. Keynes used this hypothesis frequently when he spoke about 'the mass psychology of a large number of ignorant individuals'. This is also the basis of the noise-trader approach, as the name clearly indicates. However, this precondition is not an absolute necessity, either at the theoretical or at the practical level. Bubbles can emerge without any need to assume the presence of noise-traders. Highlighting this fact is essential for theoreticians. If this were not the case, the destabilizing effects of price mechanisms could only be inferred if one assumed that most market players were irrational, that is, that they understood nothing about fundamental value. In this sort of framework, self-referential behaviour by rational spectators simply serves to magnify collective blindness; there is no possibility that they could produce the blindness all by themselves. Within such a theoretical framework, bubbles are not rational, because irrational expectations are essential in order for them to be generated. However, it is possible to go even further. We can conceive of bubbles in a market that is solely comprised of self-referential speculators, without noise traders. Such situations could indeed be defined as 'rational bubbles', but in a larger sense than those defined by Blanchard and Watson (1984).

A simple way to demonstrate this is to take a new look at the example of bearish speculation against the euro. Why did the trader choose to sell? The reason is that he was up against '50 sellers' who were too strong for him, as he stated. However, why were these people selling? This is the crux of the matter. According to the previous interpretation, it is because they believe the euro to be overvalued. It is an erroneous conception of fundamentals that causes them to sell. Faced with this fait accompli, the trader no longer has a choice. All that remains for him to do is to surrender to irrational majority opinion. Yet, this is not the correct interpretation. There is absolutely no need to assume that other players are acting differently from the trader. To explain the bearish bubble on the euro, it is not necessary to postulate the existence of stubborn fundamentalist sellers persisting in the error of their ways. Instead, we can offer an alternative interpretation, one that is more useful and which is probably closer to reality, namely that the 50 sellers have determined their own behaviour on the basis of their personal expectations with respect to majority opinion. They are also acting in a self-referential manner, in line with their own models of market opinion. In this view, each individual is as rational as the next and acts according to shared beliefs about everyone else's behaviour. Thus, we are not in the presence of 50 'committed' sellers, but 50 traders who, when they think about what everyone else is going to do, expect that they are going to sell (just as the trader interviewed by Libération does). If we were to interview them personally, they would remark that there is no use going against a market that is determined to be blind. And they would be right. The important thing here is not the value of the euro, but the fact that 'everyone believes that the market is bearish', a belief that we have called a 'market belief'. This is a second-order view, that is, a belief involving another belief. If everyone thinks that 'people are bearish', then they themselves will be sellers and the market will actually drop, validating the initial belief ex post. Economists talk about a 'self-fulfilling prophecy' when describing this last phenomenon. One is in a situation where the generalization of self-referential behaviours validates traders' initial beliefs. Taking these sorts of dynamics into account completely upsets the intuitive understanding of markets and of rationality, given that everything is happening as if collective representation were totally autonomous. By so doing, the holistic nature of the initial model is validated de facto. In this sort of situation, everyone is being perfectly rational in terms of personal beliefs, market belief valuations, and market valuations. Nonetheless, a bubble emerges. Investors who believe that they are doing nothing more than reacting to the market will feel completely vindicated in their conception of the autonomy of market opinion. This seems all the more unarguable to them, as the developments being observed find no coherent explanation in fundamental analysis.

The same conclusion can be drawn from the De Long et al. model, as long as it is extended by means of an additional hypothesis that the number of positive feedback investors is not constant, but depends upon the size of the trend being observed. This extension is presented in the Appendix. In these conditions, it can be demonstrated that when rational speculators take this particular reality into account, their demand function takes on a positive feedback aspect. In other words, it is no longer necessary to postulate the existence, outside of the circle of rational speculators, of 'irrational' individuals who are convinced of the intrinsic soundness of extrapolative anticipations. Self-referential rationality can engender this type of behaviour just by itself. Rational speculators believe that they are reacting to the presence of positive feedback investors without realizing that they themselves are also trend chasers.

Conclusion

The self-referential hypothesis tells us that the significance of a price is found in the market itself and not in the real economy. It then devises a theory where opinions, conventions and learning play a crucial role. This hypothesis differs from traditional approaches in finance on a crucial point: one can no longer assume that rationality is common knowledge. It follows that investors' behaviour is closely tied into their a priori conceptions and to their limited capabilities for learning. In other words, the market is governed by an inductive, and no longer a deductive, type of rationality.

One of the most interesting results of using the self-referential approach is that it allows for bubbles without investors having to make any subsequent errors in evaluation. These can be termed 'rational bubbles', in a larger sense than the term defined by Blanchard and Watson (1984). This is the case because the approach clearly distinguishes the personal opinions of the investors regarding the fundamental value of a security from their 'market beliefs' about future price dynamics. They may all believe that a security is worth 1, but will buy it at 10, because they believe the market value to be 10. Given this configuration, no incorrect judgments have been made, inasmuch as the security is indeed essentially worth 1, and the market does value it at 10.

Clearly therefore, the self-referential approach necessitates a thorough re-examination of the theoretical questions that have been raised concerning financial economics. The basis of this approach lies in the study of collective and individual opinions. This is not to say that market prices are indeterminable. It simply means that prices reflect the beliefs of those operating in the market – no more, no less.

Appendix

Thoughts on 'Positive Feedback Investment Strategies and Destabilizing Rational Speculation' by J. Bradford de Long, Andrei Shleifer, Lawrence H. Summers and Robert Waldmann (1990), The Journal of Finance, **XLV** (2), 379-395.

This Appendix presents some of the findings from this article and suggests avenues for extending them.

1. The article's findings

Let us assume that investors had a choice between cash, on one hand, and a risky asset which is in zero net supply on the other. The trading takes place over three periods of time. In period 3, the asset is liquidated and its dividend is announced publicly. It is equal to $\Phi + \theta$, where θ is distributed normally with mean 0 and variance σ_θ^2, and where Φ is a random variable with mean zero, and which can take on the following three values : φ, 0, and $-\varphi$. No information is provided on variable θ before period 3. The value of Φ is made public during period 2. However, from period 1 onwards, rational speculators (but only they) observe a signal ε that provides them with information on the value of Φ. This signal can be either noisy or noiseless. Both examples are analysed by the authors. It is only this latter case that will be examined in this Appendix. We therefore have $\varepsilon = \Phi$. Hence, rational investors' expectations are such that $E(p_3|\varepsilon) = \Phi$. When this information is not available, we have $E(\Phi + \theta) = 0$.

We assume that the market for this risky asset is comprised of three types of agents: fundamentalist investors represented by *f*, positive feedback traders represented by *p*, and rational speculators represented by *r*. The fundamentalist participants intervene in the market during a period *t* if the price differs from its fundamental value. We therefore have:

$$D_t^f = \alpha(FV_t^a - p_t)$$

where FV_t^a is the fundamental value expected at period *t* and p_t the market price. Positive feedback traders' interventions are a function of past price trends. In other words, on date *t*, their demand function can be written:

$$D_t^p = \beta(p_{t-1} - p_{t-2}).$$

Assume that α > β. These two categories of actors are not interested in other investors' actions. They make their decisions on the sole basis of price and fundamental value. Rational speculators act in a strategic manner, taking other actors' behaviour into account. Their utility function is such that their demand during period *t* can be written as:

$$D_t^r = \frac{(p_{t+1}^a - p_t)}{2\gamma\sigma_{t+1}^2} \tag{1.1}$$

where p_{t+1}^a is the price at *t* + 1 that the rational speculators anticipate at *t*, σ_{t+1}^2 is their evaluation of the risk during period *t* + 1, and γ measures their risk aversion. The equilibrium price is determined by the equation:

$$n_f(t)D_t^f + n_p(t)D_t^p + n_r(t)D_t^r = 0 \tag{1.2}$$

where $n_f(t)$, $n_p(t)$ and $n_r(t)$ are the numbers of the respective classes *f*, *p*, and *r*, of investors during period *t*. B. de Long, A. Shleifer, L. Summers and R. Waldmann hypothesize that positive feedback traders have a mass equal to 1, that rational speculators have a mass equal to μ, and that fundamentalist investors have a mass equal to 1 - μ. Thus, equation (1.2.) can be written:

$$(1-\mu)D_t^f + D_t^p + \mu D_t^r = 0 \tag{1.3}$$

The reasoning suggested by B. de Long, et al. works backwards from period 3. During this period, the asset's price is announced and is therefore worth Φ + θ. During period 2, the trend followers' demand function is:

$$D_2^p = \beta(p_1 - p_0).$$

The fundamentalist investors observe Φ during period 2, which means that their expectation of the fundamental value is equal to Φ. As a result, their demand function can be written as:

$$D_2^f = \alpha(\Phi - p_2)\,.$$

Finally, according to equation (1.1), the rational speculators' demand function is:

$$D_2^r = \frac{\Phi - p_2}{2\gamma\sigma_\theta^2} = \alpha(\Phi - p_2)$$

since the authors assume that $\alpha = \dfrac{1}{2\gamma\sigma_\theta^2}$.

The price during period 2 then satisfies the equation:

$$\beta(p_1 - p_0) + (1-\mu)\alpha(\Phi - p_2) + \mu\alpha(\Phi - p_2) = 0 \tag{1.4}$$

In the same way, during period 1, we find:

$$\begin{cases} D_1^f = -\alpha p_1 \\ D_1^c = 0 \end{cases}$$

Under these conditions, the market equilibrium can be written:

$$\mu D_1^r = (1-\mu)\alpha p_1 \tag{1.5}$$

To specify speculators' actions during period 1, the authors note that since they are able to resolve equation (1.4.), rational speculators know the value of p_2 with certainty (a case where the risk σ_2^2 is zero) since they have observed $\varepsilon = \Phi$. As such, once rational speculators are actually present in the market (that is to say, when $\mu > 0$), we find the following equality:

$$p_1 = p_2 \tag{1.6}$$

During period 0, the price (that is, the fundamental value) is worth 0, and no trading takes place.

Solving equations (1.4) and (1.6) with $p_0 = 0$, we find that:

$$\begin{cases} p_1 = p_2 = \dfrac{\alpha}{\alpha - \beta}\Phi = p^* > \Phi \ \text{ if } \mu > 0 \\ p_1 = 0 \ \text{ and } \ p_2 = \Phi \ \text{ if } \mu = 0 \end{cases} \tag{1.7}$$

Given our hypothesis that $\alpha > \beta$, it appears that the presence of rational speculators makes prices stray from their fundamental value. This stems from the fact that they anticipate the positive feedback traders' behaviour and seek to benefit from it. Their impact is therefore destabilizing.

2. Extension

It is possible to make the model slightly more sophisticated by abandoning the hypotheses that have been made concerning the number of investors, and by returning to equation (1.2). We then show that the preceding findings can be preserved (subsection 2.1). Subsequently, once we agree that variables n_f and n_p depend on past prices, it can be shown that rational speculators also turn into trend followers (subsection 2.2).

2.1. The price p_2 is no longer known with certainty

In this example, during period 2, the equilibrium equation can be written:

$$n_p(2)\beta(p_1 - p_0) + n_f(2)\alpha(\Phi - p_2) + n_r(2)\alpha(\Phi - p_2) = 0 \qquad (2.1)$$

This gives us:

$$p_2 = \Phi + \frac{n_p(2)}{n_f(2) + n_r(2)} \frac{\beta}{\alpha} p_1 = \Phi + kp_1 \text{ with } n_r(2) > 0 \qquad (2.2)$$

We assume that the number of rational speculators is constant, and we postulate $n_r = n$. We focus on situations in which $k < 1$. Let us assume from now on that rational speculators do not know the exact number of trend followers during period 2. In these conditions, their demand will be written in the form of equation (1.1). We can make a variety of hypotheses regarding the expected risk. Let us assume that the risk is constant and equal to ν. We can then write:

$$D_1^r = \frac{(\Phi + kp_1) - p_1}{2\gamma\nu} = h[(\Phi - (1-k)p_1] \text{ with } h = \frac{1}{2\gamma\nu} \qquad (2.3)$$

We then find that by resolving the equilibrium equation during period 1:

$$nD_1^r - n_f(1)\alpha p_1 = 0 \qquad (2.4)$$

that the price during period 1 is equal to:

$$p_1 = \frac{\Phi}{1 - k + \frac{\alpha n_f(1)}{hn}} = \frac{\Phi}{1 - k + y} \text{ with } y = \frac{\alpha n_f(1)}{hn} = f(n) \qquad (2.5)$$

We can then easily verify that p_1 is an increasing function of $n \in]0, \infty[$, such that:

$$p_1 \in \left]0, \frac{\Phi}{1-k} = p^*\right[.$$

Equation (2.5) again gives us the value p^* of equation (1.7) when we assume that $k = \alpha/\beta$ and $v = 0$. According to (2.2), one finds:

$$p_2 = \Phi \frac{1+y}{1-k+y} \in \left]\Phi, \frac{\Phi}{1-k} = p^*\right[.$$

p_2 is always greater than p_1 and is a decreasing function in *y*, or else an increasing function in *n*. For plausible parameter values, the model, when rewritten in this way, offers the same findings as those that have been obtained with the basic model. More specifically, the greater the number of rational speculators, the more the price diverges from the best forecast of the fundamental value, that is from Φ, during period 1 and during period 2. For $n = +\infty$, we again discover the preceding case, that is, p^*. For $n \to 0, p_2 \to \Phi$.

During period 1, rational speculators will purchase the share, even when its price exceeds Φ. Then, during period 2, they will sell it back to the positive feedback traders. This is how they make their profits.

Figure 1 shows three types of dynamics. The unbroken lines represent the dynamics at work without the presence of any rational speculators; the dotted lines represent the dynamics as described in the present article; and the dashed lines represents the dynamics as described in the present paragraph.

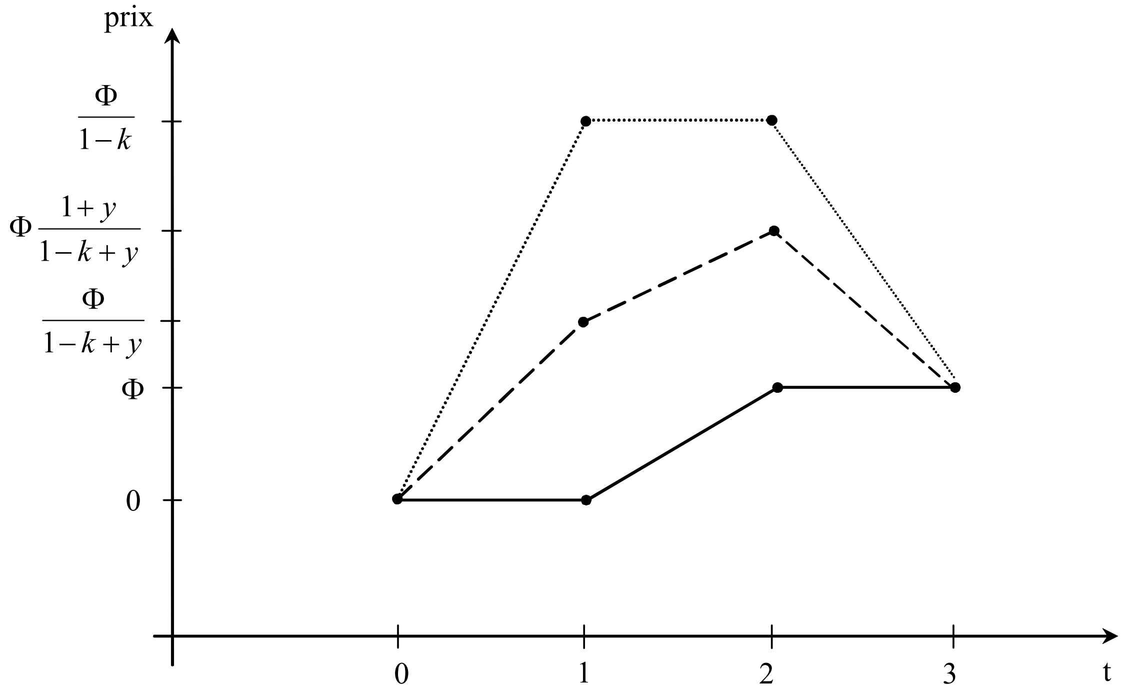

figure 1

Figure 1: Three kinds of price dynamics

2.2. Rational speculators become positive feedback traders

Based on specification (2.3) of the demand function, we can go a little further. If we assume that fundamentalist actors change strategy once they observe that a sufficiently strong price trend was manifested during the preceding period, this means that the *k* parameter is no longer constant. *k* becomes an increasing function of the price p_1. In such conditions, once rational speculators have integrated this fact into their reasoning, their demand can also become an increasing function of p_1. In other words, rational speculators in turn become positive feedback traders if they become aware of the fact that fundamentalist actors are likely to change their strategy when the price rise exceeds a certain threshold.

To illustrate this point, a small numerical example suffices. Let us assume that the behaviour of the fundamentalist investors leads to the following changes in the parameter *k*:

$$\begin{cases} k = \underline{k} < 1 \text{ if } p_1 \leq \underline{p} \\ k = ap_1 + b \text{ if } \underline{p} \leq p_1 \leq \overline{p} \\ k = \overline{k} < 1 \text{ if } \overline{p} \leq p_1 \end{cases}$$

with the following parameter values:

$$\underline{k} = 1/4,\ \underline{p} = 1,\ \overline{k} = 3/4,\ \overline{p} = 5/2,\ a = 1/3,\ b = -1/12.$$

In conditions such as these, if we make the simplifying hypothesis that the expected risk remains constant and is worth *V*, then the demand function of the rational speculators can be written:

$$\begin{cases} D_1^r = \dfrac{1}{2\gamma V}(\Phi - \dfrac{3}{4}p_1) \text{ if } p_1 \leq \underline{p} = 1 \\ D_1^r = \dfrac{1}{2\gamma V}(\dfrac{1}{3}p_1^2 - \dfrac{13}{12}p_1 + \Phi) \text{ if } \underline{p} \leq p_1 \leq \overline{p} = \dfrac{5}{2} \\ D_1^r = \dfrac{1}{2\gamma V}(\Phi - \dfrac{1}{4}p_1) \text{ if } \dfrac{5}{2} = \overline{p} \leq p_1 \end{cases}$$

This demand is strictly positive for $p_1 < 4\Phi$. It is increasing for values of p_1 lying between $\frac{13}{8}$ and $\frac{5}{2}$. It is easy to find the values for Φ, $2\gamma V$, n and $\alpha n_f(1)$ such that the equilibrium price p_1 of equation (2.4) belongs to this interval. This is shown in Figure 2, in which we have set $1/2\gamma V = n_f(1)\alpha/n = 1$ for the sake of simplicity.

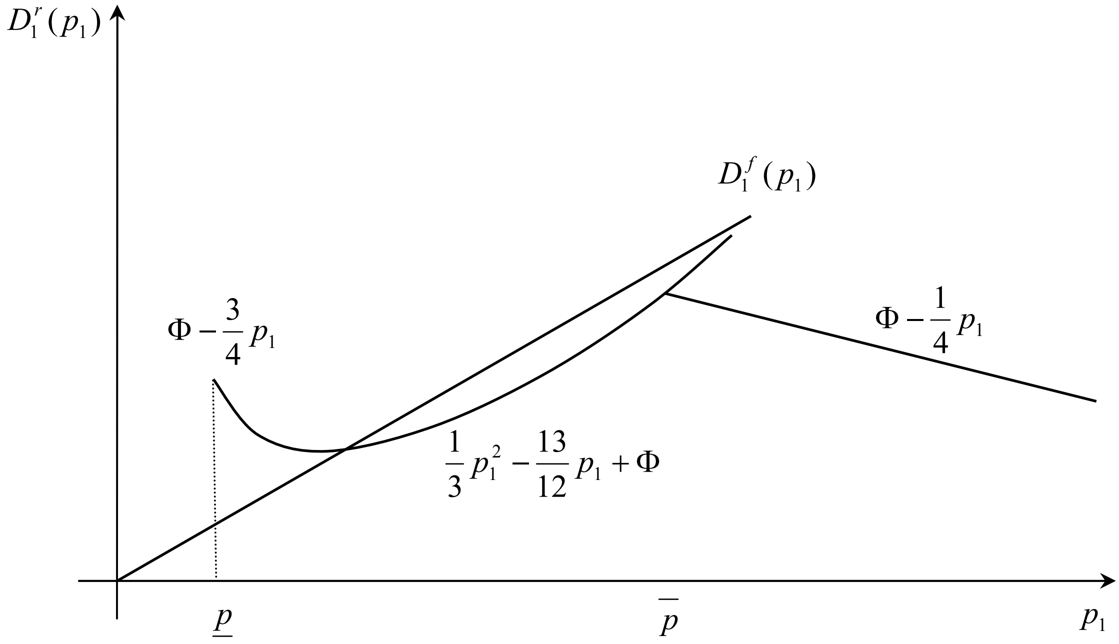

Figure 2: The demand function of rational speculators

Bibliography

Aumann, R. (1994), 'L'irrationalité dans la théorie des jeux', in A. Orléan (ed.), Analyse économique des conventions, Paris: Presses Universitaires de France, collection Economie, 43-59.

Arthur, W. B. (1994), 'Inductive Reasoning and Bounded Rationality', American Economic Review, **84** (2), 406-411.

Avery C. and P. Zemsky, (1998) 'Multidimensional Uncertainty and Herd Behavior in Financial Markets', The American Economic Review, **88** (4), 724-748.

Balley, P. (1987), La Bourse : mythes et réalités, Paris: Presses Universitaires de France.

Barberis, N., A. Shleifer and R. Vishny (1998), 'A Model of Investor Sentiment', Journal of Financial Economics, **49**, 307-343.

Banerjee, A. (1992), 'A Simple Model of Herd Behavior', Quarterly Journal of Economics, 107 (3), 797-818.

Bikhchandani, S., D. Hirshleifer and I. Welch (1992), 'A Theory of Fads, Fashion, Custom and Cultural Changes as Informational Cascades', Journal of Political Economy, **100** (5), 992-1026.

Blanchard, O. and M. Watson (1982), 'Bubbles, Rational Expectations, and Financial Markets', in P. Wachtel (ed.), Crisis in Economic and Financial Structure, Lexington, UK: Lexington Books.

Brière, M. (2000), 'Pessimisme et extrémisme: une étude des prévisions économiques de consensus', miméo, Paris: Cepremap, 27 pages.

Daniel, K., D. Hirshleifer and A. Subrahmanyam (1998), 'Investor Psychology and Security Market Under- and Over-reactions', The Journal of Finance, **LIII** (6), 1839-1885.

De Bondt, W. and R. Thaler (1985), 'Does the Stock Market Overreact?', The Journal of Finance, **XL** (3), 793-807.

De Long, J., A. Shleifer, L. Summers and R. Waldmann (1990a), 'Positive Feedback Investment Strategies and Destabilizing Rational Speculation', The Journal of Finance, **XLV** (2), 379-395.

_______. (1990b), 'Noise Trader Risk in Financial Markets', Journal of Political Economy, **98** (4), 703-738.

Dupuy, J-P (1989), 'Convention et "common knowledge"', Revue Economique, **40** (2), 361-400.

Frenkel J. and K. Froot (1986), 'The Dollar as an Irrational Speculative Bubble: The Tale of Fundamentalists and Chartists', Marcus Wallenberg Papers on International Finance, **I,** 27-55.

Hong H. and J. Stein (1999), 'An Unified Theory of Underreaction, Momentum Trading, and Overreaction in Asset Markets', The Journal of Finance, **LIV** (6), 2143-2184.

Jensen, M (1978), 'Some Anomalous Evidence Regarding Market Efficiency', Journal of Financial Economics, **6**, 95-102.

Keynes, J.M. (1973), The General Theory of Employment, Interest and Money, Collective Writings, vol. 7, Cambridge: Cambridge University Press. (1st edition 1936).

Kreps D. and R. Wilson (1982), 'Reputation and Imperfect Information', Journal of Economic Theory, **27,** 253-279.

Kreps, D. et al. (1982), 'Rational Cooperation in the Finitely Repeated Prisoners' Dilemma', Journal of Economic Theory, **27,** 245-252.

MacKay, C. (1980), Extraordinary Popular Delusions and the Madness of Crowds, New York, Harmony Books (1st edition 1841).

Mehta, J., C. Starmer and R. Sugden (1994), 'The Nature of Salience: An Experimental Investigation of Pure Coordination Games', American Economic Review, **84** (2), 658-673.

Nagel, R. (1995), 'Unraveling in Guessing Games: An Experimental Study', American Economic Review, **85** (5),1313-1326.

Noussair, C. and B. Ruffieux (2001), 'A Major Lesson from the Experimental Study of Markets: The Contrast in Efficiency Between Services and Asset Markets', mimeo, 24 pages.

Noussair, C., S. Robin and B. Ruffieux (2000), 'Price Bubbles in Laboratory Asset Markets with Constant Fundamental Values', mimeo, 27 pages.

Orléan, A. (1999), Le pouvoir de la finance, Paris: Odile Jacob.

Orléan A. (2002), 'Le jeu financier de base', miméo, 12 pages.

Rosenthal, R. (1981), 'Games of Perfect Information, Predatory Pricing and the Chain-Store Paradox', Journal of Economic Theory, **25,** 92-100.

Scharfstein, D. and J. Stein (1990), 'Herd Behavior and Investment', American Economic Review, 80 (3), 465-479.

Schelling, T. (1960), The Strategy of Conflict, Oxford: Oxford University Press.

Selten, R. (1981), 'The Chain-Store Paradox', Theory and Decision, **9,** 1978, 127-159.

Shiller, R. (1981), 'Do Stock Prices Move Too Much to Be Justified by Subsequent Changes in Dividends?', American Economic Review, **71,** 421-436.

_______. (1990), 'Speculative Prices and Popular Models', Journal of Economic Perspectives, 4 (2), 55-65.

______. (2000), Irrational Exuberance, Princeton, New Jersey: Princeton University Press.

_______. (1991), Market Volatility, Cambridge, Massachusetts and London: MIT Press.

Shleifer, A. (2000), Inefficient Markets. An Introduction to Behavioral Finance, Oxford: Oxford University Press.

Shleifer, A. and L. Summers (1990), 'The Noise Trader Approach to Finance', Journal of Economic Perspectives, 4 (2), 19-33.

Shelling, T. (1960), The Strategy of Conflict, Oxford: Oxford University Press.

Smith, V. (1988), 'Experimental Methods in Economics', in J. Eatwell, M. Milgrate and P. Newman (eds), The New Palgrave: A Dictionary of Economics, London and Basingstoke: MacMillan, vol. 2, 241-249.

Smith, V., G. Suchanek and A. Williams (1988), 'Bubbles, Crashes, and Endogenous Expectations in Experimental Spotty Asset Markets', Econometrica, 1988, 1119-1151.

Summers, L. (1986), 'Does the Stock Market Rationally Reflect Fundamental Values?', The Journal of Finance, **XLI** (3), 591-601.

Thaler, R. (ed.) (1993), Advances in Behavorial Finance, New York: Russell Sage Foundation.

End Notes

[1] This is so even if some important works contesting the hypothesis of efficiency were published prior to this event, such as the one by Shiller (1981).

[2] Let us remember what 'common knowledge' signifies. Towards this end, we introduce the term K_iP which can be read as 'agent *i* knows P'. We say that a proposition P involves common knowledge for agents 1 and 2 if: (1) P is true; (2) K_1P and K_2P are true; (3) K_2K_1P and K_1K_2P are true; and if this holds for an infinity of shared knowledge.

[3] At the end of the game, the assets are repurchased by the experimenter.

[4] On this subject, we should again refer to Noussair and Ruffieux who are extremely clear on this point: 'It is fashionable…to claim that experimental markets are irrelevant because they are much simpler in structure than markets outside of the laboratory that economists are interested in. The response to this critique is to recognize that experiments in economics are not necessarily intended to reproduce environments outside the laboratory. Many experimental economies are specified with the sole purpose of studying the conditions under which theory is pertinent. The assumptions of the theoretical model under investigation, rather than the structure of particular markets outside the laboratory, guide the design of the experimental economy' (Noussair and Ruffieux, 2001, p.2).

[5] See, inter alia, Charles MacKay's very famous analysis (1841) of the Dutch Tulip mania, a phenomenon whose very name is already highly evocative.

[6] Noussair and Ruffieux indicate 'three periods at most' (Noussair and Ruffieux , 2001, p.16).

[7] The illustration I am suggesting reveals a link between the potential degree of irrationality and the game's duration. It is therefore understandable that when one approaches the end of the game, it is not very plausible to expect the other participants to act irrationally. This could even be construed as a way of calculating what agents see as being the highest possible level of shared beliefs. Note, however, that a link of this nature is not always required: speculative reasoning such as 'I believe that the other believes that the other believes …' can be carried out instantaneously.

[8] Nagel also studies cases where *p* is above 1.

[9] T. Schelling (1960) criticized this solution. For further discussion, see Mehta, Starmer and Sugden (1994), and Orléan (1999, pp.74-81).

[10] One illustration is R. Shiller's concept of 'popular models' (1990).

[11] See Orléan (1999).

[12] In addition, note that information is not a commodity that exists per se. In reality, when agents react to a signal, it is because they expect the market to react to it. An item of information is only information insofar as agents believe that it will be information in the market's view. M. Brière (2000), when analysing the bond market, showed that the magnitude of people's reactions to signals depends on the importance granted to such signals in the prevailing model that agents use at any given moment in time. The meaning of a particular signal could therefore vary over time. This too is damaging to the fundamentalist approach. The concept of 'available information' loses meaning.

[13] Another way of making this same point consists of asserting that, as with Noussair, Robin and Ruffieux's model (2000), fundamental value has not successfully fulfilled its role, despite the following assumption: (1) of an explicit and stationary law of probability for dividends; (2) of a fundamental value that is known to everyone; and (3) of a fundamental value being forced on everyone at the end of the final period. We should stress that these three hypotheses are needed to validate backwards induction reasoning. In this kind of framework, if prices do not converge towards their fundamental value, what can be expected from situations in which the game has no end, where future dividends are totally unknown and where we have no exact knowledge of what the fundamental value is really worth?

[14] Euro-trader interviewed in Liberation newspaper, 8 September 2000, p. 24.

[15] As mentioned earlier, this stems from the fact that this particular model postulates that the fundamental value will be announced publicly during period three, and that it will serve as a basis for the liquidation of risky securities.