Rational Exuberance*

Stephen F. LeRoy
University of California, Santa Barbara

April 14, 2004

Abstract

In this article the theory of speculative bubbles is reviewed. It is argued that bubbles are a promising candidate as an explanation for the stock price runup and collapse of the 1990s in the US. The theory considers both irrational and rational bubbles, with emphasis on the latter. It is pointed out that rational bubbles—defined as the excess of security or portfolio prices over present values—can occur under conditions that are well understood. There exists an argument relying on the assumed Pareto-optimality of equilibrium that rules out rational bubbles, but it is suggested that this argument is implausible. JEL classification G12.

1 Introduction

Consider the postage stamp. As title to a future good (or, in this case, service) with monetary value, this humble object is essentially the same as a security. Its value, 37 cents, is the present value of the service (delivery of a letter) to which its owner is entitled.

Now consider a postage stamp with a minor printing error that excites stamp collectors' interest. Because of the printing error the stamp has a value of $1000. Viewing the stamp as a security suggests two possible explanations for the difference between the values of the stamp with and without the printing error: (1) the fundamental value of the stamp with the printing error is higher than that without, and (2) the stamp with the printing error has a bubble.

*Apologies to Alan Greenspan, whose famous 1996 speech raised the possibility of "irrational exuberance" in the stock market, and to Robert J. Shiller, who developed the idea in his book [57] of the same name. Thanks to seminar participants at Yale University and the University of Southern California. I have received comments from Ted Bergstrom, Christian Gilles, John Griffin, Boyan Jovanovic, Michael Magill, Martine Quinzii, Matt Spiegel, Douglas Steigerwald, Shyam Sunder and Jan Werner. Francisco Azeredo and Yongli Zhang found several errors. Thanks to Ying Sun for research assitance.

Attributing the higher value of the stamp with the printing error to its greater fundamental value would be justified if one believes that stamp collectors derive pleasure from contemplating the printing error. Alternatively, the collector may acquire status in the eyes of other collectors based on his or her ownership of the misprinted stamp, and this is the basis for the higher value. Along these lines, fluctuations in prices of collectible stamps are necessarily attributed to fluctuations in preferences. Such arguments are best seen as making an inference about utility based on reverse-engineering the price fluctuations: someone's marginal utility must be changing, or why would the price fluctuate?

George J. Stigler and Gary S. Becker [60] argued persuasively against relying on assumed preference shifts to explain such price fluctuations, especially when there exist alternative explanations that do not appeal to preference shifts. In the present context, if assets have bubbles their prices do not necessarily have the close association with agents' utilities that the above argument presumes. Specifically, in the presence of (rational) bubbles asset prices can exceed the discounted value of their payoffs or service flows, generally by a random amount. Therefore asset price fluctuations can be explained without resorting to random preferences.

It is easiest to argue against the proposition that the prices of collectibles reflect fundamentals when the objects in question clearly have minimal aesthetic value, as with misprinted postage stamps. With fine art the argument is less straightforward: someone who prefers Caillebotte to Andy Warhol might argue that the values of Caillebotte's paintings reflect their aesthetic value, whereas those of Andy Warhol's paintings reflect bubbles. Even here, however, the fact that a Caillebotte forgery that is undetectable except to an expert—and therefore presumably has the same aesthetic value as an original to anyone who is not an expert—has negligible monetary value argues that the value of a Caillebotte original is a bubble, just as with the misprinted postage stamp.

Even if bubbles exist on collectibles, there is no guarantee that they exist on assets like stocks or land, and still less that bubbles on such assets are the major cause of the price fluctuations that we see. It is possible that there are arguments against bubbles that apply to securities but not to collectibles. However, the reasoning just presented creates a presumption in favor of bubbles: if highly valued fine art objects are bubbles, why would the stock of corporations that own art objects not have a bubble component?

Ponzi schemes—pyramids in which the contributions of new investors are used to pay high returns to earlier investors—and chain letters are other possible examples of bubbles, although the interpretation of Ponzi schemes is clouded by the fact that investors have difficulty distinguishing them from highly successful genuine investments until it is too late.

The number of papers developing the economic theory of asset price bubbles has grown exponentially since the last major survey paper was published (Colin Camerer [10]; see also the *Journal of Economic Perspectives* 1990 symposium). At least equally important in motivating another visit to this topic, asset prices in the US and elsewhere in the last ten years underwent the runup and collapse that are widely associated with bubbles. Therefore this is a good time to provide a review of the economic literature on bubbles, parts of which are technical, that is as simple as possible (but, in Einstein's famous phrase, not more so).

Introductory discussion of bubbles is presented in Section 2. It is observed there that bubbles are often taken to be synonymous with irrationality. This characterization is universal in the popular press, and is also sometimes seen in academic discussions. We present, also in Section 2, introductory discussion of rational bubbles—instances of asset prices exceeding present values of future dividends in models in which agents optimize in an environment that they understand.

Analysts have proposed explanations for asset price behavior, even that of the late 1990s, that rely on fundamentals rather than bubbles. After looking at stock price data (Section 3), we review in Sections 4 and 5 some candidate explanations in terms of fundamentals, concluding that for various reasons these are difficult to credit. The remainder of the paper deals with bubbles.

In Section 6 further discussion of irrational bubbles is presented. We consider what, if anything, it means to appeal to irrationality in a substantive economic explanation of economic phenomena such as asset price fluctuations. Section 7 resumes the discussion of rational bubbles. Established professional opinion holds that there exist compelling theoretical arguments against existence of rational bubbles (Miguel Santos and Michael Woodford [52], for example). As we will see, it is true that there exists an argument that rules out bubbles under plausible parameter values. It will be suggested that this argument involves an application of rational expectations to events in the distant future that will strike many as at least controversial. Rejection of the argument, however, raises awkward methodological issues for the analysis of bubbles, as we observe in the conclusion. To these we have no answer.

The intent of this paper is to convince readers that bubbles are a viable candidate as an explanation for the volatility of asset prices, even if it is not entirely clear how bubbles should be modeled. Conceding this does not guarantee that bubbles do in fact constitute the explanation, particularly for those who do not share Stigler and Becker's reluctance to appeal to preference shifts as determinants of changes in fundamental values. As always, theoretical arguments depend ultimately on the plausibility of alternative sets of assumptions, about which there is no consensus. Ultimately it is an empirical question, and the empirical literature on bubbles is not yet well developed. Perhaps readers will turn their attention to developing empirical tests that can reliably distinguish between bubbles and other phenomena that affect asset prices, of which there is now a shortage.

2 Irrational and Rational Bubbles

In the financial media the question "Do you think internet stocks (real estate, Japanese stocks) are or were a bubble?" appears to mean that the questioner wants to know whether prices will collapse any time soon. Along these lines, a bubble means simply a big price rise that is shortly followed by an equally big price drop. Although the argument is seldom spelled out explicitly, the presumption is that bubbles should be avoided as investments because the probability of a collapse exceeds that of a further rise (allowing for the effects of risk-aversion and discounting). Under rational asset pricing, including rational expectations, such biased expectations cannot occur: absence of arbitrage implies that the expected (risk-adjusted and discounted) gain on any security or portfolio is zero. Thus in this usage bubbles are synonymous with irrationality.

The identification of bubbles with irrationality is found not only in the financial media, but also in many professional discussions: For example, Peter Garber [18] expressed the opinion that

> Bubble is one of the most beautiful concepts in economics and finance in that it is a fuzzy word filled with import but lacking a solid operational definition. Thus, one can make whatever one wants of it. The definition of bubble most often used in economic research is that part of asset price movement that is unexplainable based on what we call fundamentals.

Garber questioned the presumption implicit in these accounts that fundamentals cannot explain even such episodes as the tulip bulb speculation in Holland in the seventeenth century (Garber's book is reviewed in LeRoy [37]).

Similarly, Robert E. Hall [27] wrote

> I reject market irrationality in favor of the hypothesis that the financial claims on firms command values approximately equal to the discounted future returns.

Failure of prices to equal discounted payoffs is taken in this passage to be equivalent to irrationality, so there is no allowance for the possibility that values could exceed discounted future returns when agents are fully rational.[1] Bubbles as manifestations of irrationality are considered more fully in Section 6 below.

Bubbles can also be defined and analyzed in settings that do not involve irrationality. To define rational bubbles, begin with the definition of the rate of return r_{t+1} from t to $t+1$ on any security or portfolio:

[1] Later in his paper Hall acknowledged the possibility of rational bubbles, but dismissed them on the grounds that existence of such bubbles would violate "a fundamental efficiency condition". This argument is examined below.

$$r_{t+1} \equiv \frac{d_{t+1} + p_{t+1}}{p_t} - 1, \tag{1}$$

where d_t denotes dividends at date t and p_t is price. Taking expectations conditional on information at t and rearranging, there results

$$p_t \equiv \frac{E_t(d_{t+1} + p_{t+1})}{1 + E_t(r_{t+1})}. \tag{2}$$

Here $E_t(...)$ is short for $E(...|F_t)$, where F_t represents the information available at t. In (2) no distinction is made between subjective and objective expectations, or between the subjective expectations of different individuals. In effect, this amounts to assuming rational expectations and symmetric information. Under the assumption that conditional expected returns equal the constant r, we can write (2) as

$$p_t = (1 + r)^{-1} E_t(d_{t+1} + p_{t+1}). \tag{3}$$

Replacing t by $t + 1$ in (3) and substituting the result in (3) gives[2]

$$p_t = (1 + r)^{-1} E_t(d_{t+1}) + (1 + r)^{-2} E_t(d_{t+2} + p_{t+2}). \tag{4}$$

Repeating this operation n times results in

$$p_t = \sum_{i=1}^{n} (1 + r)^{-i} E_t(d_{t+i}) + (1 + r)^{-n} E_t(p_{t+n}). \tag{5}$$

Allowing n to go to infinity, (5) becomes

$$p_t = \sum_{i=1}^{\infty} (1 + r)^{-i} E_t(d_{t+i}) + \lim_{n \to \infty} (1 + r)^{-n} E_t(p_{t+n}), \tag{6}$$

assuming that the limits exist. Defining the first term on the right-hand side of (6) as f_t, the *fundamental value* of the security—its value based on finite-date payoffs—and the second as b_t, its *bubble,* we have

$$p_t = f_t + b_t. \tag{7}$$

The preceding analysis makes possible several preliminary observations about rational bubbles:

- By substituting (7) and the definitions of f_t and b_t in (3), there results

$$b_t = (1 + r)^{-1} E_t(b_{t+1}), \tag{8}$$

  so bubbles, if they exist at all, have the same expected return as stocks generally. In a deteministic setting this means that bubbles must grow at the interest rate.

[2] Here we use the rule of iterated expectations, which states, for example, that $E_t(E_{t+1}(d_{t+2})) = E_t(d_{t+2})$. Use of the rule of iterated expectations is valid whenever information is nondecreasing over time.

- If one assumes that a bubble will burst with constant probability at each date, its growth rate if it does not burst must be that much higher, so as to generate the expected return implied by (8). In such a setting a bubble can have positive value even though it will burst at some time in the future with probability one. (See Olivier Blanchard and Mark Watson [8] or Blanchard and Stanley Fischer [7], Chapter 5 for further discussion).[3]

- The value of any firm that pays zero dividends, and is expected never to pay dividends in the future, consists entirely of a bubble.

- Bubbles are nonnegative on any security that is freely disposable (and therefore necessarily has a nonnegative price), such as stocks.

- Bubbles can occur only in models with an infinite number of dates; otherwise a backward induction (using (8) and the terminal condition $b_T = 0$, where T is the last date) implies $b_t = 0$ for all t.

3 The US Data

Figure 1 shows the value of US equity (from the Federal Reserve Board Flow of Funds accounts) divided by GDP. Data are from the 1950s to now.[4] The salient feature of this series is its dominance by low-frequency components. Stock prices rose in the 1950s and 1960s, fell in the 1970s, rose in the 1980s and 1990s and, so far in the 21st century, have been falling, at least up to 2003. The high-frequency component of stock price variation that dominates financial reporting is seen to be of minor importance by comparison with the low-frequency variation just discussed. The runup over the period 1995-2000 is conspicuous, as is the even more rapid subsequent collapse in prices.

Figure 2 shows stock values normalized instead by National Income & Product Accounts corporate earnings. For most of the sample period the two figures look similar. To the extent that price-earnings ratios show variations similar to price-GDP ratios, the interpretation is that stock price variations cannot be viewed as simply a proportional response to parallel trends in earnings. For most of the sample period the low-frequency variation in Figure 2 is positively correlated with, but less pronounced than, that in Figure 1. Thus stock price variations are correlated with variations in earnings, with the stock price variations being of greater amplitude.

[3] This result appears to violate the basic precept from probability theory that zero-probability events can be ignored. It does not. Demonstrating this is not appropriate here; see Christian Gilles and LeRoy [20].

[4] For historical discussion of bubbles, see Charles Mackay's classic 1841 volume [41]. More recent treatments are Garber [18] for the early bubbles in Europe and John Kenneth Galbraith [17] for the 1929-1932 US stock price selloff. General introductions were provided in Charles P. Kindleberger [32], Edward Chancellor [11] and Shiller [57].

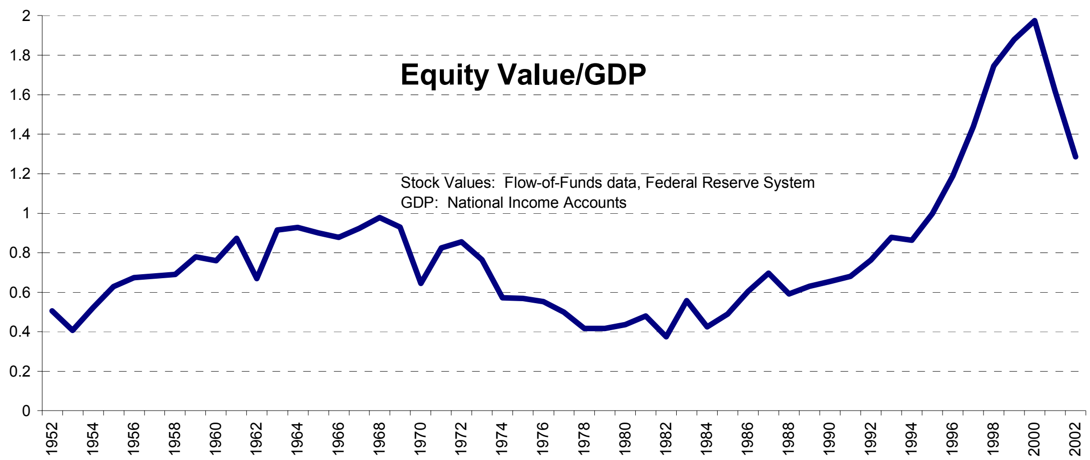

Figure 1:

The runup of stock prices in the late 1990s appears similar in the two series, suggesting that the stock price rise is not simply a proportional response to spectacular increases in earnings. Indeed, Figure 3, which displays aftertax corporate earnings as a proportion of GDP, shows that while earnings rose in the middle and late 1990s, even at their peak they were a smaller proportion of GDP than during most of the postwar period.

In contrast, the collapse in stock prices over the past several years that is conspicuous in Figure 1 has no counterpart in Figure 2. The reason is that, as Figure 3 shows, corporate earnings underwent an even more pronounced drop than stock prices as the US economy entered a recession. Consequently the price-earnings ratio rose even as the price-GDP ratio fell. Partial data for 2002 (not displayed), in fact, show a further increase in the price-earnings ratio.

4 Fundamental Explanations—Stationary Settings

Determining whether these variations in stock prices are what would be expected under the present-value model, given the behavior of corporate earnings and dividends and proxies for discount factors, requires restricting the present-value model so that it generates clear empirical predictions. A useful place to begin is the deterministic Gordon model (Myron J. Gordon [22]). Subsequently we will generalize to a stochastic version of the model.

The deterministic version of the Gordon model is generated by five assumptions:

(1) the (real) return on invested capital (earnings-price ratio) is constant over time, (2) the rate of earnings retention (equivalently, dividend payout) is constant over time, (3) retained earnings generate the same returns as preexisting capital (so that there are no opportunities for extranormal earnings), (4) the value of equity equals the discounted value of future dividends (so that there are no bubbles), and (5) the discount factor is constant over time, with the discount rate r equal to the rate of return on invested capital. Under these assumptions the dividend stream is a perpetuity that grows at rate g. Discounting at rate r, we have that p_t, the value of equity at date t, is given by

$$p_t = \sum_{i=1}^{\infty} \frac{d_{t+i}}{(1+r)^i} = \sum_{i=1}^{\infty} \frac{d_t(1+g)^i}{(1+r)^i} = \frac{d_t(1+g)}{r-g} = \frac{e_t(1+g)}{r}. \tag{9}$$

Here e_t is current earnings and

$$g = r(1-\delta) \tag{10}$$

is the growth rate of the dividend stream, where δ is the dividend payout rate (proportion of earnings paid out as dividends).[5] Thus the value of equity can be represented either as the present value of a growing dividend stream or as the present value of a constant earnings stream. The fact that p_t does not depend on δ reflects the Miller-Modigliani [46] proposition that (in this simple environment) for given current earnings the future dividend payout rate does not affect current equity value.

For the present purpose, the main shortcoming of the Gordon model is that it is deterministic. As such, it provides no insight into stock price fluctuations. To remedy this deficiency we replace the assumption that the earnings growth rate is constant with the assumption that the earnings growth rate is independently and identically distributed over time or, put differently, that earnings follow a geometric random walk.[6] We refer to the model that results from this generalization as the *stochastic Gordon model*.

The present value relation (6), together with the assumed absence of bubbles, implies that stock prices equal the discounted value of future expected dividends:

$$p_t = f_t = \sum_{i=1}^{\infty} (1+r)^{-i} E_t(d_{t+i}). \tag{11}$$

In order to determine the behavior of stock prices implied by (11) under the stochastic Gordon model it is necessary to specify how much information investors use in forming the conditional expectations $E_t(d_{t+i})$. Investors can form relatively precise estimates

[5] A firm with capital k_t at t will have retained earnings $k_t r(1-\delta)$, which equals $k_{t+1} - k_t$. If the growth rate of capital $k_{t+1}/k_t - 1$ is defined as g, then (10) results.

The rightmost expression in (9) is derived by substituting the right-hand side of (10) for g in the denominator of the fourth term of (9).

[6] The geometric random walk was found more than 40 years ago to give an accurate description of corporate earnings in the UK (A. C. Rayner and I. M. D. Little [49], W. B. Reddaway [50]).

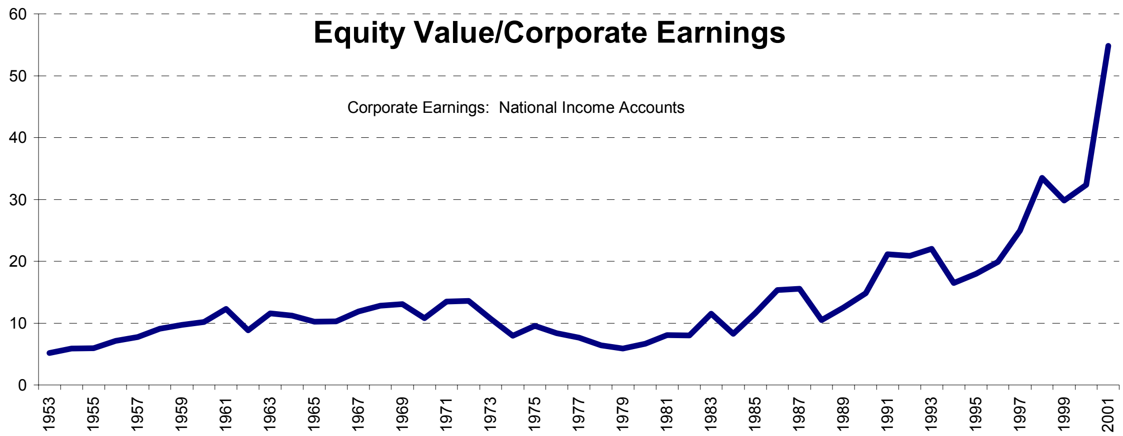

Figure 2:

of dividends in the near future, but their forecasts of dividends in the distant future are much less accurate; the question is how quickly the information accuracy decreases. One way to parametrize the stochastic Gordon model is to assume that investors can forecast dividends with perfect accuracy up to n periods in the future, but they have no information beyond $t+n$, implying that for $m > n$ investors estimate d_{t+m} by extrapolating from d_{t+n}. This all-or-nothing specification, although unrealistic, gives an easy way to generate insight about qualitative implications for the data of the assumption that investors have little (n low) or much (n high) information about future earnings and dividends.

Two extreme cases can be specified: (1) investors have complete information about future dividends ($n = \infty$), and (2) investors have no information beyond the current level of dividends that enables them to forecast future dividends ($n = 0$). When $n = \infty$ there are never any surprises, so the rate of return on stock is deterministic:

$$\frac{d_{t+1} + p_{t+1}}{p_t} - 1 = r \tag{12}$$

for all t. For $n = 0$ investors extrapolate future dividends from current dividends, so that the dividend-price ratio is deterministic. In fact, both rates of return and dividend-price ratios have nonzero unconditional volatility, so under the maintained assumption of the stochastic Gordon model, n should be taken to have intermediate values.

The assumptions used to generate the stochastic Gordon model are reasonably accurate, at least given the sparse parametrization of the model. Figure 4 shows

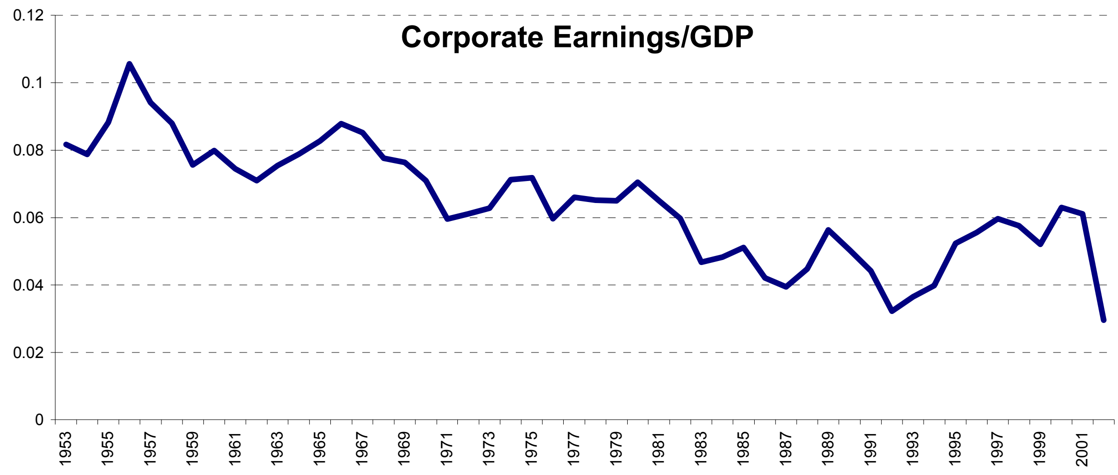

Figure 3:

the interest rate less the growth rate of earnings; the series looks like an IID process, as assumed in the stochastic Gordon model.[7] The first four autocorrelations of the earnings growth rate are 0.050, −0.145, −0.340 and −0.078 (annual data, 1958 to 2000), so one would not want to defend too strongly the assumption that earnings growth is white noise. The predominantly negative autocorrelations imply that earnings have a mean-reverting component. To the extent that investors take this mean-reversion into account in valuing stocks, the model to be presented gives an upward-biased account of stock price volatility.

The stochastic Gordon model imposes the assumption that the dividend payout rate is constant, so the autocorrelations of dividends would be the same as those of earnings. In fact the dividend-payout rate is far from constant (Figure 5), and the dividend growth rate is highly autocorrelated. However, the mean-reverting character of the dividend payout rate suggests that the failure of dividends to adjust immediately to earnings changes should not greatly distort security prices relative to the prediction of the Gordon model.

Assuming that risk premia can be taken as approximately constant, the Fisher relation (which postulates that the nominal interest rate rises and falls one-for-one with expected inflation) implies the constancy of the discount factor, as presumed in the Gordon model. Figure 6, which shows the nominal interest less the annual rate of inflation, indicates that constancy is not a bad approximation.

[7]The attractive feature of this variable is that we do not have to worry about inflation adjustment, since the inflation correction involved in figuring real earnings growth offsets that involved in

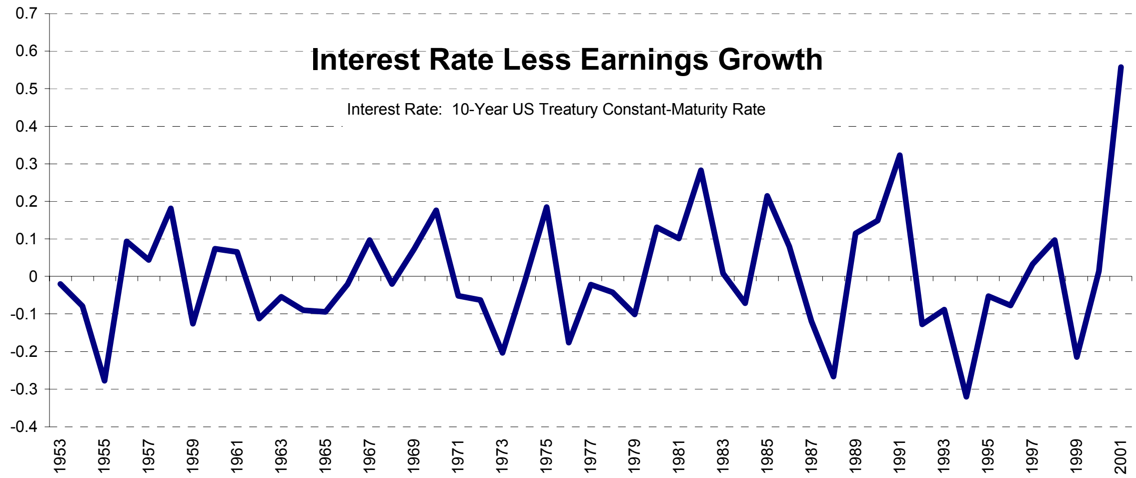

Figure 4:

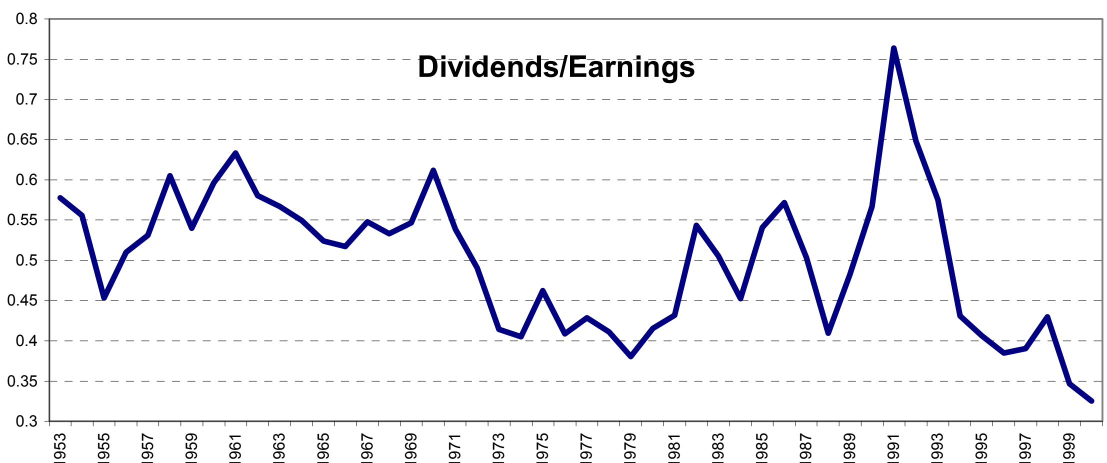

Figure 5:

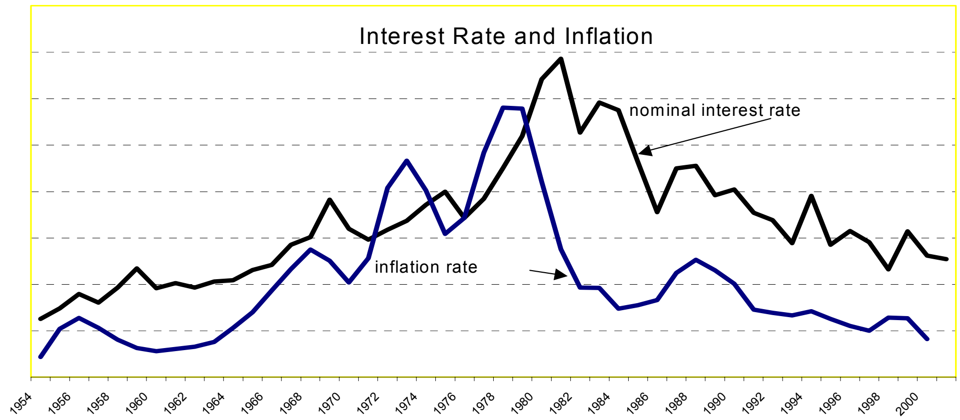

Figure 6:

So specified, the stochastic Gordon model can reproduce the volatility of the price-earnings ratio or the volatility of stock returns, but not both at the same time. Specifically, if n is high—on the order of 10 years or more—the price-earnings ratio produced by the model has fluctuations on the same order of magnitude as those of its real-world counterpart. However, in that case stock returns are predicted to have much lower volatility than we see in the data, since by assumption fluctuations in returns are attributable to earnings realizations at least n years in the future, and these are discounted almost to zero. In contrast, if n is low—one or two years—the volatility of returns can be matched, but price-earnings ratios are predicted to be nearly constant, since investors are assumed to have little information beyond current earnings on which to base estimates of future earnings. For intermediate values of n the stochastic Gordon model underpredicts the volatility of both price-earnings ratios and returns.[8]

Inspection of the diagrams makes clear that the principal source of both price and return volatility in the postwar period is the stock price runup in the 1990s and the subsequent price collapse. Can this specific episode be interpreted within the framework of the stochastic Gordon model? Under the stochastic Gordon model, price-earnings ratios will be unusually high when investors have information that leads them to predict extranormal future earnings increases. As the data indicate, earnings were in fact increasing during the late 1990s, and it is reasonable to suppose that investors were extrapolating these increases into the future. However, the increases in price-earnings ratios exceeded by orders of magnitude anything justifiable under the Gordon model, barring wildly optimistic earnings forecasts.

calculating the real interest rate.

[8] See LeRoy and William R. Parke [38] for discussion of these variance-bounds tests.

5 Fundamental Explanations—Nonstationary Settings

We see that the stochastic Gordon model reinforces the conclusion suggested by uninstructed examination of the data: generally, the volatility of stock prices and returns exceeds the volatility that can be justified in terms of underlying fundamentals. Specifically, the price runup of the 1990s and the subsequent collapse cannot be described as a typical response of stock prices to fundamentals. It is clear that no model that treats the 1990s as a typical drawing from a stationary population will succeed in explaining that episode empirically.

Fundamentals-based explanations that abandon stationarity in favor of explanations that point to the particular circumstances of the late 1990s are harder to reject. Whether due to the advent of the internet or computerization more generally, there was in fact evidence of accelerating productivity growth in the US economy, implying higher growth rates of earnings and dividends and therefore, by the Gordon model, higher price-earnings and price-dividends ratios. With the hindsight afforded by the collapse of the dot-com and tech stocks, such stories are now generally dismissed as puffery. However, the remarkably high rates of productivity growth that have occurred in the US economy generally over the past several years, even following the stock price collapse and during the recent recession, suggest that the justifications for high stock prices that circulated during the late 1990s may be more than internet hype. Of course, if accelerated productivity growth caused the price runup in the 1990s, we are left without an explanation for the price collapse, given that productivity growth has been continuing at very high rates.

A variety of nonstationary, but fundamentals-based, models have been proposed to explain both the pronounced low-frequency variation in stock prices generally and the 1990s episode. In popular discussion the price runup is often attributed to the large baby-boomer cohort saving for retirement. A recent paper by John Geanakoplos, Michael Magill and Martine Quinzii [19], following G. S. Bakshi and Z. Chen [4], noted the striking positive correlation of stock prices in the postwar period with measures of the relative size of middle-age cohorts in the population. The idea is that stock prices will be high when the middle-aged cohort is large, since they are doing most of the saving. Geanakoplos, Magill and Quinzii developed an overlapping generations model with varying population growth rates and calculated the stock price fluctuations implied by demographic shifts.

Geanakoplos, Magill and Quinzii's model implies that the price appreciation of the 1980s and 1990s was forecastable. This raises the question of why bond returns did not show an increase during that period commensurate with that for stocks, as portfolio balance implies. Also, the baby-boomer cohort was noted for exceptionally low rates of saving, suggesting that the direction of causation between personal saving and stock prices is the reverse of that implied by Geanakoplos, Magill and Quinzii's model. Finally, whatever the merits of demographic factors as explaining low-frequency stock price changes, it is difficult to connect these with the stock market selloff of the last several years.

In a series of papers Hall examined stock prices in the late 1990s (Hall [25], [26], [27], [28]). In [27] Hall began by dismissing the possibility that the stock price runup of the 1990s could reflect rational bubbles: "A fundamental efficiency condition holds that the discount rate exceeds the rate of growth of output and other quantities" (p. 3), implying that bubbles, if nonzero, increase faster than national product. He observed that "It is essentially impossible to build a model in which intelligent people believe that the value of a stock will become larger and larger in relation to all other quantities in the economy" (p. 3). Hall went on to make the case for fundamentals, drawing on several independent analytical settings that were not completely integrated. First, Hall pointed out that in a one-good production model (without adjustment costs) the value of securities equals the quantity of capital. This can be measured independently of security prices using corporate accounting data. Estimates of tangible capital as a proportion of GDP show that it is much less volatile than stock valuations. In Hall's diagrams, in fact, the two appear to be negatively correlated. Further, tangible capital has been decreasing relative to GDP since the early 1980s, making it an unlikely candidate to explain the price runup of the 1990s.

Hall then considered intangible capital, measured as the difference between security values and tangible capital. Intangible capital so defined varied from 30 per cent of GDP in the 1960s to -30 per cent in the 1970s. It rose to about 100 percent of GDP in the 1990s. Given both the timing of the price runup and the industry distribution of intangible capital, Hall connected it to computerization and the internet—"e-capital", in Hall's phrase.

Hall also offered the possibility of justifying the stock market values of the 1990s in terms of a variant of the deterministic Gordon model discussed above. He adopted the perpetuity valuation model (9), with the significant change that the discount factor and growth rate of cash flows are time-subscripted. He concluded that the high growth rate of cash flows in the late 1990s was sufficient to explain the increase in stock prices.

Hall offered these various lines as possible explanations, and only that, for the price runup. None of them seems very persuasive:

- One problem with Hall's appeal to a "fundamental efficiency condition" to eliminate bubbles is that in many infinite settings the allocations that occur in equilibrium are not Pareto optimal, so efficiency conditions are not necessarily satisfied.

- Another problem has to do with the plausibility of the argument that investors or collectors extrapolate the future price paths implied by current values of a misprinted postage stamp or a Caillebotte painting that contain bubble components to the point where these values exceed GDP. This argument is discussed further below.

- In distinguishing between tangible and intangible capital, Hall departed from the one-good setting in which securities values equal the quantity of a homogeneous capital good (unless tangible and intangible capital are perfect substitutes, in which case they can be aggregated). It is not evident from Hall's discussion how intangible capital comes into existence, except that it is clearly put in place by some process other than the corporate saving (net of depreciation) that determines changes in tangible capital.

- As Hall noted, the negative levels of intangible capital in the 1970s are difficult to interpret.

- Writing in 2001, Hall did not have to explain the subsequent evaporation of several trillion dollars of intangible capital, but it cannot be ignored today.

- Deriving the version of the Gordon model that Hall used to value stocks requires constancy of the discount factor and the expected cash flow growth rate. Hall's version has agents acting as if the future cash flow growth rate will be constant at its current value, even though that value changes over time. This is inconsistent with rational expectations. (Robert E. Lucas, in his critique [39] of econometric policy evaluation, made the same observation about Hall and Dale Jorgenson [29]).

- Finally, the basis for Hall's conclusion that the growing cash flows of the late 1990s are sufficient to justify the stock prices is not clear. In his diagrams, like those presented above, cash flows were high and growing in the 1990s, but not more so than in several earlier periods. It is difficult to believe that any calibration that reproduces the stock prices of the late 1990s will not produce wildly inaccurate predictions for the stock market in earlier periods.

Like Hall, Ellen R. McGrattan and Edward C. Prescott [43] attributed the high value of the stock market around the year 2000 to intangible capital. Rather than defining intangible capital as a residual, as Hall did, McGrattan and Prescott estimated the magnitude of intangible capital from data on corporate profits and the return on bonds. In the deterministic version of their model, this calculation is particularly simple: in equilibrium equity must have the same return as debt, making possible a direct calculation of the amount of capital required to generate observed corporate profits. Under uncertainty it is necessary to make allowance for the effects of risk aversion on the equity premium; the authors used a calibrated real business cycle model for this calculation. As would be expected from Rajnish Mehra and Prescott's earlier finding in [45] that the return to conventionally measured equity capital is much higher relative to bond returns than risk aversion justifies, this calculation led to a high estimate of intangible capital. In fact, the authors concluded that equities were approximately correctly valued in 2000. It follows that stocks were undervalued before and after 2000.[9]

A number of other papers have proposed explanations for the long-term variations in stock prices that we have seen. For example, John Laitner and Dmitriy Stolyarov [34] observed that major innovations have the effect of reducing the market value of the preexisting capital that is rendered obsolete by the innovation (see Jeremy Greenwood and Boyan Jovanovic [23] and Bart Hobijn and Jovanovic [30] for similar analysis). They suggested that the low stock prices of the 1970s may be attributable to this effect. It is difficult to believe that such effects are strong enough to account for the stock price runup and decline of the last ten years. We are led to reject the hypothesis that price variations of the amplitude that we have seen can be explained in terms of changing present values of rationally forecasted dividends.

6 Irrational Bubbles

The meaning of "rationality" and "irrationality" in economic discussion appears to have changed in recent years. In mainstream nonfinancial economic theory the received view has been that rationality is not a substantive hypothesis about the world, but rather a conceptual tool used in formulating economic models. In fact, the idea of rationality provides the basis for a working definition of an economic model as a model describing human (or, for that matter, animal) behavior that assumes consistent choice. Along these lines there is no room for irrational behavior: irrational behavior, as manifested by apparently nontransitive choices, is actually evidence of omitted costs or the like. As this argument makes clear, taking rationality as an analytical tool rather than a substantive hypothesis gives rationality a tautologous character. As part of a maintained hypothesis, no conceivable evidence can contradict it.

A generation ago financial economists fully accepted this view of rationality, if the term "market efficiency" is substituted for "rationality". In his classic survey of efficient capital markets, Eugene F. Fama [14] emphasized that market efficiency by itself has no observable implications. It can be tested only in conjunction with a particular market model.[10]

Many contemporary writers are more sympathetic to the idea of irrationality than mainstream financial economists were a generation ago. However, in making their point that irrationality plays a greater role in financial markets than hitherto believed, these writers have altered the definition of rationality. Irrationality is identified not

[9] In more recent work McGrattan and Prescott [44] emphasized the role of changing tax rates and regulatory environments in explaining equity valuation.

[10] To be sure, in the early empirical literature on market efficiency financial economists, Fama included, were not always clear about exactly what market model was embedded in the hypothesis being tested. Accordingly, they often described empirical results as favoring or not favoring market efficiency directly, a usage that conflicts with Fama's dictum.

with nontransitivity, but with the existence of agents who trade for reasons that are not modeled ("noise traders", introduced into the finance literature by Albert S. Kyle [33] and Fischer Black [6]). Along these lines bubbles are presumably defined as the difference between asset prices as they are and asset prices as they would be in the absence of noise traders, although it is difficult to find an explicit source for this.

Models in which some of the economic behavior being modeled is taken as given, as distinguished from models in which all agents are represented as optimizing subject to constraint, have been known in the received economics literature as partial equilibrium models. There was no suggestion in the received literature that partial equilibrium has anything to do with irrationality. It is difficult to see the advantage in broadening—and blurring—the definition of irrationality so that it coincides with partial equilibrium, but there is little doubt that this has happened.

In the recent literature the standard vehicle for analyzing the effects of irrationality (under the expanded definition) is a model in which there exist both noise traders and rational traders (for example, Andrei Shleifer [59]). Oddly, this setting coincides exactly with that envisioned in the early finance literature as embodying efficient markets. It was emphasized in the efficient markets literature that market efficiency does not require exclusion of noise traders as long as there are enough rational agents to dominate asset pricing (formally, this means that in equilibrium the portfolios of the rational agents are interior). Proponents of market efficiency emphasized that in the setting just described the major propositions of neoclassical finance theory, such as the risk-return tradeoff, will survive the presence of noise traders, implying nonexistence of profitable trading rules.

As noted, recent writers have adopted exactly the same analytical setup, but interpreted it as reflecting inefficient rather than efficient capital markets. Rather than identifying market efficiency with the risk-return tradeoff (equivalently, the nonexistence of profitable trading rules), they emphasized a different aspect of equilibria in models populated by both rational and irrational agents: that in general the existence of noise traders will affect asset prices in equilibrium even if the rational agents are at interior optima. This is so because if the rational agents are risk averse, they will rationally reject trades that exploit minor mispricing because of the added risk, implying that they will generally not completely eliminate the effects of noise traders on security prices (John Maynard Keynes [31], p. 157, J. Bradford DeLong *et al.* [12]).[11]

Let us specialize this discussion to bubbles. As noted above, in the literature associating bubbles with irrationality, bubbles are defined, usually implicitly, as the

[11] In the early literature on market efficiency it was presumed (although generally not stated) that the rational agents would completely arbitrage away the effects of noise trades on security prices. This conclusion, we now know, follows only if agents are risk neutral, a specification often made implicitly in early discussions of market efficiency (LeRoy [35], [36]). This is so because if all agents are risk neutral, then rational traders will bid all asset prices to levels that equate expected returns, so the rational traders completely eliminate the effect of noise traders on asset prices.

difference between asset prices as they are and asset prices as they would be in the absence of noise traders. They reflect the herd behavior of traders, who bid prices above levels that can be justified by any rational calculation. Rational agents perceive the mispricing and bet against it but, because they are risk averse, their trades are not sufficient to eliminate the mispricing.

How realistic is this story? In the 1990s US stock price runup it was easy for any trader to bet against the bubble, for example by selling NASDAQ futures or buying puts. Few perceived doing so as an attractive trading opportunity, even on a small scale, for the obvious reason that the mispricing might increase, implying losses. The few investors who made this bet and did so too early did in fact post losses, sometimes substantial. In particular, it appears that institutional traders, who presumably correspond to the rational agents of the models, typically bet with the bubble rather than against it (see, for example, John M. Griffin *et al.* [24]).

The fact that virtually all of the traders who believed that stocks were overvalued still declined to bet against the bubble might mean only that there exist very few rational agents. A more likely possibility, we argue, is that betting against the bubble is not profitable, so the fact that few traders did so does not contradict the proposition that most investors trade rationally. This points toward rational bubbles, analyzed below.

Two qualifications are necessary. First, papers like that of Dilip Abreu and Markus K. Brunnermeier [2] presented models in which rational traders bet with the bubble rather than against it. To the extent that these models are valid, failure to locate rational traders who bet against the bubble does not count against the hypothesis of irrational bubbles. Second, in asset price runups prior to the 1990s, index futures and options did not exist, implying that it was impossible to take a short position against the market as a whole. To be sure, investors could take short positions in individual issues that they believed to be overpriced. However, doing so involves foregoing the benefits of diversification. Also, taking short positions on individual issues involves the risk of being subjected to a short squeeze, which was a regular feature of early stock price manipulations (Chancellor [11]).[12] Therefore the foregoing argument against the irrationality of the 1990s bubble does not necessarily apply to earlier episodes. Also, it does not apply to chain letters and pyramid schemes, which cannot be sold short.

A number of recent papers model bubbles in a setting that does not presume full rationality and rational expectations, but also avoids the noise trader construction (that is, these papers avoid including in the model traders whose behavior is taken as given rather than explained in terms of optimization). Instead, the traders are assumed to optimize, but also to have expectations that are biased in some way. For example, DeLong *et al.* [13] analyzed a setting in which traders are systematically

[12] A short squeeze occurs when a trader buys such a large quantity of a security that short sellers are forced to close out their positions at high prices. Short squeezes are impossible to implement on index futures because futures transactions are settled in cash.

optimistic, whereas Jose Scheinkman and Wei Xiong [53] assumed a setting in which agents place more weight on their own information than its accuracy justifies.

7 Rational Bubbles

We have seen that when bubbles are modeled in settings that include noise traders, the bubble is interpreted as reflecting the effect of noise traders on asset prices. In rational-agent models there are no noise traders, so that definition is inapplicable. We saw in Section 2 that in rational-agent settings bubbles are instead defined as the difference between the actual price of an asset and the present value of its payoff.

In some settings rationality, including rational expectations, is enough to rule out bubbles so defined. We observed in Section 2 that bubbles can be ruled out in finite-time settings by a backward induction. Bubbles can also be excluded in the following infinite-time example, which is a general-equilibrium version of the deterministic Gordon model.

Example 1: Consider a deterministic exchange economy in which the representative agent has utility function

$$U(c) = \sum_{t=0}^{\infty} \beta^t c_t. \tag{13}$$

(Alternatively, one could use a utility function like the more conventional

$$U(c) = \sum_{t=0}^{\infty} \beta^t v(c_t) \tag{14}$$

which, if v is strictly concave, incorporates risk aversion. We choose the simpler (13) because risk-aversion is irrelevant in this context.) The representative agent's endowment x_t grows at a constant rate g:

$$x_t = x_0(1+g)^t, \tag{15}$$

where $\beta(1+g) < 1$ to ensure that consuming the aggregate endowment generates finite utility. With $g = 0$ we have the tree economy of Robert E. Lucas [40], specialized to a deterministic setting.

The representative agent's budget constraint is

$$\sum_{t=0}^{\infty} (1+r)^{-t}(c_t - x_t) \leq 0, \tag{16}$$

anticipating that in equilibrium the interest rate r will be constant. If there is no bubble, the equilibrium price $p(y)$ of an arbitrary payoff stream $\{y\}$ is given by

$$p(y) = \sum_{t=1}^{\infty} (1+r)^{-t} y_t \tag{17}$$

where the equilibrium interest rate r is related to preferences according to

$$\beta = (1+r)^{-1}. \tag{18}$$

If (18) failed the representative agent would reject consuming his endowment in favor of borrowing or lending, depending on the direction of inequality. This cannot occur in equilibrium.

Measured in units of date-t consumption, the date-t value of the aggregate endowment, as the capitalized value of a growing perpetuity, satisfies the Gordon equation

$$p_t = \frac{1+g}{r-g} x_t, \tag{19}$$

assuming that there is no bubble. If there were a bubble on the aggregate endowment, and that bubble had date-0 value b_0, (19) generalizes to

$$p_t = \frac{1+g}{r-g} x_t + b_0(1+r)^t, \tag{20}$$

drawing on the demonstration in Section 2 that a deterministic bubble must grow at a rate equal to the interest rate.

For this economy we can state definitively that the aggregate endowment cannot have a bubble: $b_0 = 0$. If on the contrary we had $b_0 > 0$, the representative agent would want to sell the aggregate endowment at date 0 and consume the proceeds of the sale immediately. Doing so would increase period utility at date 0 by $(1+g)x_0/(r-g) + b_0$ and would decrease discounted future utility by $(1+g)x_0/(r-g)$, for a net gain of b_0. But, due to the repesentative agent setting, there is no one to take the opposite side of this trade, so markets could not clear.

Formally, what is happening here is that in the setting just outlined optimal portfolio choice implies not only the usual first-order conditions that are necessary in finite-dimensional optimizations, but also a necessary transversality condition. A portfolio strategy consisting of a zero trade at each date (which must be optimal in any representative-agent environment) violates the transversality condition in the presence of a nonzero bubble. Therefore only $b_0 = 0$ can occur in equilibrium This argument eliminating bubbles appears to be due to Maurice Obstfeld and Kenneth Rogoff [47].

A minor modification of the setup of Example 1 produces securities with bubbles. Suppose that the aggregate endowment is owned by competitive firms, and the firms distribute earnings by repurchasing shares on the open market rather than paying dividends. In that case the shares of stock—which can be identified with the payoffs of a buy-and-hold portfolio strategy—are a pure bubble, since there is no dividend at any date. Of course, in this case a buy-and-hold portfolio strategy is not optimal, or even feasible, and the equilibrium portfolio strategy that is optimal for the representative agent (sell shares as the firms repurchase them) does not have a bubble. Double taxation of dividends creates a strong incentive for firms to avoid dividends in this way, resulting in dividend payout rates that have been dropping over time, as seen above. Given recent legislation reducing the double taxation of dividends, this trend may not continue. However, to the extent that it does continue, we will be moving toward a situation in which stock prices are pure bubbles.[13]

Following the papers of Neil Wallace [63] and Jean Tirole [62], overlapping generations models have been the vehicle of choice in studying rational bubbles. In overlapping generations models (that assume a setting of pure exchange) there is a connection between the Pareto optimality of the endowment allocation, or lack thereof, and the necessary nonexistence or possible existence of equilibria with bubbles. If the endowment allocation is Pareto optimal, bubbles cannot exist, while if it is not Pareto optimal, equilibria with bubbles can exist. This is so because economies in which the endowment allocation is Pareto optimal have positive net interest rates (or, in growing economies, net interest rates that exceed the rate of output growth) in equilibria without bubbles. Further, in overlapping generations settings existence of a bubble will decrease the consumption of the young and increase that of the old, therefore increasing the equilibrium interest rate at each date. That being so, if there were a bubble on some security (strictly, on a buy-and-hold portfolio strategy involving some security, so as to rule out bubbles due to share repurchases), the value of that security would necessarily exceed the aggregate endowment of the economy sooner or later. Transferring the security from generation to generation would therefore eventually imply negative consumption for the young, invalidating the assumed solution path.[14] This argument is the basis for Hall's objection to rational bubbles,

[13] See Christian Gilles and LeRoy [21] for related discussion.

Lucy F. Ackert and Brian Smith [3] suggested that the finding of Shiller [56] and others that stock prices are more volatile than the present-value model implies may be due to the use by these analysts of "narrow dividends" in place of the allegedly theoretically correct "broad dividends", defined as narrow dividends corrected for share repurchases and the like. In fact neither broad nor narrow dividends is more or less theoretically correct than the other—they simply describe different portfolio strategies either of which may or may not have a bubble. If price volatility is excessive under narrow dividends but not under broad dividends, the explanation may be that narrow dividends have a bubble but broad dividends do not.

[14] This argument appears to depend on the restriction that only positive consumption is admissible, as is necessarily the case with the logarithmic utility that is usually specified in analyses of overlapping generations models. In fact positivity is not needed: even with utility functions that take negative as well as positive arguments, such as negative exponential utility, solution paths for which the assumed initial security prices exceeds its fundamental value fail in finite time. See LeRoy

discussed above.

Example 2: Consider a deterministic overlapping generations model. Generation t $(t = 0, 1, 2, ...)$ is alive at dates t and $t+1$ and has utility function $\ln(x_t^t)+\ln(x_{t+1}^t)$, where x_{t+1}^t is the consumption of generation t at date $t+1$. Each generation has an endowment of one unit of consumption when young and two units when old. Under this endowment specification the endowment allocation is Pareto optimal. The 0-th generation also has an endowment of money; the question is whether this money (necessarily a bubble since it has zero dividends) can have positive value in equilibrium. In equilibrium each generation divides the value of its endowment equally between consumption when young and consumption when old, leading to

$$x_t^t = \frac{p_t + 2p_{t+1}}{2p_t} \tag{21}$$

$$x_{t+1}^t = \frac{p_t + 2p_{t+1}}{2p_{t+1}}, \tag{22}$$

$t = 1, 2, 3, ...$ Here p_t is the price of consumption at date t relative to an abstract numeraire. Combining these demand functions with the feasibility condition $x_{t+1}^t + x_{t+1}^{t+1} = 3$ results in the difference equation

$$p_{t+2} = \frac{3p_{t+1} - p_t}{2}, \tag{23}$$

which must be satisfied on any equilibrium path.

The restriction $p_0 = 1$ may be imposed as a numeraire choice, while p_1 indexes the equilibrium paths. Choosing $p_1 = 1/2$ results in the equilibrium price sequence $1, 1/2, 1/4, 1/8, ...$, which implies the autarky allocation in which there is no bubble. Setting $p_1 < 1/2$ generates equilibrium paths with positive bubbles. For example, with $p_1 = 0.48$ the bubble at date 1 has value 0.0417, implying that generation 0 can consume 2.0417 when old instead of the endowment allocation of 2.

	Date	0	1	2	3	4	5
Generation							
0		1	2.0417				
1			0.9583	2.0909			
2				0.9091	2.2222		
3					0.7778	2.8	
4						0.2	3.6667
5							-0.6667

[37].

As the accompanying table shows, the bubble increases in value, so the young of successive generations consume less and less and the old more and more. At date 5 consumption of the old exceeds the aggregate endowment, implying that consumption of the young becomes negative. This invalidates the equilibrium path.

The preceding example illustrated the fact that in infinite-time models bubbles cannot occur when the endowment allocation is Pareto optimal. Equilibrium trajectories with bubbles can be feasible over infinite time if the endowment allocation is not Pareto optimal. A simple way to demonstrate this is to repeat the analysis of the above example, but reversing the endowments of the young and old (so that young agents have two units of consumption when young and one when old). In that case the net interest rate implied by the endowment allocation is negative. The same will be true under the equilibrium allocation implied by a bubble of preassigned date-1 size, as long as the bubble is not too large. In this setting rational bubbles shrink relative to the size of the economy, implying that they are feasible in infinite time.

The question for the existence of rational bubbles in overlapping generations models is seen to be whether or not the endowment allocation is Pareto optimal. Andrew B. Abel *et al.* [1] argued that empirically average security returns exceed the average growth rate of the economy, implying that, by the argument just given, bubbles cannot exist. This argument, while formally correct, is not convincing as a proposition about the real world. It illustrates a tendency, especially pronounced in recent years, to appeal uncritically to rational expectations. The argument presumes that agents, knowing the model that describes the environment that they inhabit, can calculate the trajectory of the economy into the arbitrarily distant future. Now, representing agents as optimizing in an environment that they understand is plausible in repetitive situations, where it can be assumed that they have learned the relevant patterns. In the context of bubbles, however, we are assuming that agents somehow come to understand the meaning of an event—model failure in the future along some trajectories—that by definition they have never experienced. This reasoning, which seems altogether implausible, goes unquestioned.

In responding to such criticisms, analysts often appeal explicitly or implicitly to Milton Friedman's [15] positivist defense of optimization, which now thoroughly pervades the reasoning of trained economists. Along these lines it is replied that no one is asserting that agents can actually calculate future states of the economy as the argument presumes, but just that they act as if they could. Now, there are obviously some contexts in which Friedman's "as if" argument is acceptable: in Friedman's example the skilled billiards player acts as if he knows the laws of mechanics, so there is no harm in assuming that he actually does know them. In the billiards context there is no problem: it is hard to imagine a successful billiards player who has not figured out that the angle of incidence equals the angle of reflection. The question is whether applying similar arguments in the context of bubbles to rule out trajectories that are not feasible over infinite time is equally plausible.

The same issue has come up in capital theory. In the 1960s economists studied efficient paths of capital accumulation in the presence of many capital goods, focusing on whether these paths can be implemented as competitive equilibria (the best introduction to this literature is still Edwin Burmeister and A. Rodney Dobell [9]). Under conditions, Pareto-optimal equilibria display a turnpike property: from any initial condition the economy approaches that steady-state path on which the economy grows most quickly. In finite-time settings with a distant horizon the economy spends most of its time in the vicinity of the turnpike, and then exits the turnpike in order to satisfy an assumed terminal condition. In infinite time the economy converges toward the steady-state growth path. It follows that only one set of initial capital goods prices generates a path—that which converges to the turnpike—that is feasible in infinite time. The other paths all fail in finite time.

Karl Shell and Joseph E. Stiglitz [55], writing before it was customary to appeal routinely to rational expectations, saw no reason to conclude from this that the nonconvergent paths were infeasible. Given the absence of the futures markets that would have revealed the impending crisis, they assumed that agents had no way to foresee future problems. In the 1970s, however, the interpretation changed: economists working in the rational expectations tradition came to rely on the assumed infinite-time feasibility of equilibrium paths to eliminate indeterminacy in dynamic economic models.[15]

The apparent mispricing of Royal Dutch/Shell shares constitutes evidence in favor of bubbles.

> Example 3: Two firms, Royal Dutch and Shell, divide a common pool of earnings in preassigned ratios. Despite this, the stocks of these firms typically trade at prices that diverge from the indicated ratio, sometimes by a wide margin (Kenneth J. Froot and Emil Dabora [16]). Since the firms have fundamental values that are proportional to the ratio in which they divide earnings, the mispricing implies that at least one of the firms has a bubble. Exploiting the bubble by buying the cheaper stock and selling short the more expensive one will produce profits if the mispricing disappears. However, the price gap might instead widen, producing losses, and the existence of trading restrictions might force arbitrageurs to close out their positions at a loss. Thus the transaction is not an arbitrage, and the mispricing is not eliminated.[16] (Further, it follows that applying the

[15] These issues were discussed in Stiglitz's [61] introduction to the 1990 *Journal of Economic Perspectives* symposium on bubbles, a paper that, although not widely cited, repays careful study.

[16] As a digression, it is worth noting that the existence of limits to arbitrage has come to play a central role in supporting assertions that financial markets are inefficient (Shleifer [59], Nicholas Barberis and Richard Thaler [5]). This is a complete inversion of logic. The existence of limits to arbitrage implies just the opposite: security prices that seem inconsistent with market efficiency because of the apparent existence of arbitrage are in fact consistent with it because the indicated portfolio strategy may have to be closed out at a loss, and therefore is not a true arbitrage.

term “mispricing” in this context is a misnomer).

Similarly, the existence of discounts and premiums on closed-end mutual funds may indicate the presence of bubbles. In contrast to this, Stephen A. Ross [51] showed that the average discount approximately equals the discount warranted by management fees. However, the finding that returns on closed-end funds are considerably more volatile than the returns on the securities the funds hold (Jeffrey Pontiff [48]) is difficult to ascribe to management fees.

8 Conclusion

We have considered four categories of accounts for the price runup and collapse of the last ten years: (1) stationary fundamentals, (2) nonstationary fundamentals, (3) irrational bubbles and (4) rational bubbles. As explanations, all four categories have problems. Fundamentals-based explanations that presume a stationary setting, like the stochastic Gordon model, are clearly inadequate on empirical grounds: the data plots presented above show that corporate earnings and other explanatory variables for asset prices simply did not vary enough to justify stock price changes of the magnitude that occurred in the late 1990s.

A variety of ad hoc, but fundamentals-based, accounts for the stock price runup and collapse were considered above. To the extent that these proposed explanations point to unique aspects of the phenomena being studied, they are difficult either to corroborate or contradict. We concluded that the accounts are unpersuasive, but this may be a matter of taste.

Characterizing asset price fluctuations as reflecting irrationality, we saw, is not parallel to explaining these fluctuations in terms of either fundmentals or rational bubbles: insofar as irrationality is identified with partial equilibrium analysis, appealing to irrationality amounts to agreeing not to try to explain these fluctuations rather than proposing any particular explanation. Rational bubbles, as a particular class of models, are not subject to this problem. However, we saw that they are subject to another equally serious problem: committing to the full neoclassical paradigm produces an argument against bubbles that, although logically airtight, is simply not plausible. It is a testament to economists' capacity for abstraction that they have accepted without question that an intricate theoretical argument against bubbles has somehow migrated from the pages of *Econometrica* to the floor of the New York Stock Exchange.

Within the neoclassical paradigm there is no obvious way to derail the chain of reasoning that excludes bubbles. An alternative to the full neoclassical paradigm is to think about bubbles in a rational-agent setting—in particular, to define fundamental values using the present-value relation—but to break off the analysis arbitrarily at some point rather than following the reasoning to implausible conclusions. The problems with this alternative are obvious: how does one write down formal models in such a setting? How does one know where to break off the analysis? Which conclusions from neoclassical analysis are to be accepted?[17] We have no answers to these questions.

A defining characteristic of rational bubbles models is that agents inhabiting these models are represented as being aware that the securities they hold are trading at higher prices than are justified by their prospective dividends. They are willing to hold these securities because they rationally believe that the disparity will increase over time, providing returns that are adequate in view of the risk. Thus even though prices exceed fundamental values, there do not exist unexploited profitable trading rules. In contrast, rational-agent models that exclude bubbles represent agents as believing that the high security prices are justified as the present value of prospective dividends. There is considerable anecdotal evidence (in Shiller [57], for example) in favor of the former description, favoring the rational bubbles hypothesis, but there is also evidence that points the other way.

We argued that attributing the stock price runup to fundamentals is implausible in the absence of a showing why apparently similar earlier trends in fundamentals did not produce similar dramatic effects on security prices. This argument, of course, applies equally to bubble explanations: why did a bubble develop in the 1990s and not earlier? We have no answer to this question; the intent in this paper is merely to argue that in seeking explanations for the price runup and collapse there is no justification for limiting attention to fundamentals-based accounts.

The discussion of this paper suggests several conclusions and indicates some directions for future research:

- Assertions that theoretical and empirical evidence argue strongly against the existence of bubbles must be rejected. It is true that there exist theoretical arguments against bubbles, but we have suggested that these are implausible. Similarly, it is true that the empirical evidence of price and return volatility linked above with bubbles might have some other cause. That, however, does not constitute a showing that empirical evidence rejects bubbles.

- Despite the foregoing argument, there is no prospect that bubbles—at least rational bubbles—can dispose of all of the major anomalies of asset pricing. The equity premium puzzle of Mehra and Prescott [45], for example, deals with returns, not prices. Since prices are integrated returns and bubbles have to do with the constant of integration, existence of bubbles does nothing (at least in any direct sense) to invalidate the Mehra-Prescott finding that the equity premium is excessive.

[17] As an example, one curious implication of analyzing bubbles in the simplest overlapping generations setting under rational expectations is that they are welfare-improving: equilibria with big bubbles Pareto-dominate equilibria with little bubbles, and these Pareto-dominate equilibria with no bubbles. Is this outcome to be taken seriously as a proposition about the real world?

- Demonstrating that models without bubbles do not generate as much price and return volatility as the data appear to exhibit does not imply that models with bubbles can do so. The link between bubbles and volatility may involve sunspots (models in which equilibrium values of endogenous variables depend on extrinsic uncertainty). Sunspots, although conceptually different from bubbles, can occur under similar circumstances (non-Pareto-optimal equilibria; see, for example, Shell [54]). The link between bubbles and sunspots and the connection between these and price volatility have not yet been adequately investigated.

- If bubbles and sunspots are part of the explanation of price and return volatility, it follows that asset price fluctuations are not necessarily traceable to changes in expected cash flows and discount factors. Thus an important piece of empirical evidence that is often interpreted as conflicting with market efficiency—that stock price changes often cannot be plausibly connected to changes in fundamentals—may not necessarily support this interpretation.

- As a corollary to the preceding point, from the rational bubbles perspective the debate about capital market efficiency looks increasingly like a dispute about whether the glass is half full or half empty. This is partcularly clear in the recent *Journal of Economic Perspectives* exchange between Burton Malkiel [42] and Shiller [58]. Both describe a similar environment, one in which, in our terms, there exist rational bubbles and sunspots. Malkiel concluded that even though prices often depart significantly from fundamentals, markets are efficient because the implied trading rules are only marginally profitable, and then not for long. Shiller, on the other hand, associated market efficiency with the proposition that security price movements are traceable to corresponding changes in fundamentals, not with nonexistence of profitable trading rules. Given this definition, he concluded that markets are inefficient.

References

[1] Andrew B. Abel, N. Gregory Mankiw, Lawrence H. Summers, and Richard. J. Zeckhauser. Assessing dynamic efficiency: Theory and evidence. *Review of Economic Studies*, 56:1–20, 1989.

[2] Dilip Abreu and Markus K. Brunnermeier. Bubbles and crashes. *Econometrica*, 71:173–204, 2003.

[3] Lucy Ackert and Brian Smith. Stock price volatility, ordinary dividends, and other cash flow to shareholders. *Journal of Finance*, 48:1147 – ?, 1993.

[4] G. S. Bakshi and Z. Chen. Baby boom, population aging and capital markets. *Journal of Business*, 67:165–202, 1994.

[5] Nicholas Barberis and Richard Thaler. A survey of behavioral finance. Reproduced, University of Chicago, 2002.

[6] Fischer Black. Noise. *Journal of Finance*, 41:529–543, 1986.

[7] Olivier Blanchard and Stanley Fischer. *Lectures on Macroeconomics*. MIT Press, Cambridge, MA, 1989.

[8] Olivier Blanchard and Mark Watson. Bubbles, rational expectations, and financial markets. In Paul Wachtel, editor, *Crises in the Economic and Financial Structure*. Lexington Books, 1982.

[9] Edwin Burmeister and A. Rodney Dobell. *Mathematical Theories of Economic Growth*. Ashgate, 1993.

[10] Colin Camerer. Bubbles and fads in asset prices: A review of theory and evidence. *Journal of Economic Surveys*, 3:3–41, 1989.

[11] Edward Chancellor. *Devil Take the Hindmost: A History of Financial Speculation*. Penguin, New York, 1999.

[12] J. Bradford DeLong, Andrei Shleifer, Lawrence H. Summers, and Robert J. Waldmann. Noise trader risk in financial markets. *Journal of Political Economy*, 98:703–738, 1990.

[13] J. Bradford DeLong, Andrei Shleifer, Lawrence H. Summers, and Robert J. Waldmann. The survival of noise traders in financial markets. *Journal of Business*, 64:1–19, 1991.

[14] Eugene F. Fama. Efficient capital markets: A review of theory and empirical work. *Journal of Finance*, 25:283–417, 1970.

[15] Milton Friedman. The methodology of positive economics. In Milton Friedman, editor, *Essays in Positive Economics*. University of Chicago Press, 1953.

[16] Kenneth A. Froot and Emil Dabora. How are stock prices affected by the location of trade? *Journal of Financial Economics*, 53:182–216, 1999.

[17] John Kenneth Galbraith. *The Great Crash of 1929*. Houghton Mifflin, New York, 1979.

[18] Peter M. Garber. *Famous First Bubbles*. MIT Press, Cambridge, 2000.

[19] John Geanakoplos, Michael Magill, and Martine Quinzii. Demography and the long-run predictability of the stock market. reproduced, Yale University, 2002.

[20] Christian Gilles and Stephen F. LeRoy. Bubbles and charges. *International Economic Review*, 33:323–339, 1992.

[21] Christian Gilles and Stephen F. LeRoy. Rational bubbles as payoffs at infinity. *Economic Theory*, 9:261–281, 1997.

[22] Myron J. Gordon. Variability of price-earnings ratios. *American Economic Review*, 52:203–208, 1962.

[23] Jeremy Greenwood and Boyan Jovanovic. The information-technology revolution and the stock market. *American Economic Review Papers and Proceedings*, 1999:116–122, 1999.

[24] John M. Griffin, Jeffrey H. Harris, and Selim Topaloglu. Investor behavior over the rise and fall of Nasdaq. Reproduced, Yale University, 2003.

[25] Robert E. Hall. E-capital: The link between the labor market and the stock market in the 1990s. *Brookings Papers on Economic Activity*, pages 73–118, 2000:2.

[26] Robert E. Hall. The stock market and capital accumulation. *American Economic Review*, 91:1185–1202, 2001.

[27] Robert E. Hall. Struggling to understand the stock market. *American Economic Review Papers and Proceedings*, 91:1–11, 2001.

[28] Robert E. Hall. Dynamics of corporate earnings. Reproduced, Stanford University, 2002.

[29] Robert E. Hall and Dale W. Jorgenson. Tax policy and investment behavior. *Tax Policy and Investment Behavior*, 57:391–414, 1967.

[30] Bart Hobijn and Boyan Jovanovic. The information-technology revolution and the stock market: Evidence. *American Economic Review*, 91:1203–1230, 2001.

[31] John Maynard Keynes. *General Theory of Employment, Interest and Money.* Harcourt, Brace, New York, 1936.

[32] Charles Kindleberger. *Manias, Panics and Crashes: A History of Financial Crises.* Basic Books, New York, 1978.

[33] Albert S. Kyle. Continuous auctions and insider trading. *Econometrica*, 53:1315–1335, 1985.

[34] John Laitner and Dmitriy Stolyarov. Technical change and the stockmarket. *American Economic Review*, 2003.

[35] Stephen F. LeRoy. Risk aversion and the martingale model of stock prices. *International Economic Review*, 14:436–446, 1973.

[36] Stephen F. LeRoy. Efficient capital markets: Comment. *Journal of Finance*, 31:139–141, 1976.

[37] Stephen F. LeRoy. Review of Garber, Famous first bubbles. *Journal of Economic Literature*, XL:959–960, 2002.

[38] Stephen F. LeRoy and William R. Parke. Stock price volatility: Tests based on the geometric random walk. *American Economic Review*, 82:981–992, 1992.

[39] Robert E. Lucas. Econometric policy evaluation: A critique. In Karl Brunner and Allan H. Meltzer, editors, *Carnegie-Rochester Conference Series on Public Policy*. North-Holland, 1976.

[40] Robert E. Lucas. Asset prices in an exchange economy. *Econometrica*, 46:1429–1445, 1978.

[41] Charles Mackay. *Extraordinary Popular Delusions and the Madness of Crowds*. Farrar, Strauss and Giroux, New York, 1932.

[42] Burton Malkiel. The efficient market hypothesis and its critics. *Journal of Economic Perspectives*, 17:59–82, 2003.

[43] Ellen R. McGrattan and Edward C. Prescott. Is the stock market overvalued? *Federal Reserve Bank of Minneapolis Quarterly Review*, 24:20–40, 2000.

[44] Ellen R. McGrattan and Edward C. Prescott. Average debt and equity returns: Puzzling? Reproduced, Federal Reserve Bank of Minneapolis, 2003.

[45] Rajnish Mehra and Edward C. Prescott. The equity premium: A puzzle. *Journal of Monetary Economics*, 15:145–161, 1985.

[46] Merton H. Miller and Franco Modigliani. Dividend policy, growth, and the valuation of shares. *Journal of Business*, 34:411–433, 1961.

[47] Maurice Obstfeld and Kenneth Rogoff. Speculative hyperinflations in maximizing models: Can we rule them out? *Journal of Political Economy*, 91:675–687, 1983.

[48] Jeffrey Pontiff. Excess volatility and closed-end funds. *American Economic Review*, 87:155–169, 1997.

[49] A. C. Rayner and I. M. D. Little. *Higgledy Piggledy Growth Again*. Oxford University Press, Basic Blackwell, 1966.

[50] W. B. Reddaway. Higgledy piggledy growth. *Economic Journal*, 77:595–598, 1967.

[51] Stephen A. Ross. A neoclassical look at behavioral finance: Closed end funds. Reproduced, MIT, 2002.

[52] Manuel Santos and Michael Woodford. Rational asset pricing bubbles. *Econometrica*, 65:19–57, 1997.

[53] Jose Scheinkman and Wei Xiong. Overconfidence and speculative bubbles. Reproduced, Princeton University, 2003.

[54] Karl Shell. Sunspot equilibrium. In J. Eatwell, M. Milgate, and P. Newman, editors, *The New Palgrave: A Dictionary of Economics*. McMillan, 1987.

[55] Karl Shell and Joseph E. Stiglitz. The allocation of investment in a dynamic economy. *Quarterly Journal of Economics*, 81:592–609, 1967.

[56] Robert J. Shiller. Do stock prices move too much to be justified by subsequent changes in dividends? *American Economic Review*, 71:421–436, 1981.

[57] Robert J. Shiller. *Irrational Exuberance*. Princeton University Press, Princeton, NJ, 2000.

[58] Robert J. Shiller. From efficient market theory to behavioral finance. *Journal of Economic Perspectives*, 17:83–104, 2003.

[59] Andrei Shleifer. *Inefficient Markets—An Introduction to Behavioral Finance*. Oxford University Press, Oxford and New York, 2000.

[60] George J. Stigler and Gary S. Becker. De gustibus non est disputandum. *American Economic Review*, 67:76–90, 1977.

[61] Joseph E. Stiglitz. Symposium on bubbles. *Journal of Economic Perspectives*, 4:13–17, 1990.

[62] Jean Tirole. Asset bubbles and overlapping generations. *Econometrica*, 53:1499–1528, 1985.

[63] Neil Wallace. The overlapping generations model of fiat money. In John Kareken and Neil Wallace, editors, *Models of Monetary Economies*. Federal Reserve Bank of Minneapolis, 1980.