

Rational Exuberance*

Stephen F. LeRoy
University of California, Santa Barbara

March 31, 2003

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, can be identified with 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 now has a value of \$1000. The increase in value can be ascribed to increased value of the services the stamp provides only if collectors derive pleasure from looking at the printing error on the stamp. Barring such a farfetched argument, one must conclude that the fundamental value of the stamp remains 37 cents. The remainder of its value constitutes a bubble (defining here a bubble as the price of an asset minus the present value of its future payoff).

This argument establishes beyond reasonable doubt that bubbles exist. As regards collectibles generally, it is easiest to argue that their values reflect bubbles rather than fundamental values when the objects in question clearly have no aesthetic value whatsoever, as with misprinted postage stamps. With fine art the argument is less straightforward: someone who prefers Monet to Andy Warhol might argue that the values of Monet's paintings reflect their aesthetic value, whereas those of Andy Warhol's paintings reflect bubbles. Even here, however, the fact that a Monet 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 Monet original is a bubble, just as with the misprinted postage stamp. Its value derives primarily from its scarcity, and only secondarily from its aesthetic value.

*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 [65] of the same name. I have received comments from Christian Gilles and Douglas Steigerwald.

The fact that bubbles exist on collectibles does not imply that they exist on assets like stocks or land, and still less that bubbles on such assets are quantitatively important if they do exist. 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?

In this paper we will look at data on US stock prices and related series with a view to determining whether they can plausibly be explained without resorting to bubbles (Sections 2 and 3). Of course, our attention will be drawn to the price runup and collapse over the most recent decade. We will argue that this behavior cannot successfully be so explained. As Shiller [64] and others have argued, stock price volatility exceeds that which is consistent with the present-value relation. We interpret this as favoring bubbles.

It is true that there exist possible explanations for observed stock price volatility other than bubbles. Of these the most attractive on theoretical grounds is stochastic discount factors (Section 4); of course, this is not to deny that bubbles and stochastic discount factors can coexist. Risk aversion induces a stochastic component in discount factors, thereby generally invalidating the constant discount factor version of the present-value model used in the variance bounds tests. Specifically, risk aversion induces asset price volatility greater than that which would occur if agents were risk neutral, at least in some settings. However, the theory of consumption-based asset pricing, while generating a theoretically coherent account of how risk aversion affects asset prices, turns out to be a failure empirically.

In Section 5 the discussion is specialized to the events of the 1990s. We briefly consider and reject alternatives to the bubbles hypothesis.

The remainder of the paper deals with bubbles. Bubbles are taken to be synonymous with irrationality, universally in the popular press, but also sometimes in academic discussions. Especially given the recent ascendancy of behavioral finance, it is appropriate to digress (Section 6) to consider what, if anything, it means to appeal to irrationality in a substantive economic explanation of economic phenomena. It will be suggested that this is a dead end.

In Sections 7-10 the discussion is restricted to rational bubbles—instances of asset prices exceeding present values in models in which agents optimize in an environment which they understand. Established professional opinion holds that there exist robust theoretical arguments against existence of such bubbles. Such arguments are in fact anything but robust. We note that bubbles cannot generally be ruled out on purely theoretical grounds. In an important class of models—overlapping generations models—bubbles can be ruled out under certain values for key parameters, but the evidence is far from clear that these parameter restrictions are satisfied. Even if they are satisfied, the argument for nonexistence of bubbles in such settings requires an implausibly literal-minded application of a rational expectations argument.

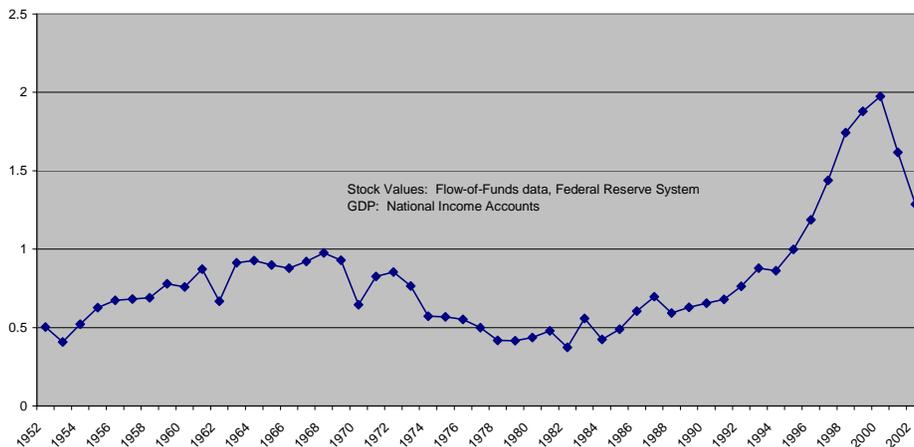


Figure 1: Equity Value/GDP

2 The Data

Figure 1 shows the value of equity (from the Federal Reserve Board Flow of Funds accounts) divided by GDP. Data are from the 1950s to now. The salient feature of this series is its dominance by low-frequency components. Stock prices rose in the 50s and 60s, fell in the 70s, rose in the 80s and 90s and, so far in the 21st century, have been falling. 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 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. One might argue that for most of the sample period the low-frequency variation in Figure 2 is less pronounced than that in Figure 1, in which case the interpretation is that stock price variations are correlated with variations in earnings relative to GDP, with the stock price variation being of greater amplitude.

The runup of stock prices in the late 1990s appears similar in the two series, suggesting that it cannot be viewed as 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 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 con-

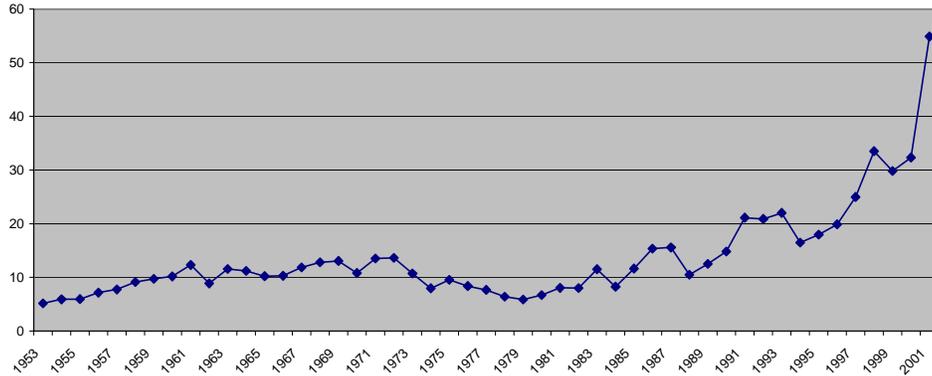


Figure 2: Equity Value/Corporate Earnings

spicuous 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.

Our first question is whether these variations in stock prices are within the range that is appropriate under the present-value model, given the variation of corporate earnings and dividends. In the next section we use a variance-bounds test to show that they are not.

3 The Stochastic Gordon Model

The point of reference in interpreting stock price data will initially be the Gordon [31] model. 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 return on invested capital 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 equal to the rate of return on invested capital. Under these assumptions it is easy to show that the present-value model implies that p , the value of equity, is given by

$$p = \frac{d}{r - g} = \frac{e}{r} \tag{1}$$

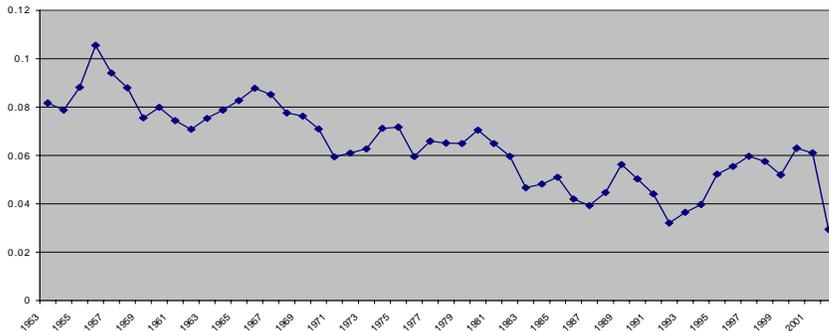


Figure 3: Corporate Earnings/GDP

Here r is the discount rate, d is the current dividend, e is current earnings and $g = r(1 - \delta)$ is the growth rate of the dividend stream, where δ is the dividend payout rate (proportion of earnings paid out as dividends). 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 (Miller and Modigliani [53]). The independence of p from δ reflects the Miller-Modigliani proposition that (in this simple environment) for given 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 return on capital 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.¹ We refer to the model that results from this generalization as the *stochastic Gordon model*.

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 here the assumption that they are 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.

Figure 4 shows the interest rate less the growth rate of earnings; the series looks like an IID process.² Figure 5 shows the dividends-earnings ratio. Obviously the

¹The geometric random walk was found more than 40 years ago to give an accurate description of corporate earnings (Little and Rayner [57], Reddaway [58]).

²The 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 exactly offsets that involved

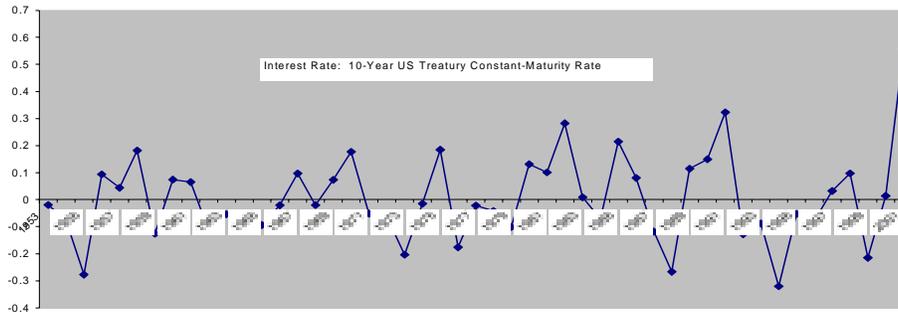


Figure 4:

assumption adopted in the stochastic Gordon model that the dividends-earnings ratio is constant is an oversimplification; dividends are well modeled as a highly smoothed function of past earnings, implying that dividends-earnings ratios will vary as earnings fluctuates. Nevertheless, the series appears to be strongly mean-reverting, so the failure of dividends to adjust fully to earnings changes should not greatly distort security prices relative to the prediction of the Gordon model.

In order to determine the behavior of stock prices implied by the stochastic Gordon model we need in addition to specify investors' information. Two extreme cases can be specified: (1) investors have complete information about future earnings, and (2) investors have no information beyond the current level of earnings that enables them to forecast future earnings. Each of these extreme cases generates a volatility bound: the first on the volatility of the price-earnings ratio, the second on the volatility of the rate of return.

The more information about the future investors have, the more volatile will be the price-earnings ratio. This will be demonstrated below, but the idea is intuitive: if investors have no future information about earnings, they will price stock at a constant multiple of current earnings, so the price-earnings ratio will have zero variance. To the extent that investors have information about future earnings realizations, they will price stocks at varying multiples of earnings. An upper bound on the variance of the price-dividend ratio will be generated in the perfect certainty case. LeRoy and Parke [46] showed³ that this variance is given by the right-hand side of the following inequality, so the variance bound is

$$V(p/e) \leq \frac{\beta^2 \delta^2 \sigma^2}{(1 - \beta^2(\mu^2 + \sigma^2))(1 - \beta\mu)^2}. \quad (2)$$

in calculating the real interest rate.

³This and several other expressions in the text differ in detail from those in LeRoy-Parke because those authors were working with price-dividend ratios rather than price-earnings ratios.

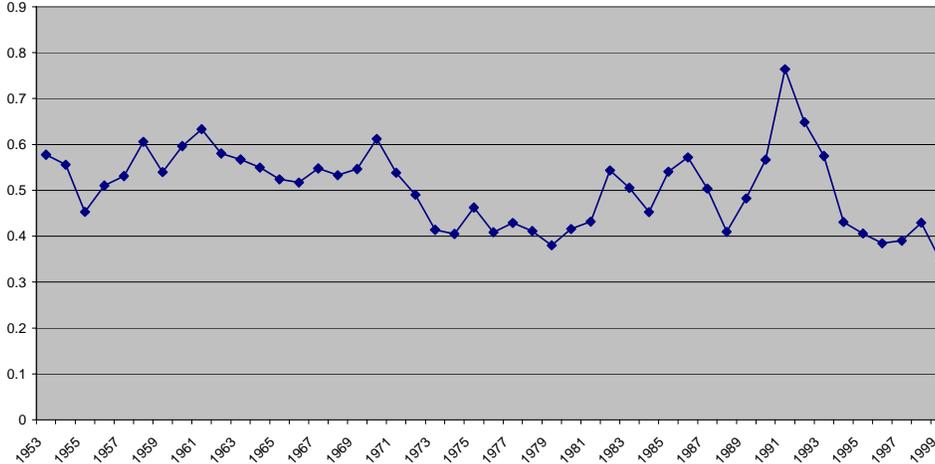


Figure 5: Dividends/Earnings

Here β is the discount rate, and μ and σ^2 are the mean and variance, respectively, of the earnings growth rate. Inequality (2) is essentially the same as the variance bound of LeRoy-Porter [47], adapted to the geometric random walk case. Inserting estimated parameters and converting to standard deviations, the upper bound on price volatility is 7.274, whereas the standard deviation of the actual price-earnings ratio is 8.945. Thus the point estimates indicate excess volatility.

Just as increasing information increases the volatility of the price-earnings ratio, it decreases the volatility of the rate of return: in the extreme case in which investors know the entire future path of earnings, the volatility of the rate of return will be zero. It follows that the case in which investors have no information beyond current earnings will generate an upper bound on the variance of the rate of return. The variance of the rate of return under the geometric random walk if investors have no information about future earnings growth rates is given by the right-hand side of the following inequality, so the analogue to (2) is

$$V(r) \leq \frac{\sigma^2}{\beta^2 \mu^2}. \quad (3)$$

Inequality (3) is the West [73] test, adapted to the geometric random walk case, and it is essentially the same as LeRoy-Porter's [47] lower bound on price volatility. Based on estimated parameter values and again converting to standard deviations, the upper bound is 0.179, whereas the standard deviation of actual returns is 0.174. Thus return volatility is almost exactly what it would be under the stochastic Gordon model if investors had no information about future earnings growth rates.

The foregoing discussion suggests that there is a tradeoff between price volatility and return volatility: high price volatility is associated with low return volatility and

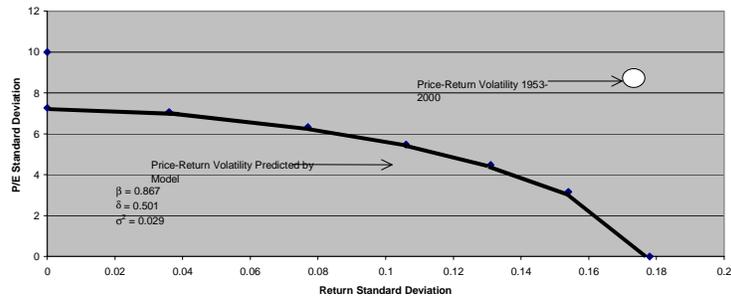


Figure 6:

vice-versa. In the setting of the stochastic Gordon model there exists a closed-form expression for the locus of pairs of price variance and return variance as a function of parameters:

$$V(p^*/e) = V(p/e) + \frac{\beta^2[V(p/e) + (\beta\delta\mu/(1 - \beta\mu))^2]}{(1 - \beta^2(\mu^2 + \sigma^2))}V(r) \quad (4)$$

Here $V(p^*/e)$ equals the upper bound of the variance of the price-earnings ratio (the right-hand side of (2)). Inserting values for $V(p/e)$ in (4) and solving for $V(r)$ allows us to plot the locus of pairs of price volatility and return volatility that are consistent with the Gordon model. Testing that $V(p/e)$ and $V(r)$ satisfy (4), as a test of an equality rather than an inequality, has better econometric properties than either of the bounds tests described above. Figure 6 shows the locus of standard deviations of the price-earnings ratios and rates of return that are consistent with estimated parameter values (standard deviations are easier to interpret than variances). The diagram displays a dot representing the estimated standard deviations from US stock data. As the diagram makes clear, the volatility measures are higher than the stochastic Gordon model implies.

The variance bounds tests have two major shortcomings: (1) they assume a constant discount factor, and (2) they exclude bubbles. In the following section we will consider the generalization to stochastic discount factors under the rubric of consumption-based asset pricing. We will follow many others in concluding that, however great the theoretical appeal of consumption-based asset pricing, it is a failure empirically. This leaves bubbles, to which the remainder of the paper is devoted.

4 Consumption-Based Asset Pricing

The variance bounds test reported in Section 3 gave clear evidence that stock prices and returns are more volatile than the stochastic Gordon model predicts, given the estimated volatility of earnings. However, the stochastic Gordon model presumes a constant discount factor. This assumption, while attractive as a simplification, is very restrictive theoretically.

In stationary exchange economies with risk-neutral agents, asset prices will equal expected payoffs discounted at the risk-free interest rate. However, when agents are risk averse, expected returns will be different for different securities depending on their risk. The theory of consumption-based asset pricing, introduced in the economics literature by Lucas [50], and in the finance literature by Breeden [10], shows how asset valuation is altered by risk aversion.

Expositions of consumption-based asset pricing can be found in a variety of sources (for example Bossaerts [9], Cochrane [15], LeRoy and Werner [48], or Campbell, Lo and MacKinlay [12]; the latter gives a very clear discussion of the empirical tests of consumption-based asset pricing).

If agents' utilities are strictly concave functions of their consumption levels, equilibrium expected returns will depend on how asset payoffs covary with aggregate consumption. Risk-free returns and risk premia, both of which are assumed constant in the stochastic Gordon model, will be stochastic, and the latter will be different for different assets. Thus there is no reason to regard the variance bounds tests as necessarily applying in a world of risk-averse agents.

In fact, there are specific reasons to expect that stock price volatility will exceed that implied by risk neutrality when agents are risk averse. When agents are risk averse they will be motivated to use portfolio transactions to smooth consumption, transferring it from high-consumption states and dates to low-consumption states and dates. In a representative agent exchange setting, they cannot do so in the aggregate. That being the case, asset prices must be such as to induce the representative agent to consume his endowment. Specifically, equilibrium asset prices will covary positively with consumption, and the more risk averse agents are, the higher the covariance must be. Thus risk aversion creates a presumption in favor of increased price volatility relative to risk neutrality. This argument was formalized by LaCivita and LeRoy [41]; a related argument is found in Grossman and Shiller [32].

As is well known, the problem with consumption-based asset pricing theory is that it appears to be a failure empirically. The implication that expected asset returns depend on covariances with consumption growth (or, in the market CAPM version, with the return on the market portfolio) appears not to be satisfied in the data. Major papers demonstrating the failure of market CAPM and consumption CAPM are Hansen and Singleton [37], Mehra and Prescott [52], Hansen and Jagannathan [38] and Fama and French [20]. Given the failure of consumption-based asset pricing theory to provide an empirically successful account of changes in expected returns,

we will proceed henceforth as if expected returns are constant.⁴

Many analysts (Cochrane [14], for example) conclude from the failure of consumption-based asset pricing theory that the quest for a successful implementation of the stochastic discount factor must continue, since we know that investors are risk averse. A variety of ever-more elaborate specifications are being tried, ranging from non-expected utility to transactions costs and various forms of market incompleteness. So far none of these modifications appears to have succeeded, but the jury is still out. For the present, we can conclude that the fact that expected asset returns do not covary with aggregate consumption means that we lose little in terms of explanatory power by suppressing the dependence of discount factors on consumption. This amounts to accepting the assumption of the stochastic Gordon model that discount factors are constant.

5 The 1990s Price Runup

One possible explanation for the price runup of the 1990s is that the advent of the internet led to a creation of new capital on the order of several trillion dollars. How this capital came into existence is unclear; it cannot have resulted from retained earnings as measured in the national income accounts because these, although somewhat higher than usual in the late 90s, were an order of magnitude lower than the stock price runup. Further, the very modest increase in aggregate corporate earnings suggests either that a large component of the return on this new capital went unmeasured or that the return per unit of capital (earnings-price ratio) suffered a sharp drop. Finally, the stock price declines of the last several years imply that this new capital disappeared even more rapidly than it had materialized. See Hall [33], [34], [35], [36] for further discussion along these lines. In the absence of any evidence in favor of this hypothesis other than the fact that a huge change in stock prices occurred unaccompanied by anything like a commensurate change in measured corporate earnings, it is difficult to evaluate it.

Another possible line would generalize the Gordon model to relax the assumption that the return per unit of capital is an IID process (Hall [35]). If the term $r - g$ in the denominator of (1) is time-varying, variations in the price-earnings ratio would result. Figure 4 displays the long-term nominal interest rate on 10-year US Treasury debt less the growth rate of corporate earnings over the previous year. The combination

⁴The alternative is to specify directly a time series model for returns that involves intertemporal dependence, but not attempt to relate the innovations in expected returns to aggregate consumption. In this setting the restriction to constant returns is clearly rejected empirically at customary significance levels (see, for example, Campbell, Lo and MacKinlay [12]). The problem with this specification is that it involves assuming not only that (conditional) expected returns are time-varying, but that investors make allowance for this dependence in pricing securities. This may not be plausible. Similar reservations about indiscriminately invoking rational expectations are recorded in Section 9 below.

of declining nominal interest rates and increasing corporate earnings imply that $r - g$ was indeed low in the years 1995-2000.

Whether the deviation from trend during these years is sufficient quantitatively to explain the stock price runup is difficult to determine without specifying how investors form expectations of earnings, and deriving an analogue to the Gordon model that is appropriate under that specification. Mehra and Prescott [52] developed a two-state generalization of the Gordon model assuming that investors have rational expectations of future earnings growth rates. That their model leads to puzzles when confronted with the data is well known. Direct implementation of a calibrated version of their model does not generate a stock price runup of anything like the requisite magnitude.⁵

6 Rationality and Irrationality

In Section 3 we identified two possible explanations for the excess volatility of stock prices and returns: stochastic discount factors and bubbles. Having rejected stochastic discount factors on empirical grounds, we are left with bubbles. The remainder of this paper will make the case that rational bubbles (defined below) are the most attractive candidate to explain stock prices. In this section we discuss the meaning of rationality in the context of bubbles.

What are bubbles? The term “bubble” is much used in the media, almost never with a clear meaning. The question “Do you think internet stocks (real estate, Japanese stocks) are or were a bubble?” is badly in need of clarification. Most simply, it appears that the questioner wants to know if prices will collapse any time soon. Along these lines, a bubble means a price increase that is shortly followed by a price decrease. Alternatively, the questioner wants to know if there is a rational basis for current valuations, so that asset prices being a bubble is synonymous with valuations having no rational basis. This usage is virtually universal in the financial media, and is adopted also in some academic treatments of bubbles.⁶

⁵This might be due to the two-state nature of Mehra-Prescott’s model, which renders it unsuitable for analysis of a specific episode.

See McGrattan and Prescott [51] for an argument that the stock price runup of the 1990s was not excessive in view of market fundamentals.

⁶For example, Garber [25] wrote: “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.”

Similarly, Hall [35] 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”, leaving no room here for the possibility that values could exceed discounted future returns even when agents are fully rational. 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.

Given the large role assigned to irrationality, somehow defined, in the behavioral finance literature that has been in the ascendancy in recent financial economics, it is worth pointing out that there exists an independent line of research on bubbles that identifies these, not with irrationality of any sort, but with security prices that exceed the (expected risk-adjusted) present values of the associated payoffs. This was the definition of bubbles that was presumed in the postage stamp example discussed in the introduction above. Bubbles so defined can occur in equilibrium under conditions that are well understood, at least in theoretical terms, as we will see below. Bubbles can be ruled out on theoretical grounds only under assumptions that, we will argue, are altogether implausible. On grounds of parsimony, therefore, no appeal to irrationality is needed to sustain the view that asset prices can systematically exceed present values.

Before presenting this discussion it is appropriate to consider irrational bubbles further. First, we note the ambiguity of the terms “rational” and “irrational” as used in the behavioral finance literature, and also the divergence between the meanings of these terms in behavioral finance and their meanings in economics. Given the importance of the idea of rationality in economic theory, together with the central role assigned to irrationality in behavioral finance, it is remarkable how little serious discussion these ideas have received.

The view of rationality implicit in mainstream nonfinancial economic theory is that rationality is not a substantive hypothesis about the world, but rather is a conceptual tool. In fact, the idea of rationality provides the basis for a good working definition of an economic model: a model describing human (or, for that matter, animal) behavior is an economic model insofar as it assumes consistent choice. Along these lines there is no room for irrational behavior: apparently irrational behavior 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 it a tautologous character: as part of a maintained hypothesis, no conceivable evidence can contradict rationality.

A generation ago financial economists fully accepted this view of rationality, if the term “market efficiency” is substituted for “rationality”: in his classic [19] survey of efficient capital markets Eugene Fama emphasized that market efficiency can be tested only in conjunction with a particular market model.⁷

Contemporary exponents of behavioral finance are much 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

⁷To 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. Also, it should be noted that Fama’s formulation of market efficiency was tautologous in a different sense from that discussed above, as I pointed out in LeRoy [43].

markets than hitherto believed, exponents of behavioral finance have altered the definition of rationality. Irrationality no longer connotes the theoretical property of nontransitivity, as would seem to follow from the definition of rationality in economic theory. Rather, irrationality is now identified with the existence of noise traders, agents who trade for reasons that are not modeled.

Models in which some of the economic behavior being modeled is taken as given have been known in the economic literature as partial equilibrium models, as distinguished from general equilibrium models in which all agents are represented as optimizing subject to constraint. There is no identification of partial equilibrium with irrationality. It is difficult to see the advantage in broadening—and blurring—the definition of irrationality as the advocates of behavioral finance have done, but there is little doubt that this has happened.

In the earlier finance literature the existence of noise traders was seen as consistent with market efficiency as long as the noise traders coexisted with rational agents. The major propositions of neoclassical finance theory, such as the risk-return tradeoff, will survive the presence of noise traders as long as the rational agents are sufficient in numbers (and wealth) so that their optima are interior.

It is true that in some of the earlier discussions of market efficiency it was presumed 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, since then all securities are perfect substitutes and rational traders will bid all assets to prices that imply equal expected returns. If agents are risk averse, the rational agents will generally not completely eliminate the effect of noise traders on security prices. This is so because risk-averse traders will rationally reject trades that exploit minor mispricing because of their risk. This result provides the basis for much of the recent enthusiasm for behavioral finance. This is so despite the fact that it can be derived using completely neoclassical (but partial equilibrium, given the assumed presence of noise traders) methods (Shleifer [66]), and despite the fact noted in the preceding paragraph that most properties of mainstream finance theory obtain even in the presence of noise traders.

In the preceding paragraphs we have argued on methodological grounds against attributing bubbles to noise traders. The same conclusion can be reached using a more pragmatic line of reasoning. The presumed irrationality of the noise traders motivates rational traders to bet against them. To be sure, as Shleifer emphasized, rational traders will not completely offset the effect of noise trades on prices, but they will always take positions against the noise traders. In the case of the 1990s stock price runup in the US, it would have been easy for any rational trader to bet against the bubble by, for example, selling NASDAQ futures short or buying NASDAQ puts. Few perceived doing so as an attractive trading opportunity, and a fair number of those who did make this bet posted spectacular losses because of bad timing. The unwillingness of apparently rational market participants to bet against the bubble indicates that the perceived return distribution was not attractive. This is equivalent

to asserting that those who took long positions in securities with bubbles were not irrational.

In models of noise traders the irrational agents have systematically worse outcomes than the rational agents in utility terms—that is what calling them irrational means. The traders are not aware of the suboptimal nature of their trades; if they were, they would trade rationally. In contrast to this account, during the late 1990s the financial press included innumerable discussions of whether current stock prices were a bubble, so it is hard to believe that traders did not at least consider that possibility. Significantly, many traders who took long positions in dot-com stocks and the like believed that stock prices exceeded fundamental values.⁸ They did not see this as implying that they were trading incorrectly, for the simple reason that they believed that the gap between price and fundamental value was at least as likely to increase as to decrease. Awareness of the existence of a bubble is inconsistent with the noise trader interpretation, but perfectly consistent with rational bubbles.

7 Rational Bubbles under Simultaneous Markets

The argument of the preceding section suggests that we are unlikely to make progress in analyzing bubbles if we put some economic behavior off limits by attributing it to noise traders. Relying on traditional general equilibrium analysis appears to be a more promising research strategy. Thus we restrict ourselves henceforth to the analysis of rational bubbles—meaning that agents optimize in environments that they understand (subject to a qualification discussed below), and in particular that they have rational expectations.

As noted above, a security has a bubble if its price exceeds the discounted value of its expected payoffs. Formally, suppose that agents are risk neutral (this involves no material loss of generality because, as is well known, risk aversion can be handled by distorting the probability measure). Then the expected returns on all securities are equal, implying that the equilibrium price p_t of a security satisfies

$$p_t = E_t \frac{d_{t+1} + p_{t+1}}{1 + r_{t+1}}, \quad (5)$$

where r_{t+1} is the interest rate from t to $t + 1$ and E_t denotes conditional expectation. Replacing t by $t + 1$ in (5), substituting in (5) and invoking the rule of iterated expectations, there results

$$p_t = E_t \left[\frac{d_{t+1}}{1 + r_{t+1}} + \frac{d_{t+2} + p_{t+2}}{(1 + r_{t+1})(1 + r_{t+2})} \right]. \quad (6)$$

⁸This is not to deny the existence of other traders who fit the noise trader description better. Many market participants subscribed to the “new economy” view that the old rules of security valuation no longer applied, so that the much higher stock values that prevailed during the late 1990s were seen as appropriate. It is less easy to describe these as rational traders.

Repeating this substitution n times yields

$$p_t = E_t \prod_{i=1}^n \frac{\bar{A}_{t+i}}{(1+r_{t+i})} \left(d_{t+i} + \prod_{j=1}^n \frac{\bar{A}_{t+j}}{(1+r_{t+j})} p_{t+n} \right) . \quad (7)$$

Taking the limit (which will exist under very weak assumptions) as n goes to infinity gives

$$p_t = f_t + b_t, \quad (8)$$

where

$$f_t = E_t \prod_{i=1}^{\infty} \frac{\bar{A}_{t+i}}{(1+r_{t+i})} d_{t+i} \quad \text{and} \quad b_t = \lim_{t \rightarrow \infty} E_t \prod_{i=1}^{\infty} \frac{\bar{A}_{t+i}}{(1+r_{t+i})} p_{t+n} . \quad (9)$$

Here f_t is the fundamental value of the security—the expected discounted value of its finite-date payoffs—and b_t is the bubble. If the bubble is zero, the price of the security equals the expected discounted value of its payoff. If, in contrast, the bubble term is positive, then securities are overpriced relative to their fundamentals.⁹ This does not imply existence of a trading opportunity because the overpricing is expected to continue in the future.

An obvious implication of this development is that (nonzero) bubbles can exist only in models that allow for an infinite future. It follows that nontrivial analysis of bubbles cannot be conducted using classical general equilibrium theory because of the finite-dimensional nature of the associated models. This complicates matters considerably: there exist more than one possible generalization of classical equilibrium theory to infinite-dimensional settings (even maintaining the assumption of competitiveness, as we do throughout). For the present purpose the important question is whether markets are characterized as simultaneous or sequential. In the complete-markets case these are equivalent when time is finite (Arrow [2]), but not in the infinite case that is relevant here. Further, the alternative specifications have different implications for bubbles. Analysis of bubbles in the simultaneous markets case is presented in this section, while sequential markets are considered in the following section.

In simultaneous-markets settings the analyst maintains the convention of finite-date general equilibrium theory that contingent claims are traded in timeless markets

⁹Negative bubbles can be ruled out on assets that can be freely disposed of, such as stocks. The reason is that in the presence of a negative bubble the present-value relation implies that the price of the security will eventually become negative with probability 1, which conflicts with free disposal. Nonnegativity of bubbles is frequently incorrectly attributed to the limited liability of corporations rather than to free disposal. Limited liability implies that the dividend of the asset is nonnegative, not that its price is nonnegative. It is possible to imagine a non-disposable asset with positive cash flow at every date but negative price, because of a negative bubble.

that, for concreteness, are best thought of as occurring before time begins (even when the models allow for an infinite past). This assumption is obviously abstract, but it is not clear that it is more objectionable in the infinite-time case than in the finite-time case.

In models incorporating simultaneous markets equilibrium, the price system is represented by a positive linear function on the commodity space. The price of an individual security is assumed to equal the value taken on by that security's payoff under the equilibrium price function, just as in the finite case in which it is routinely invoked.

In some settings the price function is representable by state prices. When that is the case the payoff of any security equals its fundamental value (the sum of its payoffs multiplied by the appropriate state prices), so that there are no bubbles. However, in infinite-dimensional spaces linear functions can take a form that does not allow representation by state prices. For example, suppose that there is no uncertainty and that time is countable. Let the price system $p(x)$ be

$$p(x) = \lim_{n \rightarrow \infty} \frac{\sum_{t=1}^n x_t / (1 + \rho)^t}{n}, \quad (10)$$

where ρ is a constant, for all x such that the limit exists. Note that this price system, although not of a conventional form, is concave and does incorporate discounting when $\rho > 0$. It has the property that any commodity bundle x that is nonzero at only a finite number of dates has zero value. The same is true of any commodity bundle that has a limiting discounted value of zero. Under price functions of the form (10), all state prices equal zero, implying that valuation cannot be identified with state prices. Since the fundamental value of any commodity bundle is identified with the limit of the sums of the values of its finite-date payoffs, the fact that these values are zero implies that all commodity bundles have zero fundamental value under the price system (10). Thus all values are bubbles.¹⁰

Modeling bubbles in this way was initially proposed by Gilles [27], and was developed in Gilles and LeRoy [28], [29]. Of course, the price system (10) is highly unrealistic because of its implication that all finitely nonzero commodity bundles have zero value. However, the price system

$$p(x) = \sum_{t=1}^{\infty} q_t x_t + \lim_{n \rightarrow \infty} \frac{\sum_{t=1}^n x_t / (1 + \rho)^t}{n}, \quad (11)$$

¹⁰For an example of an equilibrium in which (10) is the equilibrium price system, consider a deterministic representative-agent exchange model in which agent's utility function coincides with (10) and, for concreteness, whose endowment e is $e = \{e_t\} = (1 + \rho)^t$. Then the representative agent's endowment has value 1.

which combines the component (10) with a component incorporating conventional state prices, implies that the value of any commodity bundle may be decomposed unambiguously into a fundamental component and a bubble component, either of which may equal zero. This is not so obviously unrealistic.

Under this treatment bubbles are identified with the tails of commodity bundles. If the model-builder restricts preferences and/or admissible commodity bundles so that the rightmost term in (11) equals zero for all admissible commodity bundles, then there can exist no bubbles. In a general setting this involves restricting preferences and/or admissible commodity bundles so that preferences are Mackey-continuous (Bewley [5]) on the set of admissible commodity bundles. Thus ruling out bubbles amounts to restricting preferences and/or commodity bundles so that their tails make zero contribution to their present values.

In the foregoing development, whether or not a model economy has bubbles depends on the form of the price system. Under a different specification of the commodity space and price system bubbles are associated with payoffs rather than price systems (Gilles and LeRoy [30]). Under that respecification fundamental payoffs are those that occur at finite dates, whereas bubbles are payoffs that occur “at infinity”. In this alternative usage, for example, the payoff of a self-financing portfolio strategy has zero fundamental, implying that the initial cost of the portfolio strategy equals the value of the bubble under the equilibrium price system. This notion of infinity can be made precise mathematically using functional analysis (the branch of mathematical analysis that studies infinite-dimensional linear spaces and linear functions), but this is not the place for such discussion.

8 Rational Bubbles under Sequential Markets

An attractive feature of the analysis of bubbles under simultaneous markets is that the required departures from classical general equilibrium are minor, consisting of little more than respecifying commodity and price spaces to be infinite-dimensional.¹¹ However, simultaneous-markets models have some disadvantages: they presume the existence of complete markets. Also, the assumption that trading occurs simultaneously strikes many economists as implausible. Under sequential markets agents trade at each date rather than simultaneously, implying that they face separate budget constraints at each date and event.

The intuitive attractiveness of sequential markets is mitigated by the attendant assumption usually made that agents know security prices at future events (Radner

¹¹This is not to minimize the consequences for formal analysis of generalizing to infinite dimensions. For example, proving existence of equilibrium in that case involves serious mathematical complications.

It could be argued that the term “classical general equilibrium theory” should include the infinite-dimensional case, which was studied as early as the 1950s (Debreu [17]), the same time as standard general equilibrium theory as reflected in Debreu [18] was being developed.

[56]). That assumption is satisfied definitionally under simultaneous markets because the fiction of the auctioneer is maintained, but not under sequential markets. Under sequential markets the assumption is usually defended by a loose appeal to rational expectations, the plausibility of which depends on how recursive the economy is.

An essential feature of models incorporating sequential markets is that they necessarily impose trading restrictions. Otherwise the possibility of Ponzi schemes—borrowing money and rolling over the indebtedness forever—precludes existence of optima (this problem does not occur under simultaneous markets because of the integrated nature of intertemporal budget constraints in that setting). Major papers analyzing bubbles in sequential markets are Santos and Woodford [61] and Huang and Werner [39]. In the latter paper the authors compared bubbles under simultaneous markets with those under sequential markets; the two settings are surprisingly different. In sequential markets bubbles can occur only if trading restrictions rule out the arbitrage—short the security with a bubble and hold long securities that generate its dividends—that would exploit their existence.

More generally, the existence of trading restrictions under sequential markets limits the role of arbitrage in eliminating mispricing of any sort. For example, in continuous-time models the arbitrage that would exploit even a finite-time mispricing has unbounded sample paths (Liu and Longstaff [49]). Therefore the existence of bounds on portfolio values implies that the position may have to be closed at a loss before the purported arbitrage has converged. The necessary existence of such trading restrictions is routinely ignored in applied financial analysis (derivation of the Black-Scholes [7] option pricing model, for example).

Curiously, the existence of such limits to arbitrage has come to play a central role in supporting assertions that financial markets are inefficient (Shleifer [66], Barberis and Thaler [4]). 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 arbitrage is infeasible along some sample paths. For example, the apparent mispricing of Royal Dutch/Shell shares does not imply any market inefficiency because of the riskiness of the implied arbitrage.¹² In fact, the apparent mispricing is properly interpreted as evidence of the existence of a bubble on whichever of the companies is overpriced (or, possibly, both companies). A parallel argument applies to closed-end mutual funds, which display similar pricing phenomena.

Models incorporating simultaneous markets have the unattractive feature that they cannot be decentralized since budget sets extend into the indefinite future. Sequential markets, in contrast, can at least potentially be decentralized. However, the trading restrictions incorporated in simultaneous markets models frequently preclude decentralized solutions. For example, it is frequently assumed that commodity

¹²In this example two firms 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 (Froot and Dabora [24]).

bundles are finitely nonzero, that they have zero limits, or that they are bounded; these properties cannot be verified at any finite date. Thus the possibility of decentralized implementation does not afford clear grounds for preferring either market specification to the other.

Many students of bubbles do not specify clearly whether they are assuming simultaneous or sequential markets, perhaps because they presume incorrectly that the equivalence between the two in finite settings carries over to infinite settings. An example of this ambiguity is given in the next section. Of those who are clearly assuming sequential markets, many find the idea of analyzing bubbles in simultaneous markets difficult to understand or otherwise unattractive (for example, see Santos and Woodford [61], note 3). The basis for this opinion is far from clear. As already noted, the two settings are equivalent in finite settings, so a preference for one over the other must reflect some intuitive idea of how people behave when the future has no finite endpoint. One would like to see the idea that sequential markets are more plausible than simultaneous markets defended explicitly.¹³

9 Conditions that Eliminate Bubbles

As observed in the introduction, many analysts take the view that there exists a strong theoretical presumption against the existence of bubbles, particularly in finite-agent settings (for example, Santos and Woodford [61]). The simplest argument that eliminates bubbles is based on arbitrage: if under sequential markets agents can sell short arbitrary quantities of a security with a bubble and maintain the short position forever, and if they can hold long claims to the security's dividends, then they have an arbitrage. In that setting there can exist no bubble. If, on the other hand, the indicated portfolio strategy would violate trading restrictions, then there is no (feasible) arbitrage in the presence of a bubble, so the bubble can exist. In overlapping generations models it is typically assumed (often without explicit statement, as in Tirole [69] and in the discussion below) that agents can trade only when they are alive, so they cannot arbitrage away a bubble. The existence of equilibria with bubbles depends on this assumption.

One frequently encounters a purported proof that bubbles cannot occur in representative agent settings even if the agent is infinitely lived (Tirole [68]; Obstfeld and Rogoff [55]; Flood, Garber and Kaplan [22]). Assume that the representative agent maximizes

¹³Sequential markets, but not necessarily simultaneous markets, imply that agents view the doubling strategy (doubling the bet until they win) as an arbitrage: they are modeled as believing that being given the opportunity to gamble with each other forever allows each to make money with certainty (see Fisher and Gilles [21] and LeRoy [44]). This implausibility of this specification diminishes the intuitive appeal of sequential markets.

$$E \sum_{t=1}^{\infty} \beta^t U(c_t) \quad , \quad (12)$$

$\beta < 1$, where U has standard properties. Let p_t be the price of a capital good that produces the aggregate endowment c_t , as in Lucas [50]. For any n the Euler equation can be written

$$z_t = \sum_{i=1}^n \beta^i E_t(a_{t+i}) + \beta^n E_t(z_{t+n}), \quad (13)$$

where $z_t = U'(c_t)p_t$ and $a_t = U'(c_t)c_t$. Letting n go to infinity, there results

$$z_t = \lim_{t \rightarrow \infty} \sum_{i=1}^n \beta^i E_t(a_{t+i}) + \beta^n E_t(z_{t+n}) \quad . \quad (14)$$

The first term on the right-hand side of (14) is held to equal the marginal utility gain from holding the asset forever, which must equal the left-hand side. Therefore the bubble term must equal zero.

The problem with this argument depends on whether the model is interpreted as assuming simultaneous or sequential markets. In the former case the proof incorrectly generalizes from the utility function (12), for which the argument is correct, to the general representative agent case, for which it is not. For example, if the representative agent has utility function that coincides with the price system (11) the argument fails. The reason is that the argument assumes that $\sum_{i=1}^n \beta^i E_t(a_{t+i})$ converges to the utility gain from holding the asset forever. This transversality condition is satisfied for the utility function (12), but not for (11). Under the utility function (11), $\sum_{i=1}^n \beta^i E_t(a_{t+i}) = 0$ for any n , so the limiting value must also equal 0.

On the other hand, if the example is interpreted as assuming sequential markets, then there exists no equilibrium in the model as stated. This is so because the absence of trading restrictions implies that agents can conduct Ponzi schemes which, as noted above, is inconsistent with the existence of equilibria. Depending on what trading restrictions are added, bubbles may or may not be possible. For example, suppose that trading restrictions require each agent to hold at least a pro rata share of the aggregate endowment. Then bubbles can exist in equilibrium because agents cannot implement the arbitrage that would exploit the bubble. In general finite-agent settings it is known that there exists an intimate connection between what trading restrictions are assumed and whether or not bubbles can occur (Huang and Werner [39]).

While blanket assertions that bubbles cannot occur in representative-agent models are seen to be false, it is true that bubbles can occur only in highly specialized cases: Mackey-discontinuous utility functions under simultaneous markets; trading restrictions that essentially force agents to hold securities with bubbles under sequential

markets.¹⁴ Assuming sequential markets, bubbles are much more likely to occur in infinite-agent economies than in finite-agent economies. To see why, we note a basic condition that eliminates bubbles in either finite-agent or infinite-agent economies: under weak characterizations of trading restrictions (not satisfied, however, in the example of the preceding paragraph), bubbles cannot occur on securities in positive net supply if the aggregate endowment has finite value (Santos and Woodford [61], Huang and Werner [39]). The reason is that a bubble would confer wealth on some agents without creating new goods for the agents to consume, resulting in a violation of Walras' Law.¹⁵

The condition that the aggregate endowment have finite value is satisfied vacuously in finite-agent settings, but it may or may not be satisfied in infinite-agent settings. Overlapping generations models are the standard example of economies with an infinite number of agents. Bubbles cannot occur in overlapping generations economies when the allocation is Pareto-efficient. In that case the aggregate endowment has finite value, so the preceding theorem applies.¹⁶

Bubbles in overlapping generations models were first studied in the macroeconomics literature in the context of determining when intrinsically useless money could be valued in equilibrium (Wallace [70]); valued money is obviously a bubble because the fundamental value of money is zero. The fact that each generation lives for only a finite period does not prevent money from being valued because each generation passes the money to the next generation. General analysis of bubbles in overlapping generations models was first provided by Tirole [69]. Blanchard and Fischer [8] Ch. 5 is a good exposition.

Whether bubbles can occur in overlapping generations models boils down to whether the equilibrium allocation is efficient. Standard practice at this point is to cite Abel *et al.* [1], which argues that empirically the return on capital exceeds the growth rate of the economy, implying that the equilibrium allocation is in fact efficient and that therefore bubbles cannot exist.

The foregoing argument illustrates a recent tendency to appeal uncritically to rational expectations arguments, and also to rely, again uncritically, on an extreme reading of Milton's Friedman's [23] positivist defense of optimization. Let us decon-

¹⁴Kocherlakota [40] is an example.

¹⁵In relying on Walras' Law, the proof of nonexistence of bubbles is formally similar to the proof of Pareto optimality in overlapping generations models. Thus it is no accident that, a borderline case aside, Pareto optimality is equivalent to nonexistence of bubbles.

Henceforth the qualification having to do with the borderline case will be omitted.

¹⁶This result is frequently motivated using a misleading argument. In efficient equilibria the interest rate exceeds the rate of growth of the endowment. Therefore, it is argued, if a bubble existed it would eventually exceed the current endowment of the young, resulting in negative consumption. Negative consumption is infeasible under the logarithmic utility that is usually assumed.

This argument suggests incorrectly that the demonstration of nonexistence of bubbles is valid only when consumption is nonnegative. In LeRoy [45] this point is discussed in the context of an overlapping generations model in which agents have negative exponential utility, which admits negative consumption.

struct the argument that bubbles cannot exist (on securities in positive net supply) in efficient equilibria. Suppose that a bubble existed on some security. The bubble would constitute a wealth transfer to agents with positive endowments of that security. Walras' Law requires that some agent be subject to an offsetting negative wealth transfer at some date, possibly far in the future. But bubbles do not entail a negative wealth transfer at any future date, which contradicts the maintained assumption that the assumed equilibrium path is feasible into the infinite future. Realizing this, agents reprice the security in question so as to eliminate the bubble.

This line of reasoning uses the rational expectations argument that agents can calculate the trajectory of the economy into the arbitrarily distant future. Of course, it can be replied to this that the argument depends on agents acting as if they are making the calculation, not on the assumption that they actually do so. For example, following Friedman, it is not necessary that a skilled billiards player actually know the laws of mechanics, but just that he be able to play billiards as if he did. If he did not do so, he would not be a skilled billiards player.

Representing agents as optimizing in an environment that they understand is most plausible in repetitive situations. That is why billiards players practice: they encounter similar situations over and over again. 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 equilibrium paths—that by definition they have never experienced. This reasoning, which seems altogether implausible, goes unquestioned. This is especially odd given the recent enthusiasm for arguments involving irrationality in behavioral finance.

Similar issues have 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.¹⁷ 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 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 steady state—that is feasible in infinite time. The other paths all fail in finite time. Shell and Stiglitz [63], 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. It was only later that economists came to rely on the infinite-time feasibility of equilibrium paths to eliminate indeterminacy in dynamic economic models. These issues are discussed in Stiglitz's [67] introduction to the

¹⁷The best introduction to this literature is Burmeister and Dobell [11].

1990 *Journal of Economic Perspectives* symposium on bubbles, a paper which repays careful study.

Game theorists have been confronted with similar conundrums. Thoroughgoing application of rationality, and common knowledge of rationality, leads to serious conceptual difficulties. In some settings such apparently noncontroversial embodiments of rationality as Nash equilibrium are implausible (for example, Rubinstein [60]), and there is even a question of whether the ideas of rationality and common knowledge of rationality have unambiguous meaning in strategic settings involving asymmetric information. These ideas are discussed in Binmore [6]. These problems have led game theorists to rely increasingly on experimentation rather than deductive theorizing. It seems likely that experiments will be equally fruitful in the study of bubbles (see Lei, Noussair and Plott [42] and the papers cited there for experimental studies of bubbles).

The foregoing discussion raises serious methodological issues for dynamic macroeconomics. If, as argued, unswerving application of the positivist-rational expectations paradigm is implausible in the context of models incorporating an infinite future, it is at least methodologically coherent. Admitting paths that are not viable in infinite time raises awkward questions of when agents realize that the end is near, and how they behave when they learn this. We have no answers to these questions. As a stopgap, the best course of action may be to analyze bubbles within the framework of rational expectations models the equilibria of which are not assumed to be efficient.

This is not the place to pursue these lines. It seems fair to conclude, however, that the widely-encountered idea that there exists a strong theoretical presumption against the existence of bubbles is incorrect.

10 Bubbles and Sunspots

Many discussions of bubbles treat these as synonymous with sunspots.¹⁸ Sunspots occur when asset prices or other endogenous economic variables depend on non-fundamental uncertainty; bubbles occur when security prices are higher than fundamentals justify. The common presence of the term “fundamental” in these definitions misleads many into thinking that sunspots and bubbles are two equivalent terms for the same thing. In fact the ideas are completely different. “Fundamental value” in the definition of bubbles refers to the component of value given by discounting finite-date payoffs. This has no relation to “fundamentals” in the definition of sunspots, which refers to intrinsic uncertainty (uncertainty about variables like endowments that would be expected to be determinants of equilibrium values of endogenous variables).

¹⁸Major papers on sunspots are Azariadis [3] and Cass and Shell [13]. Shell [62] is a good introduction.

As the foregoing observation suggests, bubbles can occur without sunspots and vice-versa. First, it is easy to construct models in which equilibrium is unique and deterministic, but in which some or all assets have bubbles. Such models, being deterministic, do not have sunspots. Correspondingly, it is equally easy to construct finite-date models with multiple equilibria. Randomizing over these equilibria produces sunspots, but because the models have a finite number of dates there can be no bubbles.

Despite the fact that the definitions of bubbles and sunspots are unrelated, it turns out that in some models of interest the conditions under which bubbles and sunspots can occur coincide. To see this, note that deterministic overlapping generations models in which the endowment allocation is inefficient have a continuum of equilibrium paths with bubbles. Randomizing over these paths introduces sunspots. In contrast, when the endowment allocation is efficient there cannot exist bubbles, as shown above. Further, equilibrium is unique, so there are no sunspots.¹⁹ It is seen that inefficient overlapping generations equilibria produce both bubbles and sunspots, whereas neither can occur under efficient equilibria.

This result justifies the interpretation of the variance bounds tests described in Section 3 above as favoring the existence of bubbles: if conditions are such that the economy can produce bubbles, it is likely that the equilibrium randomizes over multiple equilibria, producing excess volatility of asset prices.

11 Conclusion

As the title of this paper implies²⁰, we are arguing that the best explanation for the outstanding characteristics of stock prices—high volatility of prices and returns, plus a major price runup and decline over the decade just ended—is that the stock market is best modeled as an equilibrium exhibiting both rational bubbles and sunspots. It is true that there exists a theoretical argument that eliminates bubbles, as summarized above. That line of reasoning, although formally valid, is so recondite that it is understood only by highly trained theorists. It is a testament to economists' capacities for abstraction that they take seriously the possibility that this theoretical line has somehow gravitated from the pages of *Econometrica* to the New York Stock Exchange.

Turning to empirical considerations, the volatility of security prices and returns is the clearest evidence of the presence of sunspots and bubbles. As frequently observed, it is conceivable that various forms of model misspecification could also produce the phenomena that we are attributing to bubbles, but there has been little progress in

¹⁹We are assuming a single good here. For discussion of the dimensionality of equilibrium in the general case, see Geanakoplos [26] and the papers cited there.

²⁰At this writing (February 2003) the term “exuberance” describes the stock market less well than when the project was begun during the twentieth century. However, it is worth pointing out that, despite the current negative tone, the stock market now is at about the same level as it was in 1996 when Alan Greenspan warned of irrational exuberance.

developing such alternatives. One would like to see the development of empirical tests that could distinguish between bubbles and misspecification.²¹

The contention that the bubbles are rational is controversial. How, it is argued, could any rational person buy dot-com stocks at the prices prevailing in 1999? This is, of course, 20-20 hindsight; things were not so obvious at the time. Perhaps five years from now we will say the same thing about real estate prices in southern California. Those who are confident of the existence of a class of rational traders who can reliably identify and bet against the bubbles will be comfortable with the idea that these are nonrational bubbles, since as noted rational but risk-averse traders will not completely arbitrage away the bubbles. Those who, like me, are not confident of this will prefer the rational bubbles alternative, so that bubbles do not in fact offer any trading opportunities.

Academic financial economists think about stock prices as the discounted values of their dividend streams. Practitioners accept this story “theoretically”—here “theoretically” has the usual undergraduate meaning of according to some classroom model that they have to learn—but don’t think this story has much to do with reality. This puts economists in the awkward position of believing that despite their stated skepticism in the classroom practitioners somehow come on board in terms of their actual conduct. Again we see the easy application by economists of an extreme version of Friedman’s positivism. Financial economists have more luck convincing practitioners of the merits of the idea of capital market efficiency: if there existed profitable trading rules investors would exploit them, and the resulting shifts in asset prices would eliminate the opportunities. Puzzlement reappears, however, when it is asserted that present-value pricing and capital market efficiency are in fact equivalent, so there is no justification for skepticism about present-value pricing combined with acceptance of capital market efficiency.

But the two are actually not quite equivalent: the efficient markets characterization of returns implies the present-value relation only if there are no bubbles. A world of rational bubbles conforms to practitioners’ intuitive ideas: asset prices are not accurately described as the present value of their dividend streams, but there do not exist unexploited profitable trading rules.

Rational bubbles, unlike the explanations for financial market misbehavior favored in behavioral finance, have falsifiable implications. In fact, it is easy to point to phenomena which rational bubbles cannot possibly explain. The equity premium puzzle of Mehra and Prescott [52] is a clear example. Mehra and Prescott focused on the implications of absence of arbitrage on returns. These predictions hold whether or

²¹West observed that one can test for the presence of bubbles by determining whether the coefficients of an estimated pricing model differ significantly when a no-bubble condition is imposed from the corresponding unrestricted coefficients. In [71] he found that the test just described favored the existence of bubbles. The West test was criticized on econometric grounds by Flood, Garber and Kaplan [22]. In [72] West reversed his conclusion about the likelihood of bubbles. One wonders what the outcome of the West test would be if they were updated to include data from the 1990s.

not there exist bubbles, so the presence of bubbles cannot explain why the predictions appear not to be satisfied.

Asset prices have been shown innumerable times to depend strongly on variables not easy to reconcile with the present-value relation (for example, Victor Niederhoffer [54], David M. Cutler, James M. Poterba and Lawrence H. Summers [16], Richard Roll [59]). To the extent that these dependences imply profitable trading rules, they are genuine anomalies. However, for the most part they do not do so. Therefore they should be interpreted as evidence of rational bubbles combined with sunspots.

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] Kenneth J. Arrow. The role of securities in the optimal allocation of risk bearing. *Review of Economic Studies*, pages 91–96, 1964.
- [3] Costas Azariadis. Self-fulfilling prophecies. *Journal of Economic Theory*, 25:380, 1981.
- [4] Nicholas Barberis and Richard Thaler. A survey of behavioral finance. reproduced, University of Chicago, 2002.
- [5] Truman F. Bewley. Existence of equilibria in economies with infinitely many commodities. *Journal of Economic Theory*, 4:514–540, 1972.
- [6] Ken Binmore. *Essays in the Foundations of Game Theory*. Blackwell, 1990.
- [7] Fischer Black and Myron Scholes. The valuation of option contracts and a test of market efficiency. *Journal of Finance*, 27:399–418, 1972.
- [8] Olivier Blanchard and Stanley Fischer. *Lectures on Macroeconomics*. MIT Press, Cambridge, MA, 1989.
- [9] Peter L. Bossaerts. *The Paradox of Asset Pricing: Frontiers of Economic Research*. Princeton University Press, Princeton, NJ, 2002.
- [10] Douglas T. Breeden. An intertemporal asset pricing model with stochastic consumption and investment opportunities. *Journal of Financial Economics*, 7:265–296, 1979.
- [11] Edwin Burmeister and A. Rodney Dobell. *Mathematical Theories of Economic Growth*. Ashgate, 1993.

- [12] John Y. Campbell, Andrew W. Lo, and A. Craig MacKinlay. *The Econometrics of Financial Markets*. Princeton University Press, Princeton, NJ, 1996.
- [13] David Cass and Karl Shell. Do sunspots matter? *Journal of Political Economy*, 91:193–227, 1983.
- [14] John Cochrane. Volatility tests and efficient markets: A review essay. *Journal of Monetary Economics*, 27:463–485, 1991.
- [15] John Cochrane. *Asset Pricing*. Princeton University Press, Princeton, 2001.
- [16] David Cutler, James Poterba, and Lawrence Summers. What moves stock prices? *Journal of Portfolio Management*, 15:4–12, 1989.
- [17] Gerard Debreu. Valuation equilibrium and Pareto optimum. *Proceedings of the National Academy of Sciences*, 40:588–592, 1954.
- [18] Gerard Debreu. *Theory of Value*. Wiley, New York, 1959.
- [19] Eugene F. Fama. Efficient capital markets: A review of theory and empirical work. *Journal of Finance*, 25:283–417, 1970.
- [20] Eugene F. Fama and Kenneth R. French. The cross section of expected stock returns. *Journal of Finance*, 47:427–466, 1992.
- [21] Mark Fisher and Christian Gilles. An analysis of the doubling strategy: The countable case. reproduced, Federal Reserve Bank of Atlanta, 2002.
- [22] Robert P. Flood, Peter M. Garber, and Paul Kaplan. An evaluation of recent evidence on stock market bubbles. In Robert P. Flood and Peter M. Garber, editors, *Speculative Bubbles, Speculative Attacks, and Policy Switching*. MIT Press, 1994.
- [23] Milton Friedman. The methodology of positive economics. In Milton Friedman, editor, *Essays in Positive Economics*. University of Chicago Press, 1953.
- [24] Kenneth A. Froot and Emil Dabora. How are stock prices affected by the location of trade? *Journal of Financial Economics*, 53:182–216, 1999.
- [25] Peter M. Garber. *Famous First Bubbles*. MIT Press, Cambridge, 2000.
- [26] John Geanakoplos. Overlapping generations models. In M. Milgate J. Eatwell and P. Newman, editors, *The New Palgrave: A Dictionary of Economics*. McMillan, 1987.
- [27] Christian Gilles. Charges as equilibrium prices, and asset bubbles. *Journal of Mathematical Economics*, 18:155–167, 1989.

- [28] Christian Gilles and Stephen F. LeRoy. Asset price bubbles. In John Eatwell, Murray Milgate, and Peter Newman, editors, *The New Palgrave Dictionary of Money and Finance*. Macmillan Press, 1992.
- [29] Christian Gilles and Stephen F. LeRoy. Bubbles and charges. *International Economic Review*, 33:323–339, 1992.
- [30] Christian Gilles and Stephen F. LeRoy. Rational bubbles as payoffs at infinity. *Economic Theory*, 9:261–281, 1997.
- [31] Myron J. Gordon. Variability of price-earnings ratios. *American Economic Review*, 52:203–208, 1962.
- [32] Sanford J. Grossman and Robert J. Shiller. The determinants of the variability of stock prices. *American Economic Review Papers and Proceedings*, 71:222–227, 1981.
- [33] 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.
- [34] Robert E. Hall. The stock market and capital accumulation. *American Economic Review*, 91:1185–1202, 2001.
- [35] Robert E. Hall. Struggling to understand the stock market. *American Economic Review Papers and Proceedings*, 91:1–11, 2001.
- [36] Robert E. Hall. Dynamics of corporate earnings. reproduced, Stanford University, 2002.
- [37] Lars Hansen and Kenneth J. Singleton. Stochastic consumption, risk aversion, and the temporal behavior of asset returns. *Econometrica*, 91:249–265, 1983.
- [38] Lars P. Hansen and Ravi Jagannathan. Implications of security market data for models of dynamic economies. *Journal of Political Economy*, 99:225–262, 1991.
- [39] Kevin X. D. Huang and Jan Werner. Asset price bubbles in Arrow-Debreu and sequential equilibrium. *Economic Theory*, 15:253–278, 2000.
- [40] Narayana R. Kocherlakota. Bubbles and constraints on debt accumulation. *Journal of Economic Theory*, 57:245–256, 1992.
- [41] C. J. LaCivita and Stephen F. LeRoy. Risk aversion and the dispersion of asset prices. *Journal of Business*, 54:535–547, 1981.
- [42] V. Lei, C. N. Noussair, and Charles R. Plott. Non-speculative bubbles in experimental asset markets: Lack of common knowledge of rational versus actual irrationality. *Econometrica*, 69:831–859, 2001.

- [43] Stephen F. LeRoy. Efficient capital markets: Comment. *Journal of Finance*, 31:139–141, 1976.
- [44] Stephen F. LeRoy. Infinite portfolio strategies. reproduced, University of California, Santa Barbara, 2002.
- [45] Stephen F. LeRoy. Positivity and bubbles in overlapping generations models. reproduced, University of California, Santa Barbara, 2002.
- [46] Stephen F. LeRoy and William R. Parke. Stock price volatility: Tests based on the geometric random walk. *American Economic Review*, 82:981–992, 1992.
- [47] Stephen F. LeRoy and Richard D. Porter. Stock price volatility: Tests based on implied variance bounds. *Econometrica*, 49:555–574, 1981.
- [48] Stephen F. LeRoy and Jan Werner. *Principles of Financial Economics*. Cambridge University Press, Cambridge, 2001.
- [49] Jun Liu and Francis A. Longstaff. Losing money on arbitrages. reproduced, Anderson School, UCLA, 2000.
- [50] Robert E. Lucas. Asset prices in an exchange economy. *Econometrica*, 46:1429–1445, 1978.
- [51] Ellen McGrattan and Edward C. Prescott. The stock market crash of 1929: Irving fisher was right! reproduced, NBER, 2001.
- [52] Rajnish Mehra and Edward C. Prescott. The equity premium: A puzzle. *Journal of Monetary Economics*, 15:145–161, 1985.
- [53] Merton H. Miller and Franco Modigliani. Dividend policy, growth, and the valuation of shares. *Journal of Business*, 34:411–433, 1961.
- [54] Victor Niederhoffer. The analysis of world news events and stock prices. *Journal of Business*, 44:193–219, 1971.
- [55] Maurice Obstfeld and Kenneth Rogoff. Speculative hyperinflations in maximizing models: Can we rule them out? *Journal of Political Economy*, 91:675–687, 1983.
- [56] Roy Radner. Existence of equilibrium of plans, prices and price expectations in a sequence economy. *Econometrica*, 40:289–303, 1972.
- [57] A. C. Rayner and I. M. D. Little. *Higgledy Piggledy Growth Again*. Oxford University Press, Basic Blackwell, 1966.
- [58] W. B. Reddaway. Higgledy piggedly growth. *Economic Journal*, 77:595–598, 1967.

- [59] Richard Roll. Orange juice and weather. *American Economic Review*, 74:861–880, 1984.
- [60] Ariel Rubinstein. The electronic mail game: Strategic behavior under ‘almost common knowledge’. *American Economic Review*, 79:385–391, 1989.
- [61] Manuel Santos and Michael Woodford. Rational asset pricing bubbles. *Econometrica*, 65:19–57, 1997.
- [62] Karl Shell. Sunspot equilibrium. In J. Eatwell, M. Milgate, and P. Newman, editors, *The New Palgrave: A Dictionary of Economics*. McMillan, 1987.
- [63] Karl Shell and Joseph E. Stiglitz. The allocation of investment in a dynamic economy. *Quarterly Journal of Economics*, 81:592–609, 1967.
- [64] Robert J. Shiller. Do stock prices move too much to be justified by subsequent changes in dividends? *American Economic Review*, 71:421–436, 1981.
- [65] Robert J. Shiller. *Irrational Exuberance*. Princeton University Press, Princeton, NJ, 2000.
- [66] Andrei Shleifer. *Inefficient Markets—An Introduction to Behavioral Finance*. Oxford University Press, Oxford and New York, 2000.
- [67] Joseph E. Stiglitz. Symposium on bubbles. *Journal of Economic Perspectives*, 4:13–17, 1990.
- [68] Jean Tirole. On the possibility of speculation under rational expectations. *Econometrica*, 50:1163–1182, 1982.
- [69] Jean Tirole. Asset bubbles and overlapping generations. *Econometrica*, 53:1499–1528, 1985.
- [70] 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.
- [71] Kenneth D. West. A specification test for speculative bubbles. *Quarterly Journal of Economics*, 102:37–61, 1987.
- [72] Kenneth D. West. Bubbles, fads and stock price volatility: A partial evaluation. *Journal of Finance*, 43:636–656, 1988.
- [73] Kenneth D. West. Dividend innovations and stock price volatility. *Econometrica*, 56:37–62, 1988.