Gracia, Eduard

Working Paper

On the power and weakness of rational expectations: Logical fallacies, periodic bubbles and business cycles

Economics Discussion Papers, No. 2012-27

Provided in Cooperation with:
Kiel Institute for the World Economy (IfW)

Economics
The Open-Access, Open-Assessment E-Journal

Discussion Paper
No. 2012-27 | June 7, 2012 | http://www.economics-ejournal.org/economics/discussionpapers/2012-27

On the Power and Weakness of Rational Expectations: Logical Fallacies, Periodic Bubbles and Business Cycles

Eduard Gracia

Abstract A popular interpretation of the Rational Expectations/Efficient Markets hypothesis states that, if the hypothesis holds, then market valuations must follow a random walk. This postulate has frequently been criticized on the basis of empirical evidence. Yet the assertion itself incurs what we could call 'fallacy of probability diffusion symmetry': although market efficiency does indeed imply that the mean (i.e. 'expected') path must be a random walk, if the probability diffusion process is asymmetric then the observed path will most closely resemble not the mean but the median, which does not necessarily follow a random walk.

To illustrate the implications, this paper develops an efficient markets model where the median path of Tobin's *q* ratio displays regular cycles of bubbles and crashes reflecting an agency problem between investors and producers. The model is tested against U.S. market data, with results suggesting that such a regular cycle does indeed exist and is statistically significant. The aggregate production function in Gracia (*Uncertainty and Capacity Constraints: Reconsidering the Aggregate Production Function*, 2011) is then put forward to show how financial fluctuations can drive the business cycle by periodically impacting aggregate productivity and, as a consequence, GDP growth.

JEL E22, E23, E32, G12, G14

Keywords Rational Expectations; efficient markets; financial bubbles; stock markets; booms and crashes; Tobin's *q;* business cycles; economic rents

Correspondence Eduard Gracia, 39 Palace Court, London W2 4LS, United Kingdom; e-mail: eduard.gracia@booz.com

ON THE POWER AND WEAKNESS OF RATIONAL EXPECTATIONS:

LOGICAL FALLACIES, PERIODIC BUBBLES AND BUSINESS CYCLES

"The next major bust, 18 years after the 1990 downturn, will be around 2008, if there is no major interruption such as a global war"

Fred Foldvary (1997)

1. Introduction: The Fallacy of Diffusion Symmetry

In the wake of the 2008 recession, as of every major recession for the last 150 years, the question of why the downturn happened and whether the dismal science should have predicted it has been posed again. In a way, this follows an old tradition: after all, the Long Depression in the 1870s triggered the Marginalist Revolution, the Great Depression in the 1930s gave birth to Keynesianism, and the Oil Crisis in the 1970s marked the ascendancy of the Rational Expectations Hypothesis. In its 21st Century edition, a substantial part of the challenge seems to focus on this latter hypothesis, and whether a return to adaptive expectations / bounded rationality models would improve their explanatory power respective to empirical observations. Unfortunately the debate, at least at a popular level, is vitiated by a logical fallacy: what we could call the 'fallacy of probability diffusion symmetry'.

The popular controversy goes as follows. Mainly under the neoclassical banner, defenders of the Rational Expectations / Efficient Markets hypothesis claim that, should markets not behave according to this hypothesis, they would create arbitrage opportunities that would enrich anyone clever enough to spot them, and it would therefore suffice to add a few rational players in the mix for their irrational

competitors to be driven out of business. This, they argue, means that observed market values must follow a random walk, and therefore stock market bubbles and crashes must be utterly unpredictable, for any predictable patterns mush already be discounted out. Those in the opposite camp take then this conclusion as the central contention point, against which they pose numerous examples of departures from the random walk hypothesis in the observed data, as well as names of economists who, against the mainstream opinion, were able to forecast the 2008 crisis well before it hit.

From a strictly logical perspective, however, both sides might well be wrong. Indeed, both camps are implicitly assuming that, if expectations of future market prices were rational, they would reflect the mean (or 'expected') path calculated on the basis of all the information available, and therefore, if current prices reflected the future mean path, we should expect their observed trajectory also to approximate the mean. Hence, if arbitrage ruled out any predictable patterns along the mean path, then they should also be absent from the observed path – and, therefore, proving their existence in the observed data series would also prove that expectations were not rational in the first place. Yet, sensible as it sounds, **this implication is only true if the underlying probability diffusion process is symmetric.**

An intuitive example may make this point clear. Imagine a game of triple-or-nothing: you make a bet of, say, $10, toss a coin and triple the investment if result is heads, or else lose it all if it is tails. Evidently, if the game is played only once, the distribution of results is symmetric, with a 50% probability of making triple or nothing, and with both mean (i.e. average) and median (i.e. the value that leaves 50% of the distribution on either side) being $15. Yet if we play the game, say, ten times in a row, the distribution changes: now there is a 0.098% probability of making $590,490 and a

99.9% of losing everything: in this case, the mean value is $577 (a substantial profit respect to the $10 investment), but the median is obviously zero. Under Rational Expectations, $577 is of course the investor's expected value; yet, should an external observer analyze the data series, there is a 99.9% probability that the observed value after ten tosses be nil i.e. closest to the **median path**. In other words: the prediction with the highest probability of success is not the *mean*, but the *median*[1].

It follows that, if probability diffusion processes are asymmetric (as virtually all the standard asset pricing models are), then **neither does the observation of predictable market patterns imply irrationality, nor does the Rational Expectations Hypothesis rule them out**, for there is nothing preventing them from appearing on the median path[2]. Hence, the classical papers by Fama (1965) and Samuelson (1965) postulating that valuations in a rational, arbitrage-precluding efficient market would follow a random walk remain completely valid – only, they apply solely to the mean path. This, importantly, does not imply that we are in front of a money machine: it does not mean investors can resort to those predictable patterns to 'beat the market', but, at most, to fine-tune their probability of losses. For example, in the game of triple-or-nothing, a rational investor may bet $10 and expect to end up with $577, but by not betting may avoid suffering a loss 99.9% of the times – and, by selling short (assuming this were allowed in the game), might achieve a gain 99.9% of the times even with a negative expected value of the position (more or less like the trapeze artist who makes a bit of money every night at the expense of risking life and limb). Similarly, investing in assets with, say, an abnormally high P/E ratio may not indicate

[1] See Appendix 1 for a more rigorous proof of this same proposition in the case of an asset whose returns follow a standard Gauss-Wiener process (i.e. a normally-distributed random walk).

[2] This, importantly, has really nothing to do with the "fallacy of ergodicity" highlighted by authors such as Paul Davidson (e.g. Davidson 2009): the binomial random process in this example, for instance, is totally ergodic and has a perfectly well understood probability distribution.

any form of 'irrational exuberance' on the investors' side but a rational valuation of an asset with, say, a high expected gain as well as a high probability of loss[3]. Under these conditions, to be sure, there is nothing preventing as smart analyst from repeatedly issuing a successful contrarian prediction: in the game of triple-or-nothing, for instance, our analysts could forecast a loss of $10, and be right 99.9% of the time.

There are unfortunately very few examples of papers in the literature where the association of the observed time series to the median trajectory instead of the mean plays a role in the core analysis. A good example is Roll (1992), which proves that portfolio managers who are measured against their deviation from a market index are essentially being forced to track the market median path, which is suboptimal respective to the mean. A much more direct precedent, however, is Gracia (2005), which shows how an efficient market subject to a normal random-walk perturbation may display a persistent, periodic cycle of asset valuation bubbles and crashes along the median path, even though the mean remains cycle-free[4].

The purpose of this paper is to develop a model of financial valuations in an efficient, rational expectations market such that it would result in a persistent cycle on its median path – and hence, given a representative enough sample, in the observed path. The model is directly inspired in Gracia (2005), although it has been improved to make it both more general and more parsimonious. The predictions are then tested on the basis of two closely-related indicators of U.S. stock market valuation: Robert Shiller's Cyclically-Adjusted Price-Earnings ratio (CAPE) (see for instance Shiller 2005) and Stephen Wright's long-run estimates of Tobin's *q* ratio (provided in Wright

[3] Note that this differs from the rational bubbles literature following Blanchard & Watson (1982), which requires asset values to diverge from their fundamentals – something our model does not need.

[4] For reference, this article can be downloaded from http://www.metropolis-verlag.de/Predator-Prey---An-Alternative-Model-of-Stock-Market-Bubbles-and-the-Business-Cycle/11423/book.do.

2004). In this context, testing against Shiller's CAPE metric is particularly relevant because it has proven to be a very good leading indicator of financial bubbles and therefore, if it happened to have a strong cyclical component, it would constitute *prima facie* evidence that bubbles also do. Last but not least, we resort to the production function model developed in Gracia (2011) to explain how such financial market behavior would cause business cycle fluctuations that would be consistent with various key stylized facts without implying any sort of irrational behavior[5].

This paper is structured as follows. Section 2 outlines some literature on successful market predictions, and elaborates a bit more on the questions this paper tries to answer. Section 3 provides an intuitive explanation of the model; for the sake of readability, the analytical development has been committed to Appendix 2. Section 4 then proceeds to test the model's key predictions. Subsequently, Section 5 discusses the macroeconomic implications of this model when combined with Gracia (2011). Finally, Section 6 provides a summary of the paper's main findings and conclusions.

2. Predicting the Unpredictable

As many economists kept asserting that no one could have seen this latest downturn coming[6], it became almost a popular pastime in heterodox circles to list those who had most egregiously missed it, as well as those who had most spectacularly got it right. There is certainly no shortage of the latter. A 2010 contest in the Real-World Economics Review Blog, the "Revere Award", shortlisted twelve economists who made particularly accurate predictions of the crash (Dean Baker, Wynne Godley, Michael Hudson, Steve Keen, Paul Krugman, Jakob Brøchner Madsen, Ann Pettifor,

[5] This can be downloaded from http://www.economics-ejournal.org/economics/journalarticles/2011-19.

[6] For a good survey of such assertions see for example Bezemer (2009).

Kurt Richebächer, Nouriel Roubini, Robert Shiller, George Soros and Joseph Stiglitz). Even this list has been heavily criticized due to some glaring omissions such as Marc Farber, Fred Foldvary, Fred Harrison, Michael Hudson, Eric Janszen, Raghuram Rajan, Peter Schiff or Nassim Nicholas Taleb[7]. Many if not most of these authors arrived to their conclusions by considering various financial indicators in the context of a non-neutral-money, bounded-rationality or otherwise inefficient-markets model. Moreover, by relying on cyclical market patterns some were even able to make uncannily accurate predictions even *before* the financial indicators would raise any grounds for concern. This is the case of Foldvary (1997) and Harrison (1997) who, on the basis of the 18-year cycle that Hoyt (1933) observed 80 years ago in the Chicago real estate market, forecasted the 2008 global recession more than a decade in advance. Not that this was a one-off success either: Harrison (1983), for example, already resorted to the same real estate cycle pattern to predict the property and overall business downturn in 1992 as a follow up from the one in 1974.

Even among the supporters of bounded rationality, the existence of predictable long-range cycles spanning many years or even decades is frequently regarded with skepticism. It has not always been so: until the 1930s it was commonly accepted in academic circles that economic waves existed and that they had well-defined frequencies. Nearly 150 years ago, Juglar (1862) observed a trade cycle with a wavelength of 7-11 years associated to fixed asset investment. Then, in the early 20th Century, similar findings came in quick succession: Kitchin (1923) identified a shorter, inventory-driven cycle lasting around 35-50 months (i.e. 3 to 5 years), Kondratiev (1926) a long wave lasting 45-60 years, and Kuznets (1930) an

[7] References for each one of these authors' predictions can be found, for example, in Bezemer (2009) and Gaffney (2011) as well as in the *Real-World Economics Review Blog*: http://rwer.wordpress.com/2010/03/31/shortlist-for-the-revere-award-for-economics-3/

intermediate swing lasting 15-25 years, which he associated to bursts of building activity (thus linking it to the 18-year property pattern that Hoyt 1933 would identify shortly afterwards). Yet, as Burns & Mitchell (1946) argued, under more demanding empirical tests the evidence was far from conclusive, so the view that cycles were, as Zarnowitz (1992) put it, "recurrent but non-periodic" gradually took hold.

Identifying cycles on the basis of aggregate GDP data is indeed fraught with technical difficulties, as the data series are either not long enough or not homogeneous enough. Yet stock market datasets, which are more granular, have been empirically proven to reflect mean-reversion patterns akin to business cycles. Thus, for example, Fama & French (1988) as well as Poterba & Summers (1988) identified a 3-5 year cycle in stock market returns, which is essentially the same frequency that Kitchin documented 65 years. Evidently the existence of such cycles lends credibility also to longer waves such as the 18-year property cycle Harrison and Foldvary postulate. Furthermore, those who successfully predicted the crash and did not rely on the regularity of a cycle generally managed to do this on the basis of their observation of anomalous values for certain publicly-available indicators such as the CAPE ratio (e.g. Shiller 2005) or the debt to GDP ratio (e.g. Keen 2006). Under market efficiency, the information contained in those variables should already have been discounted from the mean path.

At the same time, the fact that these predictions not only involved fluctuations in the financial markets but also in the aggregate economy poses a different question: how can a financial crash impact the real economy in a strictly rational world? How does, for instance, a liquidity shortage impact growth in a world where money is neutral? To make things more complicated, we know at least since Kydland & Prescott (1991)

that around 70% of the cycle is statistically explained by (i.e. correlated to) fluctuations of Total Factor Productivity (TFP). This has frequently been put forward as empirical evidence in favor of the Real Business Cycle view, first put forward by Kydland & Prescott (1982) and then Long & Plosser (1983), that business cycles result from technology shocks impacting TFP… Yet, as Galí (1996) and Shea (1998) point out, there is virtually no correlation between TFP fluctuations and actual technology shocks. So it does not suffice to (re)introduce imperfect markets and/or bounded rationality in the model, as is done in various New Keynesian models, to explain credit cycles: one also needs to explain how a financial shock could cause a fall in output *through* productivity[8] and not just through a fall in aggregate demand.

In sum, the questions the followings section will try to answer are two:

a. Can an efficient stock market display a periodic cycle of bubbles and crashes?

b. How could such financial phenomena impact aggregate productivity (and, through it, aggregate output) in a rational expectations economy?

Section 3 outlines how (a) could happen[9]; a rationale for (b) is offered in Section 5.

3. A Model of Periodic Stock Market Cycles

Let's imagine a market with two kinds of players:

I. Producers who manage and have control of a company's productive process.

II. Investors who contribute their resources in exchange for a return.

[8] Kiyotaki & Moore (1997), for example, assume a productivity shock is the trigger, not the result, of the credit shortage – which is inconsistent with the experience of many finance-driven recessions.

[9] See Appendix 2 for an analytical version of the reasoning in Section 3.

This of course defines an agency problem i.e. a situation where agents (producers in this case) act on behalf of principals (that is, investors) but, at the same time, can also use their privileged position and knowledge to exploit them. In the absence of any control mechanism by investors, to be sure, the information asymmetry in favor of producers would open an easy route for them to exploit investors simply by raising cash, transferring it to their own private accounts or otherwise spending it for their personal purposes and then declaring bankruptcy. Even with controls, this exploitation can take many forms. Managers, for example, may assign to themselves salaries and perks above market level, or may decide to make empire-building investments that increase their power but yield poor returns, or may use their privileged information to do some insider dealing; other employees, similarly, may also use their privileged information to shirk their duties and/or to generate efficiency rents for themselves.

To counter these dangers, investors implement a system of punishments and rewards including, for example, regulation against fraud, audits and bureaucratic controls, the threat of takeover and dismissal, etc. To be sure, these controls are neither free nor foolproof, so their implementation will be subject to a cost-benefit analysis: investors will only impose them up to the point where their marginal costs equal their marginal benefits (i.e. the expected reduction of depredation costs). Ultimately, the benchmark against which investors will gauge the firm's performance is its liquidation value i.e. whether the assets the enterprise manages would be worth more or less if sold in the market instead of being managed within the company's framework. A straightforward way to measure this cost of opportunity is the ratio of a company's market price divided by its assets' replacement value i.e. what is known as Tobin's *q*: if the ratio falls below unity, the firm is actually destroying value and therefore liquidation

becomes an attractive option[10]. Since liquidation (or, in a less dramatic version, reorganization and dismissal of the most exploitative producers) puts a hard stop to the producers' rent extraction, it represents the investors' ultimate weapon; yet, to the extent it is neither instant nor cost-free, its deterrence power is also limited.

Under these conditions, the producers' rent extraction takes place as follows. For a given production process structure, the producer has a certain degree of control that translates into a given percentage of the output being siphoned out as rents. At every given point in time, a certain number of new opportunities to modify this productive structure will randomly pop up. Other things being equal, the producers' decisions are of course biased in favor of the options that generate a higher level of rents, but their power to choose is limited by the control mechanisms imposed by investors. Rational producers will therefore pluck the goose just enough to maximize their future expected rents, which also means leaving just enough to keep investors happy. Markets will then price these companies accordingly, so that the expected growth of their value equals their discount rate less (plus) the weight of the net cash flow they distribute to (raise from) investors. The higher this price stands above the threshold value (i.e. the higher Tobin's q ratio), to be sure, the lower the probability that investors liquidate, and therefore the more freedom of action producers will have to maximize their rents within the boundaries of the investors' control mechanisms.

In an efficient market, the mean path is of course the one where both investors' and producers' expectations are fulfilled and asset values grow precisely at their market discount rate less the dividends they cash out. Yet it is not too difficult to prove (see Appendix 1) that, when a probability diffusion process is asymmetric (as is the case of

[10] Needless to say, Tobin's q is the standard metric first put forward by James Tobin (1969).

the geometric Wiener process so frequently used in financial modeling), the *median*, not the mean path, is the best approximation to the observed time series… And along the median path (as along any path different from the mean) the players' expectations consistently fail to be met, and thus constantly need to be realigned.

Hence, if one assumes a standard geometric Wiener perturbation on market valuations (i.e. a function such that the median always falls below the mean), the observed interaction of producers and investors under the conditions outlined above would result in a cycle. Indeed, as new opportunities to pluck the goose appear, producers take them only to the extent they do not expect them to bring the company's return to investors below the market discount rate – yet this leads them to overshoot more often than not, as the underlying business' *median* returns are lower than the *mean* (i.e. expected) ones. When this happens, as long as the companies' market price remains comfortably above their liquidation value, the increment in the number of liquidated companies resulting from this market correction is small. Sooner or later, however, the accumulated impact of this excess rent extraction eventually brings company valuations below asset replacement values. As a result, the firms whose q-ratio is the lowest, which are in principle those whose producers have been extracting the largest rents, are gradually liquidated (as liquidation is not an instant process), thereby eventually reducing the weight of producer rents on the whole system until growth can be resumed[11]. Since both the process of excess rent accumulation and that of asset liquidation take time, the result of their interaction may well be a periodic cycle.

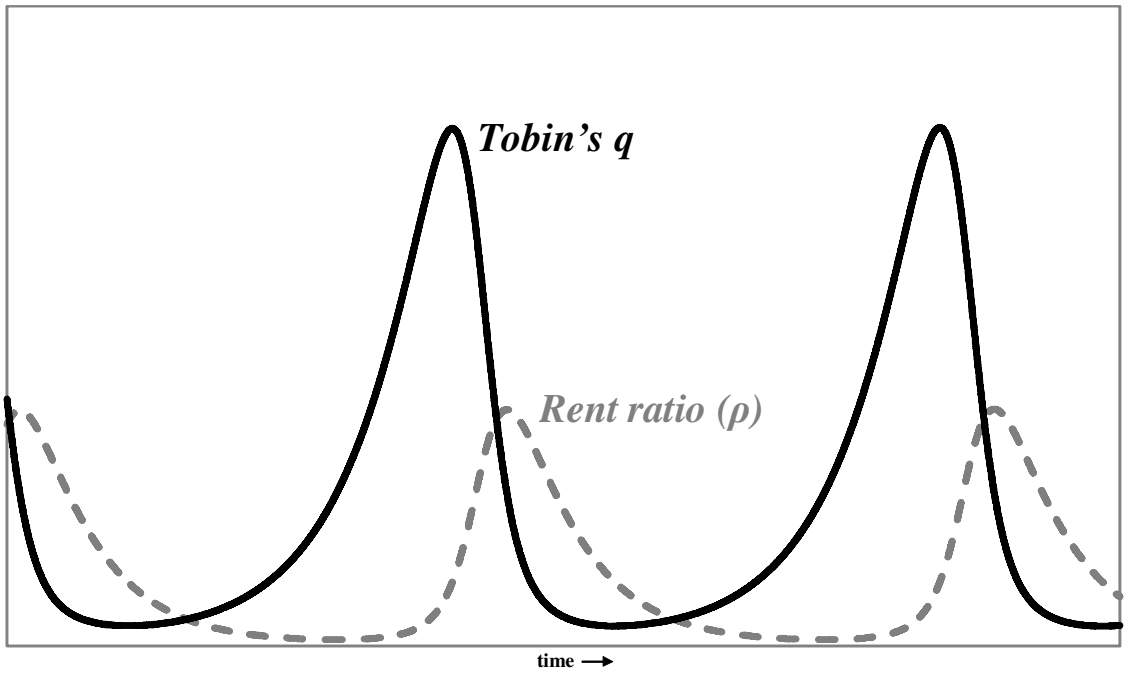
[11] To be sure, when a company's assets are liquidated they are bought by (and thus incorporated into) another company, so liquidation does not necessarily lead to a reduction in the volume of assets put to productive purposes; yet, to the extent the firms liquidated are those that extracted the highest rents, the average rate of rent extraction must come down as a result. The same outcome is of course to be expected from "low key" liquidations i.e. reorganizations where the worst-offending producers are fired and the company processes are rearranged to further limit shirking opportunities.

Appendix 2 develops the reasoning above analytically and shows how, under the stated assumptions, the median path of the company's *q*-ratio displays a periodic cycle punctuated by bubbles and crashes. Specifically, the median path is:

$$\left.\begin{aligned} \frac{dq_t}{q_t} &= \left(\pi - \frac{\sigma^2}{2}\rho_t\right)dt \\ \frac{d\rho_t}{\rho_t} &= \lambda(q_t - 1)dt \end{aligned}\right\} \quad \text{where} \quad \pi, \sigma^2, \lambda > 0 \text{ are positive parameters} \qquad (1)$$

Where 'q_t' represents Tobin's *q* ratio, 'ρ_t' is the ratio of producers' rents divided by investors' income, and 'π', 'σ^2' and 'λ' are parameters capturing, respectively, the median equity premium, the variance on equity returns and the speed of liquidation. Figure 1 illustrates how solvency and the rent ratio behave in such a model[12]:

Figure 1: Predator-prey dynamics in the model in Appendix 2 (median path)



[12] Note that, to develop this diagram, the specific values of the parameters 'π', 'σ^2' and 'λ' have not been selected to be realistic but merely to make the shape more evident to the viewer.

Five comments are worth making at this point:

a. In this model, the parameters 'π', 'σ^2' and 'λ' determine the frequency of the cycle, so that its wavelength is longer the larger they are, and becomes infinite if any of them is zero[13].

b. Regular as these waves appear, they are not deterministic. What makes them periodic is that, as the relative weight of rents climbs up, it progressively takes a smaller negative shock to trigger a solvency crisis, so that, at the point where its probability exceeds 50%, the median path starts a downturn.

c. The timing of bubbles and crashes would therefore be predictable to some extent (albeit never with certainty) both on the basis of timing since the last crash and of the behavior of key variables (e.g. Tobin's q or any related financial variable, such as CAPE, shooting above its long-term central value).

d. The fall of q as a result of the "crash" is always steeper than its increase during the preceding boom – which of course fits well the historical experience of financial bubbles (after all, they are called bubbles because they 'pop').

e. There is no need that the market price depart from the net present value of future cash flows, as in Blanchard & Watson's (1982) rational bubbles: here, the mean path may be perfectly consistent with its fundamentals, yet the periodic bubbles and crashes depicted in Figure 1 would appear all the same.

Theoretical models are not of much use, though, without empirical confirmation. In the next section we therefore test whether this model's main predictions hold.

[13] Although, incidentally, if any of them is infinite then the cycle also disappears by collapsing to the trivial solution $q_t = \rho_t = 0$.

4. Testing the Model

Separating the value of producers' rents from the marginal productivity of their services is inherently tricky, if nothing else because it is precisely their privileged knowledge that allows them to extract them from the investors, so they must necessarily be difficult for an external observer to detect. Tobin's q is, on the other hand, a metric that is not too hard to calculate through a fairly standard calculation method. Several authors (e.g. Wright 2004) argue convincingly that the value obtained through these methods may be downward-biased due to systematic overvaluation of asset values at replacement cost (more or less in the same way that a used car is not worth the same as a new one, and therefore its liquidation value is below replacement cost). Yet, although this bias would certainly impact a test of whether Tobin's q converges to unity, it poses no particular challenge if what we want to test is whether it does indeed display a cyclical behavior.

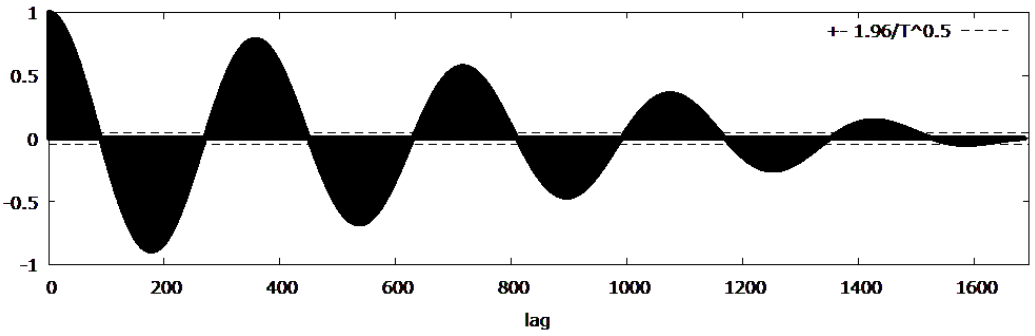
The key challenge is in finding a data time series that is granular enough yet covers a period long enough to actually test this. Indeed, however one may want to estimate them, realistic values for the parameters in expression (1) are always small (equity premiums and average annual variances are typically single-digit, and the process of liquidating or reorganizing the management of an underperforming company can also be quite time-consuming), which means cycles may well stretch over many years. If we want to test whether such cycles do actually exist, therefore, we need a sample that covers the span of the underlying cycle several times over so that it can provide a reasonable degree of confidence to accept or reject the hypothesis.

In addition, there is a question on what test to perform. To avoid any bias towards acceptance of the hypothesis, we should try to avoid any test that somehow forces a

positive answer: if nothing else, we should remind ourselves that, per Fourier's theorem, any continuous function can be decomposed into an aggregation of sinusoidal curves, so "fitting" a cyclical function to the sample, or decomposing it as a spectrum, does not necessarily prove the point either. This is also why filters and calibrations are dangerous: as it is not always clear whether the cycle actually exists in the underlying data or it is a result of the interaction between the sample and the filter.

To avoid these pitfalls, this paper will resort to a simple autocorrelation test on the unfiltered data, and demand that its correlogram follow the pattern in Figure 2:

Figure 2: Theoretical autocorrelation diagram for the *q*-ratio in Appendix 2



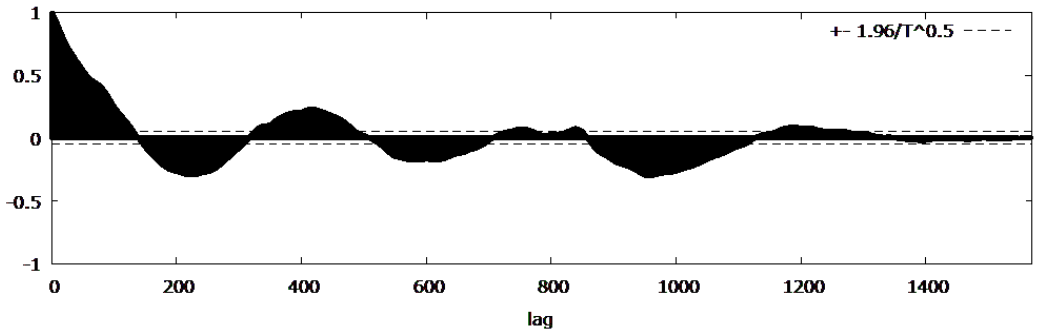
In other words, the autocorrelation diagram should display:

a. Positive and negative autocorrelation intervals alternating over time
b. Regular time lags between positive and negative autocorrelation intervals
c. Symmetric autocorrelation waves[14]
d. Decreasing autocorrelation waves as they become more distant in time (due to the increasing paucity of data as the autocorrelation lag lengthens)

[14] This is the case even if the underlying cycle is not symmetric (as is the case of the predator-prey model). The reason, intuitively, is that the autocorrelation diagram reflects both the correlation of low and high values today, so even if, say, crashes are quicker than booms, the lags compensate each other and thus the result is a symmetric wave on the correlogram.

The most complete dataset available for Tobin's *q* is perhaps Wright (2004), which spans from 1900 to 2002; Wright put forward several computations[15] of which we will take here the one he calls "equity *q*". Unfortunately an annual series such as this is not granular enough to test long-wave autocorrelations, so we resort to Shiller's Cyclically Adjusted Price Earnings (CAPE) ratio, which has a correlation coefficient of 85% with the equity *q* but provides a long monthly series[16], as a proxy. The CAPE (just a standard P/E where last 10 years' earnings average is taken as the denominator to eliminate the impact of short-term fluctuations) has proven to be a very powerful tool to identify market bubbles, as was in fact the basis of Shiller's successful prediction of both the 2008 subprime and the 2000 dot.com crashes. It is thus particularly intriguing to test whether the CAPE contains a cyclical component both due to its link to stock market bubbles and because, to the extent its behavior can be explained through a rational expectations model, it would not be necessary to resort to bounded rationality or any other form of market inefficiency as an explanatory hypothesis. In short, Shiller's CAPE yields the following correlogram (Figure 3):

Figure 3: CAPE autocorrelation diagram (monthly, Jan. 1881 to Oct. 2011)



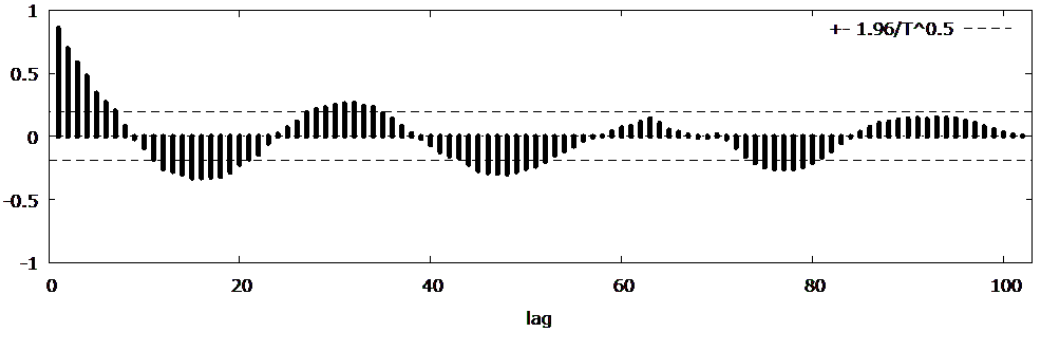
[15] The data can be downloaded from http://www.econ.bbk.ac.uk/faculty/wright/.

[16] The data are regularly updated in Shiller's website, http://www.econ.yale.edu/~shiller/data.htm. The high correlation with Tobin's *q* has been repeatedly highlighted before even in non-academic circles e.g. the Smithers & Co. website http://www.smithers.co.uk/page.php?id=34.

Allowing for the usual dampening impact of random noise, this result is entirely consistent with the theoretical pattern depicted in Figure 3. Moreover, the autocorrelations are fairly high (up to ± 0.3 i.e. 30% autocorrelation coefficient) and exceed comfortably the normal 95% statistical significance threshold. The wave this evidence highlights, incidentally, has a very low frequency: its average wavelength is in fact just over 30 years. This, as mentioned above, is consistent with the fact that the observed values of the key parameters in Appendix 2 (particularly the equity premium π and the variance on equity returns σ^2) are typically single-digit percentages, which is bound to lead to a slow-motion cycle in this model.

One could of course argue that, even though we know CAPE and the equity *q* are so closely correlated in the available samples, we do not know yet whether this cycle would be exclusive of CAPE and absent from the equity *q* itself. The simplest answer to this is of course to run the same autocorrelation test on the *q* ratio (Figure 4):

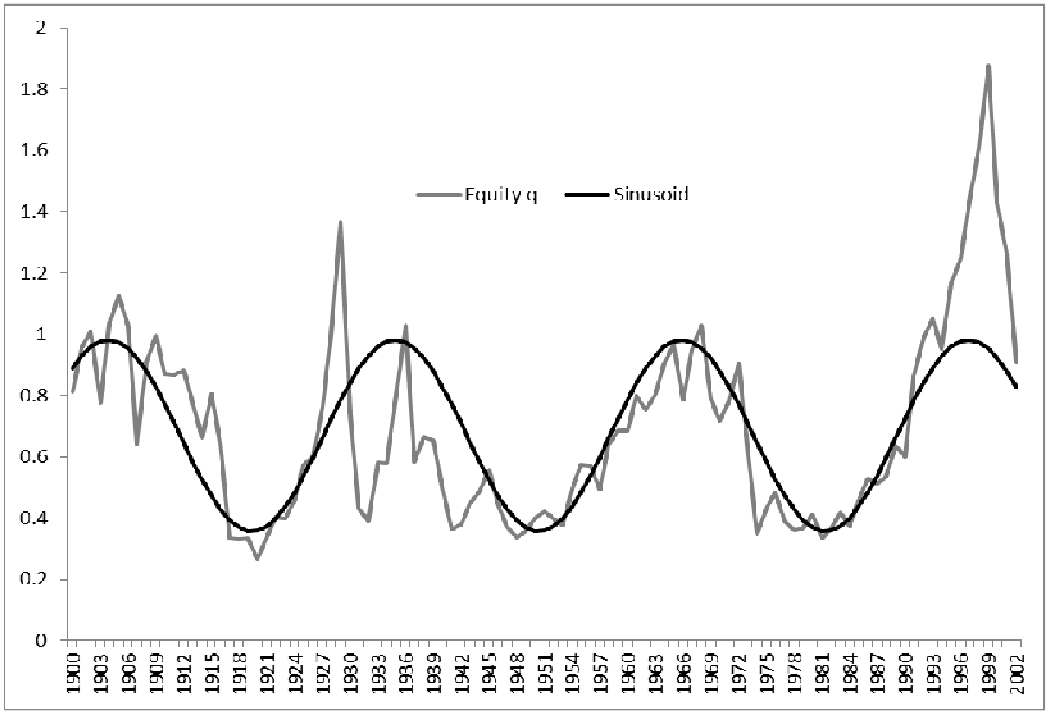
Figure 4: Equity *q* autocorrelation diagram (annual, 1900 to 2002)



Not only does the cyclical pattern reappear in this dataset, but it also matches the average wavelength of just over 30 years in the CAPE diagram, as well as the maximum autocorrelation level of *circa* $\pm 30\%$ … only, due of course to the paucity of data we discussed above, in this case the threshold for 95% confidence is very high

and therefore the correlation wave exceeds it less comfortably. Furthermore, this long-wave cycle has quite a high degree of explanatory power: in fact, a simple sinusoidal curve with a 31-year wavelength (i.e. the cycle time suggested by Figures 3 and 4) has a 72% correlation coefficient respective to the equity *q* series[17] – or, what is the same, can be said to explain more than half ($R^2 = 52\%$) of its observed variation (Figure 5):

Figure 5: Equity *q* fitted with 31-year-cycle sinusoidal wave



This poses another question: although we know the correlogram yields symmetric waves even against an asymmetric wave, if a simple sinusoidal curve already has such a high explanatory power, how do we know the underlying waves are not sinusoidal anyway? After all, with a very small change in the basic assumptions of Appendix 2 one could have turned the model's median path into a sinusoidal (i.e. symmetric)

[17] Note, importantly, that the sinusoidal curve is NOT used here as a device to 'prove' the existence of a cycle (this was the purpose of the autocorrelation diagrams in Figures 3 and 4), but simply to estimate how much of the equity *q* path could be explained through it.

wave[18]. The evidence, however, suggests that the actual path behaves just as in a predator-prey wave, where the fall is steeper (and therefore also shorter in time) than the climb-up. Indeed, according to Wright's time series there was an increase of the equity *q* in 57% of the years in the sample and a fall in only 43% of them. In fact, even if we restrict this calculation to the period 1900 to 1993 (as the tracking sinusoidal wave appears to reach in 1993 the same point the cycle as in 1900), the average of climb and fall years still holds: 57% upwards and 43% downwards.

In sum, *the evidence supports the hypothesis that the bubbles and crashes observed over at least the last 100+ years could largely be explained through an efficient-markets, predator-prey cycle model.* The underlying model itself might not necessarily be exactly the one put forward in this paper (to begin with, a model such as Gracia 2005 would behave in a similar way), but the hypothesis that a long-wave cycle with steeper slopes downwards that upwards (i.e. with a predator-prey shape) explains a substantial part of the observed equity *q* behavior cannot be rejected.

Two side comments are probably appropriate at this point:

a. Although this long cycle is consistent with the model in this paper (for the equity premium and return variance values are usually single-digit over long periods of time), it is worth mentioning once again that its identification is not model-dependent. The finding of any cycle under these conditions is hence all the more robust, as it was not "imposed" on the data through any filtering or regression that could make something visible when it actually does not exist.

[18] Specifically, it would suffice to replace the expression $d\rho_t = \lambda(q_t - 1)\rho_t dt$ in Assumption 7 with the expression $d\rho_t = \lambda(\ln q_t)dt$ to turn the median path into a pure sinusoidal wave.

b. The wavelength of just over 30 years is, on the other hand, quite different from that of most classical studies: it certainly does not match the classical waves identified by Kitchin (3-5 years), Juglar (7-11 years), Kuznets (15-25 years) or Kondratiev (45-60 years). Although the original purpose of this paper was not to test, let alone to impose, any particular cycle wavelength, it is of course legitimate to wonder why this might be the case. Whether this 31-year cycle might result from the summation of others or represent something entirely different is, however, not a question we can answer on the basis of the evidence above and will therefore remain out of this paper's scope.

5. Considerations on Aggregate Productivity Impact

The next question is how these oscillatory financial phenomena could, under rational expectations, have an impact not only on GDP growth but also on TFP (which, per Kydland & Prescott 1992, typically explains more than 70% of the business cycle). The standard Cobb-Douglas production function does not lend itself very well to this (unless modified through ancillary assumptions, of course), as its TFP growth rate is exogenous *ex hypothesi*. The answer is, conversely, quite straightforward on the basis of the aggregate production function put forward in Gracia (2011). As this function has been empirically proven to be a better specification of aggregate U.S. GDP behavior (and the Cobb-Douglas function was in fact rejected as a valid specification in a non-nested model comparison test), it is entirely legitimate to resort here to it

Both the analytical development and an extensive discussion of the rationale for this aggregate function are available online[19], so here we will only outline its intuitive logic. The key difference respective to the Cobb-Douglas function is that Gracia

[19] This can be downloaded at http://www.economics-ejournal.org/economics/journalarticles/2011-19.

(2011) does not assume perfect competition and therefore allows for the existence of economic rents[20]. Economic rents may result from any form of resource supply constraint: physical availability (Ricardian rents), control over resources (monopoly rents), information asymmetry (agency rents), etc. By definition, they constitute fixed costs (or fixed assets) above marginal costs (or income), so they cannot be instantly flexible to unexpected changes in demand volumes: one cannot change the capacity of a plant from one day to the next. A positive demand shock, therefore, will face higher unit costs (i.e. lower productivity) as extra production is subject to capacity constraints, whereas a negative shock will not reduce costs (i.e. increase productivity) in the same proportion due to the impact of fixed costs. Hence, if we consider a closed Walrasian economy where Say's Law applies so that an increase in one product's demand must equal a decrease in another product's, any unexpected change in the composition of demand, swapping demand from one product to another, will result in a fall of productivity. This fall will be more severe the higher the weight of rents over total income (for rents represent the economic translation of capacity constraints) and the higher the variability of demand composition. In other words: rents are a measure of the productive system's rigidity so, for a given level of demand uncertainty, the higher this rigidity, the lower the system's productivity will be.

This links up directly with the rent-driven cycle model we developed in Section 3 and Appendix 2. In the good times, producers (and, in general, anyone with more or less control of the means of production) gradually find ways to increase their rent extraction, which of course means they impose additional rigidity on the productive process. In the initial stages, the weight of these rents over the rest of the income

[20] By introducing ancillary it is of course possible to develop (as the literature does extensively) a Cobb-Douglas-based model where economic rents exist; yet, per Gracia (2011), this raises methodological questions related to the aggregate production function's microfoundations.

produced (i.e. the producers' rent ratio ρ_t) does not grow very quickly, so its impact on productivity is also small. Yet, as Figure 1 shows, when the bubble finally bursts and Tobin's q starts its downward spiral, the rent ratio shoots up because, as the model assumes prices adapt instantly to the changing conditions, their variable portion (i.e. the marginal cost) reduces faster than the fixed costs (that is, the economic rents). Hence, as the rent ratio goes up, other things being equal, productivity must go down.

In short, the prediction is that financial recessions will impact output growth through the fall in aggregate productivity they cause. We can now resort to the production function in Gracia (2011) to provide a visual image of how this would happen. Analytically, the function is defined as follows:

$$\overline{Y}_t = \overline{A}_t \overline{H}_t^{\frac{1}{1+\overline{\rho}_t}} \tag{2}$$

Where $\overline{Y}_t$ represents aggregate GDP, $\overline{H}_t$ aggregate working hours, $\overline{A}_t$ the productivity factor and $\overline{\rho}_t$ the aggregate average rent ratio[21], and where, in turn, the productivity factor is such that:

$$\frac{d\overline{A}_t}{\overline{A}_t} = \left(\overline{\Gamma}_t + \frac{\overline{\sigma}_t^2 - \overline{\delta}_t}{2} \overline{\rho}_t \right) dt + \overline{s}_t d\overline{W}_t \tag{3}$$

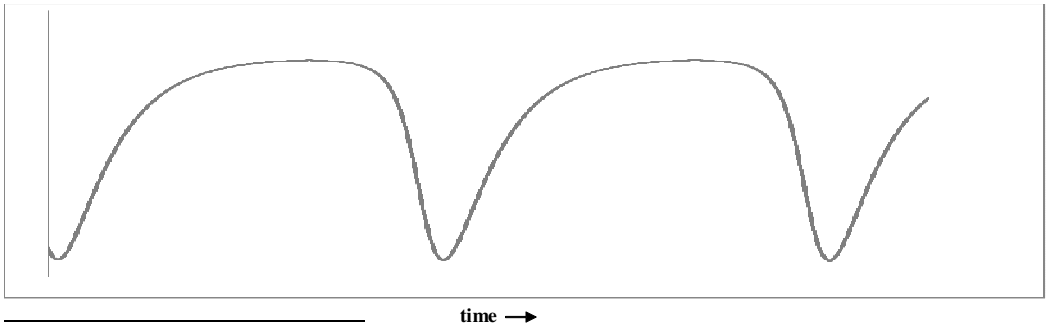
Where $\overline{\Gamma}_t dt + \overline{s}_t d\overline{W}_t$ is the technology-driven component of productivity growth (which, as in the standard literature, is broken down into a deterministic growth rate $\overline{\Gamma}_t$ plus a serially-uncorrelated Gaussian perturbation component $d\overline{W}_t$ whose standard

[21] This ratio includes not only producers' rents but also any other type of economic rents in the overall economy: in fact, Gracia (2011) tested the model against another type of rent (namely, the risk-free interest rate net of monetary dilution) which was easier to measure. For convenience, in the remainder of this section we will assume all the rents other than producers' rents to be constant, but of course in a more sophisticated analysis such an assumption logically ought to be relaxed.

deviation is $\bar{s}_t$), $\bar{\sigma}_t^2$ is the non-technology-driven portion of overall demand growth variance and $\bar{\delta}_t$ is the weighted sum of the individual product demand variances[22].

Hence, for example, if this were a closed Walrasian economy with a known resources endowment, overall demand would only vary with productivity shocks, which means that $\bar{\sigma}_t^2 = 0$, whereas the weighted sum $\bar{\delta}_t$ would be positive, as it can only be nil if if each and every product demand is deterministic. This means that, at least in a closed economy (as would be the whole world, or an economy whose foreign sector is proportionally small, such as the USA or the EU) the factor $\frac{\bar{\sigma}_t^2 - \bar{\delta}_t}{2}$ should have a negative sign so that (as we had anticipated above), the higher the rent ratio $\bar{\rho}_t$, the lower the rate of growth of productivity. Hence, if we assume technology growth as well as the difference $\left(\bar{\sigma}_t^2 - \bar{\delta}_t\right)$ to be constant, then the median rent we depicted in Figure 1 translates into the following path of aggregate productivity (Figure 6):

Figure 6: Productivity Growth Rate over the Median Path (illustrative)



[22] Technically it is defined as an aggregate $\bar{\delta}_t \equiv \sum_{j=1}^{m} \alpha_{j,t} \delta_{j,t}$ of all the products / production units $j \in \{1...m\}$, where $\alpha_{j,t}$ represents the relative share of output of every individual production unit and $\delta_{j,t}$ is defined as $\delta_{j,t} \equiv \sigma_{j,t}^2 \frac{\alpha_{j,t}}{\beta_{j,t}}$, and where, in turn, $\sigma_{j,t}^2$ is the portion of the demand variance of each product / production unit *j* that is not driven by a technology shock and $\beta_{j,t}$ the relative share labor input corresponding to each one of those products / production units.

In other words, productivity must behave over time just like the rent ratio curve in Figure 1 (only upside down) i.e. displaying a cycle with falls that are steeper than the previous or subsequent climb up periods, and deeper respective to the trend than the booms stand above it. This behavior is consistent with the findings of Neftçi (1984), Sichel (1993), Ramsey & Rothman (1996), Verbrugge (1997) or Razzak (2001), who conclude that business cycles present "deepness" (i.e. recessions tend to fall deeper than expansions are tall respective to the trend) and "steepness" (i.e. the fall into recession is steeper than the climb up back to expansion).

One is reminded at this point of Simon Johnson's views on crises in emerging economies, which he largely based on his experience as chief economist of the IMF (e.g. Johnson 2009 or Johnson & Kwak 2010). Johnson explains that, despite the wide diversity of their triggering events, economic crises always look depressingly similar, because they all result from powerful, privileged elites overreaching in good times to maximize their rents, but resisting the pressure to cut back on them when their excessive risk taking results in a credit crisis. Johnson also makes a very strong case that the U.S. 2008 credit crisis presents exactly the same profile, with the U.S. financial sector playing the role of the privileged elite. This is of course, to a large extent, the sort of mechanism portrayed in Appendix 2 as well as in Gracia (2005).

Furthermore, these crises are usually associated to low productivity growth rates during as well as, to a lesser extent, in the years immediately before the recession – just as Figure 1 would suggest. This link is so well established that the observation of low productivity growth among the so-called East Asian "tiger" economies during their expansion period until the '90s (e.g. Young 1992, 1994 & 1995, or Kim & Lau 1994) provided the rationale for Krugman's (1994) prediction that their expansion

would come to an abrupt halt – as it did indeed in the 1997-98 financial crisis. When the crisis hit, its root cause was found in the "crony capitalism" that politically well-connected elites practiced in those countries to maximize their rent extraction.

More recently, the same failure to grow productivity during a boom, leading to loss of competitiveness and hence excessive debt accumulation to finance a large current account deficit, has been highlighted, e.g. in a recent World Bank special report (Gill & Raiser 2012), as the driver of the financial crisis' severity in Southern European Eurozone countries (Greece, Italy, Portugal and Spain). This report finds, for instance, that the low *quality* of government services (low rule of law, low government effectiveness, low control of corruption) correlates strongly with these countries' sluggish productivity growth. Such institutional failures, of course, translate into rent extraction opportunities for various power groups: wasteful public investments to benefit well-connected construction moguls, abnormally high salaries and perks for public servants, rigid labor laws, guild privileges and heavy business regulation leading many companies outside the government-favored circle to prefer staying small (even at the expense of productivity) to avoid some of that burden.

6. Conclusions and final remarks

The **first objective** of this paper was to highlight a serious fallacy (what we called the **Fallacy of Probability Diffusion Symmetry**) that underpins much of the debate around the Rational Expectations Hypothesis by leading to the erroneous implication that, if markets are efficient, then regular, predictable bubbles and crashes cannot exist. This implication, it has been shown, is only valid for symmetric probability diffusion processes, which are rarely, if ever, posed in modern finance models – although the fallacy itself all-too-frequently goes unacknowledged.

The **second objective** was to develop a **rational-expectations model of periodic financial bubbles driven by an agency conflict between producers and investors** along the lines of Gracia (2005), test its likelihood against U.S. stock market data, and then extend it to macroeconomic business cycles by combining it with the aggregate production function put forward in Gracia (2011).

Both have modeling implications. If the rational expectations debate is vitiated by the fallacy of diffusion symmetry, it may be unnecessary to resort to bounded rationality or price stickiness to explain many of the phenomena put forward as evidence of market inefficiency. On the contrary, it may make sense to review some classical rational expectations propositions (e.g. money and debt neutrality, Ricardian equivalence, even perfect market clearing) to assess under what conditions they would still hold along the median path. Ultimately, the mean path be as it may, if the median trajectory is the one with the highest likelihood to match reality, then it arguably makes sense to target it instead of the mean as the primary objective of economic policy – which might, in turn, vindicate a number of macroeconomic policies that a simplistic interpretation of efficient markets may have seemed to invalidate.

Rational expectations is a hypothesis with high explanatory power but, by ignoring the fallacy of diffusion symmetry, its implications may have been oversold beyond what the theory granted and the observations supported. Neither does market efficiency necessarily entail random walk observations nor does their absence imply that rationality is somehow 'bounded'. Recognizing this may lead to reviewing some popularly-accepted corollaries of market efficiency – yet, paradoxically, the result may also vindicate rational expectations above any bounded rationality alternatives.

*****

APPENDICES

APPENDIX 1

The purpose of this appendix is to show how an empirical analysis of a time series generated by a geometric Gauss-Wiener diffusion process (i.e. a geometric Brownian motion) should be expected to yield an observed growth rate closer to the median, not the mean, path of the distribution.

Consider an asset P_t whose market value follows a geometric Brownian motion i.e.:

$$\frac{dP_t}{P_t} = \mu dt + \sigma dZ_t \tag{1.1}$$

Where μ, σ are constants and Z_t is a linear Brownian motion $dZ_t = \varepsilon_t \sqrt{dt}$ such that $Z_0 \equiv 0$, and the white noise variable ε_t is normally-distributed, i.e., $\varepsilon_t \sim \mathrm{N}[0,1]$. Then, of course, the expected (or "mean") growth rate of the asset's market value is:

$$\mathrm{E}_0\left[\frac{dP_t}{P_t dt}\right] = \mu + \sigma \mathrm{E}_0\left[\frac{dZ_t}{dt}\right] = \mu \tag{1.2}$$

We can now resort to Itô's lemma to obtain the general expression of P_t:

$$d(\ln P_t) = \frac{dP_t}{P_t} - \frac{1}{2}\left(\frac{dP_t}{P_t}\right)^2 = \frac{dP_t}{P_t} - \frac{(\mu dt + \sigma dZ_t)^2}{2} = \frac{dP_t}{P_t} - \frac{\sigma^2}{2} dt$$

$$\frac{dP_t}{P_t} = d(\ln P_t) + \frac{\sigma^2}{2} dt = \mu dt + \sigma dZ_t$$

$$d(\ln P_t) = \mu dt - \frac{\sigma^2}{2} dt + \sigma dZ_t \Leftrightarrow \ln P_t = \left(\mu - \frac{\sigma^2}{2} \right) t + \sigma Z_t + \text{constant}$$

$$P_t = P_0 e^{\left(\mu - \frac{\sigma^2}{2} \right) t + \sigma Z_t} \tag{1.3}$$

Since Z_t is normally-distributed and therefore symmetrical, the diffusion process of P_t must be asymmetrical, and the growth rate of its median path (i.e., the path that would cut across the distribution leaving 50% of the probability on each side) must be $\mu - \frac{\sigma^2}{2}$ i.e. different from that of the mean path.

If we now had a sufficiently long time series of empirical observations of P_t, we could calculate its average growth rate through a logarithmic regression under the following specification:

$$\ln P_t = a + bt + u_t \tag{1.4}$$

Where *a* and *b* represent the regression parameters and u_t the series of residuals. It is therefore immediate that, assuming the data series is long and representative enough, the results of this regression should be expected to approximate the following:

$$\left.\begin{aligned} a &= \ln P_0 \\ b &= \mu - \frac{\sigma^2}{2} \\ u_t &= \sigma Z_t \end{aligned}\right\} \tag{1.5}$$

In other words, the result of the empirical analysis should be expected to be a growth rate equal to that of the *median*, not the *mean*.

Furthermore, the result would be the same if, instead of estimating the parameters of a logarithmic regression, we calculated the average continuous growth rate along the time series. Indeed: if, given a sample covering the time interval $t \in [0,T]$, the average continuous growth rate g is defined as a magnitude such that:

$$e^{gT} \equiv \frac{P_T}{P_0} \Leftrightarrow g \equiv \frac{\ln P_T - \ln P_0}{T} \tag{1.6}$$

Then, in our case, it yields the following expected result:

$$g \equiv \frac{\ln P_T - \ln P_0}{T} = \frac{\left(\mu - \frac{\sigma^2}{2}\right)T + \sigma Z_T}{T}$$

$$\mathrm{E}_0[g] = \left(\mu - \frac{\sigma^2}{2}\right) + \frac{\sigma}{T}\mathrm{E}_0[Z_T] = \left(\mu - \frac{\sigma^2}{2}\right) \tag{1.7}$$

Which is, once again, the *median* growth rate of the distribution, not the *mean*.

In sum, the behavior of its observed path may, on average, be very different from this expected equilibrium without this posing any challenge to the efficient markets hypothesis.

Q.E.D.

APPENDIX 2

The purpose of this appendix is to develop in analytical form the reasoning that in Section 3 of the main text was presented in an intuitive, discursive way.

Definitions and Assumptions

Consider a firm whose total market value at instant *t* is K_t, and that holds productive assets whose replacement cost at market prices would be $\tilde{K}_t$. The enterprise has two types of stakeholders: investors (i.e. principals) who hold ownership of the assets, and producers (i.e. agents) who control and manage those assets on their behalf. Hence, the net value added Y_t the entity generates (i.e. the sum of the return it generates for labor and capital or, what is the same, the difference between its sales revenue and the market price of its non-labor inputs) may be broken down into four portions:

1. The marginal cost of labor $m_t L_t$,
2. The agency rents R_t extracted by producers,
3. The cash flow C_t paid to investors (negative if the firm is raising capital) and
4. The remaining value added, which is reinvested as retained earnings.

Net value added is therefore defined as follows:

$$Y_t \equiv m_t L_t + R_t + C_t + \frac{dK_t}{dt} \qquad (2.1)$$

Of course this means the investors' profit ('Π_t') is equal to $\Pi_t \equiv C_t + \frac{dK_t}{dt}$.

We now designate as K_t^* the capitalized value of the producers' control of the productive process i.e. of the "asset" represented by their ability to extract rents for themselves[23]: Therefore, the firm's overall rate of return r_t is:

$$r_t \equiv \frac{Y_t - m_t L_t}{K_t + K_t^*} \equiv \frac{\Pi_t + R_t}{K_t + K_t^*} \tag{2.2}$$

We also represent as $\tilde{r}_t$ the cost of opportunity of the assets $\tilde{K}_t$ i.e. the return they would yield if invested outside the structure of the company in a risk-free form – e.g. lending them for a fixed rent. In this context, $\tilde{C}_t$ represents the cash investors would extract from those assets while invested at a risk-free rate i.e.:

$$\tilde{C}_t \equiv \tilde{r}_t \tilde{K}_t - \frac{d\tilde{K}_t}{dt} \tag{2.3}$$

Finally, we define Tobin's q ratio as the market value of the company weighted by the replacement market value of its assets:

$$q_t \equiv \frac{K_t}{\tilde{K}_t} \tag{2.4}$$

And the producer rent ratio ρ_t as the value of the producers' rents divided by the value remaining for the investors in the company:

$$\rho_t \equiv \frac{K_t^*}{K_t} \tag{2.5}$$

Now we introduce the following assumptions:

[23] The asset that would be represented by the marginal cost of labour $m_t L_t$ has no financial value in this framework, as its cost would always equal its income.

1. **Efficient Market Valuation:**

The company's rate of return r_t equals the market rate for assets of equivalent risk exposure, so its equity market valuation is such that, $\forall t \geq T$:

$$\mathrm{E}_T\left[r_t K_t\right] = \mathrm{E}_T\left[C_t + \frac{dK_t}{dt}\right] \tag{2.6}$$

Where $\mathrm{E}_T[\circ]$ indicates the mean value per the information available at instant *T*.

2. **Wiener Perturbation**

The company's rate of return follows a Wiener diffusion process with drift, i.e.:

$$r_t dt = \mu dt + \sigma dZ_t \tag{2.7}$$

Where $\mu, \sigma > 0$ are positive parameters, Z_t is a Brownian motion $dZ_t = \varepsilon_t \sqrt{dt}$ such that $Z_0 \equiv 0$ and the white noise ε_t is normally-distributed, i.e., $\varepsilon_t \sim \mathrm{N}[0,1]$.

3. **Constant Risk-Free Rate:**

The risk-free rate at which assets could be lent/borrowed is constant i.e.:

$$\tilde{r}_t = \tilde{r} \qquad (\text{ where } \tilde{r} \to \text{constant }) \tag{2.8}$$

4. **Observable market valuations:**

The market values K_t, $\tilde{K}_t$ and K_t^* are observable at the instant *t* where they take place, i.e., $K_t, \tilde{K}_t, K_t^* \in I_t$ (where I_t is the set of information available) i.e. $\forall t \geq T$:

$$\lim_{T\to t} \mathrm{E}_T[K_t] = K_t \qquad \text{and} \qquad \lim_{T\to t} \mathrm{E}_T[\tilde{K}_t] = \tilde{K}_t \qquad \text{and} \qquad \lim_{T\to t} \mathrm{E}_T[K_t^*] = K_t^* \qquad (2.9)$$

Comment: This is quite intuitive, as these are all stock variables i.e. they are defined as values at a point in time (instead of flows between a point in time and the next).

5. Observable cash flows:

The cash flows paid to investors (C_t and $\tilde{C}_t$) and to producers (R_t) are known at the instant t in which they take place i.e. $C_t, R_t \in I_t$ (where I_t represents the set of information available at time *t*) or, what is the same, $\forall t \geq T$:

$$\lim_{T\to t} \mathrm{E}_T[C_t] = C_t \qquad \text{and} \qquad \lim_{T\to t} \mathrm{E}_T[\tilde{C}_t] = \tilde{C}_t \qquad \text{and} \qquad \lim_{T\to t} \mathrm{E}_T[R_t] = R_t \qquad (2.10)$$

Comment: In other words, the risk of the return being different from expected at any given point in time is borne by the retained earnings $\frac{dK_t}{dt}$.

6. Consistent Preference for Cash:

The investors' willingness to sweat cash from their investments is the same for the company K_t and for the assets $\tilde{K}_t$ i.e.:

$$\frac{C_t}{K_t} = \frac{\tilde{C}_t}{\tilde{K}_t} \qquad (2.11)$$

Comment: There is quite a wide range of utility functions that would produce this result. As an example, Appendix 3 shows its derivation from a standard functional form (a time-additive discounted expected utility function with unity time elasticity).

7. Linear Investors' Controls Function:

The degree of investors' control on the producers' ability to increase their rent extraction over time follows a linear function dependent on the investors' relative gain or loss of value respective to investing outside the company i.e.:

$$\frac{d\rho_t}{\rho_t} = \lambda(q_t - 1)dt \qquad (2.12)$$

Where λ represents a positive constant.

Comment: This postulate combines three intuitive ideas:

a. Liquidation is not an instant process (hence λ is assumed finite) but, the more investors find they are losing by not liquidating (i.e. the smaller Tobin's q), the more companies will be reorganized or liquidated to cut down their rents.

b. On the flip side, of course, the larger Tobin's q the least likely are investors to liquidate or to impose heavy controls, so more opportunities will pop up over time for producers to increase the rents they extract.

c. On balance, liquidations will dominate when $q_t < 1$ (i.e. when it is more profitable for investors to liquidate), whereas producers will have more room to expand their rents when $q_t > 1$.

Although the functional form in (2.x) has been chosen primarily because of its simplicity, it can be justified intuitively if we assume that both the probability of liquidation and that of increased rent extraction opportunities are distributed according to an exponential function.

Analytical Development

Combining Assumption 1 (i.e. expression 2.6) with Definition (2.1) we find, $\forall t \geq T$:

$$\mathrm{E}_T[r_t K_t] = \mathrm{E}_T[\Pi_t] \quad \Leftrightarrow \quad \mathrm{E}_T[r_t K_t^*] = \mathrm{E}_T[R_t] \tag{2.13}$$

Which, per Assumptions 4 and 5 (i.e. expressions 2.9 and 2.10), becomes, for the special case $T = t$:

$$\mathrm{E}_t[r_t] K_t^* = R_t \tag{2.14}$$

If we now combine Definitions (2.1), (2.2) and (2.5):

$$R_t + C_t + \frac{dK_t}{dt} \equiv r_t\left(K_t + K_t^*\right) \equiv r_t(1+\rho_t)K_t \tag{2.15}$$

This, combined with expression (2.14), becomes:

$$\mathrm{E}_t[r_t]\rho_t K_t + C_t + \frac{dK_t}{dt} = r_t(1+\rho_t)K_t$$

$$\frac{dK_t}{dt} = r_t K_t - C_t + (r_t - \mathrm{E}_t[r_t])\rho_t K_t \tag{2.16}$$

By simple inspection, we can see that, along the expected path, the impact of q_t will be fully discounted out, for, if we write the expected value of (2.16) at point *t* and then apply Assumption 2 (i.e. expression 2.7), we obtain:

$$\mathrm{E}_t\left[\frac{dK_t}{dt}\right] = \mathrm{E}_t[r_t K_t - C_t] + \underbrace{(\mathrm{E}_t[r_t] - \mathrm{E}_t[r_t])}_{=0}\rho_t K_t$$

$$\mathrm{E}_t\left[\frac{dK_t}{dt}\right] = \mathrm{E}_t[r_t]K_t - C_t = \mu K_t - C_t \tag{2.17}$$

Which, after integration, yields the familiar Net Present Value formula i.e., $\forall t \geq 0$:

$$\mathrm{E}_0[K_t] = Ae^{\bar{r}t} + \int_0^{\infty} \mathrm{E}_0[C_t]e^{-\mu t}dt \tag{2.18}$$

Where *A* represents an integration constant.

This, to be sure, does not imply that the size of agency rents has no impact on the asset value, but simply that, in an efficient market, their expected impact has already been discounted from the asset market value at instant $t = 0$ and therefore, as long as the observed path matches the initial expectations, no further adjustment is necessary.

The *median* path, conversely, can be derived from expression (2.16) by applying the rule we developed in Appendix 1 i.e. since the median return is $\left(\mu - \frac{\sigma^2}{2}\right)$ then the median path of expression (2.16) is as follows:

$$\mathrm{Median}_t\left[\frac{dK_t}{dt}\right] = \left(\mu - \frac{\sigma^2}{2}\right)K_t - C_t - \frac{\sigma^2}{2}\rho_t K_t \tag{2.19}$$

At the same time, by combining Definition (2.3) with Assumptions 3, 4 and 5 (i.e. with expressions 2.8, 2.9 and 2.10) we obtain the (deterministic) path of $\tilde{K}_t$ i.e.:

$$\frac{d\tilde{K}_t}{dt} = \tilde{r}\tilde{K}_t - \tilde{C}_t \tag{2.20}$$

Hence, if we differentiate Definition (2.4) according to Itô's lemma we obtain that:

$$\frac{dq_t}{q_t} = \frac{dK_t}{K_t} - \frac{d\tilde{K}_t}{\tilde{K}_t} + \underbrace{\frac{1}{2}\left(\frac{d\tilde{K}_t}{\tilde{K}_t}\right)^2}_{=0} \tag{2.21}$$

Whose median path is, applying Assumption 6 (i.e. expression 2.11):

$$\text{Median}_t\left[\frac{dq_t}{q_t}\right] = \left(\mu - \frac{\sigma^2}{2} - \tilde{r}\right)dt - \frac{\sigma^2}{2}\rho_t dt - \underbrace{\left(\frac{C_t}{K_t} - \frac{\tilde{C}_t}{\tilde{K}_t}\right)}_{=0}dt$$

$$\text{Median}_t\left[\frac{dq_t}{q_t}\right] = \left(\mu - \frac{\sigma^2}{2} - \tilde{r}\right)dt - \frac{\sigma^2}{2}\rho_t dt \tag{2.22}$$

For simplicity, we will designate by π_t the equity premium, which, for the median path in expression (2.22), will be the constant value $\pi = \mu - \frac{\sigma^2}{2} - \tilde{r}$.

On this basis it is now possible to close the dynamic system representing the median path by combining expression (2.22) with (2.12) in the following final expression:

$$\left.\begin{aligned} \frac{dq_t}{q_t} &= \left(\pi - \frac{\sigma^2}{2}\rho_t\right)dt \\ \frac{d\rho_t}{\rho_t} &= \lambda(q_t - 1)dt \end{aligned}\right\} \tag{2.23}$$

Which is equal to expression (1) in Section 3.

Q.E.D.

Comment: Note that, if all three parameters $\pi, \sigma^2, \lambda > 0$ are all finite and positive, then expression (2.23) belongs to the family of Lotka-Volterra predator-prey dynamic

systems[24]. We know of course that the variance σ^2 is positive by definition (as it is the square of a real number) and, per Assumption 7, λ is positive *ex hypothesi* (a negative value would mean that investors are more likely to liquidate the higher the *q*-ratio, which makes no sense). The median risk premium π, conversely, could theoretically also be zero or even negative (if investors are risk averse, risky assets will offer a positive risk premium along the mean path, but not necessarily along the median). Nevertheless, at least in countries that have been both politically stable and financially sophisticated for a very long time, such as Britain or the USA, historical equity returns have been above low-risk interest rates more than 50% of the time. Hence, as long as the starting values of q_t and ρ_t are positive, we can say that under Assumptions 1 to 7, *if the median equity premium is positive, then the median path of Tobin's q ratio will follow a Lotka-Volterra predator-prey cycle* such as the one plotted in the simulation in Figure 1 in the main text.

[24] Lotka-Volterra predator-prey dynamic systems were originally developed in the context of biological studies (Lotka 1925, Volterra 1926) analysing the evolution of predator and prey populations in an ecosystem (hence their name). In economics, their best known instance of usage of a predator-prey process in economic modelling is of course Goodwin (1967).

APPENDIX 3

The purpose of this appendix is to show how Assumption 6 in Appendix 2 can be derived from a standard representative consumer utility function within the parameters most usually applied in mainstream literature.

In the following example, we will assume that the representative consumer intends to maximize a von Neumann-Morgenstern time-additive discounted utility function with unity inter-temporal elasticity of substitution, i.e.:

$$\max \mathrm{E}_0\left[\int_0^{\infty}\left(\ln C_t\right)e^{-\beta t}dt\right] \tag{3.1}$$

Where $\beta > 0$ is constant.

This maximization is then subject to the budget constraint:

$$K_0 = \mathrm{E}_0\left[\int_0^{\infty} C_t \frac{M_t}{M_0}dt\right] \tag{3.2}$$

Where M_t represents a martingale such that:

$$\frac{dM_t}{M_t} \equiv -\left(\bar{r}dt + \sigma dZ_t\right) \Leftrightarrow M_t \equiv e^{-\left(\bar{r}-\frac{\sigma^2}{2}\right)t-\sigma Z_t} \tag{3.3}$$

Thus, the first-order condition for the resolution of this problem is, $\forall t \geq 0$:

$$\frac{\partial U\left(C_t\right)}{\partial C_t}e^{-\beta t} = \frac{e^{-\beta t}}{C_t} = \Lambda e^{-\left(\bar{r}-\frac{\sigma^2}{2}\right)t-\sigma Z_t}$$

$$\Lambda = C_t^{-1} e^{\left(\bar{r} - \frac{\sigma^2}{2} - \beta\right)t + \sigma Z_t} \tag{3.4}$$

Where Λ represents the Lagrange multiplier. As this applies $\forall t \geq 0$, then:

$$\frac{C_t}{C_0} = e^{\left(\bar{r} - \frac{\sigma^2}{2} - \beta\right)t + \sigma Z_t} \Leftrightarrow \frac{dC_t}{C_t} = (\bar{r} - \beta)dt + \sigma dZ_t \tag{3.5}$$

If we now use this to replace into the budget constraint (3.2) we obtain:

$$K_0 = \mathrm{E}_0\left[\int_0^\infty C_t \frac{M_t}{M_0} dt\right] = \mathrm{E}_0\left[\int_0^\infty C_0 e^{-\beta t} dt\right] = \frac{C_0}{\beta}$$

$$\frac{C_0}{K_0} = \beta \tag{3.6}$$

Hence, for any point in time *t* taken as a reference, under this utility function the ratio $\frac{C_t}{K_t}$ equals the constant β irrespective of the rate of return of the underlying asset, and therefore the ratio will apply all the same if the asset is $\tilde{K}_t$, hence:

$$\frac{C_t}{K_t} = \frac{\tilde{C}_t}{\tilde{K}_t} = \beta \tag{3.7}$$

Q.E.D.

BIBLIOGRAPHY

Bezemer, Dirk J. (2009) 'No One Saw It Coming': Understanding Financial Crisis through Accounting Models. *MPRA Paper 15892*, June

Blanchard, Olivier J. & Watson, Mark W. (1982) Bubbles, rational expectations, and financial markets. *Crisis: the Economic and Financial Structure*, ed. P. Wachtel. Lexington, MA: Lexington

Burns, Arthur F. & Mitchell, Wesley C. (1946) *Measuring Business Cycles*. National Bureau of Economic Research

Davidson, Paul (2009) Can future systemic financial risks be quantified? Ergodic vs nonergodic stochastic processes. *Brazilian Journal of Political Economy*, vol. 29, nº 4 (116), pp. 324-340, October-December

Fama, Eugene (1965) The Behavior of Stock Market Prices. *Journal of Business* 38, pp. 34–105

Fama, Eugene & French, Kenneth (1988) Permanent and Temporary Components of Stock Returns. *Journal of Political Economy*, Vol. 96, No. 2, pp. 243-273

Foldvary, Fred (1997) The Business Cycle: A Georgist-Austrian Synthesis. *American Journal of Economics and Sociology* 56(4), pp. 521-541

Gaffney, Mason (2011) An Award for Calling the Crash. *Econ Journal Watch*, Vol. 8, Number 2 (May), 185-192

Galí, Jordi (1996) Technology, Employment and the Business Cycle: Do Technology Shocks Explain Aggregate Fluctuations? *NBER Working Paper No. 5721*, NBER

Gill, Indermit S. & Raiser, Martin (2012) Golden growth: restoring the luster of the European economic model. *The World Bank*

Goodwin, Richard M. (1967) A Growth Cycle. *Socialism, Capitalism and Economic Growth: Essays Presented to Maurice Dobb* (Feinstein, Charles ed.), Cambridge University Press, pp. 54-58

Gracia, Eduard (2005) Predator-Prey: An Alternative Model of Stock Market Bubbles and the Business Cycle. *Intervention Journal of Economics*, Vol. 2, No. 2 (September), pp. 77-105

Gracia, Eduard (2011) Uncertainty and Capacity Constraints: Reconsidering the Aggregate Production Function. *Economics: The Open-Access, Open-Assessment E-Journal*, Vol. 5, 2011-19

Harrison, Fred (1983) *The power in the land: An inquiry into unemployment, the profits crisis, and land speculation*. Universe Books

Harrison, Fred (1997) The Coming 'Housing' Crash. *The Chaos Makers* (Frederic J. Jones & Fred Harrison), Othila Press

Hoyt, Homer (1933) *One Hundred Years of Land Values in Chicago*, Arno Press

Johnson, Simon (2009) The Quiet Coup *The Atlantic Online* (May)

Johnson, Simon & Kwak, James (2010) *13 Bankers: The Wall Street Takeover and the Next Financial Meltdown*. Pantheon

Juglar, Clément (1862) *Des Crises commerciales et leur retour periodique en France, en Angleterre, et aux Etats-Unis*. Guillaumin

Keen, Steve (2006) The Lily and the Pond. *Interview reported by the Evans Foundation*, at http://evatt.org.au/news/445.html

Kim, Jong-Il & Lawrence Lau (1994) The Sources of Economic Growth of the East Asian Newly Industrialized Countries. *Journal of the Japanese and International Economies* 8, pp. 235-71

Kitchin, Joseph (1923) Cycles and Trends in Economic Factors. *Review of Economics and Statistics* 5 (1), pp. 10–16

Kiyotaki, Nobuhiro & Moore, John (1997) Credit Cycles. *Journal of Political Economy* 105 (2), pp.211–248

Kondratiev, Nikolai (1926) Die langen Wellen der Konjunktur. *Archiv fuer Sozialwissenschaft und Sozialpolitik* 56/3, pp. 573–609

Krugman, Paul (1994) The Myth of Asia's Miracle. *Foreign Affairs* November/December, pp. 62-78

Kuznets, Simon (1930) *Secular Movements in Production and Prices. Their Nature and their Bearing upon Cyclical Fluctuations*. Houghton Mifflin

Kydland, Finn E. & Prescott, Edward C. (1982) Time-to-Build and Aggregate Fluctuations *Econometrica*, Vol. 50, No. 6, pp. 1345-1370

Kydland, Finn E & Prescott, Edward C (1991) Hours and Employment Variation in Business Cycle Theory. *Economic Theory*, vol. 1(1), pages 63-81, January

Long, John B. & Plosser, Charles I. (1983) Real Business Cycles. *Journal of Political Economy*, Vol. 91, No. 1, pp. 39-69

Lotka, Alfred J. (1925) *Elements of physical biology.* Williams & Wilkins Co.

Neftçi, Sahlin N. (1984) Are Economic Time Series Asymmetric over the Business Cycle? *Journal of Political Economy* 92, pp. 307-328

Poterba, James M. & Summers, Lawrence H. (1988) Mean Reversion in Stock Returns: Evidence and Implications. *Journal of Financial Economics*, Vol. 22, No. 1, pp. 27-59

Ramsey, J. B. and Rothman, P. (1996) Time irreversibility and Business Cycle Asymmetry. *Journal of Money, Credit, and Banking* 28, pp. 1-21

Razzak, Weshah A. (2001) Business Cycle Asymmetries: International Evidence. *Review of Economic Dynamics* 4, pp. 230-243

Roll, Richard (1992) A Mean/Variance Analysis of Tracking Error *Journal of Portfolio Management*, Vol. 18, No. 2, pp. 7-19

Samuelson, Paul (1965) Proof That Properly Anticipated Prices Fluctuate Randomly. *Industrial Management Review* 6: 41–49

Shea, John (1998) What Do Technology Shocks Do? *NBER Working Paper No. 6632*, National Bureau of Economic Research

Shiller, Robert (2005) *Irrational Exuberance*. (2nd Edition) Princeton University Press

Sichel, Daniel E. (1993) Business Cycle Asymmetry: Deeper Look. *Economic Inquiry* 31, pp. 224-36

Tobin, James (1969) A general equilibrium approach to monetary theory. *Journal of Money Credit and Banking*, Vol. 1, No. 1 pp. 15-29

Verbrugge, Rineke (1997) Investigating Cyclical Asymmetries. *Studies in Nonlinear Dynamics and Econometrics* 2, 15-22

Volterra, Vito (1926) Variazioni e fluttuazioni del numero d'individui in specie animali conviventi. *Memoria della Reale Accademia Nazionale dei Lincei.* Ser. VI, vol. 2

Wright, Stephen (2004) Measures of Stock Market Value and Returns for the U.S. Nonfinancial Corporate Sector, 1900-2002. *Review of Income and Wealth*, Series 50, Number 4 (December), pp. 561-584

Young, Alwyn (1992) A Tale of Two Cities: Factor Accumulation and Technical Change in Hong-Kong and Singapore. *NBER Macroeconomics Annual* 7. National Bureau of Economic Research.

Young, Alwyn (1994) Lessons from the NICs: A Contrarian View. *European Economic Review* 38, pp.964-73

Young, Alwyn (1995) The Tyranny of Numbers: Confronting the Statistical Realities of the East Asian Growth Experience. *Quarterly Journal of Economics* 110 (August), pp. 641-80

Zarnowitz, Victor (1992) *Business Cycles: Theory, History, Indicators and Forecasting*. University of Chicago Press

*****