The US Business Cycle: Power Law Scaling for Interacting Units with Complex Internal Structure

Paul Ormerod (pormerod@volterra.co.uk)

Volterra Consulting Ltd

The Old Power Station

121 Mortlake High Street

London SW14 8SN

November 2001

Paper to be presented at the conference in honour of Gene Stanley's 60th birthday, 'Horizons in Complex Systems', Messina, Sicily, December 2001.

Abstract

In the social sciences, there is increasing evidence of the existence of power law distributions. The distribution of recessions in capitalist economies has recently been shown to follow such a distribution.

The preferred explanation for this is self-organised criticality. Gene Stanley and colleagues propose an alternative, namely that power law scaling can arise from the interplay between random multiplicative growth and the complex structure of the units composing the system.

This paper offers a parsimonious model of the US business cycle based on similar principles.

The business cycle, along with long-term growth, is one of the two features which distinguishes capitalism from all previously existing societies. Yet economics lacks a satisfactory theory of the cycle. The source of cycles is posited in economic theory to be a series of random shocks which are external to the system.

In this model, the cycle is an internal feature of the system, arising from the level of industrial concentration of the agents and the interactions between them.

The model - in contrast to existing economic theories of the cycle - accounts for the key features of output growth in the US business cycle in the twentieth century.

1. Introduction

A distinctive feature of the Western market economies is the short-run fluctuations in output around trends of slow but persistent growth over time. This paper offers a new, complexity theory-based explanation of these short-term fluctuations in output, often referred to in economics as the 'business cycle'. I address issues both in complex systems and in economic theory.

The frequency of economic recessions in 17 Western economies over the 1871-1994 period is considered in [1]. Power law relationships give a good description of the frequency of recessions in terms of duration in these countries (defining a recession as a year in which real output growth is negative), although there are deviations from this behaviour when one considers the frequency of occurrence of large recessions.

This is one example of the increasing evidence in the social sciences of the existence of power law distributions. The preferred explanation for these results, both in the social and biological sciences, is self-organised criticality [for example, 2]. However, Gene Stanley and colleagues comment that 'it is difficult to imagine that for all these diverse systems, the parameters controlling the dynamics spontaneously self-tune to their critical values'. In a study of the distribution of firm sizes in the US [3], they propose an alternative mechanism based on a) the complex evolving structure of the units making up individual firms and b) an evolution of these units according to a random process.

In this paper, I propose an agent-based model of the business cycle based on similar principles. The individual agents in this model are firms, which operate at different scales of output. The fluctuations in the level of output of each agent arise from a) interactions with other agents and b) random processes which reflect the inherent uncertainty facing each agent in interpreting information.

The properties of this model provide a good approximation to both the power law distribution of recessions and to other key features of Western business cycles. Our specific focus in this paper is in fact on the US business cycle in the twentieth century, the US having been by some margin the largest and most important economy in the world throughout this period.

The vast majority of economic theories of the business cycle postulate that the cause of the cycle is a series of random shocks which are external to the economic system - 'exogenous' in the jargon of economics (see [4] for a detailed survey of the literature to 1990, and [for example, 5] for a discussion of the most recent theoretical approach). A recent empirical study concluded that the evidence is not consistent with the hypothesis that output fluctuations are caused by technological shocks, the most widely cited potential source in the economics literature of such shocks [6]. In our model, the cycle is intrinsic to the system and does not rely on external shocks - it is 'endogenous' in economic terminology.

More generally, mainstream economic models of the business cycle appear to be unable to replicate the time-series properties of US business cycle data, as the economics profession itself has noted [for example, 7, 8]. Further, they do not generate the observed power law-like behaviour of the duration of recessions.

Section 2 considers the key features of the US business cycle. Section 3 sets out the theoretical model, and Section 4 presents the properties of the model compared to those of the actual data. Section 5 gives a brief conclusion.

2. Key empirical features of the US business cycle

Perhaps the most influential definition of the business cycle was given in the National Bureau of Economic Research study in 1946 [9] : 'Business cycles are a type of fluctuation found in the aggregate economic activity of nations that organise their work mainly in business enterprises: a cycle consists of expansions occurring at about the same time in many economic activities, followed by similarly general recessions, contractions, and revivals which merge into the next expansion phase of the cycle; the sequence of changes is recurrent but not periodic; in duration cycles vary from more than one year to ten or twelve years; they are not divisible into shorter cycles of similar character with amplitudes approximating their own'.

There are three key points in this definition:

- the cycle arises primarily through the activities of 'business enterprises' i.e. firms
- the cycle is determined only weakly in the frequency domain
- changes in output in individual sectors of the economy or individual firms tend to be positively correlated over the cycle

The first point is important both in itself and in its empirical implications. A substantial proportion of the output of the American economy, and indeed that of the other Western market economies, is accounted for by a small number of very large firms. The first major phase of mergers, acquisitions and expansions which led to this situation took place in the decades immediately around 1900. By the time of the First World War, the wave of corporate restructurings had been consolidated, and firms capable of operating on a global scale then dominated the US economy for the first time (for example, [10, 11]).

There is, of course, a very large number of small firms whose contribution to total output is substantial[1], and the average size of business organisations in the closing decades of the twentieth century fell quite markedly [12]. Nevertheless, a key feature of the US economy is a concentration of output amongst a small number of firms. For example, over 80 per cent of the total war production of America in the Second World War was carried out by just 100 firms [13].

The annual growth rate of real (i.e. net of inflation) aggregate output (GDP) of the US is plotted in Figure 1.

[1] In 1990, there were around 20 million business enterprises in the United States, most of which had annual receipts below $25,000 [12]

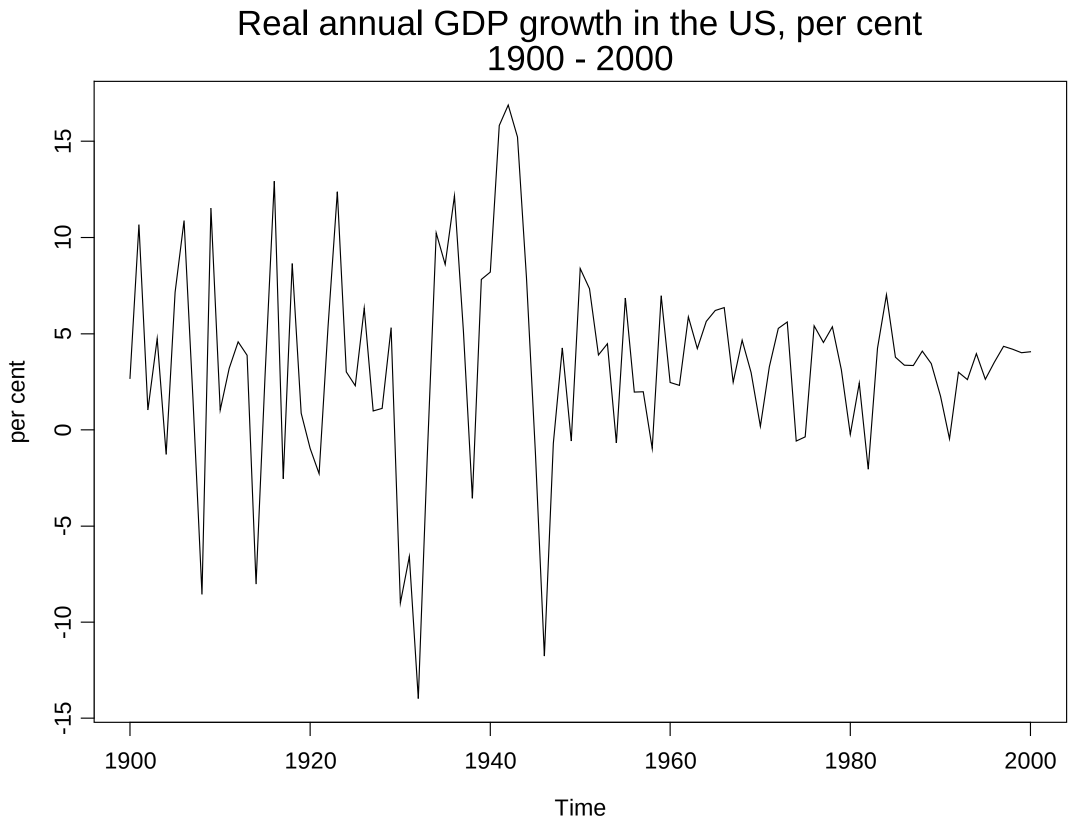

Figure 1 *Time series plot of real (i.e. net of inflation) aggregate output growth in the US 1900-2000.*

The dramatic collapse in output in the Great Depression in the 1930s is clear. Less well known are the enormous increases during the Second World War, and the subsequent sharp fall in the adjustment to a peace-time economy in 1946. The amplitude of the business cycle has clearly been lower in the second half of the century compared to the first. However, this is no guarantee that a major collapse can no longer happen. In the early 1990s, for example, in both Sweden and Finland GDP fell by more than it had done in the early 1930s.

The lack of any distinctive periodicity in the data is apparent. Nevertheless, there is a certain amount of structure, as is shown by analysis in the time domain using the autocorrelation function of the data and in the frequency domain using its Fourier transform pair, the power spectrum. The choice of the sample period over which to estimate these is a matter of judgement. The whole century from 1900 to 2000 could be used, or alternatively the start date could be 1910, by which time the massive wave of mergers referred to above had been consolidated. A case can also be made for starting in 1929, which is the earliest date at which official estimates of US GDP exist[2].

The estimates of the autocorrelation function and the power spectrum of real annual GDP growth vary somewhat depending on the sample period which is used, but their qualitative structure is the same:

- low positive first-order autocorrelation (around 0.4)
- lower, but in general statistically significant, negative autocorrelation at lags 3-5
- other lags of the autocorrelation usually not significantly different from zero
- a weak concentration of the power spectrum at frequencies between 6 to 9 years

The final property of the business cycle is one which is not in the economics literature, but is reported in [1]. A power law distribution provides a good description of the relationship between the frequency and size of the duration of recessions. This finding is robust with respect to different ways of defining a recession. The most unequivocal definition is a year in which real GDP growth is negative, and it is this which is used here.

An important refinement is that a power law fitted to the data excluding those recessions which lasted only one year gives a larger absolute value of the estimated exponent. Further, this power law fits the data even more closely than when the one year observations are included. Using data from 17 capitalist economies, the estimated exponent of the power law is very robust with respect to the sample period chosen, both including and excluding recessions which lasted just for one year. Including one year recessions, the exponent is around 1.65 ± 0.20, and excluding these it is some 3.00 ± 0.06.

[2] The source for pre-1929 data is [13]

For the United States on its own, the estimated exponent varies slightly depending upon the data period which is used, but it is around -2.10 ± 0.20. In the case of the United States, the longest recession in terms of duration has only been four years, so it is not meaningful to carry out a least squares regression excluding recessions of only one year, since this would leave only three observations.

3. An agent-based model of the business cycle

The principle source of the business cycle is in the activities of firms, as noted above. We therefore assume that our model is populated by N individual firms. The model evolves on a step-by-step basis, and in each step, or period, each firm decides two things:

- its rate of growth of output for that period
- its degree of optimism or pessimism about the economic conditions in which it is operating - 'sentiment' for short

In economic terms, the model is very Keynesian in spirit. Agents do not follow complicated behavioural rules involving multi-period optimisation as they do in many economic mainstream theoretical models of the cycle (for example, [5 - 8]). Agents in this model follow simple rules of thumb.

Each agent sets its rate of growth of output in period t by:

$$x_i(t) = (1 - \alpha)x_i(t - 1) + \alpha[Y(t - 1) + \varepsilon_i(t)] \qquad (1)$$

where $x_i(t)$ is the rate of growth of output of agent i in period t and Y is the overall sentiment of all agents (the weighted sum of the levels of sentiment of the N individual agents). Information about Y can be obtained readily by reading, for example, the *Wall Street Journal* or the *Financial Times*. The views of company chairmen and CEOs are widely publicised, and there is a great deal of commentary on the current state of the economy.

The variable $\varepsilon_i(t)$ plays a crucial role in the model. This is a random variable drawn separately for each agent in each period from a normal distribution with mean zero and standard deviation sd_1. Its role is to reflect both the uncertainty which is inherent in any economic decision making and the fact that the agents in this model, unlike mainstream economic models which are based on the single representative agent, are heterogenous.

The implications of any given level of overall sentiment for the growth rate of output of a firm differs both across the N agents and over time. Firms are uncertain about the precise implications of a given level of sentiment for the exact amount of output which they should produce. Further, the variable Y is based upon an interpretation of a range of information which is in the public domain. Agents again differ at a point in time and over time in how they interpret this information and in consequence the value which they attach to Y.

The sentiment of the i th agent is determined by the following:

$$y_i(t) = (1 - \beta)y_i(t - 1) - \beta[X(t - 1) + \eta_i(t)] \quad (2)$$

where X is the overall rate of growth of output of the economy (the weighted sum of the x_i), and where η_i is drawn from a normal distribution with mean zero and standard deviation sd_2.

The coefficient on X(t - 1), β, has a negative sign, again reflecting the Keynesian basis of the model. Keynes never articulated a formal theory of the business cycle. In chapter 22 of [14], however, he wrote that: ‘By a cyclical movement we mean that as the system progresses in, e.g., the upward direction, the forces propelling it upwards at first gather force and have a cumulative effect on one another but gradually lose their strength until at a certain point they tend to be replaced by forces operating in the opposite direction;

which in turn gather force for a time and accentuate one another, until they too, having reached their maximum development, wane and give place to their opposite'. A mathematical approximation to this description is, of course, that of a simple oscillator, and hence the negative sign on X(t - 1) in (2).

The variable $\eta_i(t)$ again reflects agent heterogeneity and uncertainty. At any point in time, each agent is uncertain about the implications of any given level of X(t - 1) for its own level of sentiment. A further practical point is that, although estimates of X are provided in the national accounts of the economy, they are both estimated with potential error and are subject to future revision.

Even at the risk of over-emphasising the point, it is worth repeating that in each time period firms do *not* share the same ε and η. The variables ε and η are *not* degrees of uncertainty which are common to all firms, but each firm in each period has its own ε and η. In other words, ε and η must *not* be regarded as a common, exogenous shock which all firms experience.

The importance of these two variables can be seen as follows. It can be shown [15] that an approximate solution of (1) and (2) for the rate of growth of total output, X, is given by:

$$\Delta^2 X(t) + (\alpha + \beta) \cdot \Delta X(t) + 2\alpha\beta \cdot X(t) = F \qquad (3)$$

The left-hand side of (3) represents a damped pendulum, and the right-hand side a forcing term *F*, where

$$F = \alpha\Delta\bar{\varepsilon}(t) + \alpha\beta\{\bar{\varepsilon}(t) + \bar{\eta}(t)\} \qquad (4)$$

and the bar represents the (weighted) sum of the components. In other words, in the absence of uncertainty and therefore the ε and η variables, in general the system gives rise to damped oscillations.

Equation (4) reveals a key property of the model. *F* is a random variable, so the larger is the number of agents in the model, N, by the rule of adding random variables the lower will be the variance of *F*. In other words, for any given values of sd_1 and sd_2, the standard deviations of ε and η, the greater is N the lower will be the range of fluctuations of the aggregate output series, X.

An essential part of the calibration of the model against the data is to approximate the distribution of values over which output growth has fluctuated. A good approximation can always be obtained by increasing the values of sd_1 and sd_2 as N is increased. However, the higher the values of sd_i, the lower is the mean correlation between the rate if growth of output of the individual agents in the model, the x_i.

Along with the power law distribution of recessions, the most distinguishing feature of the data is the fact that ' a cycle consists of expansions occurring at about the same time in many economic activities, followed by similarly general recessions' [9]. In other words, the larger the value of N, the harder it becomes for the model to reproduce both the range of the data and the fairly strong positive correlations between the output growth of the individual agents. It is the fact that output is concentrated in a small number of large firms which enables these two to be reconciled. I return to this point in section 4 below.

4. Properties of the theoretical model

Because of the stochastic nature of the model, each of the results discussed below refers to 500 separate solutions of the model each carried out over 500 periods, for any given set of values of the parameters (N, α, β, sd_1 and sd_2).

There are five features of the data which the model needs to approximate:

- the distribution of total output growth over time i.e. the range of values of the data
- fairly strong positive correlations between the output growth of individual agents over time

- a power law distribution of recessions
- the structure of the autocorrelation function of total output growth
- the structure of the power spectrum of total output growth

After a considerable amount of experimentation, it was found that with a value of α close to 1 and β close to 0.5, the latter three features above can be approximated for any given values of N and sd_i. A value of α close to 1 is very much in keeping with the Keynesian economic spirit of the model, in which output growth is decided by sentiment. In fact, in all the results reported below, $\alpha = 1$ and $\beta = 0.4$.

The autocorrelation function and the power spectrum are very similar for different values of N. And the exponent of the power law on the distribution of recessions is also very similar, once the sd_i are calibrated appropriately.

The results obtained with N = 100 are typical, and are as follows.

Table 1 The distribution of actual and model-generated output growth

	Min	**1st quartile**	**Mean**	**3rd quartile**	**Max**
Actual	-13.98	+1.00	3.33	5.51	16.88
Model	-6.70	+1.09	3.33	5.53	13.39

The model numbers refer in each case to the average across 500 solutions e.g. -6.70 is the average of the minimum value recorded in each of 500 solutions, each carried out over 500 steps. The mean value of the data represents the long-run growth rate of the US economy. The model in this paper is of the cyclical fluctuations around long-run growth, so this mean value is added to the output of the model, which would otherwise have a mean value of zero.

The bulk of the data, between the 1st and 3rd quartiles, can be accounted for by the model, whilst at the same time approximating other features of the actual data.

However, a small number of extreme events appear to be due to exogenous shocks rather than to the endogenous mechanism of the cycle. There are 3 observations which lie above the mean maximum output growth rate generated by the model: 1941, 1942 and 1943. In these years, there was a massive growth of government spending as the economy was geared to war production. These can be thought of as a common shock to the growth rates of individual agents. There are 5 observations below the mean minimum of the model. One of these, 1946, is due to the run-down of government spending at the end of the war, and is also a common shock to agent output growth. Two relate to years in the Great Depression, 1930 and 1932, and two relate to earlier, one-off years of recession, 1908 and 1914. These can perhaps be thought of as common downward shocks to agent sentiment.

Over the 500 simulations, the model generates recessions of up to six years in duration. A power law gives a reasonable fit to the data, with the standard error of the equation being 1888 compared to 6143 of the dependent variable in the regression. The exponent on a power law relationship fitted to this data is -1.53 ± 0.33. This compares to -2.10 ± 0.20 on the actual US data, and -1.65 ± 0.20 on data across 17 Western economies. The model value is statistically significantly different from the estimated US value at $p = 0.05$, but not from the wider data set. However, in a qualitative sense the model is similar to the actual experience of the US economy.

The conformity of the model with the data is reinforced when recessions lasting just one year are excluded. A least squares fit gives an exponent of -3.72 ± 0.30. This compares to -3.00 ± 0.06 from the relationship estimated across 17 economies. As with the actual data, the absolute value of the power law is substantially higher when recessions lasting just one year are excluded. Further, again as with the actual data, the power law provides a very accurate description of the model-generated data, with the standard error of the equation being just 269 compared to 3538 of the dependent variable.

Figure 2 plots the power spectrum of the model data. The spectrum is concentrated at frequencies between some 5.25 and 7.75 years, slightly shorter than the 6 to 9 years of the actual data.

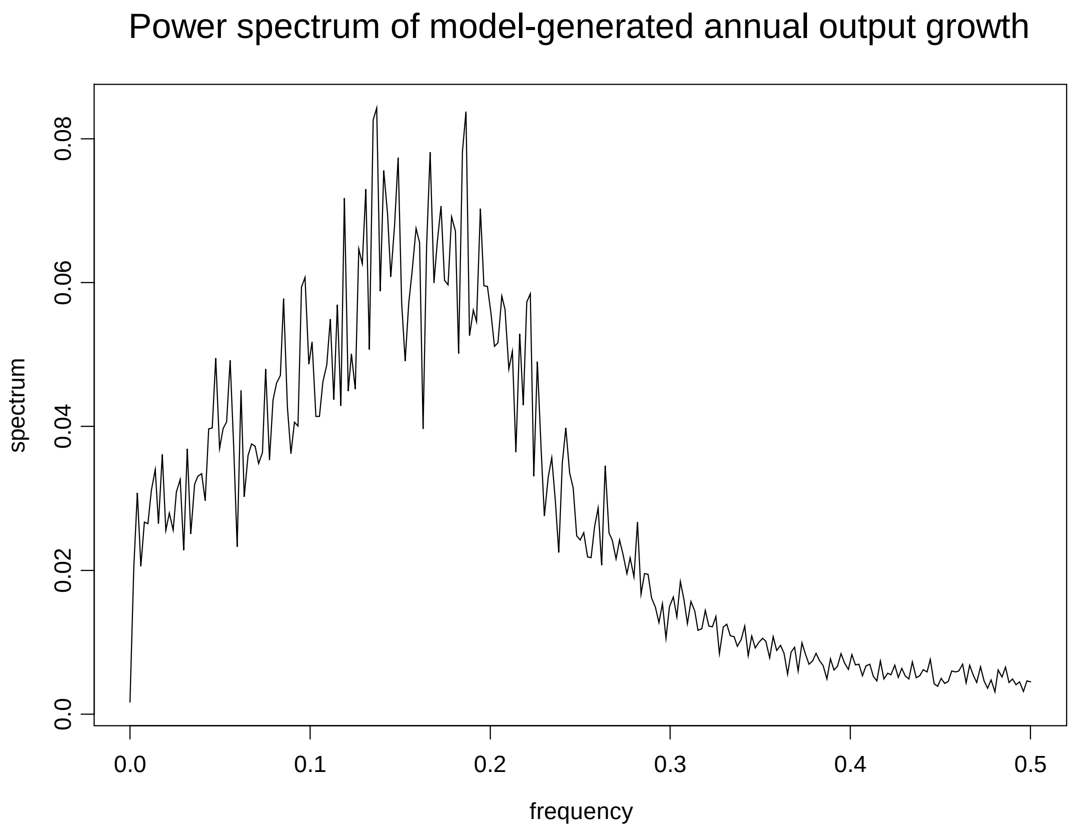

Figure 2. Power spectrum of 500 solutions over 500 steps of the annual output growth in the model, X. $N = 100$, $\alpha = 1$, $\beta = 0.6$, $sd_1 = 0.0375$, $sd_2 = 0.525$.

The autocorrelation function of total output growth is also qualitatively very similar to, though not identical to, that of the actual data. Table 2 shows the average values across 500 solutions of the model from 1 through 9 years.

Table 2 Autocorrelation function of model generated total output growth

Lag	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**
Average value	+0.42	-0.15	-0.26	-0.10	+0.04	+0.06	+0.02	-0.02	-0.02

Like the actual data, the model exhibits:

- low positive autocorrelation at lag 1
- low negative autocorrelation at lags 3 and 4
- autocorrelation not significantly different from zero at higher lags

However, the model generates negative values at lags 2 - 4, whereas the actual data exhibits them at lags 3 - 5. Further, the absolute value at lag 4 tends to be the highest of these three in the actual data, and at lag 3 in the model data.

Finally, and importantly, the model exhibits positive correlations between the rates of output growth over time of the individual agents. The minimum correlation of output growth between the 100 agents in the model average across 500 solutions is 0.307, the average of the mean correlation is 0.433, and that of the maximum correlation 0.546.

Overall, the model is capable of approximating a set of rather subtle properties of the actual US experience. It is not perfect, but the model as it stands is of course highly parsimonious.

The above results are generated by the model populated by 100 firms. A similar set can be obtained for most of the above for different values of N, by choosing suitable values of the sd_i. However, when this is done, the higher is N, the weaker becomes the correlation between the output growth of the individual agents. This is shown in Figure 3, which plots for different values of N the mean value of the correlation coefficients between the growth rates of individual agents, averaged over 500 solutions of the model.

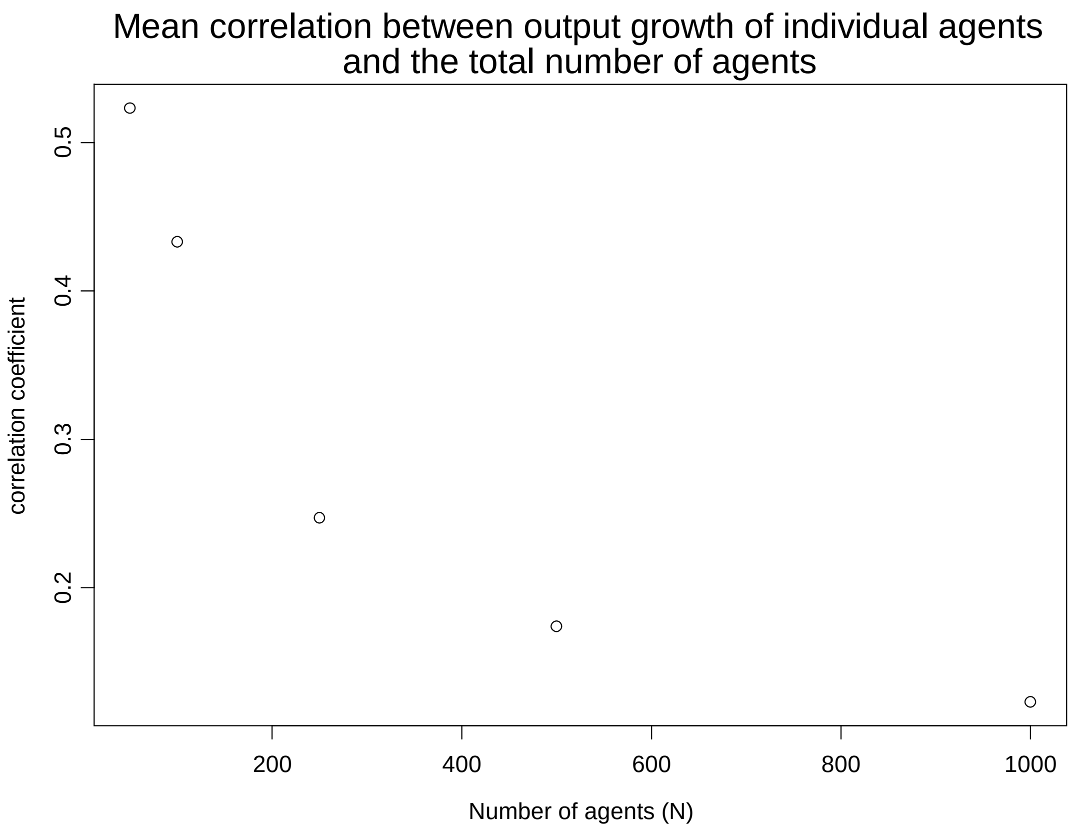

Figure 3. The mean value of the correlation coefficient between the output growth rates of N agents over 500 steps of the model, averaged over 500 solutions. N = 50, 100, 250, 500 and 1000; sd_i calibrated in each case to replicate the range of fluctuations in aggregate output between the first and third quartiles of the distribution.

It is possible to increase the correlation between agents for any given N by reducing the values of the sd_i. However, when this is done the model becomes unable to approximate both the range over which the actual data moves, and the power law behaviour of recessions. In particular, for N = 1000 (and even more so for higher values), recessions become very infrequent and are almost entirely confined to a duration of just one year.

The model therefore suggests that the business cycle arises because of, first, the different scales on which individual firms produce and, second, the interactions between these firms. It is the concentration of output amongst a relatively small group of firms which gives rise to the business cycle.

The model is of course calibrated against the properties of the data during the twentieth century. Arguably, the experience since 1950 has been different from that of pre-1950, as Figure 1 might suggest. The reduced scale of fluctuations in output may be due in part to reduced volatility in sentiment (expectations that governments will intervene to try to prevent mild recessions turning into major ones are more prevalent in the second half of the century than the first). But it may also be due to the fact that an increasing proportion of total US output is being produced by small companies since around 1960.

5. Conclusion

A striking feature of the Western market economies is the short-run fluctuations in output around trends of slow but persistent growth over time. This paper offers a new, complexity theory-based explanation of these short-term fluctuations in output, often referred to in economics as the 'business cycle'.

The frequency distribution of recessions (defined as being when the rate of growth of output is negative) in the business cycle in 17 leading Western economies is well described by a power law for much of the data. This finding applies to both the duration and the magnitude of recessions. However, there are systematic deviations from this behaviour when one considers the frequency of occurrence of large recessions.

Power law distributions are common in complex systems and are often attributed to the self-organised criticality of the system. An alternative explanation for such observed behaviour is in the complex structure of the units composing the system.

The economy is made up of a large number of firms, which are diversified. In other words, it consists of heterogeneous agents, which interact in complex ways. Further, different firms operate at different scales of output. We propose a model of the business cycle in which these factors are capable of generating aggregate output growth data which follows the power law-like behaviour of recessions which exist.

The paper shows that it is not necessary to rely upon the parameters of a system spontaneously self-tuning themselves to their critical values in order to give rise to power law behaviour. We offer an alternative theoretical account in which such behaviour arises from the complex structure of the sub-units of a system (in this case the firms as sub-units of the whole economy) and the interactions between them.

The paper also makes a contribution to economic theory. The mainstream models of the business cycle find it extremely difficult to replicate the time-series properties of US business cycle data, as the economics profession itself has noted. Further, they do not generate the observed power law-like behaviour of the duration and magnitude of recessions. The model described here therefore represents an advance on existing economic models.

References

1. P.Ormerod and C.Mounfield (2001), 'Power law distribution of duration and magnitude of recessions in capitalist economies: breakdown of scaling', *Physica* **A**, 293, 573-582

2. B.Drossell (2001), 'Biological evolution and statistical physics', cond.mat/ 0101409, forthcoming in *Advances in Physics*

3. L.A.N Amaral, S.V.Buldyrev, S.Havlin, M.A.Salinger, and H.E.Stanley (1998) 'Power law scaling for a system of interacting units with complex internal structure', *Phys. Rev. Lett.*, **80**, 1385-1388

4. A.W.Mullineux (1990), *Business Cycles and Financial Crises*, Harvester Wheatsheaf, New York

5. M.Eichenbaum (1995) 'Some comments on the role of econometrics in economic theory', *Economic Journal*, **105**, 1609-21

6. J.Gali (1999), 'Technology, employment, and the business cycle: do technology shocks explain aggregate fluctuations?', *American Economic Review*, **89**, 249-271

7. T.Cogley and J.M.Nason, (1995), 'Output dynamics in real business cycle models', *American Economic Review*, **85**, 492-511

8. J.J.Rotemberg and M.Woodford (1996), 'Real business cycle analysis and forecastability', *American Economic Review*, **86**, 71-89

9. A.F.Burns and W.C.Mitchell (1946), *Measuring the Business Cycle*, NBER

10. A.Chandler (1990), *Scale and Scope: the Dynamics of Industrial Capitalism*, Harvard University Press

11. L.Hannah (1999) 'Marshall's "trees" and the global "forest": were "giant redwoods" different?' in N.R.Lamoreaux, D.M.G.Raff and P.Temin, eds., *Learning by doing in markets, firms and countries*, National Bureau of Economic Research

12. G.R.Carroll and M.T.Hannan (2000), *The Demography of Corporations and Industries*, Princeton University Press

13. R.Overy (1995), *Why the Allies Won*, Jonathan Cape, London

14 J.M. Keynes (1936), *The General Theory of Employment*, Macmillan

15. P.Ormerod (2001), 'The Keynesian micro-foundations of the business cycle: some implications of globalisation' in *What Global Economic Crisis?*, eds., P Arestis, M. Baddeley and J.S.L. McCombie ,Basingstoke; Palgrave