PHYSA 8255
pp: 1-13 (col.fig.: Nil)
PROD. TYPE: COM
ED: Preetha
PAGN: Usha.K – SCAN: Roopa

ARTICLE IN PRESS

ELSEVIER
Available online at www.sciencedirect.com
SCIENCE DIRECT®
Physica A ■■■ (■■■■) ■■■–■■■
PHYSICA A
www.elsevier.com/locate/physa

Information cascades and the distribution of economic recessions in capitalist economies

Paul Ormerod

Volterra Consulting Ltd, The Old Power Station, 121 Mortlake High Street, London SW14 8SN, UK

Received 24 November 2003; received in revised form 8 April 2004

UNCORRECTED PROOF

Abstract

We consider in this paper the distribution of the cumulative size of recessions in 17 capitalist countries over the period 1871–1994, using data on annual percentage changes in real GDP. A recession is defined as a year in which GDP growth is negative, and the cumulative change is the change from peak to trough during a recession period. We examine both the whole sample and different partitions of the data.

The null hypothesis that the size distribution of recessions follows an exponential distribution is never rejected at the conventional level of statistical significance, $p = 0.05$. However, there are always a small number of large recessions, no matter how the data is partitioned, which are not well fitted by a least-squares regression of the log of size on the rank of size. In other words, in a qualitative sense we see a bimodal distribution of recessions, with an exponential fit to the bulk of the data, and a second peak describing a small number of very large recessions.

A previously published agent-based economic theory model of the business cycle, calibrated purely on US data, is able with no change in its parameters to generate an exponential distribution of the size of recessions very close to that which is actually observed. In this model, information flows between agents on a completely connected network.
© 2004 Elsevier B.V. All rights reserved.

PACS: ■; ■; ■

Keywords: ■; ■; ■

1. Introduction

The phenomenon of information cascades [1] is widely observed in economic and social systems. Ref. [2] notes that small initial shocks can cascade to affect or disrupt

E-mail address: pormerod@volterra.co.uk (P. Ormerod).

0378-4371/$ - see front matter © 2004 Elsevier B.V. All rights reserved.
doi:10.1016/j.physa.2004.04.109

large systems that have proved stable with respect to similar disturbances in the past. Examples include financial markets [3], the commercial success or failure of films [4], and the diffusion of crime [5].

Watts [2] offers a general theoretical model of a possible explanation in terms of a sparse, random network of interacting agents whose decisions are determined by the actions of their neighbours according to a simple threshold rule. Two regimes are identified in which the network is susceptible to very large cascades that occur very rarely. When cascade propagation is limited by the connectivity of the network, a power law distribution of cascades is observed. But when the network is highly connected, the size distribution of cascades is bimodal, with an exponential tail at small cascade size and a second peak at the size of the entire system corresponding to a single global cascade.

The purpose of this paper is to present evidence on the distribution of the size of recessions in the advanced capitalist economies, which are obvious examples of a networked system. A theoretical model is described in which information flows between agents on a completely connected network, and which offers an initial explanation of the observed distribution.

2. The background

Firms in the economy are connected in two distinct ways. First, there are the technological connections which arise from the need for companies in any particular sector to use as inputs into their production the outputs of other industries. An example is the motor vehicle industry, which uses materials produced by other industries to make cars and trucks. Input–output tables in the national economic accounts describe these connections at the level of the industry.[1] Ref. [6] shows that in a modern developed economy, the connections between industries are rather extensive.

The second type of connection arises from the impact on any given firm of the decisions and opinions about the future of other firms in the economy. Keynes [7], for example, attached considerable importance to the state of long term expectations, or sentiment, for the decisions by companies to invest. The more optimistic the overall climate of opinion, the more likely it is that a firm will carry out an investment decision. Information about the general climate of expectations is readily available from the media, through newspapers such as the *Wall Street Journal*, for example. This is certainly the case as far as large companies are concerned. In other words, it is as if all agents are connected to, say, Microsoft or General Motors in terms of gathering information about the business climate. The information they gather may not be perfect, but a large amount of it is available.

So on both the technological and informational criteria, capitalist economies appear to be highly connected.

3. The data

Annual percentage changes in real GDP in 17 capitalist economies from 1871 to 1994 are available from Ref. [8]. The individual countries are Australia, Austria,

[1] Similar data are not available at the level of the firm.

UNCORRECTED PROOF

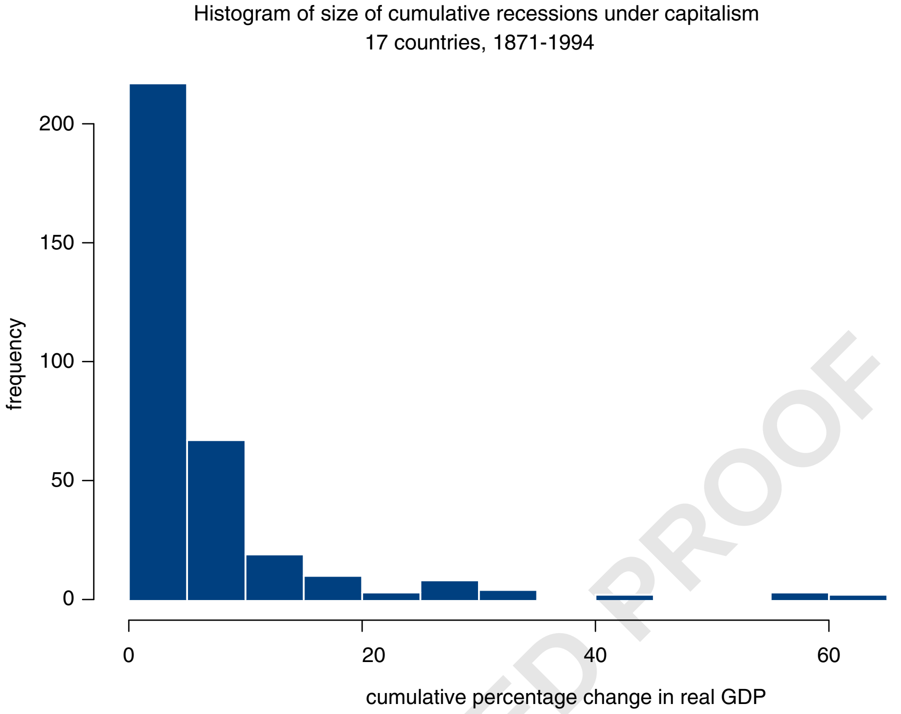

Fig. 1. Histogram of cumulative absolute fall in real GDP from peak to trough, 336 recessions in 17 capitalist countries 1871–1994.

UNCORRECTED PROOF

Belgium, Canada, Denmark, France, Finland, Germany, Italy, Japan, Netherlands, New Zealand, Norway, Sweden, Switzerland, the UK and the US.[2]

There are formidable problems involved in constructing such data series, many of which are discussed in Ref. [8]. It is important for readers with a natural science background to appreciate that, outside financial markets, almost all economic data is estimated and the margin of error around any individual data point can be substantial. This is the case even with modern estimates of data for the most recent years.

Recessions can be measured in a number of ways. An unequivocal and widely accepted way of doing so is to define a recession as being a period of at least 1 year (given that annual data is being used[3]) in which the rate of growth of real GDP is less than zero. Across the 17 countries, we find 336 examples of recessions over the 1871–1994 period.

In this paper, we analyse the cumulative size of recessions. In other words, the percentage change in real GDP from peak to trough. The size of recessions has varied

[2] Data for Japan is only available from 1886 and for Switzerland from 1901.

[3] Data at a higher frequency such as quarterly does not exist for periods before World War Two, and is only gradually available across the countries since then.

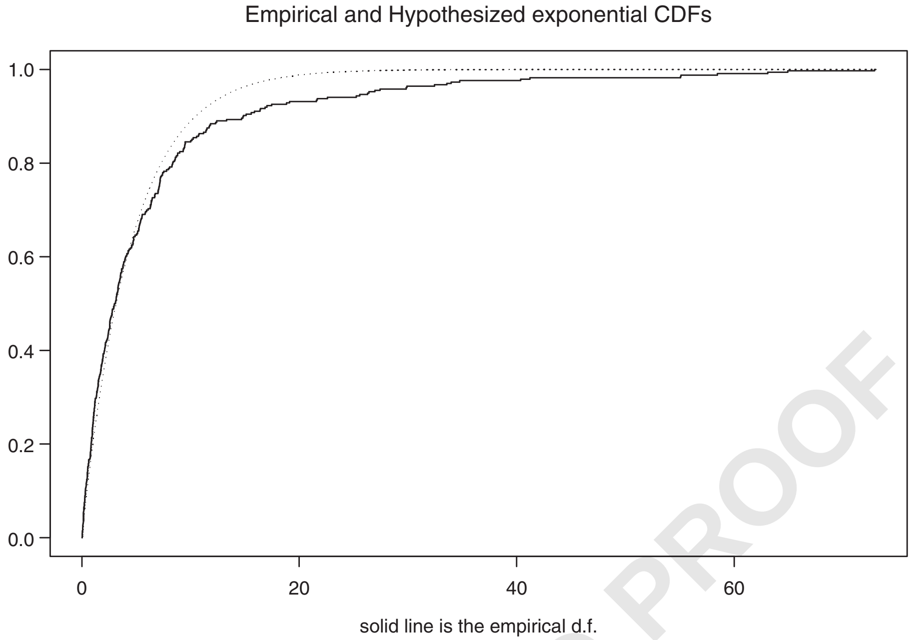

Fig. 2. Empirical distribution function of cumulative size of recessions (solid line) and hypothesised exponential cumulative distribution function with rate=0.22.

dramatically, from recessions in which output hardly fell at all to a few in which it fell by more than 50 per cent. We use the absolute value of that data, for ease of exposition (Fig. 1).

Most recessions have been small, with 81 out of the total of 336 observations having a value of less than 1 per cent. There have been a small number of catastrophic falls, with 12 observations being over 30 per cent. Most of these relate to the end and immediate aftermath of war in the defeated countries in the world wars, but Canada and the United States both experienced falls of over 30 per cent in the Great Depression of the 1930s.

4. Results

Fig. 2 compares the empirical distribution function of the sample with a hypothesised cumulative distribution function which is an exponential with rate parameter=0.22.

On a formal Kolmogorov–Smirnov test, the null hypothesis that the data are distributed exponentially with rate=0.22 is rejected at $p = 0.092$. In other words, the null hypothesis is not rejected at the standard level of statistical significance $p = 0.05$. This rate, and the other empirically calibrated ones described below, was chosen by obtaining an initial calibration by means of a plot as in Fig. 2, and then searching around the neighbourhood of the initial rate in steps of 0.01 and choosing the rate at which the p-value is maximised.

UNCORRECTED PROOF

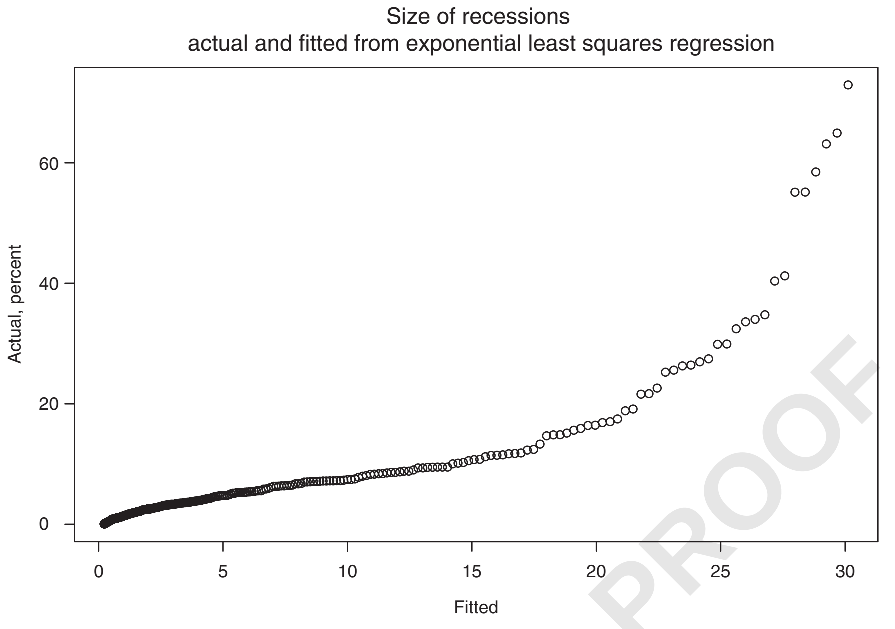

Fig. 3. Actual size of recessions and values fitted from exponential least-squares regression of log of size on rank. The relationship up to around values of 20 is described by a 45° line, but the large values are clearly outliers.

However, it is the bulk of the data which is essentially described by an exponential, and the tail is fitted quite poorly. In the actual data, there are 313 out of the total of 336 observations with a value of less than 20, and 308 in the values fitted by a least-squares exponential fit to the rank of the data. However, the maximum value predicted by the regression is 30.1, and we noted previously that there are 12 actual observations above this level.

Excluding the observations with values above 20 makes little difference to the estimated coefficient on the cumulative size. Using the whole sample it is 0.0147 with an equation standard error of 0.374 and excluding the observations with values above 20 it is 0.0142 with an equation standard error of 0.367.

In other words, the bulk of the data is fitted well by an exponential relationship between size and rank of size, but there is evidence of a bimodal distribution with a small number of very large recessions being outliers. This is summarised in Fig. 3.

Most of the very large observations refer to the collapse of output in some occupied countries in World War Two, and the defeated countries at the end of both world wars. The sample was therefore divided into two groups, which for descriptive purposes we refer to as ‘allies’ and ‘rest’. The former group comprises those countries which were neither defeated nor occupied during the world wars, plus the neutral countries of Sweden and Switzerland. In other words, Australia, Canada, New Zealand, Sweden, Switzerland, the UK and the US. The rest are in the ‘rest’ group.[4]

[4] This is not to suggest that France, for example, did not at some point fight against Germany, but it was occupied.

UNCORRECTED PROOF

Table 1
Size of cumulative recessions 1871–1994, absolute per cent, allied countries and the rest

	Min.	1st quartile	Median	3rd quartile	Max.
Allied	0.04	1.4	3.2	7.0	40.4
Rest	0.03	1.0	2.6	7.3	72.9

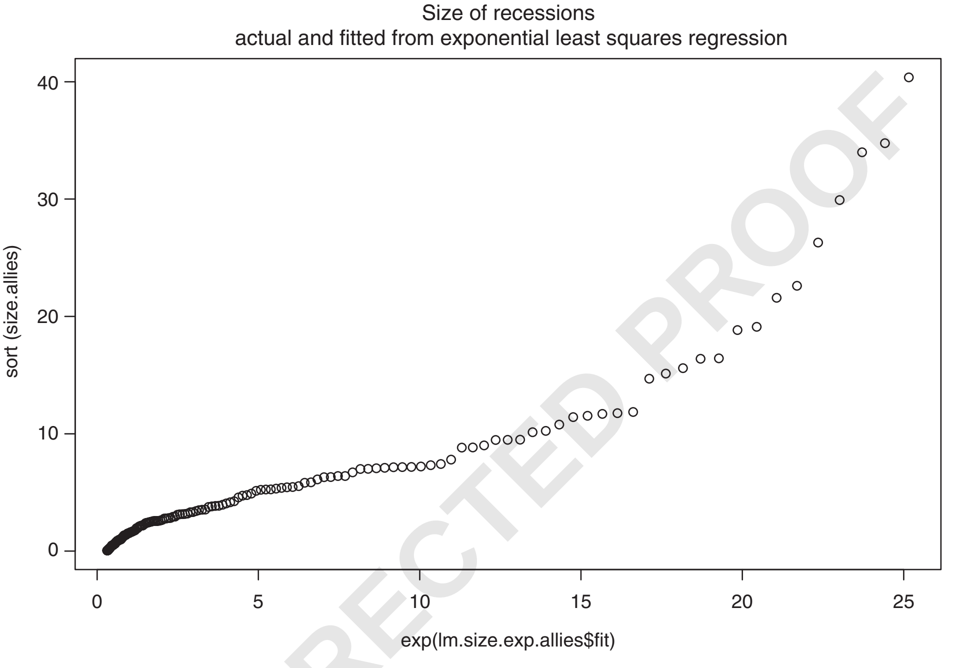

Fig. 4. Actual size of recessions and values fitted from exponential least-squares regression of log of size on rank, allied group of countries. The relationship up to around values of 20 is described by a 45° line, but the large values are clearly outliers.

UNCORRECTED PROOF

There are 149 observations in the ‘allies’ group and 187 in the ‘rest’ group. The bulk of the distribution of the data in these two groups is very similar (Table 1).

On a formal Kolmogorov–Smirnov test, the null hypothesis that the two distributions are the same is rejected at $p = 0.071$, above the conventional level of significance at $p = 0.05$.

However, it is in the tail of the distributions where the experience of the two sets of countries differ. The allies group has only 7 observations above 20 per cent with a mean value of 29.9, and the rest group has 16 such observations with a mean value of 41.1.

In neither case can the null hypothesis that the data are distributed exponentially be rejected. With the allied group, a rate of 0.22 is only rejected at $p = 0.77$, although

Table 2
Size of cumulative recessions, absolute per cent, 1871–1994 and 1950–1994

	Min.	1st quartile	Median	3rd quartile	Max.
1871–1994	0.03	1.0	3.1	7.1	72.9
1950–1994	0.03	0.5	1.5	2.9	14.8

for the rest a rate of 0.25 is rejected at $p = 0.053$, so this is much closer to rejection on conventional criteria.

Fig. 4 plots the allied group actual values and those fitted from a least-squares regression of the log of size against rank. The extreme values in the tail for this group do not arise from war-related events, such as the physical destruction of much of the capital stocks of Germany and Japan in 1945. But it is clear that even in this group, there are a small number of observations in the tail which are not described well by an exponential fit to the data.

Another way of examining the data is to divide the sample into pre- and post-World War Two, or more precisely from 1950, when the immediate economic impact of the war had faded and pre-war living standards had been surpassed in almost all countries.[5] In general, the variability of output growth from year to year has been less in the latter than in the former period. The standard deviation of the annual percentage change in real GDP across the 17 countries from 1871 to 1949 is 6.72, from 1871 to 1939 it is 5.10. But from 1950 to 1994 it is 2.86. Table 2 compares the post-1950 experience with that of the full sample.

Fig. 5 compares the empirical distribution function of the 1950–1994 sample with a hypothesised cumulative distribution function which is an exponential with rate parameter=0.42.

The null hypothesis that the data are exponentially distributed is only rejected at $p = 0.73$ on a Kolmogorov–Smirnov test. There is, however, an outlying observation, relating to the recession in Finland in the early 1990s when output fell by almost 15 per cent (Fig. 6).

Finally, we can consider the size distribution of recession according to their duration. There is a number of ways in which the distribution of duration of recessions can be described [9–11], none of which are completely conclusive, given that the longest recession observed lasts for only 7 years. Fig. 7 plots the relationship between cumulative size and duration of recessions.

Overall, there is a positive correlation between the two of 0.58. But clearly, even in short recessions, falls in output can be very considerable. The null hypothesis that the size follows an exponential distribution cannot be rejected for recession lasting 1, 2, 3 and more than 3 years, although the rate parameters differs between them (Table 3).

For recessions lasting both 1 and 2 years, the same problem regarding the fit of an exponential to the tail remains. The bulk of the data is fitted well, but there is in each case a small number of large outlying observations. One and two year durations

[5] Except Germany and Japan, which reached 1938 living standards in 1953.

UNCORRECTED PROOF

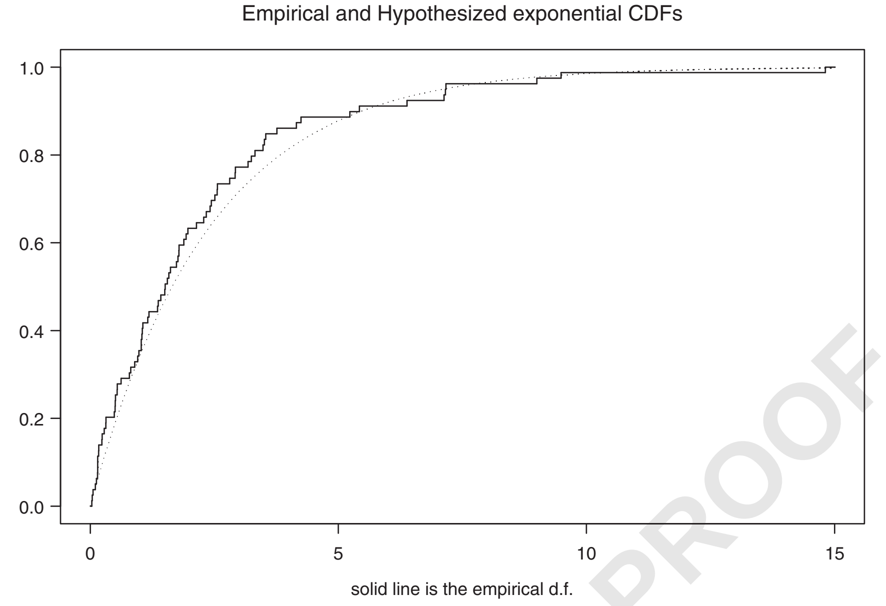

Fig. 5. Empirical distribution function of cumulative size of recessions 1950–1994 (solid line) and hypothesised exponential cumulative distribution function with rate=0.42.

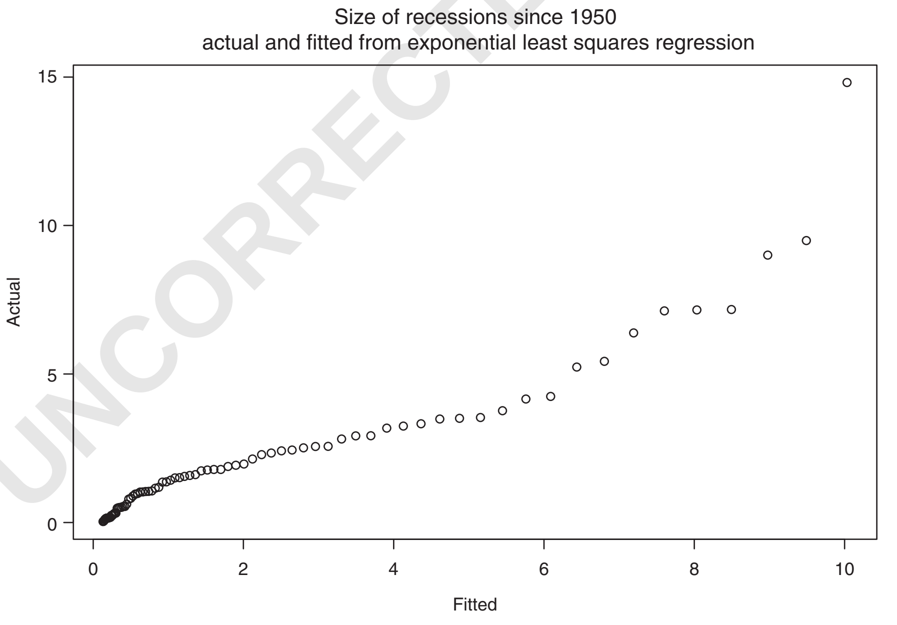

Fig. 6. Actual size of recessions and values fitted from exponential least-squares regression of log of size on rank, 17 capitalist countries, 1950–1994.

UNCORRECTED PROOF

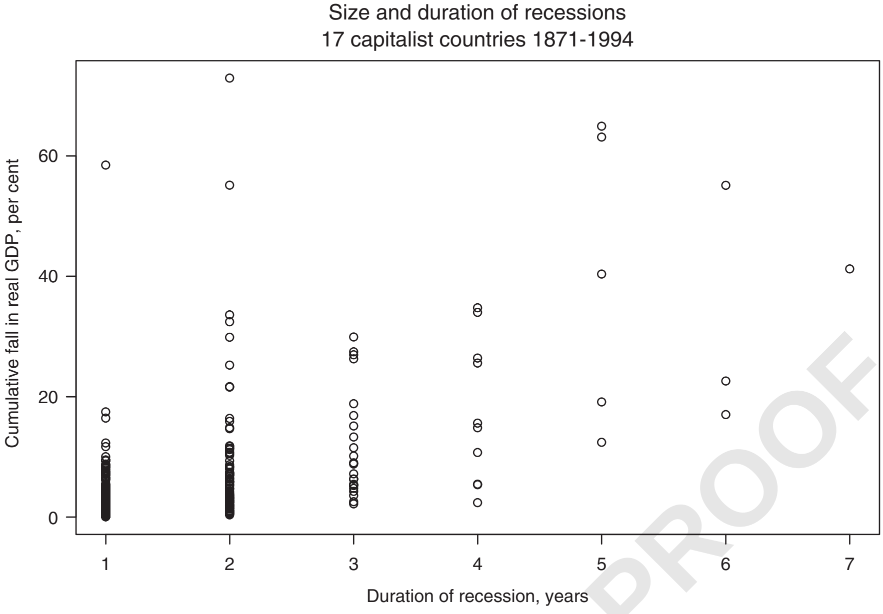

Fig. 7. Cumulative size of recession and duration, 17 capitalist economies, 1871–1994.

Table 3
Null hypothesis that recessions of different durations follow an exponential distribution

Duration (years)	Rate parameter	p-value at which null hypothesis rejected on K–S test	No. of obs.
1	0.36	0.24	208
2	0.15	0.22	88
>2	0.05	0.73	42

account for 296 out of the total of 336 observations. For recessions lasting three years or more, the outlier problem is less acute. The actual and predicted values from a least-squares regression of the log of size on rank, when plotted lie close to the 45° line even in the tail, although even here the three very largest recessions are on average some six percentage points larger than predicted.

Overall, then, there is strong evidence that the bulk of the observations on the cumulative size of recessions in advanced capitalist economies follow an exponential distribution. However, although the null hypothesis of an exponential distribution is never rejected at the conventional level of statistical significance, there are a number of large recessions which are not at all well described by an exponential least-squares fit of size on the rank of size. There is clearly a distinct element of bi-modality in the distribution, because the latter point holds no matter how the data set is partitioned, the only partial exception to this being with the small number of recessions lasting three or more years.

UNCORRECTED PROOF

5. An initial theoretical model

Capitalist economies are highly connected networks. In such circumstances, when it is as if the degree of connectivity between agents is high, Ref. [2] suggests that a bimodal distribution of cascades will be observed, with an exponential tail at small cascade size, and a second peak with a much larger, global cascade size. This appears to be very close to what we actually observe.

A model based on economic theory which is able to approximate the size distribution of recessions is described in Ref. [12]. An agent-based theoretical model was developed to account for the business cycle in the United States economy over the 1871–1994 period. The model was developed to account for the key features of the annual changes in real US GDP, both greater than and less than zero, and was not tested at the time by comparing the size distribution of the recessions which it generated with the actual distribution.

The model is populated by firms, and is able to approximate closely both the time and frequency domain characteristics of annual US GDP growth, as well as exhibiting positive cross-correlation between output growth of individual agents, which is regarded in economic theory as *the* distinguishing feature of the business cycle.

The theoretical basis of the model is developed from Keynes [13]. It has strong affinities with Ref. [2], in that agents (firms) decide their rates of growth of output with respect to the general level of the sentiment of other agents about the future. In other words, information flows across a network.

The model evolves on a step-by-step basis. In each step, each agent decides its rate of growth of output for that period, and its level of sentiment about the future.

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

$$x_i(t) = [Y(t-1) + \varepsilon_i(t)]\,, \tag{1}$$

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). Eq. (1) implies that in making the decision on output, it is as if all agents are on a completely connected network. Each agent considers the sentiment of all other agents in making its decision.

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 heterogeneous.

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.

UNCORRECTED PROOF

The model is completed by the equation describing how each agent determines its level of sentiment:

The sentiment of the ith 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 [12, Chapter 22].

The coefficients calibrated against US data, β, sd_1 and sd_2 were left unchanged. The only change made was to set the mean value of the output of the model at 2.0, which is the average annual real growth rate across all capitalist economies 1871–1994, instead of the 3.3 which is the growth rate of the US economy.

The model was run for 30,000 steps, which generated 4317 recessions. In the actual data, there are 2074 observations and 336 recessions. The proportions of recessions to total observations are 0.144 for the model and 0.162 for the actual data. Even using the model parameters calibrated on US data, the model is able to describe the range of the bulk of the data well. The 1st quartile, mean and 3rd quartiles of the actual data are, respectively, -0.1, 2.0 and 4.4. The model generates -0.1, 2.0 and 4.1. However, the extreme values, both in recession and in recoveries (almost exclusively after wars), of the actual data are not captured by the theoretical model.

The null hypothesis that the distribution of the size of the 4317 recessions generated by the model follows an exponential distribution is only rejected at $p = 0.77$ on a Kolmogorov–Smirnov test. However, the rate parameter which gives this is 0.26, compared to the 0.22 of the actual data. The null hypothesis that the model output and the actual data follow the same distribution is rejected at $p = 0.006$.

Most of the very large recessions which are observed relate to periods either during or at the end of the two world wars. In the theoretical model above, the cycle is endogenous. In other words, it arises from the internal properties of the model and does not rely on the existence of external shocks. We might plausibly regard large scale destruction of domestic capital stock by enemy action as an external shock.

We therefore introduced downward external shocks to output (but not directly to sentiment) which, unlike $\varepsilon_i(t)$, are common to all agents in the period in which they occur. The actual data covers some 120 years, in the course of which there have been two periods of serious external shock to some, but not all, economies. The shocks are introduced every 60 steps and last for a single step each time. The size is drawn from a uniform distribution on [-20,-30]. The parameters sd_1 and sd_2 were re-calibrated slightly to 0.032 and 0.42 so that the output of the model continued to match the range of the bulk of the data between the first and third quartiles. The external shocks generated cumulative recessions of over 40 per cent fall in output.

With the model adjusted in this way, the null hypothesis that the actual size distribution of recessions is the same as the model generated distribution is not rejected at $p = 0.11$ on a Kolmogorov–Smirnov test.

So this simple model, using parameters calibrated purely on US data, is able to replicate qualitatively the size distribution of recessions in 17 capitalist economies.

UNCORRECTED PROOF

A refinement, to take account more explicitly of the idea in Ref. [2], would be to introduce sparsity into the connections between agents in terms of how they set their individual output growth. At present, each agent takes account of the sentiment of all other agents. Perhaps more realistically, agents will take particular account of agents in their own or related industries, and less account of agents operating in different sectors of the economy.

6. Conclusion

We consider in this paper the distribution of the cumulative size of recessions in 17 capitalist countries over the period 1871–1994, using annual percentage changes in real GDP. A recession is defined as a year in which GDP growth is negative, and the cumulative change is the change from peak to trough during a recession period.

We examine different partitions of the data. We divide the sample into the victorious plus neutral countries in the world wars, and the rest. Very large output falls characterise many of the occupied and defeated countries. The overall variability of output growth has been much less since 1950, and we consider the distribution of the post-1950 recessions. And we examine the distribution of recessions of different duration.

The null hypothesis that the size distribution of recessions follows an exponential distribution is never rejected at the conventional level of statistical significance, $p=0.05$.

An agent-based economic theory model, previously published and calibrated against US real GDP growth data [12], generates an exponential size distribution of recessions, although the rate parameter is slightly different from the data across the 17 capitalist economies. In the model, the rate of growth of output of each agent is determined simply by the overall level of sentiment about the future held by all agents. In other words, it is as if the agents are on a fully connected network across which information flows. Each agent takes account of the sentiment of all other agents.

Although the null hypothesis that the size distribution of recessions follows an exponential distribution is not rejected, there are always a small number of large recessions, however the data is partitioned, which are not well fitted by a least-squares regression of size on the rank of size. In other words, in a qualitative sense we see a bimodal distribution of recessions, with an exponential fit to the bulk of the data, and a second peak describing a small number of very large recessions.

Ref. [2] describes a general theoretical model with no explicit economic content of a random network of interacting agents whose decisions are determined by the actions of their neighbours according to a simple threshold rule. This suggests that when the network is highly connected, a bimodal distribution of cascades will be observed. There will be an exponential tail at small cascade size and a second peak at the size of the entire system corresponding to a single global cascade. This seems qualitatively similar to what is observed. A combination of Refs. [2,12] may enable a general explanation of recessions to be given which explains the facts even more closely.

UNCORRECTED PROOF

Acknowledgements

I am very grateful to an anonymous referee of *Physica A* for comments on a previous draft.

References

[1] S. Bikhchandani, D. Hirshleifer, I. Welch, J. Polit. Economy 100 (1992) 992–1006.
[2] D.J. Watts, Proc. Nat. Acad. Sci. 99 (2002) 5766–5771.
[3] A. Kirman, Bank of England Quarterly Bulletin, Bank of England, UK, 1995.
[4] A.S. De Vany, W.D. Walls, Econ. J. 106 (1996) 1493–1514.
[5] P. Ormerod, C. Mounfield, L. Smith, Non-linear Modelling of Crime in the UK, Home Office Research Series, Home Office, London, UK, 2003.
[6] W. Cook, P. Ormerod, Physica A (2003), submitted for publication, pre-print available at www.volterra.co.uk.
[7] J.M. Keynes, The General Theory of Employment, Interest and Money, Macmillan, New York, 1936.
[8] A. Maddison, Monitoring the World Economy 1820–1992, OECD, Paris, 1995.
[9] P. Ormerod, C. Mounfield, Physica A 293 (2001) 573.
[10] I. Wright, cond-mat/0311585, 2003.
[11] M. Ausloos, J. Miskiewicz, M. Sanglier, cond-mat/0403143, 2004.
[12] P. Ormerod, Physica A 314 (2002) 774–785.
[13] J.M. Keynes, The General Theory of Employment, Macmillan, New York, 1936.

UNCORRECTED PROOF