Twentieth Century Shocks, Trends and Cycles in Industrialized Nations

Herman K. van Dijk*
Econometric Institute and Tinbergen Institute, Erasmus University Rotterdam, Burg. Oudlaan 50, 3062PA Rotterdam, The Netherlands,
*(*hkvandijk@few.eur.nl*)*

Report of the Econometric Institute 2004-01

Abstract. Using annual data on real Gross Domestic Product per capita of seventeen industrialized nations in the twentieth century the empirical relevance of shocks, trends and cycles is investigated. A class of neural network models is specified as an extension of the class of vector autoregressive models in order to capture complex data patterns for different countries and subperiods. Empirical evidence indicates nonlinear positive trends in the levels of real GDP per capita, time varying growth rates, switching behavior of individual countries with respect to their position in the distribution of real GDP per capita levels over time. Such evidence presents challenges for economic modelling, forecasting and policy analysis for the long run.

Keywords: Nonlinear trends, shocks, cycles, neural networks

* The author is indebted to Johan Kaashoek for several helpful discussions on the topic of this paper and for his assistance in preparing graphical and empirical results. Helpful comments from Henk Don, Lennart Hoogerheide, and two anonymous referees substantially improved an earlier version of this paper. Any errors remain, of course, the author's responsibility.

1. Introduction

The "business cycle" refers to the alternating states of expansion and contraction of the economies of industrialized nations which occurred periodically in the twentieth century. It is a concept of key interest for all economic actors and policy makers. Consumption and saving decisions of private individuals; investment, production and sales decisions of the industrial sector; and monetary and fiscal policy decisions of the banking and government sector are all based on forecasts of future developments of economic variables, which depend, to a large extent, on the state of the business cycle. For instance, consumer expenditures on durable goods are lower in a recession than in an expansion; investment plans of firms may be postponed when a recession is forecasted and government expenditures may be increased to counter a recession.

The characterization of the business cycle and the analysis of its properties have been the subject of innumerable studies, ever since the seminal work on the statistical testing of business cycles by Tinbergen (1939) and the empirical analysis by Burns and Mitchell (1946). The business cycle has been measured and modelled at the international level, the national level, the industry level and in the financial sector.

The research of Tinbergen on key features of the business cycles in the years before the second World War may be characterized as data driven. There existed already several theories on periodic movements of economic variables, see e.g. Haberler (1937) but extensive empirical research was lacking. Tinbergen's statistical contribution was pioneering in the period before World War II. Tinbergen extended his empirical work and concentrated his research efforts on model building and finding explanatory variables for the business cycles. Given the lack of good data for many countries and given the events during World War II, that was not an easy task. Some details on Tinbergen's work in that early period may be found in Tinbergen and Polak (1950).

Economists today can extend some features of Tinbergen's early empirical study with much longer series of economic data from many more countries. In this context, one may argue that the importance of the business cycle has been overvalued in the twentieth century. A few large shocks to the industrialized economies may have influenced economic behavior and growth more than, short term, cycles. For instance, the devastation due to World War I and II had a profound effect on the time series behavior of real Gross Domestic Product per Capita for countries like Germany and Japan. Adjustment processes of several decades took place within these countries. Such considerations lead to the issue of a proper description and modelling of long run trends in economic variables. Using annual data on real Gross Domestic Product

(GDP[1]), measured in constant PPP dollars per capita, of seventeen industrialized nations for the period 1900-2002, three related issues are discussed in this paper:

- The relevance of deterministic trends versus stochastic trends and their effect on the cycle;
- The importance of shocks versus cycles;
- The existence of long run nonlinear trends.

In order to describe long run behavior, that is economic behavior over a century, it appears a sensible strategy to make use of a wide class of models. Here we make use of the class of neural network models. Neural networks are flexible models for handling complex data patterns of economic variables. This feature has led to the diffusion and implementation of neural network models in the fields of economics and econometrics; see e.g. Gallant and White (1989), Kaashoek and Van Dijk (2003) and White (2000). We will use the class of neural network models as a generalization of the class of linear autoregressive models. Since the work by Sims (1980), vector generalizations of linear autoregressive models have become popular in the empirical econometric literature dealing with time series data like GDP. It will be argued in section 2 of this paper that a nonlinear generalization of autoregressive models is important for long run analysis. For more empirical work in this area of nonlinear modelling of economic series we refer to Granger and Teräsvirta (1993), Teräsvirta (1994) and Franses and Van Dijk (2000) and the references cited there.

Before the specification of neural network models, several stylized facts on the development of the levels and the growth rates of GDP in industrialized nations are discussed in section 2. From a simple graphical and descriptive analysis it is seen that (i) there are nonlinear positive trends in the level of GDP of industrialized nations in the twentieth century; (ii) growth rates exhibit shocks, mainly due to World War I and II; (iii) there is increasing variability in growth rates in the first half of the century and decreasing variability in the second half; (iv) many countries switch substantially from position in the distribution of the level of GDP during the twentieth century; (v) there exists convergence after the oil-price shock. Summarizing, several structural data features seem to be time-varying.

Neural network models are introduced in section 3 in order to analyze these data patterns. The exposition is in terms of a nonlinear

[1] Henceforth, real Gross Domestic Product per capita will be abbreviated as GDP

generalization of (vector) autoregressive models. It is shown in appendix 5 that a linearized neural network model may be interpreted as an autoregressive model with time varying coefficients. Then a link can be made between the class of neural network models and the flexible class of state-space models; see e.g. Harvey (1989).

In section 4 empirical results are presented using estimated neural network models. The presentation of the empirical evidence is organized in the same way as the results in section 2. In section 5 conclusions and some final remarks are presented.

We emphasize that the scope of this paper is limited to reporting empirical evidence on noteworthy stylized facts and to indicate that such facts pose challenges for economic modelling, forecasting and policy analysis for the long run.

2. Stylized Facts

With respect to the issue of deterministic versus stochastic trends, one may argue that many economists until the 1970s were of the opinion that the long term growth path of the economy is exponential - that is, the log of this path is linear - and that deviations from this path lead to temporary cyclical behavior. These deviations might be reduced in length and amplitude by a sensible counter-cyclical policy. This was also the approach of Tinbergen. The interaction of such macro-economic variables like consumption, investment, output and employment was studied, the models were specified in a Keynesian way. The practice of counter-cyclical policy was to use government expenditures to stimulate investment and consumption in a recession in order to create employment. In this context stochastic shocks have a temporary effect. However, there also exists the Schumpeterian view that stochastic shocks, mostly due to the introduction of new technology, have a permanent effect. Such shocks are the driving force of western economies based on the market mechanism, see Schumpeter (1939). This view is also supported by Lucas (1977).

The two different views may be explained using a simple time series model where output is connected to its own lagged values. The two views are shown in Figure 1 for the log of GDP of the USA. In panel (a) one observes that a linear trend gives a good indication of the increase in log of GDP in the USA in the twentieth century. The model appears to be trend stationary (TS). In panel (b) one may observe however that a so-called stochastic trend describes the variation in the series much better than the linear deterministic trend. The model appears to be a difference stationary (DS) model and as a result panel (c) suggests

that there was no major business cycle in the 1930s. The depression of the thirties is apparently a consequence of unexpected shocks (for more details and a technical explanation we refer to a paper by Schotman and Van Dijk (1991)). Clearly, different ways of detrending yield different cyclical behavior. When one is interested in long run analysis, the result may be sensitive to the *linear* specification of the trend. We will investigate in the next section the relation between a nonlinear trend and the implied cycle using the levels of GDP.

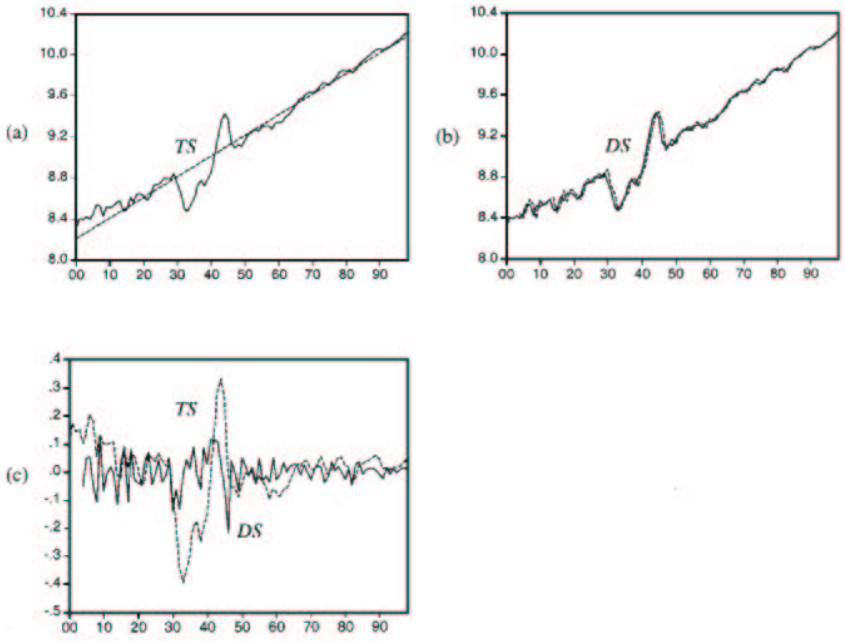

Figure 1. (a) USA log of real GDP per capita, and deterministic trend (TS = trend stationary); (b) USA log of real GDP per capita and stochastic trend (DS = difference stationary); (c) Residuals of deterministic trend model (dashed) and stochastic trend model (solid)

As an exploratory step in the data analysis of all countries some results on the dynamic behavior of levels and growth rates of GDP are shown in Figure 2 and Tables 1 and 2. The data are taken from the well-known Maddison data set and from the OECD data bank (www.sourceoecd.org)[2]. We emphasize that henceforth we make use of levels of GDP since these are a relevant indicator of the wealth of

[2] The GDP data are measured in constant PPP dollars. Using the PPP hypothesis implies limitations with respect to comparability of data over time and between countries.

a nation. We summarize the main data features in the following five points:

– *nonlinear positive trends*

The most important fact is the substantial increase in the level of GDP for all countries. From Figure 2 and Table 2 it is seen that, roughly speaking, the level of GDP doubled every period of thirty years, which is the equivalent of a generation. However, this increase is not constant over time and over countries. Nonlinear effects may be observed in the sense that the increase in levels is fast after World War II for many countries but also slows down in the nineties, in particular for Germany and Japan. This is seen in Figure 2 and Table 1 where results for subperiods are reported.

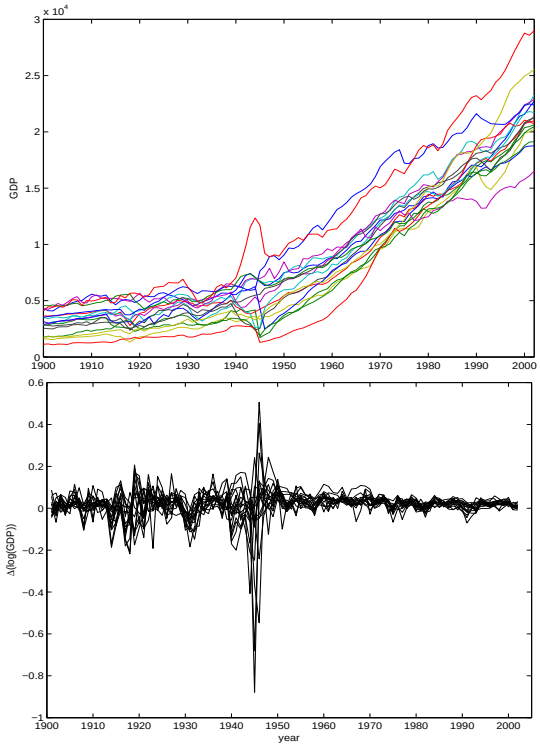

Figure 2. Levels and growth rates of GDP of all 17 countries

– *growth rates exhibiting shocks and catch-up effects*

Figure 2 shows clearly that shocks to growth may be interpreted as outliers and that these occur mainly due to World War I and II. To

a lesser extent the so-called Roaring Twenties and the Depressed Thirties are responsible for varying growth patterns. Note, from Table 1, the substantial growth of Japan, Austria, Italy, France and the Netherlands after World War II. This may be interpreted as a so-called catch-up effect.

– *changing variability of growth rates*

The bottom of Figure 2 and Table 1 show that there is an increase in variability of growth rates in the first half of the twentieth century and a decrease in the second half. This is an indication that the amplitudes (and lengths) of possible cycles are not constant during the twentieth century. The lower level of growth for most countries in the nineties compared with the growth in the period after World War II is also noteworthy.

– *winners and losers and switching behavior*

Table 2 shows that the level of GDP has on average increased sevenfold. The USA is the absolute winner in this league of 17 nations in the twentieth century. Norway shows the most dramatic positive move: from low position 14 in 1900 it moved upwards to the second position in 2000. In addition to the USA and Canada, the countries in North-Western Europe have done very well. Relatively speaking, the UK and New Zealand are big losers. We note that Germany ends up at the one but lowest position in 2002. This is mainly due to the reunification of East- and West Germany in 1991. We emphasize, however, that after the top four in the third column of Table 2, there is a large group of countries in which the levels of GDP differ only little. The results of the subperiods indicate switching behavior with respect to the wealth of nations in the twentieth century, where we measure wealth as level of GDP. This may also be deduced from the subperiod analysis presented in Table 1.

– *convergence*

It is clear that during the whole period of the twentieth century, convergence[3] in levels of GDP of all industrialized nations did not occur. It appears, see the final row of Table 1, that convergence improved in the subperiods after the oil price shock of 1973.

[3] Convergence is here defined as σ-convergence. For details on different ways of defining convergence, we refer to Paap and Van Dijk (1998).

Table 1. Mean and standard deviation of growth rates, $\Delta(log(GDP))$, in percentages

Mean of $\Delta(log(GDP))$ in percentages

country	1901-2002	01-13	14-18	19-28	29-38	39-45	46-72	73-89	90-02
Australia	1.63	1.90	−1.75	0.21	0.82	2.30	2.26	1.92	2.22
Austria	1.92	1.41	−6.09	3.15	−0.27	−10.35	6.76	2.43	1.81
Belgium	1.71	0.94	−7.76	3.44	−0.61	−1.55	3.69	2.20	1.73
Canada	2.09	3.25	−0.02	0.67	−1.28	6.43	2.44	2.25	1.49
Denmark	2.02	2.00	−2.46	2.54	1.85	−1.84	3.78	1.76	1.73
Finland	2.47	1.81	−8.65	4.89	2.82	−0.56	4.21	2.84	1.38
France	1.96	1.47	−7.49	3.61	0.07	−7.86	5.90	2.03	1.38
Germany	1.75	1.54	−3.91	2.93	1.73	−2.42	3.61	1.95	1.21
Italy	2.34	2.78	5.59	−0.82	0.95	−7.79	6.21	2.72	1.38
Japan	2.84	1.24	3.69	1.63	2.06	−8.54	7.83	3.02	1.10
Netherlands	1.78	0.85	−3.77	4.14	−0.85	−9.57	5.80	1.67	2.01
New Zealand	1.31	1.39	−0.95	0.23	2.28	0.99	1.98	0.96	1.25
Norway	2.62	1.96	−1.80	2.58	3.07	−1.05	4.02	3.01	2.63
Sweden	2.07	1.45	−4.01	3.15	2.56	2.34	3.14	1.77	1.49
Switzerland	1.81	1.34	−2.32	4.39	0.34	2.76	3.12	0.99	0.46
UK	1.46	0.70	2.07	−0.73	1.56	1.69	1.91	2.19	1.70
USA	1.91	1.99	1.30	2.04	−0.69	9.25	1.13	2.17	1.76
Mean	1.99	1.65	−2.26	2.91	0.97	−1.52	3.99	2.12	1.58

Standard deviation of $\Delta(log(GDP))$ in percentages

country	1901-2002	01-13	14-18	19-28	29-38	39-45	46-72	73-89	90-02
Australia	3.50	3.42	6.97	3.86	4.15	6.46	2.54	1.95	1.39
Austria	10.54	2.22	7.60	6.52	6.74	34.78	5.18	1.81	1.25
Belgium	4.68	0.77	10.39	9.74	2.97	7.35	1.95	2.14	1.37
Canada	5.31	4.69	8.40	8.51	8.92	6.80	2.92	2.27	2.23
Denmark	4.18	1.47	6.06	5.44	2.56	10.10	3.25	2.16	1.31
Finland	4.79	2.46	7.46	8.05	4.20	6.06	2.67	2.03	3.79
France	7.31	4.19	11.37	11.13	5.19	11.14	7.20	1.46	1.23
Germany	8.84	1.97	8.08	11.34	8.16	15.40	12.11	1.74	1.17
Italy	6.11	4.05	4.98	7.20	4.09	11.41	4.89	2.12	1.08
Japan	8.42	4.32	6.78	4.99	4.98	26.95	2.47	1.96	1.72
Netherlands	7.75	2.46	4.73	6.65	3.97	15.19	9.45	1.73	1.32
New Zealand	4.95	5.74	1.20	6.59	7.36	4.29	5.57	2.57	2.49
Norway	4.00	1.58	5.86	6.64	4.69	6.92	3.05	1.98	1.29
Sweden	3.10	2.86	5.52	3.60	3.79	3.17	1.95	1.44	2.42
Switzerland	3.84	0.69	5.30	3.28	3.31	9.56	2.91	2.51	1.51
UK	3.25	2.24	2.86	6.22	3.27	5.59	2.09	2.26	1.56
USA	5.78	5.54	8.62	4.36	9.93	8.04	5.74	2.44	1.51
Mean	5.67	2.99	6.60	6.12	5.14	11.14	4.47	2.04	1.69

Table 2. Initial ordering, average of yearly changes and final ordering of levels of GDP

Initial order (1900)	Average yearly change	Final order (2002)
UK (4593)	USA (2.42)	USA (29002)
New Zealand (4320)	Norway (2.31)	Norway (25584)
Australia (4299)	Canada (1.99)	Canada (23283)
USA (4096)	Denmark (1.94)	Denmark (22897)
Belgium (3652)	Japan (1.90)	Australia (22742)
Netherlands (3533)	Switzerland (1.83)	Switzerland (22378)
Switzerland (3531)	Sweden (1.83)	Netherlands (21710)
Germany (3134)	Finland (1.81)	Sweden (21369)
Denmark (2902)	Australia (1.79)	France (21225)
Austria (2901)	France (1.78)	Belgium (20982)
France (2849)	Netherlands (1.76)	Japan (20710)
Canada (2758)	Austria (1.72)	Austria (20628)
Sweden (2561)	Italy (1.69)	UK (20466)
Norway (1762)	Belgium (1.68)	Finland (20274)
Italy (1746)	UK (1.54)	Italy (19135)
Finland (1620)	Germany (1.52)	Germany (18755)
Japan (1135)	New Zealand (1.18)	New Zealand (16520)
mean values: 3023.06	1.81	21627.06

Descriptive analysis is useful as a first step in a modelling strategy. As a next step we specify in the next section a flexible class of models in order to capture the complex data patterns listed above. We emphasize that we use levels of GDP. Application of a log transformation did not affect the nonlinear patterns found using the neural network models used in this paper.

3. Exploring dynamic patterns by neural network models

In this section a class of neural network models is specified as a generalization of the class of (vector)autoregressive models. Also, a connection is made between the class of neural network models and the class of state-space models.

Neural networks are flexible models for handling complex data patterns of economic variables. This feature has led to the diffusion and implementation of neural network models in the fields of economics and econometrics; see e.g. Kaashoek and Van Dijk (2003).

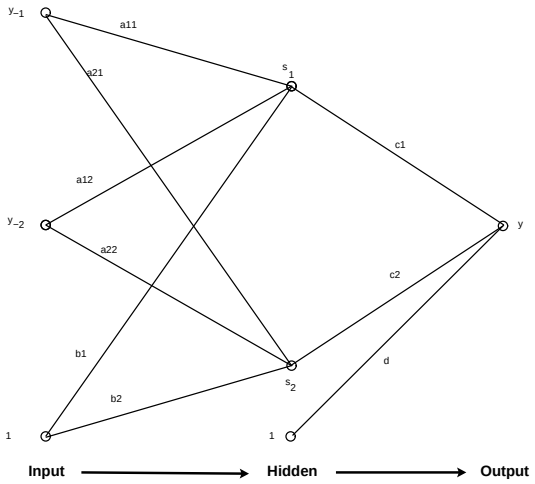

Figure 3. Graph of a neural network

Consider the graph in Figure 3. This graph represents a so-called *feed forward* neural network with 7 cells (or nodes) which are partitioned into three particular groups or layers. The 'input' layer contains cells which transmit or *forward* weighted values of lagged dependent variables and a weighted value of the constant as signals to the so-called 'hidden' layer. This 'hidden' layer forwards these signals, with again particular weights attached, together with a weighted constant signal to an output cell. The value at the output cell is denoted by y. This network transmits signals from the left (input) to the right (output) along vertices connecting the cells. This explains the use of the term *feed forward*.

Let values of the input cells be denoted as y_{-i}, $i = 1, \cdots, I$, let the weight between input cell i and hidden cell h be a number a_{ih}, $(h = 1, \cdots, H)$ and the weight of the constant input signal be b_h. Then at hidden cell h will arrive a signal with strength:

$$s_h = \sum_{i=1}^{I} a_{ih} y_{-i} + b_h \tag{1}$$

If no transformation takes place at the hidden layer cells, the signals arriving at the hidden layer are transported to the output cells along vertices with weights c_h. With inclusion of a constant hidden signal with weight d, a signal will arrive at the output cell with strength or value:

$$y = \sum_{h=1}^{H} c_h s_h + d \tag{2}$$

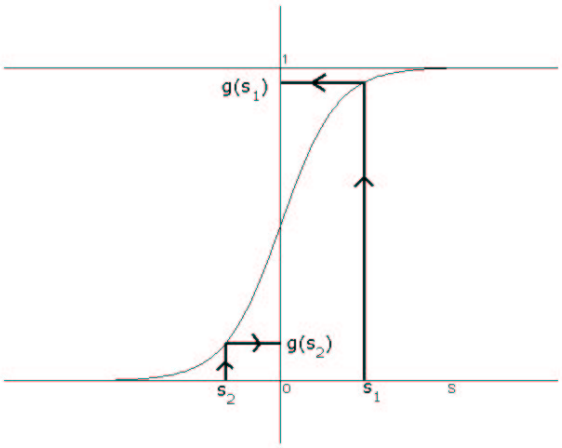

Figure 4. Graph of a logistic function

Now equation (2) is clearly a linear equation in y, its past history $(y_{-1}, ..., y_{-I})$ and a constant term:

$$y = \sum_{i=1}^{I} \phi_i y_{-i} + \phi_0 \tag{3}$$

where $\phi_i = \sum_{h=1}^{H} c_h a_{ih}, i = 1, \cdots, I$ and $\phi_0 = \sum_{h=1}^{H} c_h b_h + d$. This equation is a linear autoregressive model with dependent variable y and a constant term. However, one cannot distinguish between the effect of the input cells and the effect of the hidden cells on the dependent variable y. Hence, if nothing happens at the hidden layer cells this particular feed forward neural network is nothing more than a not very smartly specified linear autoregressive model.

Actually, in a neural network, signals arriving at the hidden cells are transformed by some non-linear transformation, the so-called 'activation' function. This transformation is the same for all hidden cells. The activation function is a monotone increasing and bounded function, e.g.:

$$g(s_h) = \frac{1}{1 + e^{-s_h}} \tag{4}$$

which is the well known logistic function, with range $(0, 1)$; see Figure 4.

A particular value of the logistic function indicates the extent to which a hidden cell is activated. A signal is transmitted from hidden cell h:

$$\tilde{s}_h = g(s_h) \tag{5}$$

and equation (2) changes to:

$$y = \sum_{h=1}^{H} c_h \tilde{s}_h + d \tag{6}$$

Thus H linear combinations, each from I input signals, are carried through the logistic activation function to give H values of 'hidden' signals. Next, linear combinations of the H 'hidden' signals and a constant hidden signal are carried through to form the output y.

We note that the neural network describes the situation at one moment in time and it indicates a deterministic relation. By adding a subscript t and an error term ε_t one can write the whole network in matrix notation as:

$$y_t = c'\tilde{s}_t + d + \varepsilon_t \tag{7}$$

$$\tilde{s}_t = G(s_t) = G(A'Y_{t-1} + b) \tag{8}$$

where $A = (a_{ih})$ is a $I \times H$ matrix of weights of the input values and b is now a vector. $Y_{-1} = (y_{-1}, ..., y_{-I})'$ is a $I \times 1$ vector of input values which contains the past history of y and $s = (s_1, ..., s_H)'$ is a $H \times 1$ vector of input signals. Further, G is a vector function composed of identical activation functions $G(s) = [g(s_1), ..., g(s_H)]'$; $\tilde{s}$=$(\tilde{s}_1, ..., \tilde{s}_H)'$ is a $H \times 1$ vector of hidden signals and $c = (c_1, ..., c_H)'$ is a $H \times 1$ vector of weights given to the hidden signals.

We note a certain analogy with a state space model. In that case $\tilde{s}_t$ is the unobserved component in the so-called measurement equation (7). When one inserts $\tilde{s}_{t-1}$ as explanatory variable in equation (8) one obtains a recurrent neural network which is a special case of a state-space model; see Harvey (1989). It is a topic for further research to compare the flexibility and estimability of neural nets compared with state-space models.

An example is a neural network with two inputs and two hidden cells. The graph of this neural network is given in Figure 3 and it represents a model with the following mathematical specification:

$$y_t = d + \frac{c_1}{1 + e^{-a_{11}y_{t-1} - a_{12}y_{t-2} - b_1}} + \frac{c_2}{1 + e^{-a_{21}y_{t-1} - a_{22}y_{t-2} - b_{12}}} \tag{9}$$

if $\tilde{s}_h$ is defined as output of hidden layer cell h, then one can write

$$y_t = d + c_1\tilde{s}_1 + c_2\tilde{s}_2 \tag{10}$$

hence

$$y_t = d(1 - \tilde{s}_1 - \tilde{s}_2) + (c_1 + d)\tilde{s}_1 + (c_2 + d)\tilde{s}_2 \tag{11}$$

and it follows that the range of y_t is bounded. We note that this model is closely related to a threshold autoregressive model; see e.g. Teräsvirta (1994).

One may also add a deterministic function of time to a neural network model. Such functions of time, like a simple linear trend, play an important role in dynamic autoregressive models. Also, in a neural network model, time dependency can be included explicitly. We give two examples where the first is a rather general formulation (see equation (12)) and in the second (see equation (13)) the time component is specified in an additive way:

$$y_t = nn(t, y_{t-1}, y_{t-2}) \tag{12}$$

$$y_t = nn_1(t) + nn_2(y_{t-1}, y_{t-2}) \tag{13}$$

where $nn_1(t)$ represents a neural network with input variable t only, and $nn_2(y_{t-1}, y_{t-2})$ is a neural network with lagged input variables.

In appendix 5 it is shown that a linearized version of equation (13) can be written as a linear autoregressive model with time-dependent coefficients. Thus all models based on equation (13) possess derived features such as stability measures, cycles and trend components, that are all time dependent.

More details are given in Appendix 5.

Flexibility and pruning of a network

The flexibility of three layer feed forward neural networks is well documented; see e.g. Gallant and White (1989), Hecht-Nielsen (1987) and White (2000). This flexibility follows from the so-called 'universal approximation' property of neural nets. This approximation rests on Kolmogorov's superposition theorem; see Kolmogorov (1957) and Hecht-Nielsen (1987). The Kolmogorov theorem provides, however, a justification for the existence of an approximation in just the opposite way from the one that is needed by a practical user of neural nets. That is to say, in the approximation theorem the number of layers and cells are given while the specific form of the activation function is unknown. For the practical user of neural nets one encounters the opposite case: the activation function is known, to some extent, but the number of hidden cells has to be determined. Thus, while the number of output cells is fixed by 'which series are to be modelled', the number of input cells I and, more importantly, the number H of hidden layer cells, are not fixed beforehand. This adds to the flexibility of neural network modelling but it has its price. Adding more hidden cells, may result in overfitting. Instead of filtering out noise processes, these are included

in the estimated model. As a result, bad predictive behavior occurs; see Bishop (1995). A procedure for reducing the size of the network is explained in Kaashoek and Van Dijk (2003) and (2002).

Estimating parameter values

A generally accepted optimization principle is:

$$\min_{\theta} \sum_{t} ||y_t - \hat{y}_t(\theta)||^2 \tag{14}$$

where $||.||$ is the Euclidean norm. For the case of a neural net $\theta = (A, b, c, d)$ and it follows that the criterion function:

$$\sum_{t} ||y_t - c'G(A'Y_{t-1} + b) - d||^2 \tag{15}$$

has to be minimised over (A, b, c, d). Well-known methods for numerical optimization are the simplex method and the BFGS method, a gradient method; see Press et al. (1988). These methods have been successful in estimating neural network parameters.

Dynamic properties of neural nets

Stability analysis in an autoregressive model is essentially static, that is, constant over time, while the nonlinear equivalent given by a neural net model, is by nature dynamic. The linearization of a neural network involves time dependent coefficients (see Appendix 5) and hence dynamic quantities like the Lyapunov exponents, which characterize stability, are time varying; for more details see Kaashoek and Van Dijk (1994).
Dynamic behavior of the neural network models, as defined in equations (7) and (8), may be analyzed as follows. One generates in sample, and possibly out of sample, dynamic forecasts, denoted as $\hat{y}_t$ and referred to as *orbits*, by means of:

$$\hat{y}_t = nn(\hat{y}_{t-1}) \tag{16}$$

where nn is the relevant neural network function. So, an orbit is generated by iterating the estimated neural network equation without its error term, that is, by choosing $\varepsilon_t = 0$ for all t.

Some orbits (together with the actual data) are depicted in Figure 9 and Figure 10 . The initial value $\hat{y}_1$ is taken from actual data, $\hat{y}_1 = y_1$. The graphs indicate that in all cases a nonlinear trend gives a very good fit to the data. We suggest that our analysis may be used as a

first step in a more detailed analysis of a nonlinear parametric form of a time series model.

4. Empirical Results

4.1. Nonlinear Trends and Shocks

In the first and third columns of Figure 5 the data of GDP for all seventeen nations are shown together with their orbits which are in this case equal to an $nn(t;1)$ approximation[4]. The orbits are denoted by the dotted lines. In the second and fourth columns the residuals $GDP-nn(t;1)$ are presented. The orbits exhibit a nonlinear pattern for several countries in the sense that fast growth occurs after World War II and a saturation phenomenon appears in the nineties. This feature may be observed for countries like Austria, Belgium, France, Germany, Japan, Netherlands, New Zealand, Sweden and Switzerland. The rapid growth after World War II for countries like Japan and - to a lesser extent - Germany appears to be a catch-up effect towards a long run path of lower average growth. We note that this catch-up effect is for Germany known as the "Wirtschaftswunder". The graphs in the second and fourth columns indicate that the behavior of GDP in deviation of a nonlinear trend shows cyclical but non-constant patterns. Thus it is of interest to investigate a possibly, time varying cyclical behavior. Further, for many countries one needs a shock component to explain the large outliers in these series. This holds in particular for countries that were actively involved in either World War I or World War II or in both; these include Austria, Belgium, Finland, France, Germany, Italy, Japan, Netherlands, UK and USA.

4.2. Time Varying Cyclical Behavior

4.2.1. *The case of USA and Germany*

Using a rolling window of size 30, the differences $GDP - nn(t;1)$ for the countries USA and Germany are fitted on an autoregressive model with 5 lags. These static models proved to have two cycles. For each of these cycles lengths of periods of oscillation and amplitudes (= norms) were calculated. The results are shown in Figure 6a-6b. At the end of the period the oscillations with largest period (upper graphs for each country) become shorter. The results should be interpreted with care; they appear to be sensitive to the lag length chosen. Also,

[4] Henceforth, the notation $nn(x,\ldots;k)$ indicates a neural net where $x,\ldots$ are the input variables and k is the number of hidden cells.

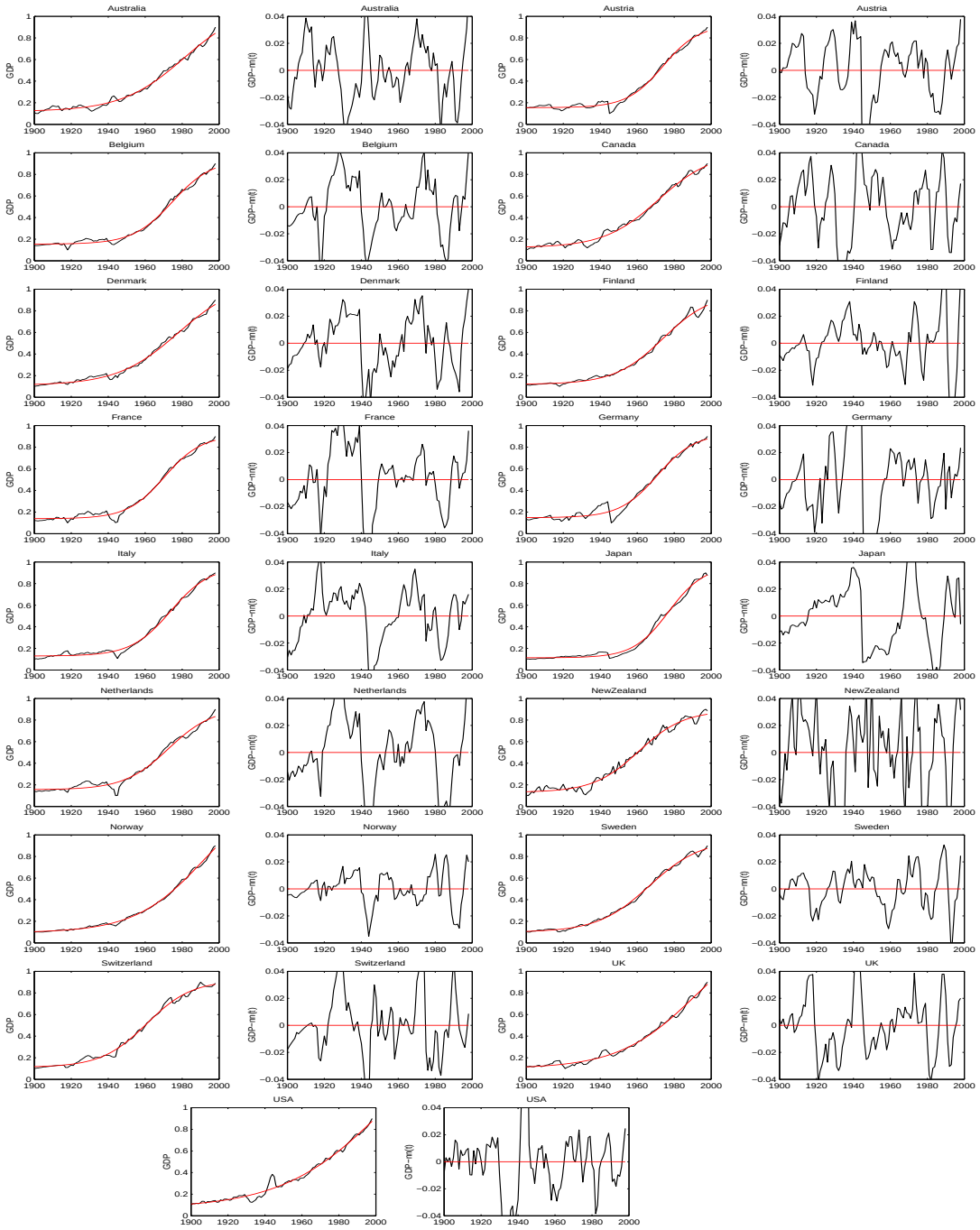

Figure 5. GDP and $nn(t;1)$ (left), and residuals $GDP - nn(t;1)$ (right)

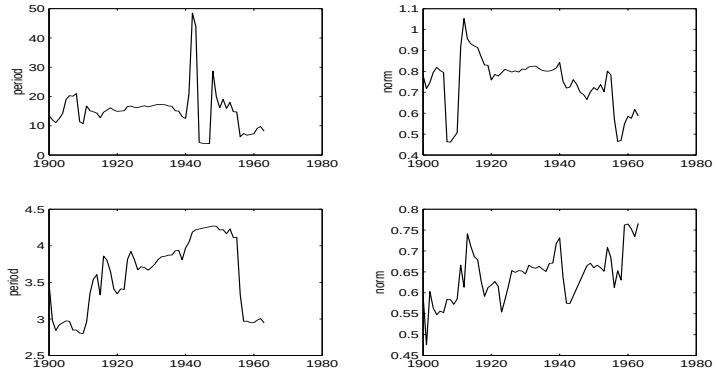

Figure 6a. Germany: varying periods of oscillation and amplitudes

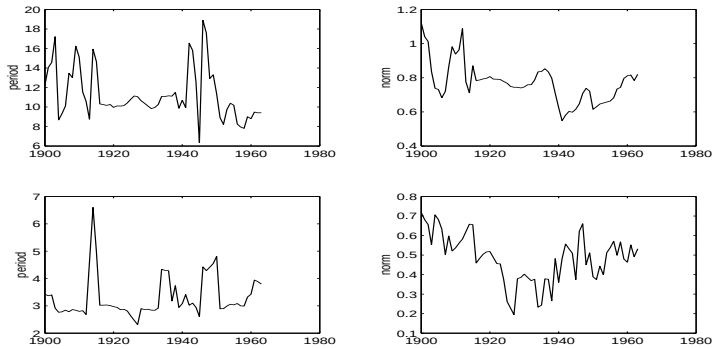

Figure 6b. USA: varying periods of oscillation and amplitudes

when a neural network is fitted to the difference between GDP and the $nn(t;1)$, similar results appear as presented in Figure 6a-6b. These results are also rather sensitive to the specification of the nn model. There is a need for proper inference on this topic.

4.2.2. *Leading and Lagging Cyclical behavior*
In order to investigate whether some countries, notably the USA, may be leading in cyclical behavior over the whole century, some results are shown in Figures 7a-7b. We compare the USA and the UK in panel (a), left graph; the USA and Germany in panel (a), right graph; Germany and France in panel (b), left graph, and Germany and UK in panel (b), right graph. The residuals $GDP - nn(t;1)$ are plotted as in the second and fourth columns of Figure 5, for each of the countries mentioned. It is seen that the UK and Germany suffered less from a recession then the USA during the thirties and fifties. The USA expanded more in the forties due to the war effort. In a similar way one may observe that Germany has expanded more during the thirties compared to

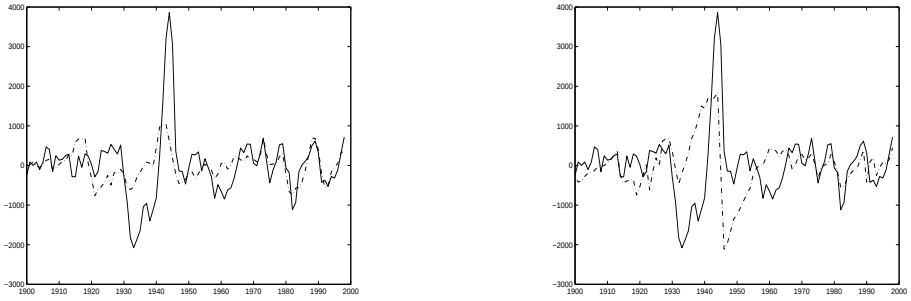

Figure 7a. Residuals $GDP - nn(t;1)$: USA (solid) against (left) UK (dashed) and (right) Germany (dashed)

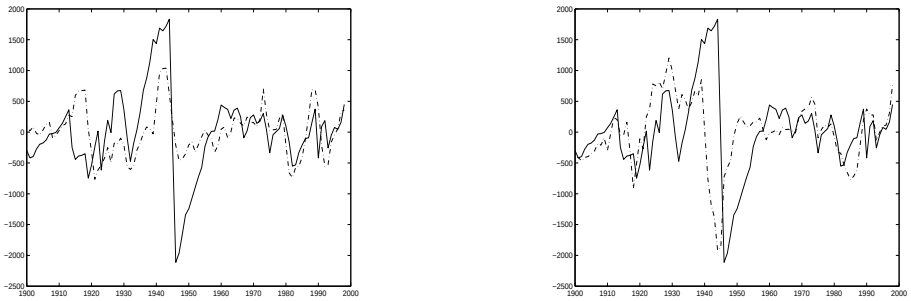

Figure 7b. Residuals $GDP - nn(t;1)$: Germany (solid) against (left) UK (dashed) and (right) France (dashed)

France and the UK. After World War II, France recovered earlier than Germany and the UK had a less severe recession than Germany. A more general conclusion is that before World War II, the USA and the European countries (UK and Germany) are out of phase with respect to recession and expansion while after the oil price shock of 1973, all four countries follow a more or less common cyclical pattern. Thus, there is some evidence of convergence to a common world cycle of industrialized nations during the twentieth century. Common cycles have been investigated by Vahid and Engle (1993, 1997) and more recently by Hecq, Palm and Urbain (2000a, 2000b). It is a topic of further research to investigate whether the empirical results on cyclical behavior over a century, reported in the present paper and based on the neural network approximation - filtering the nonlinear trend -, also hold in the model specifications and inferential procedures proposed by these authors.

4.3. Convergence Results

In Figure 8 the following data are shown. First, consider a rolling window of size 30; for each window the R^2 of GDP of each country with respect to GDP of the USA is calculated. At the end of the period, the countries with lowest R^2 are New Zealand and Switzerland and the next 'worse' cases are Sweden and Finland. The major feature of this figure is that after World War II countries moved in a similar way as the USA with respect to the level of GDP. The parallel behavior of the levels is a major indication of GDP growth convergence of industrialized nations after 1960. Evidence of growth convergence is presented in Barro (1991), Baumol (1986) and Sala-i-Martin (1994). However, using a much

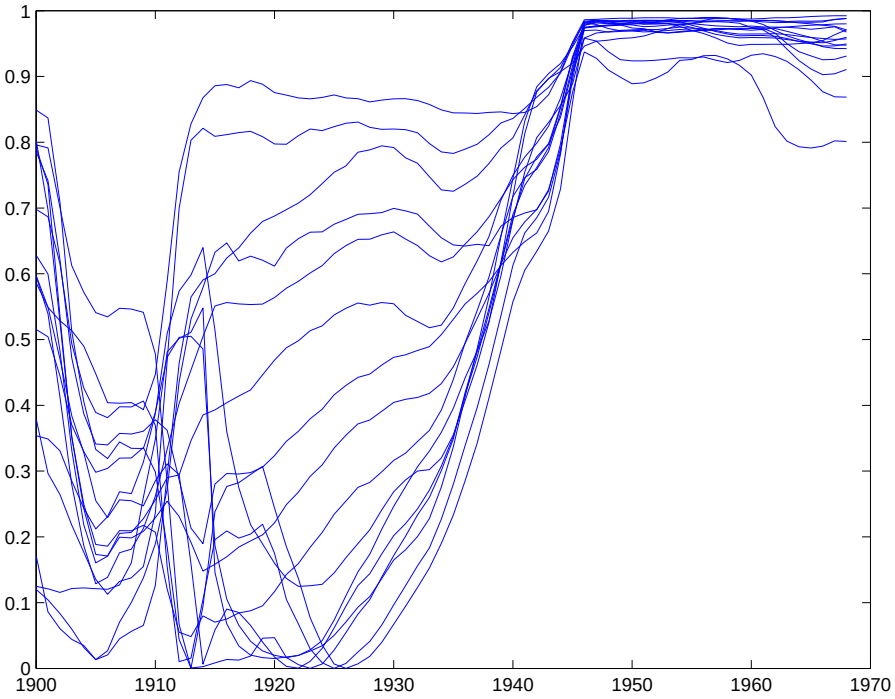

Figure 8. R^2 of GDP of each country with respect to GDP of USA (window of 30 years starting in 1900)

larger panel of eighty nations - including developing nations - for the post World War II period, Quah (1993) and Paap and Van Dijk (1998) find substantial evidence of bimodality in the distribution of the levels of GDP. It would be interesting, if possible, to extend the data of those eighty countries to the period before World War II in

order to investigate the issue of persistent bimodality in the distribution of the levels of GDP. It is of interest to investigate the behavior of GDP for a subset of countries. We compare the USA and Germany; leading industrialized countries. In Figure 9 we compare the orbits of the USA and Germany with their respective data patterns. For each of the two countries a neural net $nn(t;3) + nn(y_{t-1}, y_{t-2}; 1)$ is estimated and these estimates are used to predict their own orbit as well the orbit of the other country. In the upper left hand corner one finds the USA data pattern compared with the USA orbit; in the upper right hand corner the German orbit, with initial values taken from USA data, is compared with the USA data pattern. In the lower left hand corner the German data pattern is compared with the USA orbit, with initial values taken from German data; and finally in the lower right hand corner the German data pattern is compared with the German orbit. It is seen that the orbits can track the data patterns rather precisely for their own countries. From the off-diagonal figures it is seen that the USA was doing much better in terms of GDP during and right after World War II and that Germany had a spectacular catching-up process from the fifties onwards. Recently, Germany has a slow down in growth due to the reunification and the USA is leading again.

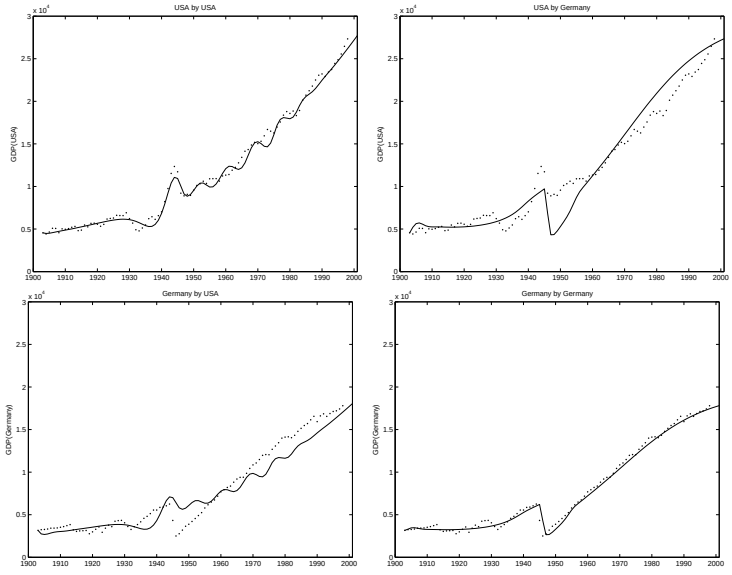

Figure 9. Orbits (solid lines) of USA and Germany GDP data, based on neural network estimates compared with actual GDP data (dashed)

In a similar way, we compare Norway and Sweden. Using Swedish GDP data, a neural net $nn(t;3) + nn(y_{t-1}, y_{t-2}; 1)$ is estimated. The estimated network is used for one period predictions and orbits calcu-

lations of Norwegian data. In Figure 10 the estimates and orbits are compared with real Norwegian data. The leading position of Sweden is recently reduced and Norway has made a spectacular progress in the nineties.

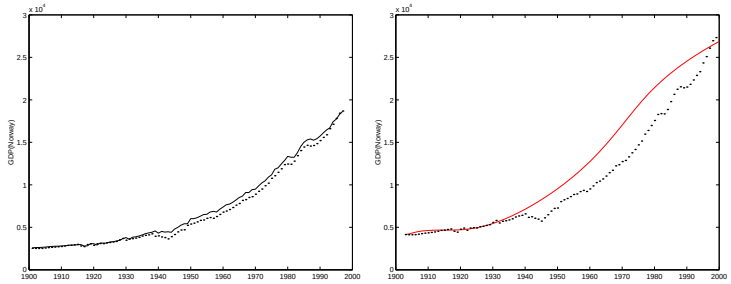

Figure 10. Actual GDP of Norway (dashed) compared with one period prediction (left) and orbit (right) of Norway using a neural network estimated with Swedish GDP data

5. Conclusions

A class of neural network models has been specified for the analysis of nonlinear dynamic patterns in economic time series of seventeen industrialized nations in the twentieth century. The level of real Gross Domestic Product per capita of these nations has, on average, increased seven-fold. However, this increase in levels has not been monotone over time and over countries. Substantial empirical evidence has been obtained of time varying growth rates. Many countries also switch substantially from position in the distribution of the level of GDP during the twentieth century. Growth convergence appears to occur after the sixties. Further, there appears to be evidence of a common world cycle since the early seventies. Shocks to the economies like World War I and II appear to be more important than cyclical behavior. Given the descriptive character of the present paper, one should be careful in drawing strong conclusions. However, it seems fair to argue that the empirical evidence presented in this paper constitutes interesting challenges for economic modelling - finding relevant explanatory variables and meaningful theories for nonlinear processes -, for economic forecasting and policy analysis for the long run.

Appendix A. Linearization of neural network function

We investigate in this appendix the connection between the class of linear autoregressive models and the class of neural network models as defined in equation (13):

$$y_t = nn_1(t) + nn_2(y_{t-1}, y_{t-2}) \tag{17}$$

At each time point t_0, the linearization in t of $nn_1(t)$ is just $a(t_0)+b(t_0)t$, and the autoregressive part $nn_2(y_{t-1}, y_{t-2})$ may be linearized as follows:

$$\begin{aligned} nn_2(y_{t_0-1}, y_{t_0-2}) + (\frac{\partial nn_2}{\partial y_{t-1}}|_{t=t_0})(y_{t-1} - y_{t_0-1}) \\ + (\frac{\partial nn_2}{\partial y_{t-2}}|_{t=t_0})(y_{t-2} - y_{t_0-2}) \end{aligned} \tag{18}$$

Hence, the linearization at $t = t_0$ of equation (13) can be written as:

$$\begin{aligned} y_t &= a^*(t_0) + b(t_0)t + (\frac{\partial nn_2}{\partial y_{t-1}}|_{t=t_0})(y_{t-1} - y_{t_0-1}) \\ &\quad + (\frac{\partial nn_2}{\partial y_{t-2}}|_{t=t_0})(y_{t-2} - y_{t_0-2}) \\ &= a^*(t_0) + b(t_0)t + C_1(t_0)y_{t-1} + C_2(t_0)y_{t-2} \end{aligned} \tag{19}$$

where $a^*(t_0)$ equals $a(t_0) + nn_2(y_{t_0-1}, y_{t_0-2})$.
One can solve the differential equation (19). The characteristic equations from the resulting linearization should have time-dependent real roots and/or time-dependent complex roots. Thus we emphasize that all models may have a time varying trend and may have time varying cycles.

References

Barro, R.J., 1991, Economic growth in a cross section of countries, *Quarterly Journal of Economics*, 106, 407-443.

Baumol, W.J., 1986, Productiviy growth, convergence, and welfare: What the long-run data show, *American Economic Review*, 76, 1072-1085.

Bishop, C.M., 1995, *Neural Networks for Pattern Recognition*, Clarendon Press, Oxford.

Burns, A.F. and W.C. Mitchell, 1946, *Measuring Business Cycles*, NBER , New York.

Franses, P.H.B.F. and D.J.C. van Dijk, 2000, *Nonlinear time series models in empirical finance*, Cambridge University Press, Cambridge.

Gallant, A.R. and H. White, 1989, There exists a neural network that does not make avoidable mistakes, in *Proc. of the International Conference on Neural Networks, San Diego, 1988*, IEEE Press, New York.

Granger, C.W.J. & T. Teräsvirta, 1993, *Modelling Nonlinear Economic Relationships*, Oxford University Press Inc., New York.

Haberler, G., 1937, 1963, *Prosperity and Depression: A Theoretical analysis of Cyclical Movements*, 5th ed. New York Atheneum; originally published by the League of Nations (1937).

Harvey, A.C., 1989, *Forecasting, structural time series models and the Kalman filter*, Cambridge University Press, Cambridge.

Hecht-Nielsen, R., 1987, Kolmogorov mapping neural network existence theorem, in *IEEE First International Conference on Neural Networks*, 11-13.

Hecq, A., F.C. Palm and J.-P. Urbain, 2000a, Permanent-Transitory Decomposition in VAR Models with Cointegration and Common Cycles, *Oxford Bulletin of Economics and Statistics*, 62, 511-532.

Hecq, A., F.C. Palm and J.-P. Urbain, 2000b, Testing for Common Cyclical Features in Nonstationary Panel Data, in B.H. Baltagi (ed.) *Advances in Econometrics: Nonstationary Panels, Panel Cointegration and Dynamic Panels*, Vol. 15, JAI Press, 131-160.

Kaashoek, J.F. and H.K. van Dijk, 1994, A neural network applied to the calculation of Lyapunov exponents, *Econometric Reviews*, 13(1), 123-137.

Kaashoek J.F. and H.K. van Dijk, 2002, Neural Network Pruning Applied to Real Exchange Rate Analysis, *Journal of Forecasting*, 21, 559-577.

Kaashoek J.F. and H.K. van Dijk, 2003, Neural Networks: an Econometric Tool, in D.E.A. Giles (ed.) *Computer-Aided Econometrics*, Marcel-Dekker, New York, 351-384.

Kolmogorov, A.N., 1957, On the representation of continuous functions of many variables by superposition of continuous functions of one variable and addition, *American Mathematical Montly Translation*, vol.28, 1963. (Russian original in *Doklady Akademii Nauk SSSR*, 144.)

Lucas, R.E., 1977, Understanding business cycles, In: R.E. Lucas, 1989, *Studies in business cycle memory*, Basil Blackwell, Oxford.

Paap, R. and H.K. van Dijk, 1998, Distribution and mobility of wealth of nations, *European Economic Review*, 42, 1269-1293.

Press, W.H., B.P. Flannery, S.A. Teukolsky and W.T. Vetterling, 1988, *Numerical Recipes*, Cambridge University Press, Cambridge.

Quah, D., 1993, Empirical cross-section dynamics in economic growth, *European Economic Review*, 37, 426-434.

Sala-i-Martin, X., 1994, Cross-sectional regression and the empiries of economic growth, *European Economic Review*, 38, 739-747.

Schotman, P. and H.K. van Dijk, 1991, On Bayesian roots to unit roots, *Journal of Applied Econometrics*, 6, 387-401.

Schumpeter, J.A., 1939, *Business cycles: A theoretical, historical, and statistical analysis of the capitalist process*, New York, MacGraw-Hill.

Sims, C., 1980, Macroeconomics and reality, *Econometrica* 48, 1-48.

Teräsvirta, T., 1994, Specification, estimation, and evaluation of smooth transition autoregressive models, *Journal of the American Statistical Association*, 89, 208-218.

Tinbergen, J., 1939, *Statistical testing of business-cycle theories, Volume 1: A method and its application to investment activity; Volume 2: Business cycles in the United States of America, 1919-1932*, Geneva: League of Nations.

Tinbergen, J. and J.J. Polak, 1950, *The dynamics of business cycles, A study in economic fluctuations*, The University of Chicago Press.

Vahid, F. and R.F. Engle, 1993, Common Trends and Common Cycles, *Journal of Applied Econometrics*, 8, 341-360.

Vahid, F. and R.F. Engle, 1997, Codependent Cycles, *Journal of Econometrics*, 80, 199-221.

White, H., 2002, A reality check for data snooping *Econometrica*, vol. 68, no. 5, 1097-1127.