ON THE PERSISTENCE OF ECONOMIC FLUCTUATIONS: FURTHER EVIDENCE FROM A CLASSICAL NONPARAMETRIC APPROACH*

Marco BIANCHI
University of Bologna, Bologna, Italy

The paper presents a nonparametric approach, based on kernel density estimators, to assess and compare persistence of output fluctuations in Belgium, Canada, France, Germany, Italy, Japan, the United Kingdom and the United States. The results obtained are interpreted to be compatible in general with those already derived from the literature, but with some advantages in terms of very easy computation and economic interpretation.

*This article is a resume of Chapter 4 of my master thesis that I did at the Catholic University of Louvain, Louvain-la-Neuve, Belgium. I am indebted to prof. Henri Sneessens of the Department of Economics for guiding me during the thesis work and to prof. Wolfgang Härdle of the Center for Operation Research & Econometrics for introducing me to nonparametric techniques and for providing me at the same time with the "XploRe" package by which all computations in the paper are drawn. Only the author, of course, is responsible for errors. Financial support of the "Unipol Assicurazioni" of Bologna is greatly acknowledged.

1. Introduction

In a recent articl, John Y. Campbell and N. Gregory Mankiw compute two different measures of persistence, suggested in Campbell and Mankiw (1986, 1987) and Cochrane (1988), to find that, for most industrialized countries, "*it is hard to reject the view that real GNP is as persistent as a random walk with drift*".[1] As already recognized in Nelson and Plosser (1982), this implies that output series do not show in general a trend reverting behaviour to the *levels*, having instead a tendence to react to shocks by reverting to the preceding *rate of growth*. At the same time, however, recent empirical evidence as in Pierre Perron (1989) contrasted the opposite view that macroeconomic fluctuations might be viewed indeed as stationary movements around a deterministic trend *if* a change in the levels or in the means of the series are allowed in correspondence of some 'exceptional' events. In this paper, we use a traditional nonparametric approach, based on kernel density estimators, to provide further evidence supporting an intermediate position between these two extreme views.

The plan of the paper is: a concept of persistence is defined in Section 2. In Section 3 a short introduction of nonparametric techniques based on kernel estimators is provided. In Section 4 and 5 the results are presented and discussed, while Section 6 briefly summarizes and concludes.

2. Permanent shocks and persistence in time series

Campbell and Mankiw (1989) define an exogenous shock to be persistent if it lasts for a 'long' time in the future. In our context, we refer equivalently to the same concept interpreting a persistent shock as one capable of *permanently changing the long-run expectation (the trend component) of the series*. This interpretation of persistence, due originally to Beveridge and Nelson (1981), has the merit of stressing what is the major consequence of a permanent shock, namely a break in the steady state path of the series affected by the shock.[2]

Based upon this simple idea, we consider the series artificially generated by the model $y_t = \alpha + \beta t + \varepsilon_t, t = 1, T$, with T=30, β=0.05 for t=1,....,14 and 0.02 otherwise, α=12 for t=1,....,14 and 12.45 otherwise, and ε=random error with zero mean and finite variance. The graph of this series is

[1] Campbell-Mankiw (1989) "International Evidence on the Persistence of Economic Fluctuations", *Journal of Monetary Economics*, 23, p. 319.

[2] Stock and Watson (1988) speak in this case of *variable* trends.

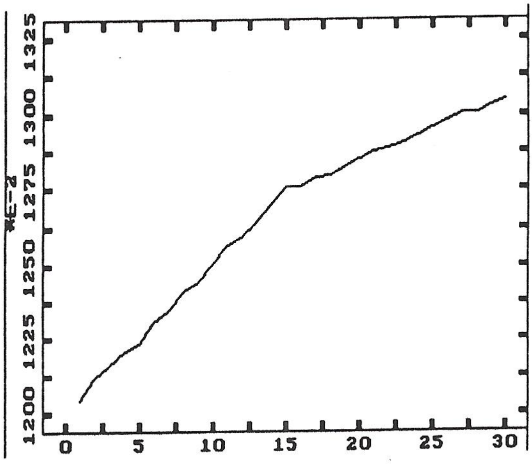


Figure 1. Simulated series generated by the model $y_t = \alpha + \beta t + \varepsilon_t, t = 1, T$. T=30, β is 0.05 for t=1,...,14 and 0.02 for t=15,...,30; α is 12 for t=1,...,14 and 12.45 for t=15,...,30.

and clearly reflects the imposed break in trend in correspondence of the time t=15.

If the series had been observed for instance for real GDP at constant prices in Italy during the period 1960-1989, the empirical evidence might be summarized by saying that output grew on average at 5% until 1974, but then a permanent shock occurred at that date (for example an oil shock) which determined a slower growth rate up to now. The main issue in the next Section is a simple procedure to detect statistically both the number of breaks in the trend of a series and their relative degree of persistence over time.

3. How to measure persistence by kernel density estimators

In statistics, kernel density estimation is a standard nonparametric technique commonly used to approximate a "true" distribution function $f(x)$ starting from a sample information $\{X_i\}_{i=1}^n$ of i.i.d. observations. Centering around each observation X_i a kernel function $K_h(u) = h^{-1}K(x/h)$ and averaging the values of this function at any given x, the kernel estimator is derived in its general form as

$$\hat{f}_h(x) = (nh)^{-1} \sum_{i=1}^{n} K(u) \qquad (1)$$

where $u = \frac{x - X_i}{h}$, with $h > 0$.

The estimated density $\hat{f}_h(x)$ depends both on the specific kernel function $K(u)$[3] and on the smoothing parameter h, called the bandwidth. Taking for instance a *Quartic* kernel

$$K(u) = \frac{15}{16}(1 - u^2)^2 I(|u| \leq 1), \tag{2}$$

$I(\cdot)$ represents an indicator function such that, for every fixed point estimate x and every fixed interval of length h centered around zero, the function has value 1 if the distance $x - X_i$ falls into the interval, and value 0 otherwise.

Since for a wider h more points fall into the bin, which imply then to average among a greater number of observations, the parameter h controls in practice the amount of smoothness, i.e. 'too small' (very near to zero) values of h should imply 'too noisy' representations of the data, while 'too big' values of h would make the estimated density 'too flat'.[4]

Intuitively, the optimal amount of smoothing is the value of h such that the estimated density is neither too noise nor too smooth. Mathematically, this value is given by the h^* such that the *Integrated Minimum Squared Error*

$$\int \mathrm{MSE}(\hat{f}_h(x))dx = (nh)^{-1} \| K \|_2^2 \int f(x)dx + \frac{h^4}{4}(\mu_2(K))^2 \int (f''(x))^2 dx + o((nh)^{-1}) + o(h^4) \tag{3}$$

[Härdle (1990a) p. 65] is asimptotically minimized, that is when the trade-off between the bias and the variance of the estimator is optimally balanced.

Assuming the unknown density function $f(x)$ twice differentiable, the optimal bandwidth is given by [Härdle (1990a) p.66]

$$h^* = \left(\frac{\| K \|_2^2}{\| f''(x) \|_2^2 (\mu_2(K))^2 n} \right)^{1/5} \tag{4}$$

where $\| K \|_2^2 = \int K^2(u)du$, $\| f''(x) \|_2^2 = \int f''(x)^2 dx$ and $\mu_2(K) = \int u^2 K(u)du$.

Since the equation implies the knowledge of $f''(x)$ to compute the optimal bandwidth h^*, two main strategies are provided in practice to the researcher to identify the correct amount of smoothing. These strategies are called respectively the <<*better rule of thumb*>> criteria and the <<*least squares*

[3]The kernel function, in general, can take various forms, but it has to be a function symmetric around zero and integrating to one.

[4]In a certain sense, the smoothing parameter h plays in this framework the same key role played by the spectral window k in Campbell and Mankiw (1989). As in Campbell and Mankiw a value of k too small relatively to the sample size T will obscure the trend reversion manifested in higher autocorrelations, while a value too large will tend to find an excessive trend reversion, in our case a value of h too large is cause of a grave inability to detect the different shocks and their relative size. [We refer to the Appendix for a concrete example].

cross-validation>> method.[5]

By the first criteria, $f(x)$ is assumed to belong to the class of the normal distributions (which allows to compute $\| f''(x) \|_2^2 = \sigma^{-5} \frac{3}{8\sqrt{\pi}} \approx 0.212\sigma^{-5}$), implying the expression

$$h^* = 1.06 \cdot \min\left(\hat{\sigma}, \frac{\hat{R}}{1.34} \right) n^{-1/5} \tag{5}$$

with $\hat{\sigma}$ estimator of σ and $\hat{R} = X_{[0.75n]} - X_{[0.25n]}$.[6]

By the second method, the *Integrated Squared Error*

$$d_I(h) = \int (\hat{f}_h - f)^2 (x) dx = \int \hat{f}_h^2(x) dx - 2 \int (\hat{f}_h f)(x) dx + \int f^2(x) dx, \tag{6}$$

is minimized with respect to h after having estimated, through a procedure called <<*Leave One Out-Estimate*>> [Härdle (1990a) pp. 108-109], the only term that has to be estimated from the data.[7]

Applying the two techniques to the artificial series previously generated, fitting for example the data set to an appropriate software such as the XploRe package,[8] we would get as results:

[5] For major details on this point, we refer directly to Härdle (1990a, 1990b).

[6] $\hat{R}$ is the interquantile range accounting for a more robust estimate of the standard deviation. This is because the rule is called the 'better' rule of thumb; it accounts for the possible presence of outliers.

[7] This term is $\int \hat{f}_h^2(x)dx$, being in (6) $\int \hat{f}_h^2(x)dx$ computable from the data and $\int f^2(x)dx$ not dependent on h.

[8] "Xplore 2.0: a Computing Environment for eXploratory Regression and Data Analysis", C.O.R.E, Catholic University of Louvain, Louvain-la-Neuve, Belgium.

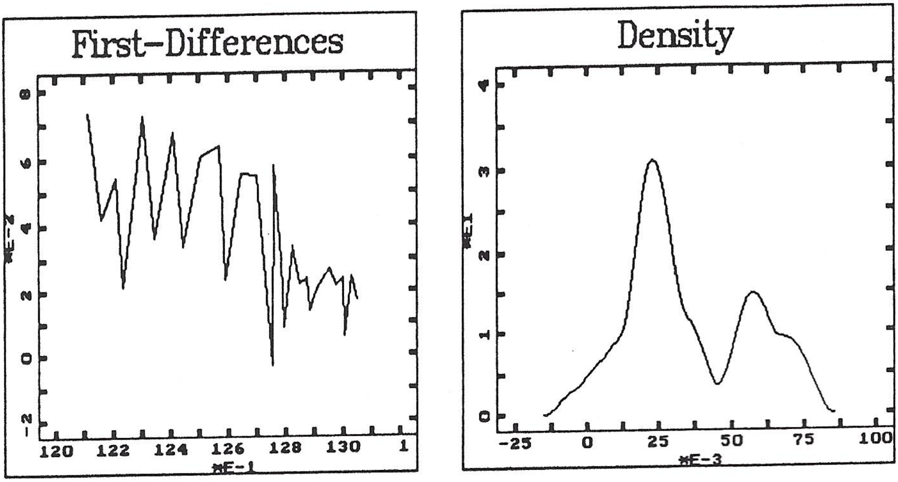


Figure 2. First differences of the simulated series (left) and associated density estimation. Smoothing parameter $h^* = 0.011$ selected through the 'better rule of thumb' criteria, Quartic kernel.

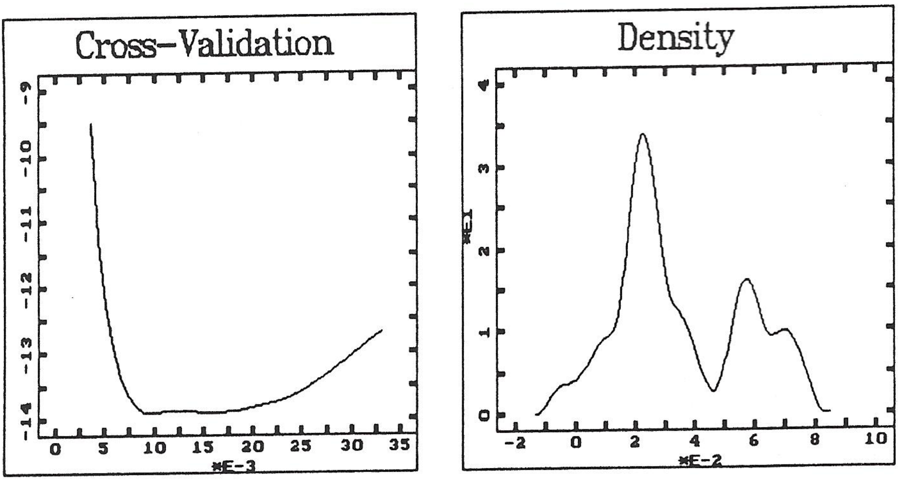


Figure 3. Least Squares Cross-Validation for bandwidth selection (left) and associated density estimation. Smoothing parameter $h^* = 0.00932$ is for a Quartic kernel.

As expected, the shape of the estimated density reveals in both cases the two steady state rates, suggesting moreover a greater persistence for the 2% steady state rate, which occurred a greater number of times in the series. The small peak in correspondence of the 7% rate is surely due to the stochastic component introduced in the simulation, and might be interpreted as a transitory shock if the series were a macroeconomic variable. As a general rule, counting the number of 'significative' modes in the estimated density (two in figure 3), breaks in trend in the series are exactly this number minus one, while the relative persistence of the shocks is given by the relative size of the densities associated to the different modes.

4. Single-country results

In this Section, we estimate persistence in the post-war GDP series of 8 countries[9] by kernel density estimators. As in Campbell and Mankiw (1989), who derive their estimates starting from the sample autocorrelations of the integrated process, we estimate the density distributions for the same stationary process; in our framework, however, the information is not given by the sample autocorrelations, and the smoothing parameter h is the optimal bandwidth computed through the better rule of thumb criteria; moreover, the outcome is a curve in place of a scalar (the factor of persistence V^k, or its lower bound $A^k(1)$, in the terminology of Campbell and Mankiw).

For the different countries, we find the optimal (both theoretically and in practice) bandwidths listed in Table 1.

TABLE 1: BANDWIDTHS THEORETICALLY OPTIMAL AND SELECTED IN PRACTICE

Country	Theoretically optimal	Optimal in practice	Difference
Belgium	0.011	0.011	-
Canada	0.009	0.012	+0.003
France	0.009	0.010	+0.001
Germany	0.009	0.010	+0.001
Italy	0.013	0.012	+0.001
Japan	0.018	0.018	-
United Kingdom	0.008	0.010	+0.002
United States	0.011	0.010	+0.001

Note: Optimal bandwidths in practice are derived repeating the estimation for various bandwidths in the neighbour of the theoretically optimal ones, selecting then the 'optimal in practice' on the basis of visual inspection of the density curves.

The estimated densities for the different countries are then:[10]

[9] These countries are: Belgium, Canada, France, Germany, Italy, Japan, the United Kingdom and the United States. The data set, taken from the OECD National Account, Main Aggregates, Volume 1, yearly data from 1960 to 1988, is in terms of product per head in US dollars at the price levels and exchange rates of 1985. Before the analysis, the series are transformed in logs and then first differenced.

[10] We suggest the following legend for a clearer understanding of the pictures: * = 'first-size' highly persistent shock; ˙ = 'second-size' highly persistent shock; - = 'negative' transitory shock; ~ = 'around zero' growth rate transitory shock; + = 'positive' transitory shock.

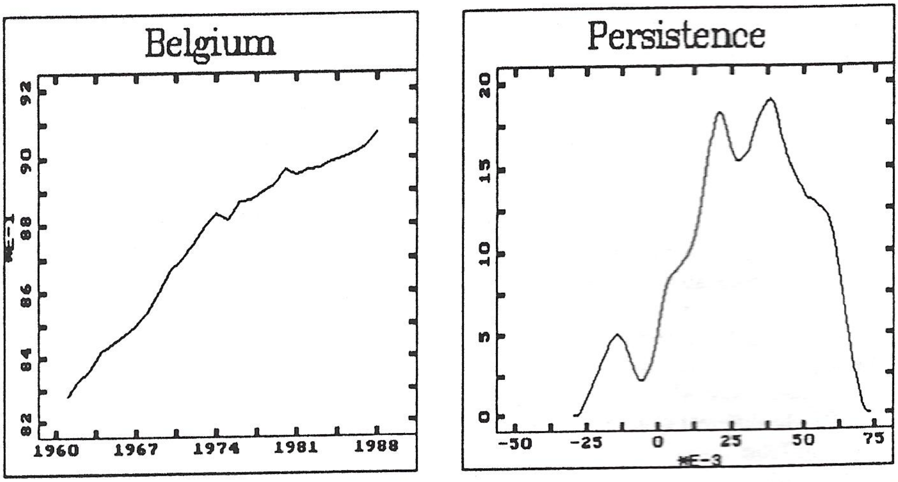


Figure 4. GDP series in logs (left) and density estimation of the differenced process (right). Smoothing parameter is 0.011 for a Quartic kernel.

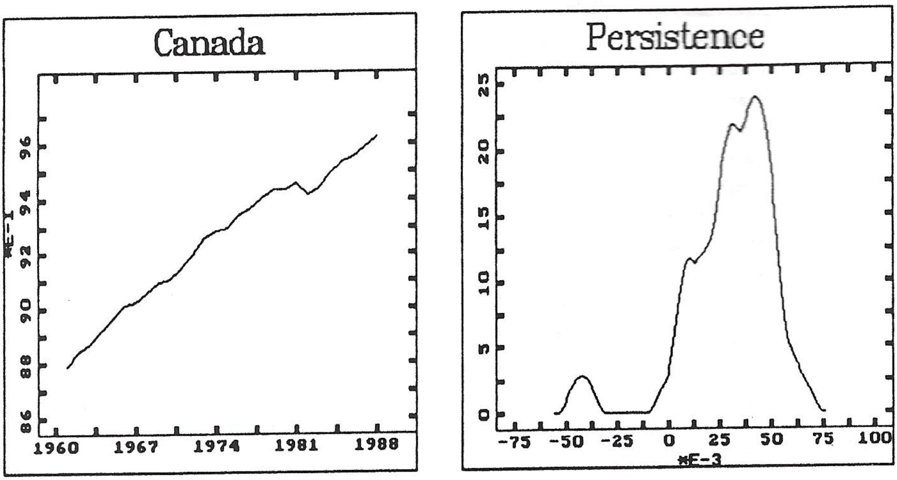


Figure 5. GDP series in logs (left) and density estimation of the differenced process (right). Smoothing parameter is 0.012 for a Quartic kernel.

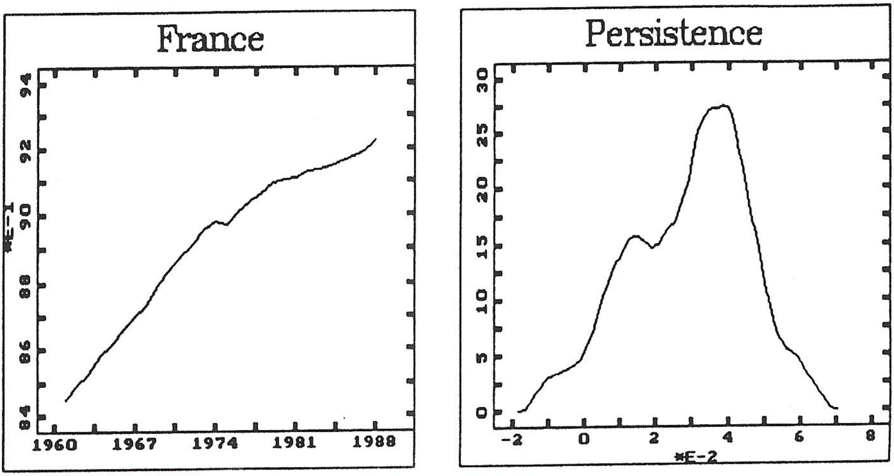


Figure 6. GDP series in logs (left) and density estimation of the differenced process (right). Smoothing parameter is 0.010 for a Quartic kernel.

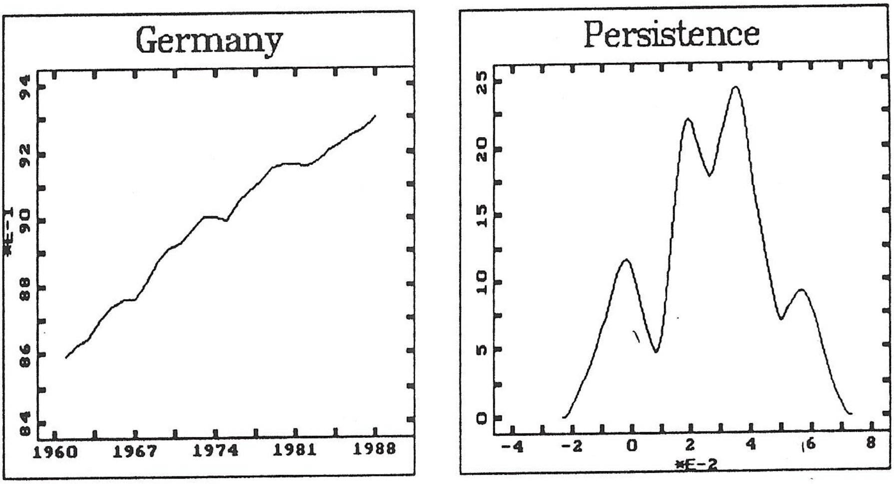


Figure 7. GDP series in logs (left) and density estimation of the differenced process (right). Smoothing parameter is 0.010 for a Quartic kernel.

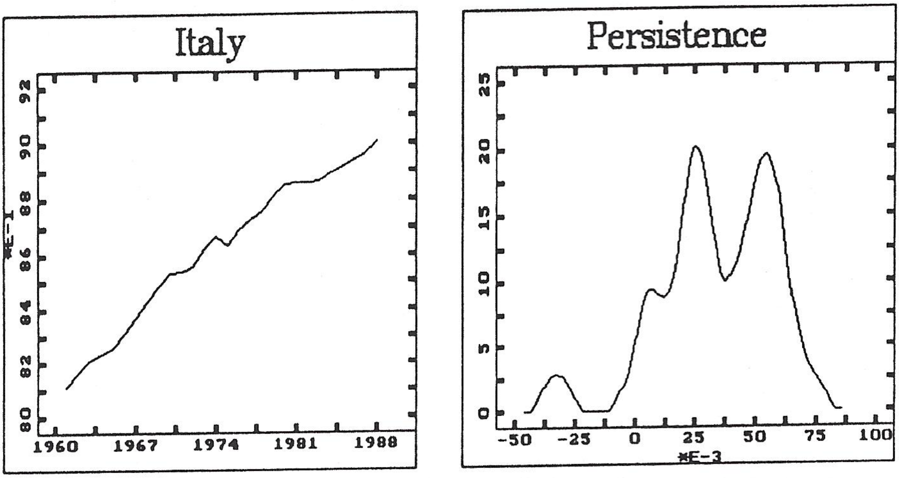


Figure 8. GDP series in logs (left) and density estimation of the differenced process (right). Smoothing parameter is 0.012 for a Quartic kernel.

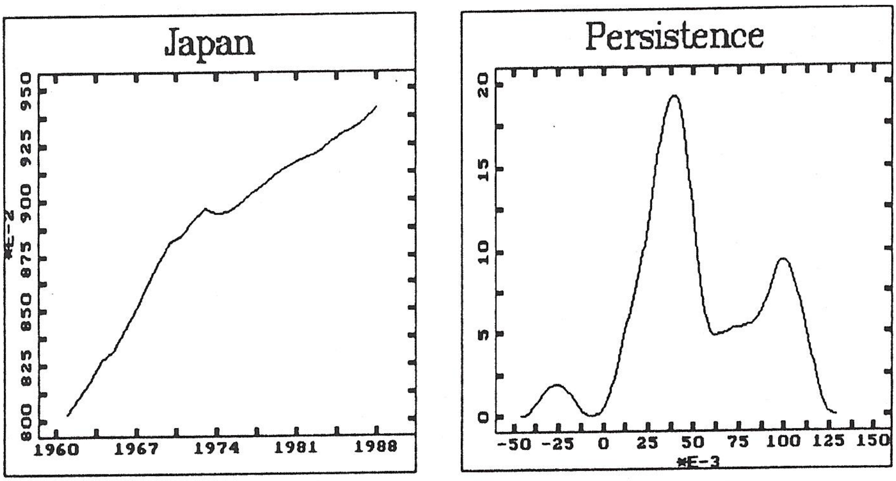


Figure 9. GDP series in logs (left) and density estimation of the differenced process (right). Smoothing parameter is 0.018 for a Quartic kernel.

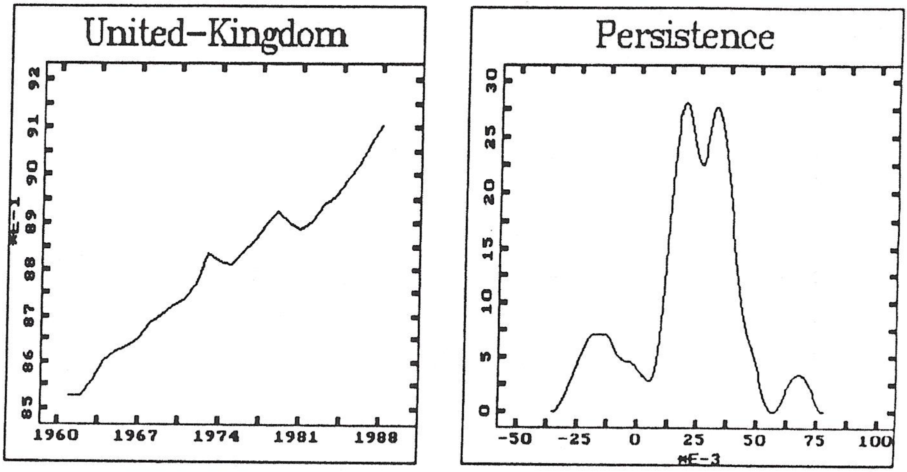


Figure 10. GDP series in logs (left) and density estimation of the differenced process (right). Smoothing parameter is 0.010 for a Quartic kernel.

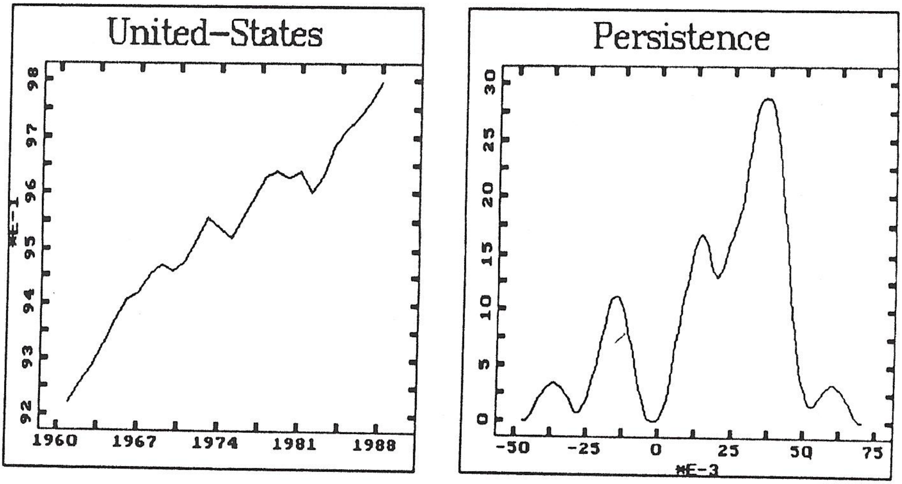


Figure 11. GDP series in logs (left) and density estimation of the differenced process (right). Smoothing parameter is 0.010 for a Quartic kernel.

and the main results might be summarized as follows:

BELGIUM: The growth was rather sustained (between 4% and 5% per year) until the first oil shock in 1973; then a break in the trend of the series occurred which constrained the economy to grow slower (around 2%) at least until 1986. The last observations seem to signal a certain improvement.

CANADA: The economy grew stable around a 3-4% per year until 1978-79; then the system arrested in 1980 and even regressed in 1981, but starting from 1983 the old trend path were already re-established.

FRANCE: The economy grew prosperous (4%) until 1974; then the first oil shock caused a downward break in the series, even if one year after the shock there was again a tendence to revert to the original trend; in fact, only the second oil shock in 1977 broke definitively the old steady state path and the economy grew since then at a lower (1%) rate. The last observations suggest however a tendence to a recuperation.

GERMANY: The post-war dynamic of output is characterized by phases of great acceleration alternated to phases of stagnation. The most prosperous period occurred between 1967 and 1973, when the rate of growth was around 5%; a slower 3% characterized the early sixties and the period from 1975 to 1978. Stagnations occurred instead around 1966, and in 1974 (first oil shock) and 1978 (second oil shock); this last shock was the most persistent of all and, after that, the tendence of the system was to grow to a 1.5% rate per year.

ITALY: The golden age of the Italian economy run in the period 1965-1970, but the first oil shock caused in 1974 a negative - even if quite transitory - growth rate (-3%); a slower rate (2.5%) occurred afterwards, and, starting from 1980, the system passed through zero growth for 3 or 4 years. The recentest tendence is to a new acceleration again, around a 2% per year.

JAPAN: The growth was surprisingly sustained (between 7.5% and 10% per year) until the first oil shock; afterwards the tendence was a slower 3% forever.

UNITED KINGDOM: The United Kingdom grew on average around a 2-3% per year until 1971; then, the positive tendence to a stronger growth (around 6%) manifested in 1972 and 1973 was completely dampened from the unexpected first oil shock, which was paid, transitorly from 1974 to 1976, in terms of negative rates (around -2%). Afterwards, the old steady state was re-established until the year 1979, followed by a new temporary recession. The last tendence is a clear sustained growth around a 3% per year.

UNITED STATES: The most persistent rate of growth in the post-war has been for this country around 3% per year, which occurred stably between 1961 and 1966 and from 1982 until 1988. In the middle, the two oil shocks and other shocks broke the trend repeatedly, but in general not for many periods of time.

5. Cross Country Results

The main feature of the post-war dynamic is output of several industrialized countries has been without doubt the presence of the two oil shocks occurred in 1973 and 1977. In general, most countries considered in the analysis were affected by these shocks, even if they were not affected exactly in the same way and the reactions were different to some extent too. Countries such as Japan and Italy, for example, very poor of energetic fonts, paid the shocks both in terms of transitory negative rates of growth and in terms of persistently lower rates of growth (this persistence is observed in Japan since the first oil shocks, while in Italy only the second oil shock broke definitively the trend). In France, only the second oil shock had permanent effects, while in Belgium it was the first one that occured in 1973. In Canada, the oil shocks did not change the steady state path at all and even the United Kingdom was affected only transitorly, while a more persistent shock occured in 1981. In Germany and in the United States a greater number of shocks had non-negligible effects on the output dynamics, i.e. for these very flexible economies many transitory shocks revealed to be much more persistent compared with the other countries.

6. Conclusions

The main goal of this paper has been to show how nonparametric kernel density estimators may provide a very flexible and useful tool to estimate persistence in macroeconomic time series. With respect to the nonparametric approach suggested in Campbell and Mankiw (1989), kernel estimators do not to require a particularly big sample size[11] and the smoothing parameter is less arbitrary selected. The graphical outcome implicit in the estimating procedure, moreover, greatly facilitate the interpretation of the results and allow to detect in a simple but rigorous way both the number and the relative size of the various shocks.[12]

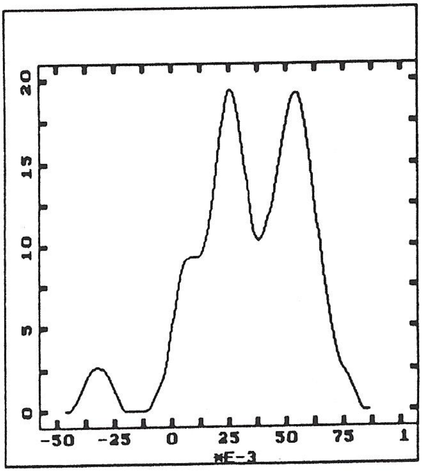
The results obtained may be interpreted as supporting the intermediate position between the two extremes views of stochastic and deterministic trends. If from one hand the finding of Perron (1989) that only a few number of shocks had large permanent effects on the output series in the postwar are confirmed from our analysis, on the other hand many temporary shocks were not 'so transitory' in the series, having a tendence to persist, in many cases, even up to four or five years.

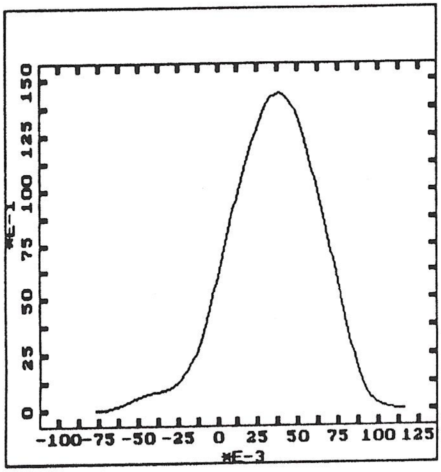
[11] This is required instead in Campbell and Mankiw (1989), since both measures of persistence, what they call $\hat{A}^k(1)$ and $\hat{V}^k$, tend to find an 'excessive' trend reversion if the spectral window k is taken 'too' large relatively to the sample size T.

[12] To assess the relative size of the different shocks has been the goal of some recent papers, such as Shapiro and Watson (1988), Blanchard and Quah (1989), Evans (1989), Blanchard (1989).

Appendix

For Italy we estimated the density of the integrated process for output through the optimal bandwidth $h^* = 0.012$. If we had used a bandwidth $\bar{h} = 0.04 > h^*$ we would have got another estimated curve. The two cases are compared as





and the estimated curve to the right (when we 'oversmooth' the density) clearly fails to detect the peak in the support [-0.05, -0.025] and let the two modes in the support [0.015, 0.075] collapse into a unique mode. The main consequence of a 'wrong' value of h is a grave inability to detect parsimoniously the number of shocks in the series and (hence) their relative size.

References

Beveridge, S. and C.R. Nelson, 1981, A new Approach to Decomposition of Economic Time Series into Permanent and Transitory Components with Particular Attention to Measurement of the Business Cycle, *Journal of Monetary Economics* 7, 151-174.

Blanchard, O.J., 1989, A Traditional Interpretation of Macroeconomic Fluctuations, *American Economic Review* 79, 1146-1164.

Blanchard, O.J and D. Quah, 1989, The Dynamic Effects of Aggregate Demand and Supply Disturbances, *American Economic Review* 79, 655-673.

Campbell, J.Y. and N.G. Mankiw, 1987a, Are Output Fluctuations Transitory?, *Quarterly Journal of Economics* 102, 857-879.

Campbell, J.Y. and N.G. Mankiw, 1987b, Permanent and Transitory Components in Macroeconomic Fluctuations, *American Economic Review Papers and Proceedings* 77, 111-117.

Campbell J.Y., N.G. Mankiw ,1989, International Evidence on the Persistence of Economic Fluctuations,, *Journal of Monetary Economics*, 23, pg. 319-333.

Cochrane, J.H., 1988, How Big is the Random Walk in GNP?, *Journal of Political Economy* 96, 893-920.

Evans, G.W., 1989, Output and Unemployment Dynamics in the United States: 1950-1985, *Journal of Applied Econometrics* 4, 213-237.

Härdle, W., 1990a, *Smoothing techniques in Theory and S*, (CORE, Catholique University of Louvain, Belgium).

Härdle, W., 1990b, *Applied Nonparametric Regression*, (Cambridge University Press).

Härdle, W., 1989, Xplore 2.0: a Computing Environment for eXploratory Regression and Data Analysis, (CORE, Catholique University of Louvain, Belgium).

Perron, P., 1989, The great crash, the oil price shock, and the unit root hypothesis, *Econometrica* 6, 1361-1401.

Shapiro, M.D. and M.W. Watson, 1988, Sources of Business Cycle Fluctuations, N.B.E.R. Annual Report, 111-157.

Stock, J.H. and M.W. Watson, 1988, Variable Trends in Economic Time Series, *Journal of Economic Perspectives* 2, 147-174.

```
@ Gauss program to generate the artificial series @
t1=seqa(1,1,14);
fn func1(t1)= 12+0.05*t1;
y1=func1(t1);
t2=seqa(15,1,16);
fn func2(t2)=12.45+0.02*t2;
y2=func2(t2);
z=y1|y2;
e=rndn(30,1)/100;
y=z+e;
output on;
y;
```

the data

1	12.04161	
2	12.1154	0.073787
3	12.15757	0.042175
4	12.21185	0.054283
5	12.23346	0.021602
6	12.30619	0.072729
7	12.34258	0.036399
8	12.41042	0.067832
9	12.44402	0.033604
10	12.50436	0.060343
11	12.56839	0.064024
12	12.5917	0.023317
13	12.6473	0.055595
14	12.70199	0.054697
15	12.7597	0.057701
16	12.75643	-0.00326
17	12.7855	0.029066
18	12.79388	0.00838
19	12.82736	0.033485
20	12.84935	0.021988
21	12.87305	0.023695
22	12.88645	0.013404
23	12.90561	0.019156
24	12.92902	0.023414
25	12.95602	0.027
26	12.97731	0.02129
27	13.00133	0.024022
28	13.00701	0.005675
29	13.0311	0.024088
30	13.04797	0.016876

```
/*===========================================================================*/
/* Gauss program to select the optimal bandwidth by the better rule
   of thumb criteria. */
/*===========================================================================*/
/*-------------------------------------------------------------*/
@ Initialization: load the xplore workunit containing the three
series (sample time (1961,962, etc.), GDP in logs and GDP in first
differences) @
    n = 28;
    loadm z[n,3] = c:\xplore\data\belgium.dat;
    y=z[.,3];
/*-------------------------------------------------------------*/
@ Ordering data from the smallest to the highest @
/*-------------------------------------------------------------*/
x = sortc(y,1);
/*-------------------------------------------------------------*/
@ Computing the interquantile range @
/*-------------------------------------------------------------*/
R = x[3/4*n,1] - x[1/4*n,1];
rs = R/1.34;
/*-------------------------------------------------------------*/
@ Computing the standard deviation @
/*-------------------------------------------------------------*/
s = stdc(x);
/*-------------------------------------------------------------*/
@ Derivation of the better rule of thumb applying the formula @
/*-------------------------------------------------------------*/
if rs < s; min = rs; endif;
if rs > s; min = s; endif;
/*-------------------------------------------------------------*/
@ Final output @
/*-------------------------------------------------------------*/
h0 = 1.06*min*(1/n)^(1/5); "The optimal bandwidth is:" h0;
/*===========================================================================*/
```