THE DYNAMICS OF AGGREGATE POLITICAL POPULARITY: EVIDENCE FROM EIGHT COUNTRIES*

Revised, December 1988

David Byers, James Davidson† and David Peel

Cardiff Business School,
Aberconway Building,
Colum Drive,
Cardiff CF1 3EU

Abstract

This paper extends previous analyses of aggregate political popularity (partisanship) data by Box-Steffensmeier and Smith (1996) for the US, and Byers, Davidson and Peel (1997) for the UK. These studies independently found that the time series of poll ratings are well modelled by fractionally integrated processes. Here, the analysis is conducted for 26 political parties in eight different countries, and the results obtained are on the whole closely in line with the ones cited above. As in the earlier studies, we find in many of our cases that the estimated fractional integration parameter *d* is close to 0.7. This implies that popularity is highly persistent and a nonstationary process, but that it is also mean-reverting eventually. Most of the time series are also found to be pure fractional noise, effectively uncorrelated after fractional differencing, so that the *d* parameter alone accounts for the dependence. As well as offering added support for theories of political allegiance based on a certain distribution of the attributes of commitment and pragmatism in the voting population, these findings have important implications for the explanation of political support using time series data.

J.E.L. classification D72, C22

Keywords : Opinion polls, political popularity, partisanship, fractional process, long memory

---

* We thank the following for their assistance in obtaining the data used in this study: Knut Kalgraff Skjek and Norwegian Social Science Data Services, Per Nielsen of Danish Data Archives, John Hughes of Gallup Canada Inc, The Canadian Embassy, Christopher Anderson, and Yoshinori Suzuki.
† Corresponding author. Email: DavidsonJE@cardiff.ac.uk.

1. Introduction

Box-Steffensmeier and Smith(1996) (henceforth B-SS) and Byers, Davidson and Peel (1997) (henceforth BDP) have recently independently proposed very similar models to explain the time series of aggregate political popularity (partisanship), as recorded by monthly or quarterly opinion polls. They report evidence, for the US and UK respectively, that aggregate partisanship is well described by a "fractionally integrated" time series process.

A fractionally integrated process is one that exhibits long memory, with persistent local trends, but which nonetheless eventually reverts to the mean. The degree of persistence is measured by a real-valued parameter d, lying on the unit interval. At the one extreme, $d = 0$ represents the short memory case. If $d > 0.5$, the process is not wide-sense stationary, having infinite variance. And at the other extreme, $d = 1$ corresponds to the ordinary integrated process, familiarly known as a random walk, which is well known not to revert to the mean but to eventually wander arbitrarily far from the starting point. Random walk processes have been extensively exploited in macroeconomic modelling in recent years, but the fractional model has found fewer applications to date, since it is relatively difficult to motivate.

However, B-SS and BDP each derive variants of a model that explains the fractional property as consequence of aggregating heterogeneous poll responses. In these models, d measures the distribution in the voting population of a certain individual characteristic, which B-SS call persistence of party identification, and BDP describe in terms of commitment versus pragmatism.[1] These are attributes that might well be supposed to depend on the political culture, traditions, and attitudes of voters, and to vary from one party to another (e.g. between left and right),[2] and also from one country to another. In the event, both sets of authors report estimates of d of very similar magnitude. Thus, BDP obtain 0.765 for the UK Conservative Party, and 0.717 for the Labour party, while B-SS obtain 0.804 for Republican support and 0.698 for the Democrats.

We find this similarity to be of considerable interest. Republicans and Democrats, after all, represent very different sorts of political coalition from Conservatives and Labour in the UK, within a very different constitutional framework. There are differences in the data sets and modelling techniques in the two studies which we de-

[1] The B-SS data measure party identification (the actual question asked is not stated) while the BDP data are polls of voting intentions "if there were an election tomorrow". Although the latter question arguably invites a more pragmatic response, we would argue that in practice something very similar is being measured in each case.

[2] It might be thought that in a two-party system, support for one party is just the obverse of support for the other, but this need not be the case. In the UK, for example, there is a substantial buffer of small parties between the main ones including Liberals and nationalists, not counting the ‘Don't Knows’.

scribe in the sequel, but we show that the estimates are broadly comparable in spite of these. The similarity further encourages the speculation entertained by BDP, that *d* might represent a stable constant of mass political behaviour in democratic societies. If it were true, it would seem to have important consequences for understanding and predicting electoral phenomena. Such findings are relevant, as B-SS point out, to the extensive literature on party identification, see Allsop and Weisberg (1988), Green and Palmquist (1990), and Mackuen, Erickson and Stimson, (1989, 1992). In addition, as pointed out by BDP, the long memory attribute has important statistical implications for empirical analyses that links popularity to economic and other explanatory variables; see e.g. Frey and Schneider (1978), Pissarides (1980), and Kramer (1983).

Clearly, a comparative study of a number of countries could throw further light on these issues. It would also be of interest to apply the model to minority parties, as well as the mainstream parties examined in the cited studies. In this paper, we examine data for a further 26 different political parties in eight countries, to determine whether the results obtained for the US and UK hold more generally. The paper is organised as follows. In Section 2 we briefly summarise the BDP model and its relationship with the B-SS model, and point out that the predictions of the two models for aggregate data are effectively the same for the data studied. Section 3 reports our new findings. Section 4 concludes the paper with some reflections on the implications for future research.

2. The Aggregate Popularity Model

The BDP model is based on the idea that voters fall into two stylised categories that they call, respectively, committed voters and ‘floating’ voters. Support for a given political party is determined in the first case mainly by conviction or group solidarity, and so is relatively insensitive to the current performance of the party. The second group is more pragmatic in outlook, and for these voters, support is driven mainly by performance. It follows that the future voting behaviour of the second group is typically less predictable from current behaviour than the first group. The degree of persistence of aggregate support depends on the distribution of these attributes in the voter population.

The BDP model assumes that the log-odds in favour of voter *i* supporting a given party is described, apart from a deterministic component, by an autoregressive process driven by news. In other words, if p_t^i represents the probability of voter *i* supporting the party at time *t* then

$$\log \frac{p_t^i}{1 - p_t^i} = C^i + y_t^i \qquad (1)$$

where

$$y_t^i = a^i y_{t-1}^i + e_t^i. \qquad (2)$$

The term C^i is interpreted by BDP as time-varying, to capture the effect of the election cycle, which is a significant phenomenon in British politics, but for the sake of the present argument we assume it to be a constant. We focus on the dynamic equation (2), which measures the degree of persistence of party support in the face of 'news', whose effect on the individual is measured by e_t^i.[3]

Assuming e_t^i to be a serially uncorrelated process, the case $a^i = 1$ in (2) corresponds to a random walk process, which evolves with high probability towards $+\infty$ or $-\infty$, so that the probability of support p_t^i tends to unity or zero under (1). Thereafter, it changes only rarely. We take this case to represent the behaviour of committed voters. On the other hand, $a^i < 1$ implies a reversion to the mean, and hence of p_t^i migrating (in the particular case $C^i = 0$) to 1/2, in the absence of news. Because of the nonlinearity of the logistic transformation, support is also a lot more volatile in this case, in the face of the same news, than it is in the unit root case. This case represents the shorter 'memory' of pragmatic voters. The a^i are assumed to be distributed in the voting population over the interval [0,1] according to the beta(u,v) density, where u and v are constant parameters, and $0 < v < 1$. For a suitable choice of v, this distribution can concentrate a significant part of the probability mass very close to 1. Since the beta is a very flexible functional form, the distribution can assume a range of shapes on the rest of the interval, depending on the parameters. It can be approximately uniform, for example.

Let $\bar{X}_t$ represent the arithmetic average of N independent binary (0-1) opinion poll responses, sampled from the population at time t, such that $100\bar{X}_t$ is the usual percentage support measure. Consider the time series properties of $\log[\bar{X}_t/(1-\bar{X}_t)]$ when t represents a succession of time periods (months or quarters). We can show that this variable converges in probability as $N \to \infty$ to the same limit as $\bar{C} + \bar{y}_t$, the mean of the right-hand side of (1), where $\bar{C}$ is converging to a constant and

$$\bar{y}_t = N^{-1} \sum_{i=1}^{N} y_t^i .$$

Note that $\bar{y}_t$ is a random variable in the limit, being a function of news variables that all voters observe, although the individual effects are averaged out. The key result,

[3] News in this context is any piece of information that might lead voters to modify their choices. These could include economic indicators, as well as changes of party leadership, policy, and the like. BDP develop a model of the relationship between the micro-shocks ε_t^i and the innovations in the aggregate data. They assume that each individual reacts to publicly available news in their own way, so that, for example, interest rate rises are good news for lenders and bad news for borrowers. This aspect of the model is not essential to the present development, and readers are referred to BDP for the details.

due to Granger (1980), is that under a beta(u,v) distribution for the a^i, the time series representation of $\bar{y}_t$ approximates (large N) to a process of the form

$$\bar{y}_t = \sum_{k=o}^{\infty} a_k \bar{e}_{t-k} \tag{3}$$

where $a_k = O(k^{-v})$, and $\bar{e}_t$ is a shock process depending on news. Simply stated, this says that averaging a mixture of stable autoregressions and near-unit root processes yields in the limit a moving average process whose coefficients decline hyperbolically. This process has high persistence, or 'long memory', but is nonetheless mean-reverting for $v > 0$. The hyperbolic-decline property is shared by the fractionally integrated or ARFIMA(p,d,q) class of processes, which take the form

$$x_t = (1 - L)^{-d} u_t \tag{4}$$

where u_t is a stationary ARMA(p,q) process, with $d = 1 - v$.[4] The ARFIMA model, plus a possible deterministic component, is accordingly proposed as a plausible model to represent the time series of $\log[\bar{X}_t / (1 - \bar{X}_t)]$. When d is close to 1 the series is accordingly more persistent, as we would expect since the parameter v is close to 0 when the distribution of the a^i is concentrated near to 1. The degree of persistence of the aggregate process therefore depends on the proportion of committed voters in the population.

The B-SS model is derived along very similar lines to the BDP model, and is in effect a 'linear probability' version of the model. It also assumes that measured popularity is the aggregate of heterogeneous autoregressive processes, and invokes Granger's result. However, it does not specify a nonlinear mapping of the popularity measure onto the unit interval, which BDP obtain by the inverse-logistic transform from y_t^i back to p_t^i. As a consequence, it is not possible to interpret their party affiliation variable x_{it} (see B-SS page 568) as the probability of support. Indeed, they think of 'affiliation' as being distributed on (–∞, +∞), so that it is functionally similar to BDP's 'log-odds in favour of support' variable. Notwithstanding, B-SS then fit their model to data for percentage support, acknowledging this as a weakness of their approach (B-SS page 572).

However, while the difference between the two models is obviously important at the micro level, it should be negligible at the aggregate level provided aggregate percentage support does not stray too near the extremes. The logistic transform is nearly linear in the mid-range, and hence a time series model that explains $\log[\bar{X}_t / (1 - \bar{X}_t)]$ will also explain percentage support, $100\bar{X}_t$. This condition will be satisfied for major parties such as the US Republicans and Democrats that B-SS analyse, who attract

[4] The expression in (4) is to be interpreted by taking the infinite-order binomial expansion of $(1–L)^{-d}$, where L represents the lag operator, defined by $Lu_t = u_{t-1}$.

roughly half the electorate each on average. However, on our interpretation their model might not perform so well with minority parties, attracting less than 20% support say, such as exist in European countries.

Given these caveats, we can validly compare the results obtained in each study, as quoted in the Introduction. B-SS use quarterly data for the period 1953-1992, whereas BDP use monthly data for the period 1960-1995.[5] The different frequencies of observation in the two studies are not important here, for it is a well-known property of the fractional model that *d* is invariant to time aggregation.[6] In fact, BDP also report directly comparable *d* estimates from quarterly data (taking every third observation) as respectively 0.601 (Conservatives) and 0.693 (Labour). The differences can be simply explained in terms of the loss of sample information in the quarterly estimates.

3. Empirical Analysis

We examined time series of percentage support for a total of 26 political parties in 8 countries. The countries are Australia, Canada, Denmark, Germany (FDR), Japan, Netherlands, Norway and Sweden. All of the series are monthly, with lengths ranging from 156 to 480 observations, and in each case these are the longest time period for which a complete set of data is available. Table 1 lists the names of the parties, the acronym we use to identify them, sample periods, and mean support over the sample. The series are based on polls by the local Gallup Organisations of the countries concerned, except for Japan where the data were taken from the Yomiuri Newspaper. With the exception of Japan, for which the information is not available, the data represent answers to the standard question: ‘Which party would you vote for if there were to be a Parliamentary Election tomorrow?’. Following the BDP model, the series analysed are the logistic transforms of the measured percentage support series divided by 100.

Tables 2, 3 and 4 show the results of fitting the ARFIMA(p,d,q) model to our series. The only additional transformation applied to these series prior to estimation is centring by subtraction of the sample mean, corresponding to treating $\overline{C}$ as a constant.[7]

---

[5] Both the BDP and B-SS estimates were computed by the time-domain maximum likelihood estimator due Sowell (1992), using Fortran code supplied by Fallaw Sowell.

[6] It is known (Davydov 1970) that after scaling by $T^{1/2-d}$ where T = sample size, a nonstationary fractional process converges weakly to a correlated stochastic process called fractional Brownian motion. This has the property of "self-similarity", or invariance to changes of time scale (see Mandelbrot and van Ness, 1968). The fractional process therefore possesses the self-similarity property, to within the relevant order of approximation.

[7] Some preliminary attempts were made to fit an election cycle to some of the series as in the BDP study, but no significant effects were found. There is reason to think that this phenomenon is peculiar to the UK. In any case, removing the cycle was found to make very little difference to the estimated values of *d* for the UK in the BDP study (see BDP, Table 3).

To estimate these models, one has a choice of several techniques including the time domain exact maximum likelihood (ML) estimator due to Sowell (1992), as used by BDP and B-SS, and the frequency-domain approximate ML estimator, or ‘Whittle likelihood’. In samples of reasonable size, the frequency domain approach has been shown superior to exact ML when the population mean of the process is unknown; see Cheung and Diebold (1994), and also Hauser (1997). In fact, both methods were available to us, and in all those cases where both estimators were computed for comparison, the differences between the two estimates were found to be small.[8] The large-sample properties of the estimator are only known for $-1/2 < d < 1/2$, the range for which the process is stationary with an invertible moving average representation, and our estimates were therefore computed from the simple-differenced data. To recover the original *d*s, 1 has been added to the reported estimates. We have to choose appropriate values for *p* and *q* in each case, and this is done on the basis of the Schwarz information criterion (SIC), whose values for six alternative models are shown in Table 2.[9]

To interpret the findings summarised in this table we should first reprise the argument of BDP, that according to the aggregation model in which opinion is driven by ‘news’, u_t in (4) ought to be a white noise process. Suppose we assume that a ‘plausible’ beta(u,v) density is one that exhibits a concentration of probability mass close to 1, but is more or less uniform over the rest of the unit interval. Such a scheme, at least, requires the minimum of supplementary hypotheses about voter behaviour.[10] If $v = 0.3$, the value of u that most nearly reproduces this pattern is 0.45 (see BDP Figure 1). BDP then consider the effect of back-solving the Granger aggregate process as a fractional difference, with $d = 1 - v$. They show that among beta(u,0.3) distributions, the one that comes the closest to reconciling an ARFIMA(0,d,0) aggregate process with uncorrelated news has $u = 0.45$! In other words, the twin hypotheses of uncorrelated shocks, and voter behaviour like BDP's Figure 1, lead to the prediction that the ARFIMA(0,d,0) process should account for the autocorrelation of the popularity series; not exactly, but to a good approximation.

This is just what we find. Table 2 shows that the simple ARFIMA(0,d,0) is in fact the preferred model on the SIC criterion (indicated by *) in all but four of our cases. In

[8] The Whittle log-likelihood function was programmed by ourselves in Ox (Doornik, 1998) and optimised using the Ox supplied BFGS search algorithm. The Sowell time domain estimator has also been programmed in Ox by Doornik and Ooms (1998). We were therefore able to compare these algorithms directly.

[9] Table 3 shows the ‘log-likelihood’ version of the SIC, such that larger is better. Note that BDP (Table 4) gave the ‘sum of squares" version, with sign reversed.

[10] With different values of *u* we either find probability mass concentrated close to 0 as well as 1, or otherwise away from 0. Both of these phenomena would require an explanation not supplied by the ‘committed-floating’ voter model.

Table 3, we report the estimates of *d* for this model for all the 26 series, together with two statistics measuring residual dependence. These are the Box-Pierce 'Q' statistics[11] of order 12, for the levels and the squares of the residuals, respectively. The second of these statistics can be thought of primarily as testing for autoregressive conditional heteroscedasticity (ARCH). This is a form of persistence in the volatility of the series that does not contradict unpredictability in mean. While we cannot eliminate the latter form of dependence by choice of ARFIMA specification, the two tests in conjunction have power to detect general forms of serial dependence, and indicate the degree to which the model represents a complete description of the data. Rejections at the nominal 5% large-sample significance level are indicated with a *. The overall success of the one-parameter model in accounting for autocorrelation is remarkable, with only one case (The Danish RVN) showing a decisive rejection. Three of the other cases showing some evidence of autocorrelation are also ones in which the SIC is dominated by more general models, and estimates of alternative specifications for these four cases are given in Table 4.

Two notable features of the results stand out. Considering first the cases in Table 3 where the model appears adequate, the majority of the estimates lie fairly close to the range of values spanned by the BDP and B-SS estimates, and none significantly below. There are a number of estimates lying closer to or above unity, although there does not seem to be any systematic connection between the estimated *d* and an obvious distinguishing factor such as the size of the party, in terms of average support. These cases arise mainly with the smaller sample sizes, notably the Netherlands. In the larger samples, especially Germany and Denmark, the estimated *d*s are all in the range 0.65-0.85, and appear as true fractionals. This fact is significant, since it is well known to be difficult to distinguish long memory processes from true unit root processes in short samples, and 200 observations rates as a short sample in this context. Thus, given what is known about the properties of the estimates, the evidence presented here does not contradict the hypothesis that these series are generated with a value of *d* in the region of 0.7-0.8.

On the other hand, it is notable that the four cases in Table 3 for which the ARFIMA(0,*d*,0) model appears inadequate are all parties whose average support is small (see Table 1). A different model appears necessary to account for these cases, but what this model ought to be does not emerge clearly from the results. One interpretation of the evidence is that we should treat these as ARIMA(0,1,1) (unit root) processes. There are certainly grounds to suppose that small parties might have a

[11] The Q(*m*) statistic is defined as the sum of the first *m* squared sample autocorrelations multiplied by the sample size. Under the null hypothesis that the series is uncorrelated, the statistic has the chi-squared distribution with *m* degrees of freedom in large samples.

higher proportion of committed supporters, and less floating support, than big parties. In the Granger aggregation model, the aggregate process (4) will have d = 1 if the point unity has positive probability mass in the distribution of the a^i.

However, a unit root process has implausible implications for the long run, since without the mean reversion property, support must eventually tend to either 0% or 100% with probability 1.[12] Also note that, while the similarity of the estimates in the different series is quite striking, the standard errors are now rather large. This illustrates the known fact that except in very large samples, it can be difficult to identify such higher order models. Although the values of d differ quite widely in the different estimated models, the differences in likelihood are small, and the trade-off of efficiency for bias in these enlarged models may not be favourable. While these cases are suggestive, they do not in our judgement constitute a decisive rejection of BDP's hypothesis that the simple one-parameter model, as in Table 3, should describe these series adequately.

4. Conclusion

The findings presented in this paper, together with those reported previously, point to the conclusion that the dynamics of political support in modern democracies have remarkably similar characteristics. We can summarise the typical case as a highly persistent, nonstationary process that nonetheless reverts to the mean; but more precisely, we can say that the dependence is typically explained by a one-parameter time series model, the ARFIMA(0, d, 0).

Whether the model of aggregation over heterogeneous voters actually does account for these findings is naturally a more speculative matter, but none of the evidence we have been able to adduce contradicts it. In particular, we find the parsimony of the fitted models highly suggestive, for the reasons discussed in BDP. The possibility of explaining mass political behaviour in terms of the distribution of individual attributes such as commitment and pragmatism in the voting population is obviously an attractive one. Cross-section studies of individual behaviour, as advocated by B-SS in their conclusion, would probably be the best way to throw further light on this issue.

Some recommendations and caveats for future time series research also emerge from our work. There is an inherent danger in trying to 'explain' popularity by regression analyses using economic variables. It has been shown (Tsay and Chung 1996) that the

---

[12] There are of course cases, such as Communist parties in Western Europe, whose support has effectively collapsed to 0. These may be the exceptions that prove the rule, since they also enjoyed a high level of commitment from their supporters. More typically, democratic parties trim their policies so as to hold their share of the vote when times change, and so avoid such a fate.

‘spurious regression’ phenomenon best known in connection with unit root processes (Granger and Newbold 1974) extends to fractional processes. Variables such as inflation and interest rates have also been shown to be fractional, if not unit root processes, see for example Baillie, Chung and Tieslau (1996). Therefore, simple regression *t* values may be misleading. One alternative is to model the fractional difference process as a function of news variables, and an exercise on these lines for the UK is reported in BDP. Another might be to attempt a generalised cointegration analysis, to determine whether the regression residuals are I(0). The theory of testing in this framework is not yet well understood, but is currently under investigation (see for example Marmol 1998). We hope to develop this approach in future work.

References

Allsop, D. and H. F. Weisberg (1988) ‘Measuring Change in Party Identification in an Election Campaign’, *American Political Science Review* 32 (November), 996-1017

Baillie, R.T., C. F. Chung and M. A. Tieslau (1996) ‘Analysing inflation by the fractionally integrated ARFIMA-GARCH model’ *Journal Of Applied Econometrics*, 11, 23-40.

Box-Steffensmeier, J. M. and R. M. Smith, (1996) ‘The Dynamics of Aggregate Partisanship”, *American Political Science Review* 90 (September), 567-580

Byers, D., J. Davidson, and D. A. Peel, (1997) ‘Modelling Political Popularity: An Analysis of Long-range Dependence in Opinion Poll Series’, *Journal of the Royal Statistical Society* Series A, 160, 471-490

Cheung. Y.-W. and F. X. Diebold (1994) ‘On maximum likelihood estimation of the differencing parameter of fractionally integrated noise with unknown mean‘, *Journal of Econometrics* 62, 301-316

Davydov, Yu. A. (1970), ‘The invariance principle for stationary processes’ *Theory of Probability and its Applications* XV, 3, 487-498

Doornik, J.A. (1998), *Object-Oriented Matrix Programming using Ox 2.0*, London: Timberlake Consultants Ltd and Oxford: www.nuff.ox.ac.uk/Users/Doornik.

Doornik, J.A. and Ooms, M. (1998), A Package for Estimating, Forecasting and Simulating Arfima Models, Oxford: www.nuff.ox.ac.uk/Users/Doornik.

Frey, B. S. and F. Schneider, (1978) ‘A politico-economic model of the United Kingdom’, *Economic Journal* 88, 243-253.

Granger, C. W. J. (1980) ‘Long memory relationships and the aggregation of dynamic models’, *Journal of Econometrics* 14, 227-238

Granger, C. W. J. and R. Joyeux, (1980) 'An Introduction to Long Memory Time Series Models And Fractional Differencing' *Journal Of Time Series Analysis* 1 (1) 15-29.

Granger, C. W. J. and P. Newbold (1974) 'Spurious regressions in econometrics' *Journal of Econometrics* 2, 111-120

Green. D. P. and B. Palmquist, (1990), 'Of Artifacts and Partisan Stability', *American Political Science Review* 34, (August) 872-902.

Hauser, M. A. (1997) 'Maximum likelihood estimators of ARFIMA models: a Monte Carlo study', Working Paper, University of Statistics and Business Administration Department of Statistics, Vienna.

Kramer, G. H. (1983) 'The Ecological Fallacy Revisited: Aggregate versus Individual-level Findings on Economics and Elections and Sociotropic Voting', *American Political Science Review* 77 (March), 92-111.

Kwiatkowski, D., P. C. B. Phillips, P. Schmidt and Y. Shin (1992) 'Testing the null hypothesis of stationarity against the alternative of a unit root' *Journal of Econometrics* 54, 159-178.

Lo, Andrew W. (1991) 'Long-term memory in stock market prices', *Econometrica* 59, 5, 1279-1313

MacKuen, M. B., R. S. Erickson, and J. A. Stimson, (1989) "Macropartisanship", *American Political Science Review* 83 (December), 1125-42.

MacKuen, M. B, R. S. Erickson, and J.A. Stimson, (1992)"Question Wording and Macropartisanship", *American Political Science Review* 86 (June) 475-81

Mandelbrot, B. B. and J. W. van Ness (1968) 'Fractional Brownian motions, fractional noises and applications', *SIAM Review* 10, 4, 422-437

Marmol, F. (1998) 'Spurious regression theory with nonstationary fractionally integrated processes' *Journal of Econometrics* 84, 233-250

Newey, W. and K. West (1987) 'A simple positive definite heteroscedasticity and autocorrelation consistent covariance matrix', *Econometrica* 55, 703-705

Pissarides, C. (1980) "British Government Popularity and Economic Performance" *Economic Journal* 90, 569-581

Phillips, P. C. B. and P. Perron (1988) 'Testing for a unit root in time series regression', Biometrika, 75, 2, 335-46

Sowell, F. (1992) 'Maximum likelihood estimation of stationary univariate fractionally integrated time series models', *Journal of Econometrics* 53, 165-188

Tsay, W.-J. and C-F. Chung (1996) 'The spurious regression of fractionally integrated processes', Michigan State University, Economics Working Paper (revised, April 1996).

Table 1
Basic Data

Party	Acronym	Sample Period	Sample Size	Mean % Support
Australia		78:01 - 90:12	156	
Liberal Party of Australia	LPA			46.26
National Party of Australia	NPA			43.85
Australian Labour Party	ALP			7.13
Canada		74:09 - 95:12	256	
Liberal Party of Canada	LPC			40.65
Progressive Conservative Party	PCP			31.65
New Democratic Party	NDP			19.64
Denmark		60:01 - 92:10	394	
Socialdemokratiet i Danmark	SDD			35.32
Konservative Folkparti	KFP			17.13
Venstre	VEN			16.00
Socialistik Folkparti	SFP			8.64
Radikale Venstre	RVN			6.71
Germany		50:01 - 89:12	480	
Christlich-Demokratische Union	CDU			44.59
Socialdemokratische Partei Deutschlands	SPD			40.20
Freie Demokratische Partei	FDP			8.13
Japan		78:06 - 95:12	211	
Liberal Democratic Party (Jiyu Minshu-to)	LDP			42.77
Social Democratic Party (Shakai i Minshu-to)	SDP			12.34
Netherlands		79:01 - 92:09	165	
Partij van der Arbeid	PVA			30.43
Christien Democratisch Appèl	CDA			30.28
Democraten 66	D66			9.31
Norway		77:08 - 96:12	233	
Det Norske Arbeider Parti	DNA			37.68
Høyre	HØY			25.91
Sweden		67:01 - 95:12	348	
Sveriges Socialdemokratika Arbetarepartiet	SSA			44.00
Moderata Samlingspartiet	MSP			19.30
Centrpartiet	CPS			14.88
Folkpartiet Liberalna	FPL			11.49
Vänsterpartiet	VNP			4.94

Table 2

Maximum Whittle-Likelihood Estimation: SIC Values[13]

	ARFIMA(p,d,q) Model: 0, d, 0	1, d, 0	2, d, 0	0, d, 1	0, d, 2	1, d, 1
Australia						
LPA	486.44*	485.61	483.10	485.39	482.98	483.10
NPA	490.67*	489.79	487.29	488.83	486.37	487.32
ALP	368.56*	367.94	365.76	368.38	365.86	365.86
Canada						
LPC	605.82*	603.05	602.37	603.05	601.87	602.09
PCP	565.33*	563.98	562.05	562.80	561.07	561.97
NDP	580.29	580.42	581.85	582.80*	580.16	580.12
Denmark						
SDD	1196.1*	1193.2	1190.4	1193.2	1190.5	1191.4
KFP	1030.2*	1029.6	1027.0	1030.1	1027.3	1027.4
VEN	1101.9*	1099.8	1097.7	1100.2	1099.1	1099.1
SFP	821.86*	820.88	817.92	820.58	818.15	817.90
RVN	835.34	843.93	843.27	846.52*	843.64	843.98
Germany						
CDU	1410.3*	1409.1	1406.1	1409.2	1406.1	1406.3
SPD	1389.5*	1386.4	1383.9	1386.4	1384.0	1383.4
FDP	1080.8	1082.8	1080.6	1083.5*	1080.5	1080.5
Japan						
LDP	564.25*	62.83	560.27	562.59	560.28	560.27
SDP	551.87*	550.91	548.23	550.82	548.21	550.64
Netherlands						
CDA	559.41*	556.98	554.50	556.96	555.11	554.44
PVA	514.49*	511.94	509.39	511.94	509.39	509.56
D66	432.03*	429.78	428.01	429.96	428.26	428.36
Norway						
DNA	603.27*	600.66	598.27	600.68	598.23	598.48
HØY	586.75*	584.16	581.44	584.16	581.44	581.43
Sweden						
SSA	1150.0*	1148.4	1146.25	1149.0	1147.3	1147.2
MSP	1079.1*	1076.6	1073.9	1076.5	1073.9	1074.2
CPS	981.38*	978.62	975.86	978.60	975.86	976.18
FPL	913.18*	910.71	908.98	910.56	908.74	908.74
VNP	779.88	781.45	780.08	784.30*	781.40	781.40

[13] Computed from differenced data. A * marks the maximum value in each row..

Table 3

Maximum Whittle-Likelihood Estimates of the ARFIMA(0,d,0) Model[14]

	Estimate (Std. Error)	Box-Pierce Q (12 lags) Levels	Box-Pierce Q (12 lags) Squares
Australia			
LPA	0.959 (0.085)	7.21	7.57
NPA	0.757 (0.087)	7.53	10.45
ALP	0.666 (0.060)	10.50	8.59
Canada			
LPC	0.882 (0.055)	12.32	40.59*
PCP	1.018 (0.058)	10.79	38.57*
NDP	0.798 (0.049)	27.41*	28.60*
Denmark			
SDD	0.689 (0.036)	6.83	57.03**
KFP	0.854 (0.045)	20.89	60.87**
VEN	0.743 (0.040)	17.68	10.74
SFP	0.760 (0.028)	16.87	90.46**
RVN	0.655 (0.062)	48.61**	38.30**
Germany			
CDU	0.799 (0.035)	16.57	70.32**
SPD	0.738 (0.036)	10.62	8.35
FDP	0.632 (0.038)	20.18	18.37
Japan			
LDP	0.898 (0.045)	6.10	19.03
SDP	0.789 (0.050)	22.47*	16.21
Netherlands			
PVA	1.029 (0.058)	2.86	6.06
CDA	0.922 (0.062)	12.90	6.89
D66	1.129 (0.047)	6.22	30.24*
Norway			
DNA	0.704 (0.048)	23.57*	38.80**
HØY	0.620 (0.046)	8.88	10.50
Sweden			
SSA	0.968 (0.067)	11.89	12.77
MSP	0.953 (0.044)	13.06	8.40
CPS	0.826 (0.037)	17.28	24.84*
FPL	1.040 (0.038)	11.48	6.08
VNP	0.730 (0.060)	23.27*	19.23

[14] Computed from differenced data, with 1 being added back to the reported *d* values. The Q statistics are chi-squared with 12 d.f. in large samples when the series are uncorrelated, * denotes rejection of the null hypothesis at the nominal 5% level, and ** at the nominal 1% level.

Table 4: [15]

Maximum Whittle-Likelihood Estimates of Selected ARFIMA(p,d,q) Models

	ARFIMA	Estimate (SE)	Box-Pierce Q (12 lags) Levels	Squares
Canada				
NDP	2,d,0	D: 1.088 (.0500) AR1:–0.405 (0.033) AR2:–0.239 (0.019)	10.86	21.75*
	0,d,1	D: 1.180 (0.178) MA1: –0.515 (0.185)	15.40	20.22
Denmark				
RVN	2,d,0	D: 0.908 (0.074) AR1: –0.434 (0.080) AR2: –0.143 (0.068)	17.38	23.76*
	0,d,1	D: 1.069 (0.319) MA1: –0.587 (0.327)	21.07	22.85*
Germany				
FDP	1,d,0	D: 0.759 (0.054) AR1: –0.214 (0.067)	10.79	19.94
	0,d,1	D: 0.855 (0.119) MA1: –0.319 (0.136)	9.91	21.49
Sweden				
VNP	1,d,0	D: 0.864 (0.078) AR1: –0.230 (0.064)	14.58	17.49
	0,d,1	D: 1.185 (0.178) MA1: –0.578 (0.139)	6.66	15.68

[15] See footnote to Table 4.