

Support for Governments and Leaders: Fractional Cointegration Analysis of Poll Evidence from the UK, 1960-2004.

James Davidson, University of Exeter* David Byers, Cardiff Business School
David Peel, Lancaster Business School

Revised, July 2005

Abstract

Two versions of a fractionally cointegrating vector error correction model (FVECM) are presented. In the case of regular cointegration, linear combinations of fractionally integrated variables are integrated to lower order. Generalized cointegration is defined as the case where the cointegrating variables may be fractional differences of the observed series. The concepts are applied to a model of poll data on approval of the performance of prime ministers and governments in the UK.

1 Introduction

A recurring issue in the study of democratic government is the relationship between the governing party and its leader, and the interactions between voter approval of party policies and leadership. Charisma of the leadership, or lack of it, is often believed to play a significant role in shaping the view of government performance held by uncommitted voters in particular. A number of questions present themselves for analysis. For example, do popular policies reflect in the leader's rating? Can a popular leader generate support for his/her government despite unpopular policies? Or, is there a two-way interaction between the perceptions of the leader and the government? Also, how do leaders in recent history rate in the charisma stakes, measured by the gap between personal support and government support?

There now exist long runs of monthly opinion poll data for the UK with which an objective answer to these questions might be attempted. In particular, we have Gallup and MORI polls, together covering the period since September 1960 to the present, which report the percentage of positive responses by voters to the questions, "Do you approve the performance of the Prime Minister?" and "Do you approve the government's record?" We refer to these series in the sequel as PM and Gov.

However, the statistical analysis of these series is not a trivial issue. Analysis by the present authors of the closely related monthly poll data on voting intentions (Byers, Davidson and Peel 1997, 2000, 2002, henceforth referred to as BDP) shows that these latter series appear fractionally integrated, with an integration parameter, d , in the neighbourhood of 0.75. They are covariance nonstationary, but mean-reverting and therefore not random walks. We show in the next section that very similar models provide a parsimonious description of PM and Gov, with similar values of d . These facts pose a problem for conventional time series methods. The nonstationarity

*Corresponding author. Address: School of Business and Economics, University of Exeter, Exeter EX4 4PU.
Email: James.Davidson@exeter.ac.uk

invalidates a conventional regression or simple VAR analysis, based on correlations. However, the alternative strategy of a conventional cointegrating VAR is invalidated by the fact that simple differencing of the data will result in over-differencing. The concept of cointegration between fractionally integrated series cannot be formulated in the usual way. However, we show in this paper that a fractionally cointegrating VAR model can be used to analyse these series, and yield evidence on the issues listed above.

The fractional integration model has been quite widely employed by political scientists as a way of capturing the characteristics of series such as opinion polls and other indices of political interactions. In addition to our own work cited above and the related work of Box-Steffensmeier and Smith (1996), see for example Clarke and Lebo (2003) which analyses UK approval series related to our own, and Lebo and Moore (2003) which analyses indices of foreign policy interactions in a similar manner. These inferential procedures involve fractionally differencing series to $I(0)$ and running regressions in the differences, also including putative cointegrating residuals as explanatory variables. Box-Steffensmeier et al. (2004) is another study in this vein, studying the persistence of the ‘gender gap’ in partisanship as a cointegrating residual. However, these studies do not attempt a true fractional multivariate analysis. A study closer in spirit to our own is Dueker and Startz (1998), which models US and Canadian bond rates using a bivariate ARFIMA process. We comment further on this work below.

The paper is organized as follows. Section 2 reviews the cointegration concepts we propose to employ, and distinguishes between the cases of regular and generalized cointegration. Section 3 then sets out the bivariate model of leader and government approval, discusses the data set, and carefully considers the treatment of data features such as break points. Section 3 reports the results, both for our complete data period covering the periods of office of eight different prime ministers, and also for the longest unbroken prime-ministerial regime, that of Margaret Thatcher. Section 5 reports tests for the existence of cointegration, and also of “full” cointegration (such that the cointegrating residuals are $I(0)$) using bootstrap methods developed by the first author in previous work, and Section 6 concludes the paper. Some background information on fractional differencing is contained in the Appendix.

2 Fractional Cointegration

One of the difficulties with modelling relationships between fractional series is that the order of integration is a real number, and once this number is allowed to be different from unity, there may be no strong reason from economic or behavioural considerations to require it to be the same for all series. This poses a difficulty in view of the requirement that cointegrating equations balance. Thus, if $x_{1t} \sim I(d_1)$ and $x_{2t} \sim I(d_2)$, these variables can only cointegrate by the accepted definition, that a linear combination exists integrated to lower order, if $d_1 = d_2$.

With a view to developing a formal (not behavioural) approach to this issue, we consider two alternative models (or classes of model). They are, respectively, the *regular* fractional cointegration model

$$[\mathbf{B}(L)\mathbf{\Delta}(L) - \boldsymbol{\alpha}(L)\boldsymbol{\beta}'L](\mathbf{x}_t + \boldsymbol{\Phi}\mathbf{D}_t) = \mathbf{u}_t \quad (2.1)$$

and the *generalized* fractional cointegration model

$$[\mathbf{B}(L)\mathbf{\Delta}(L) - \boldsymbol{\alpha}\boldsymbol{\beta}'L](\mathbf{x}_t + \boldsymbol{\Phi}\mathbf{D}_t) = \mathbf{u}_t. \quad (2.2)$$

These can be seen as generalizations of the model proposed by Granger (1986) who, however, had integer orders of integration in mind. We define the components of these expressions before discussing them further. In each case, \mathbf{x}_t ($N \times 1$) is the vector of observed variables. \mathbf{D}_t ($S \times 1$)

is a vector of exogenous variables, typically dummies, and Φ ($N \times S$) a matrix of coefficients. For convenience in what follows we write

$$\mathbf{w}_t = \mathbf{x}_t + \Phi \mathbf{D}_t \quad (2.3)$$

so that the existence of possible exogenous factors can be left implicit. \mathbf{u}_t ($N \times 1$) will typically be assumed to be a martingale difference process. $\mathbf{B}(L)$ ($N \times N$) is an optional matrix polynomial with roots outside the unit circle to represent short run effects. Next,

$$\Delta(L) = \text{diag}\{(1-L)^{d_1}, \dots, (1-L)^{d_N}\} \quad (2.4)$$

is the vector fractional differencing operator where d_1, \dots, d_N are nonnegative parameters (not necessarily distinct). The Appendix contains some additional details on the practical implementation of the differencing operator $(1-L)^d$.

$\alpha(L)$ and β (resp. α and $\beta(L)$) are matrices of dimension $N \times R$. The special lag polynomial matrices defined for these models take the respective forms $\alpha(L) = \{\alpha_{ji}(L)\}$ and $\beta(L) = \{\beta_{ji}(L)\}$ where

$$\alpha_{ji}(L) = \alpha_{ji}(1-L)^{d_j - b_{ji}} \quad (2.5)$$

$$\beta_{ji}(L) = \beta_{ji}(1-L)^{d_j - b_{ji}} \quad (2.6)$$

where d_1, \dots, d_N are as in (2.4) and b_{11}, \dots, b_{NR} are additional parameters, not necessarily distinct, and the constant matrices $\alpha = \{\alpha_{ji}\}$ and $\beta = \{\beta_{ji}\}$ are in each case of rank R .

First consider (2.1). The conventional cointegrating VAR similar to Johansen (1988, 1991) is obtained when $d_1 = \dots = d_N = b_{11} = \dots = b_{NR} = 1$. Instead, we shall assume $\frac{1}{2} < d_j < \frac{3}{2}$ and $0 \leq b_{ji} \leq d_j$. The basic principle here is that for a set of fractionally integrated variables $x_1 \sim I(d_1), \dots, x_N \sim I(d_N)$, possibly integrated to different orders d_i , there exist R independent linear combinations of the data, $\beta' \mathbf{w}_t$, that are integrated to lower orders than their constituent series. If

$$\beta'_i \mathbf{w}_t \sim I(d_j - b_{ji}) \quad (2.7)$$

the i th error correction term can drive the equation for $x_{jt} \sim I(d_j)$, satisfying the requirement that the equation balances, with both sides $I(0)$. Note that $d_j \neq d_k$ for $j \neq k$ is permissible, provided $d_j - b_{ji} = d_k - b_{ki}$. Also, the cointegrating residuals are allowed to exhibit long memory. There is no requirement that $b_{ji} = d_j$, which case we refer to as ‘‘full cointegration’’. In the presence of exclusion restrictions on the β coefficients, even $b_{ji} = b_{jl}$ is inessential. Long-memory ($d_j > b_{ji}$) and short-memory ($d_j = b_{jl}$) cointegrating relations, containing different variables, could in principle coexist in the same equation. One restriction that must be observed is that the two largest d_j values in a cointegrating relation must match. If the highest integration order is unique, this variable clearly cannot cointegrate with anything else, and the corresponding row of β must be 0. However, if $d_1 = d_2 > d_3$ (say), the third variable can quite feasibly appear in a cointegrating relation whose residual may be $I(d_3)$, but may equally well be $I(d)$ for $d < d_3$, since the third variable can (conceivably) cointegrate with an $I(d_3)$ combination of variables 1 and 2.

Next, consider (2.2). This represents in a slightly more general form of what has been described in previous work (Davidson 2004b) as generalized cointegration. Here the equation for variable j is driven by one or more $I(0)$ linear combinations of fractional differences

$$(1-L)^{d_1 - b_{1i}} x_{1,t-1}, \dots, (1-L)^{d_N - b_{Ni}} x_{N,t-1}, \quad i = 1, \dots, R.$$

In this case, while the observed series themselves are not (necessarily) cointegrated, fractional differences of them may be. This device allow variables integrated to different orders to be

cointegrated, and does not require restrictions on β . This notion of generalized cointegration is not one that has been widely entertained from a modelling perspective in the literature to date, but it can be thought of as generalizing the multicointegration framework, in which (say) the differences of an I(2) variable are cointegrated with an I(1) variable (see Granger and Lee 1989). Clearly, if $d_j - b_{ji}$ is the same for each j , then we are not able to distinguish the two cases, and we treat the model as of class (2.1) by default.

Our models may be compared with the bivariate setup of Dueker and Startz (1998). Their approach differs from ours, in that they construct a bivariate ARFIMA process, or in other words, a vector-ARMA in which the variables are respectively the d th fractional difference of the US bond rate, and the $d - b$ th fractional difference of a linear combination of the two series, for $d \geq b > 0$. Here, d , b and the cointegrating coefficient are of course treated as additional unknown parameters in constructing the Gaussian likelihood function. Their model is therefore, in effect, a reparameterization of our regular cointegration model, although ours admits somewhat more generality in allowing for cross-equation dynamics, but also has the advantage of generalizing the conventional vector-error correction setup of Johansen (1988). This permits direct tests of hypotheses such as Granger noncausality, for example.

Nonetheless, we need to emphasize that so far our setup is purely formal, with no consideration of the economic or behavioural context of these models. We should of course point out that models (2.1) and (2.2) can be specified on a computer to simulate systems having all the integration/cointegration characteristics we have described. From this point of view, they have validity as possible descriptions of data. However we should also recognise a major difficulty with permitting fractional processes in a behavioural modelling framework. Most models of dynamic interactions in the economic sphere are conceived as linear or nonlinear difference equations of *finite* order. This is true of costly adjustment, rational expectations, and any other explicitly dynamic economic mechanism. The only model in popular use in which lags of infinite order play an essential role is the unit root model. This can be thought of as embodying a concept of stock accumulation. An increment of physical capital, or human capital, or population, has a permanent effect on the system. However, this is a special case in which the weights are equal for all lags. Fractional processes, having declining but non-summable lag weights, have been successfully introduced into economic models only through aggregation over agents. The best known contribution here is Granger (1980), although Robinson (1978) gave a similar result independently. In these models, the micro-processes describing individual behaviour can still be cast into the conventional finite difference equation framework. Long memory may nonetheless arise through aggregation of the responses of heterogeneous agents.

An application of the Granger approach can be found in BDP, where the authors conceive of opinion polls as an aggregate of heterogeneously distributed binary responses. By the law of large numbers, this aggregate is distributed similarly to the aggregate of the conditional expected responses (probabilities of answer “Yes”) which can be modelled as evolving through time in different ways in response to new information. Voters are often thought of as either “committed” or “floating”, where the former vote by conviction, and the latter according to perceived performance. These propensities can be modelled in the former case by unit root processes (probability of “Yes” changing only in response to new information) and in the latter case by stationary autoregressive processes (probability of “Yes” moves away from the mean of 0.5 in response to new information, but otherwise reverts to it.) By combining these different cases in suitable proportions, represented by a Beta distribution of autoregressive coefficients on the unit interval, BDP apply Granger’s result to show that the aggregate should behave like a nonstationary long memory process.

We assume implicitly in this paper that the approval series under study are generated in the same kind of way. We emphasize this in order to clarify the role of the cointegrating FVECM

model in this application. In econometric research there is frequently an ambiguity, intended or otherwise, in the assumed role of the model equations. Sometimes they are taken explicitly to represent the behaviour of “representative agents”, although the difficulties inherent in modelling the aggregation process, especially when micro-behaviour is heterogeneous, tend to be overlooked. More often, nothing at all is said about aggregation, and the behavioural basis of the model fitted to macroeconomic data is left unspecified. Some readers may well impose a representative agent story of their own on the estimates, others may treat the equations as merely descriptive.

It therefore behoves us to say explicitly that our FVECM is intended to be descriptive of the relationship between the aggregate processes, and there is no suggestion that it describes the decisions of individual agents. The hypothesis to be studied here is that there is a long-run relationship between approval of the government and approval of the leader. If this relationship is broadly invariant across voters, it will be reflected in the co-movement of the aggregates, and we hypothesize that this aggregate relationship can be adequately described by a FVECM. This structure will enable us to test the leading hypotheses of interest, such as: is there partial or full cointegration, or even generalized cointegration? Does one measure Granger-cause the other, or are they interdependent? Which Prime Ministers have been the more idiosyncratic? And so forth.

3 The Approval Model

The approval series PM and Gov span the period September 1960–May 2004 (525 monthly observations)¹. Following the BDP analysis they are analysed in “log-odds” form, as

$$x_t = \log \frac{X_t}{100 - X_t}$$

where X_t denotes the reported percentage of positive responses in each month. This maps the data points from $[0,100]$ to the whole real line, and renders a nonstationary representation with infinite variance more credible, although in practice it has little effect on the results unless the series approach close to 0 or 1.

Let

$$w_{jt} = x_{jt} - \sum_{k=1}^8 \phi_k D_{kt}, \quad j = 1, 2 \quad (3.1)$$

where the D_{1t}, \dots, D_{8t} are 0-1 dummy variables to represent the periods of office of the eight Prime Ministers who have served over the sample period: Harold Macmillan, Lord Home, Harold Wilson, Edward Heath, James Callaghan, Margaret Thatcher, John Major and Tony Blair. Omitting possible vector-autoregressive terms for clarity (and also because they prove insignificant in practice) the models to be estimated take one of the following forms: either the regular cointegration form

$$\begin{aligned} (1-L)^{d_1} w_{1t} &= \alpha_1 (1-L)^{d_1-b_1} (w_{1,t-1} - \beta_1 w_{2,t-1}) + u_{1t} \\ (1-L)^{d_1} w_{2t} &= \alpha_2 (1-L)^{d_1-b_1} (w_{1,t-1} - \beta_1 w_{2,t-1}) + u_{2t} \end{aligned} \quad (3.2)$$

or the generalized cointegration form,

$$\begin{aligned} (1-L)^{d_1} w_{1t} &= \alpha_1 ((1-L)^{d_1-b_1} w_{1,t-1} - \beta_1 (1-L)^{d_2-b_2} w_{2,t-1}) + u_{1t} \\ (1-L)^{d_2} w_{1t} &= \alpha_2 ((1-L)^{d_1-b_1} w_{1,t-1} - \beta_1 (1-L)^{d_2-b_2} w_{2,t-1}) + u_{2t}. \end{aligned} \quad (3.3)$$

¹The data are from Gallup up to March 2001, and from MORI thereafter.

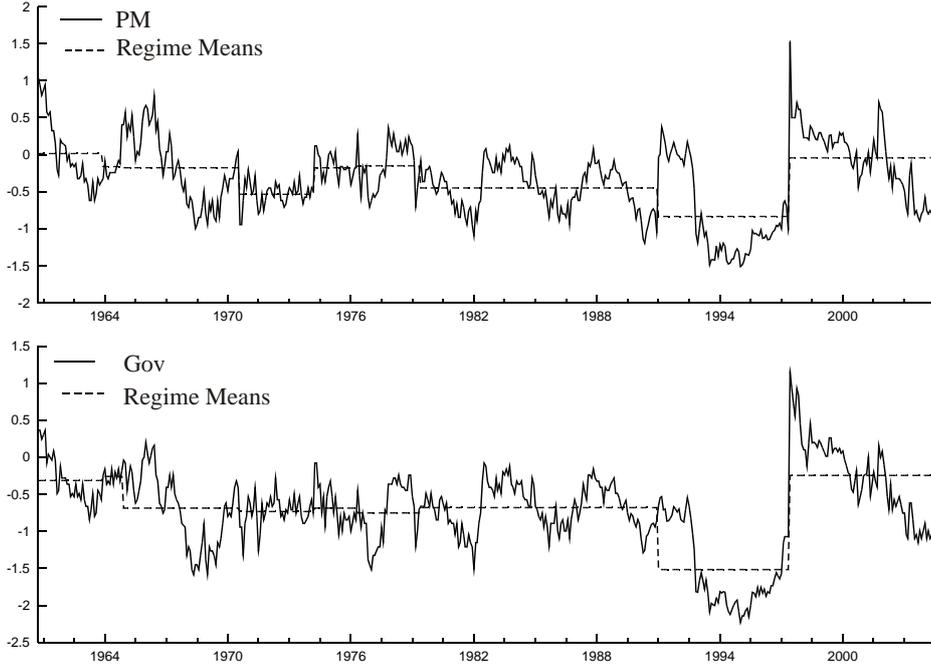


Figure 1: The approval series, showing regime means

For reasons to be discussed, it is assumed in formulating the likelihood function that $(u_{1t}, u_{2t})' = u_t$ is independently distributed as $t(\nu, \Sigma)$, or in other words that

$$f(u_t) = \frac{\Gamma((\nu + 1)/2)}{\Gamma(\nu/2)\sqrt{\pi(\nu - 2)|\Sigma|}} \left(1 + \frac{u_t' \Sigma^{-1} u_t}{(\nu - 2)} \right)^{-(\nu+1)/2}. \quad (3.4)$$

In Figure 1 the approval series (in log-odds form) are superimposed on the fitted values of a regression on the dummies, representing the average approval shown the leader and government over each episode.

A natural objection to this time series modelling approach is that the support for one prime minister is not comparable with that for his/her successor or predecessor, and therefore to treat PM and Gov as a single contiguous time series is not legitimate. Large jumps do indeed occur at the dates when personnel change, most notably in May 1997 when Tony Blair took over from John Major. While there are obvious difficulties in conducting the analysis on a set of disconnected short series, this critique is strictly correct, and ideally we should like to observe a single long run of one leader and party. Indeed, in the next section we do fit the model to the single longest leadership regime in the sample, that of Margaret Thatcher. However, the leadership regimes are consecutive, and the political and economic fortunes of successive governments are related by the progress of historical events. Combining the regimes in a suitable manner can make use of this extra information. Moreover, note that the object of the analysis is not to explain the popularity of leaders as such, but to model the relationship between leaders and governments in the public perception. There are grounds to argue that the dynamics of this latter relation should have a fundamental character, independent of individuals and regimes once the special characteristics of each has been accounted for.

However, the “splice points” between government regimes still present a special problem. These naturally feature as outlying disturbances, if otherwise unaccounted for. It’s important to note that jumps in popularity are associated with major events as well as changes of personnel.

Sample	Sept. 1960 - May 2004
Log-likelihood	1141
Schwarz SC	1063
Hannan-Quinn SC	1095
Akaike SC	1116
Error Correlation	0.807
$\sqrt{\text{Student's } t \text{ DF}}$	3.177 (0.29)
β_1 (Eql. Relation)	-1.189 (0.122)
LM Statistic, equal ds	$\chi^2(2) = 0.736$

Table 1: System Results - Regular Cointegration

Two notable instances are in May-June 1982, coinciding with the Falklands war, and September 1992, when Britain was forced out of the ERM. Since the dates are known, some system of additional dummy variables offers a possible solution, to the splice point problem, but one that complicates the model excessively. However, a more elegant solution is simply to model the shock distribution as fat-tailed. By maximizing the likelihood function derived from the Student's t distribution, instead of the Gaussian, we can effectively limit the influence of outlying observations. The likelihood based on (3.4) gives relatively small weight to outliers, and bases the parameter fit mainly on the body of the distribution. This has an effect similar to censoring the transition observations, while not discounting them altogether.

A second objection to be considered is that switches of regime of a certain frequency are well-known to create spurious long memory (Diebold and Inoue 2001). We should therefore entertain the possibility that the series are generated as a succession of stationary short-memory episodes, with occasional shifts of mean. The answer to this objection is two-fold. First, the fractional integration parameters are estimated after inserting regime dummies to pick up shifts of mean, and the nonstationary character of the series remains strongly evident. Second, these estimates are very similar to those obtained by BDP, modelling support for individual parties over much the same period, after allowing for the effect of the election cycle by a system of dummies. That swings of party support are a long memory phenomenon is something that has been attested in a range of different studies of differing democratic countries; see Box-Steffensmire and Smith (1996), BDP (1997, 2000) and Dolado et al. (2003), inter alia.

4 Results

Two estimated models of the bivariate process are reported, respectively of type (2.1) and type (2.2). The full results, except for the shift dummy coefficients, are shown, respectively, in Tables 1 and 2, and Tables 3 and 4. In each case, the first table shows the system-wide statistics, including the equilibrium relation, and the second, the equation-by-equation details. The residuals are plotted in Figure 2 and the residuals from the equilibrium relation in Figure 3. ²

In the regular cointegration model, since there are just two variables in the system their orders of integration must be equal. Accordingly, this restriction is imposed in the estimation. However, this implies that the orders of integration of the error correction terms must also match to ensure cointegration, so this restriction is also imposed. In the generalized cointegration case there is no requirement for any of the integration orders to match across variables. Therefore, no restrictions are imposed in this case and the model has two extra free parameters. The models are of course

²Estimations were performed using the Ox package Time Series Modelling Version 4. See Davidson (2005) and Doornik (1999).

	PM Equation	Gov Equation
d_1	0.908 (0.033)	*
b_1	0.542 (0.179)	*
α_1	-0.152 (0.124)	0.201 (0.085)
σ	0.149 (0.005)	0.158 (0.006)
R^2	0.886	0.9105
Jarque-Bera	341	31
Box-Pierce(12)	8.508	12.92
Box-Pierce ² (12)	7.545	17.07

Table 2: Equation Results, Regular Cointegration (except Dummies)

Sample	Sept. 1960 - May 2004
Log-likelihood	1142
Schwarz SC	1057
Hannan-Quinn SC	1092
Akaike SC	1115
Error Correlation	0.808
$\sqrt{\text{Student's } t \text{ DF}}$	3.172 (0.297)
β_1 (Eq. Relation)	-1.151 (0.94)
b_1 (PM)	0.532 (0.331)
b_2 (Gov)	0.560 (0.156)

Table 3: System Results - Generalized Cointegration

	PM Equation	Gov Equation
d	0.902 (0.141)	0.916 (0.145)
α_1	-0.170 (0.124)	.200 (0.117)
σ	0.149 (0.005)	0.158 (0.006)
R^2	0.886	0.910
Jarque-Bera	343	31
Box-Pierce(12)	8.21	13.35
Box-Pierce ² (12)	7.59	17.07

Table 4: Equation Results, Generalized Cointegration (except Dummies)

	PM Equation	Gov Equation	Equilibrium Relation
Macmillan (Conservative)	1.0192(0.045)	0.4468 (0.096)	0.48796
Home (Conservative)	1.0674 (0.043)	0.6758 (0.115)	0.26391
Heath (Conservative)	0.3532 (0.204)	0.0755 (0.169)	0.26346
Thatcher (Conservative)	1.7868 (0.093)	1.2373 (0.229)	0.31573
Major (Conservative)	2.7205 (0.111)	1.5379 (0.236)	0.89209
Wilson (Labour)	1.1814 (0.074)	0.6015 (0.176)	0.46626
Callaghan (Labour)	1.63119 (0.093)	0.97093 (0.191)	0.47685
Blair (Labour)	5.1146 (0.216)	3.9721 (0.387)	0.39216

Table 5: UK Prime Ministers: Regime Intercepts (Regular Cointegration)

identical if the orders of integration are the same. The LM statistic for the equality restrictions shown in Table 1 gives good grounds for thinking this to be the case, and the model selection criteria also favour the more restrictive model. In most respects the estimates are very close, as would be expected.

To interpret these results we assume that cointegration exists, in the sense that $b_1 = b_2 > 0$, and either $\alpha_1 < 0$, or $\alpha_2 > 0$, or both. However, note that we cannot straightforwardly test the hypothesis that $\alpha_1 = \alpha_2 = 0$ due to the so-called ‘‘Davis problem’’ (Davis 1977). Under the null hypothesis, b_1 is unidentified. Therefore, we postpone the issue of testing the existence of cointegration until the next section, and here simply assume (justifiably as we show) that it does, in other words, assume that $\alpha_1 < 0$ and/or $\alpha_2 > 0$. In fact, both of these conditions appear to apply, with the error correction coefficients having similar absolute magnitude. The estimated b_1 is significantly positive, but also significantly smaller than the estimated d_1 . In other words, there does not appear to be *full* cointegration, in the sense that the cointegrating residuals are $I(0)$. These residuals are plotted in Figure 3. Fitting an ARFIMA(p, d, q) model³ to this series yields $d = 0.55$. We observe the phenomenon of a cointegrating vector that is nonstationary or nearly so, while integrated to a lower order than its constituents. Also note that the estimated degrees of freedom for the Student’s t distribution are low (equalling the square of the fitted parameter, around 13.7) which indicates that the transition periods giving rise to large residuals (see Figure 2) are given relatively small weight in fitting the parameters.

Next, we ask what the model implies about the relationship between the perceptions of government and leadership during the forty year span covered by the sample. The intercept shifts for each leader’s period of tenure are estimated for each equation, and Table 5 shows the point estimates from the regular cointegration model. The statistics of greatest interest are the implicit shifts in the cointegrating vector, $\phi_{1j} - \beta_1 \phi_{2j}$ for $j = 1, \dots, 8$ with $\beta_1 = 1.18$, shown in the third column of the table. Observe that most prime ministers rate higher on average than the governments over which they preside, providing a measure of the charisma factor. However note (as a matter of interest primarily to students of British political history) that Edward Heath (who signed the Treaty of Rome in 1973, and subsequently fought and lost a disastrous battle with the trades unions) ranks lowest in charisma. By contrast, John Major’s government never recovered from the disastrous fall-out from the ERM in 1993, but Major was, interestingly, noticeably more popular than the government he led. Tony Blair and Mrs Thatcher, both often regarded as high on charisma, both have a middling ranking. It is notable that the extremes, high and low, were both cases of governments that came to a bad end, losing the subsequent election. Disagreements over policy may be the decisive factor in both damaging governments, and driving wedges between government and leader.

³The best model, selected by the Schwarz criterion, was ARFIMA(0, d ,0).

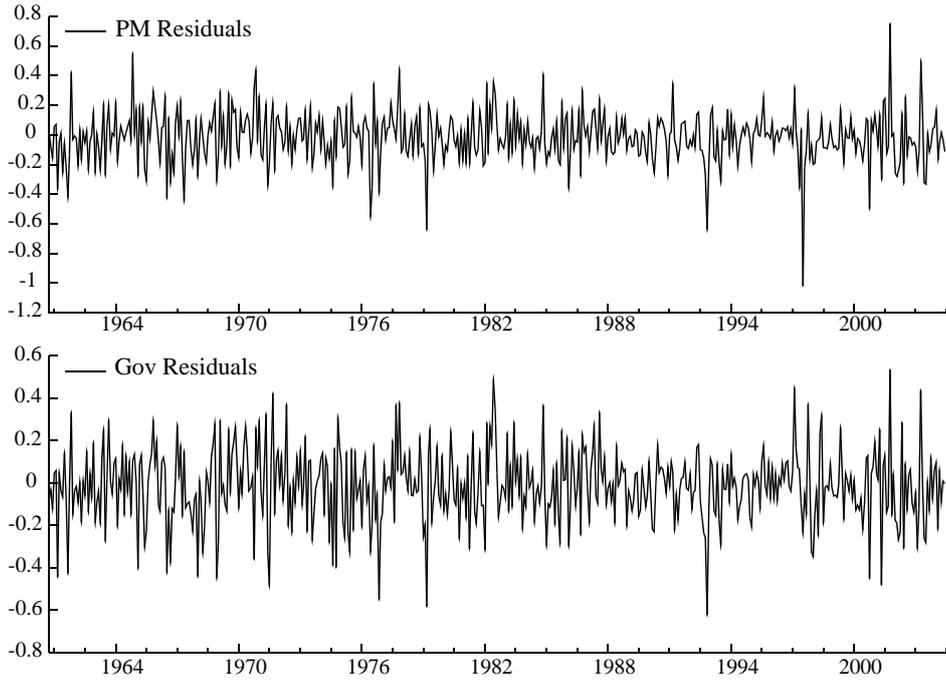


Figure 2: Residuals, regular cointegration model

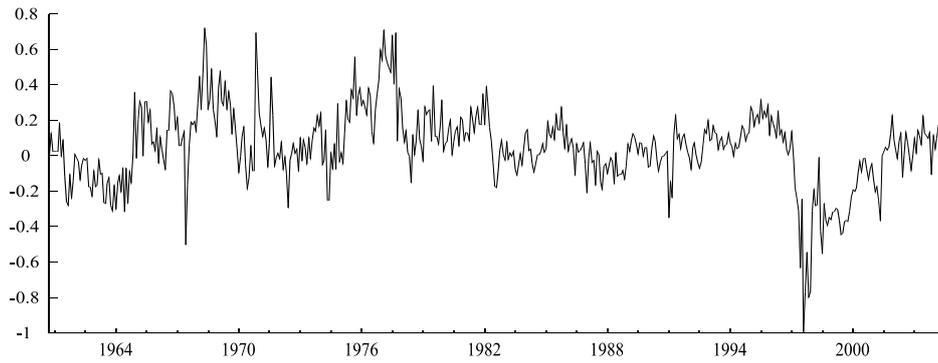


Figure 3: Equilibrium Relation, regular cointegration model

Sample	May. 1979 - November 1990
Log-likelihood	370
Schwarz SC	343
Hannan-Quinn SC	353
Akaike SC	359
Error Correlation	0.881
$\sqrt{\text{Student's } t \text{ DF}}$	3.87 (0.97)
β_1 (Eql. Relation)	-1.191 (0.206)
LM Statistic, equal ds	$\chi^2(2) = 1.069$

Table 6: System Results (Regular Cointegration) -Thatcher subperiod

	PM Equation	Gov Equation
Intercept	-0.033 (0.173)	-0.310 (0.114)
d_1	0.751 (0.108)	*
b_1	0.627 (0.099)	*
α_1	-0.391 (0.174)	-0.095 (0.577)
σ	0.131 (0.009)	0.147 (0.011)
R^2	0.803	0.748
Jarque-Bera	1.23	0.761
Box-Pierce(12)	5.86	5.39
Box-Pierce ² (12)	14.96	6.16

Table 7: Equation Results (Regular Cointegration) - Thatcher subperiod

Finally, we report estimates for the subperiod May 1979 – November 1990, corresponding to the leadership of Margaret Thatcher - the longest continuous leadership regime in the sample (139 months). These results (the regular cointegration model, except that the dummies are replaced by a simple intercept) are shown in Tables 6 and 7. Note that these have been computed without using presample data to form lags, so the data relate solely to the period indicated. The results and goodness of fit are generally similar to the full sample, supporting the view that combining the regimes is broadly legitimate. There are also differences suggesting that further modelling of the regimes could be desirable, although the shortage of observations could make this difficult in practice. While the smaller estimated d might be accounted for by the smaller sample, the main differences are the fuller cointegration, and the fact that the error correction coefficient is larger in the PM equation while effectively vanishing in the Gov equation. By this account, approval of Thatcher’s government reflects on her, but her personal support does not drive the view of her government. This would make sense if the gap between leader and party is accounted for by the charisma factor.

5 Testing Cointegration and Noncointegration

As remarked above, there is a problem in testing cointegration by testing the significance of the equilibrium relation as a driving component of the fractional system. Two parameters in each equation have to be different from zero, and the integration parameters are unidentified when the error correction loadings coefficients α_j are zero, under one form of the null hypothesis. It is true that $\alpha_1 = \alpha_2 = 0$ is not a prerequisite for non-cointegration, which follows from $b_1 = b_2 = 0$ alone. However, taking the computed standard errors at face value, the EC coefficients are not well very determined, so we need a different approach to testing the non-cointegration hypothesis.

Conventional non-cointegration tests such as the Dickey-Fuller will not serve, because their null distributions are subject to the restriction $d_1 = 1$. However, an alternative is provided by the residual-based bootstrap tests developed in Davidson (2002, 2004a, 2004b). The procedure adopted in these tests is to compare a test statistic based on least squares residuals with its null distribution under the hypothesis that the series in question are non-cointegrated. Tests for the null hypothesis that cointegration exists are also derived. In either case, the null distribution may feature short-run dependence of the $I(0)$ fractional differences, and the dummy variables can be included in the test regression. These different data features are modelled in the Monte Carlo simulation of the tests, which also entails estimating the d of the simulated series by Whittle ML for each replication, so as to mimic the true null distribution as closely as possible.

An important feature of the implementation of these tests is that only the normalized variable of the putative cointegrating set is resampled, the remainder being held fixed at their sample

values. This approach reduces computational cost and the number of parameters needing to be estimated, and also limits the hazards of mis-specification. However, to be able to condition on jointly determined variables, it is necessary to augment the lags used in the VAR modelling of the differences with a number of leads, to allow for the possibility of Granger causality of the conditioning variables by the conditioned variables. This is similar to the approach of Saikkonen’s (1991) estimator of cointegrating vectors. As discussed in Davidson (2004a), the number of leads required in fractionally integrated models should generally exceed the number of lags, to ensure a correctly sized test. However, by suppressing the leads in the bootstrap simulations, the procedure also provides a test of the stronger joint hypothesis, of noncointegration and Granger noncausality of the conditioning variables by the normalized variable. We call these hypotheses the “weak” null and “semi-weak” null, respectively. The term “strong null” is reserved for the hypothesis that there is both noncointegration and zero correlation with the fractional differences; in effect, that the series are totally independent. To ensure the tests have power, it is important not to decide the order of leads and lags by considerations of sample fit. As with any test whose performance is dependent on unknown bandwidth-type settings, the important consideration should be to determine the sensitivity of the outcome to different choices.

To test the hypothesis of noncointegration two statistics are considered, the Durbin-Watson statistic (equivalent from the bootstrap perspective to the unaugmented Dickey-Fuller) and the usual F statistic for significance of the cointegrating variables. Neither of these statistics is asymptotically pivotal, which explains why the latter test has not been entertained previously in a cointegration context. However, the properties of the F test complement those of the DW, and using them together provides added sensitivity in the detection of cointegration. The fact that these statistics are not asymptotically pivotal means that the simple bootstrap does not supply asymptotic refinements. This defect can be overcome by use of the double bootstrap, as detailed in Davidson (2004a), but that paper warns of a trade-off between errors of estimation and specification, so it should not be assumed that the double bootstrap results are more reliable; rather, taking all results together provides a check on robustness.

There is a similar implementation of tests for the null hypothesis of full cointegration. Recall that this is the hypothesis of an $I(0)$ linear combination of the test variables. These are implemented by a bootstrap version of the Shin (1994) test, based in turn on the KPSS test of stationarity (Kwiatkowski et al., 1992). Similarly to the noncointegration tests, a challenge in implementation is to strike a balance between capturing the null distribution accurately, and avoidance of overfitting that would reduce power. In this case, the main requirement is to impose a stable autoregressive root on the simulation model for the cointegrating residuals, according to the requirements of the null hypothesis. This can be done either by simply projecting the cointegrating vector onto leads and lags of the $I(0)$ fractional differences of the data, referred to as the “ $I(0)$ regression”, or solving the implicit cointegrating ECM to yield a generating equation that can be restricted to stability. The restriction imposed to enforce the null is that the autoregressive root is set to the smaller of its estimated value and 0.5.

Table 8 shows the results for tests of noncointegration, based on 1000 bootstrap replications. The tests require d_j values for each variable in the set to use for differencing the data, and these have been taken from the best ARFIMA(p, d, q) models of the series after projecting them onto the regime dummies (in other words, embodying the restriction of the null hypothesis). By the Schwarz criterion, the best choice is $p = q = 0$, and the associated d_1 estimates are 0.909 for PM and 0.865 for Gov. In the generalized cointegration case, the variable PM has been fractionally differenced by the factor 0.044 so that the two series are eligible for cointegration.

Table 9 shows the results of testing the null hypothesis of cointegration.

The tables show the computed bootstrap p -values, or in other words, the positions of the actual statistics in the bootstrap distributions. The striking feature of the results is that both

Null Hypothesis	F		DW	
	Regular	Double	Regular	Double
Regular (non-)cointegration:				
$I(d)$, weak null	0	0.11	0.005	0.005
$I(d)$, semi-weak null	0	0.14	0.006	0.014
Generalized (non-)cointegration:				
$I(d)$, weak null	0	0.13	0.012	0.014
$I(d)$, semi-weak null	0	0.12	0.006	0.01

Table 8: Noncointegration Tests: p -values

Null Hypothesis	Solved ECM		I(0) Regression	
	Regular	Double	Regular	Double
Regular cointegration	0.001	0.007	0	0.002
Generalized cointegration	0	0.006	0	0.008

Table 9: Shin Test: p -values

null hypotheses are rejected. The F test results for the double bootstrap are equivocal, but note that these are not necessarily more precise than the simple bootstrap, since specification error may play an offsetting role.⁴ These results reinforce the conclusion formed from the system analysis, that we have a case of partial but not full cointegration.

6 Conclusion

This paper proposes two variants of a vector error-correction model (VECM) for fractional processes. These models allow the generation of series that are fractionally integrated to arbitrary orders and embodying a range of cointegrating relationships. The generalized cointegration variant can be fitted unrestrictedly to an arbitrary set of fractional processes. A simple example has been exhibited, using political opinion poll data for the UK, in which fractional cointegrating relationships are estimated and tested.

Whether this type of model may prove more generally useful in macroeconomic modelling is a question that has to be left for future research. One of the obvious difficulties is that most such modelling has been founded, explicitly or implicitly, on the representative agent assumption. The forms of the equations tend to be viewed as related, at some level, to individual decision making. Economics has so far thrown up very few models of the fractional cointegrating type, both because fractional integration is a relatively novel concept in economics and, perhaps more importantly, because such models must involve infinite-order difference equations, and it is difficult to devise simple behavioural models with this property. However, if macroeconomic relationships are conceived as being generated through aggregation over heterogeneous agents, it is much easier to see how long memory can arise. In this context, the use of these models to revisit some of the standard applications in the macroeconomic lexicon might well prove a fruitful exercise.

⁴In the Monte Carlo experiments reported in Davidson (2004a), in the absence of specification error, the double bootstrap rejected more frequently than the regular bootstrap,

A Appendix: Fractional Differencing

For $-\frac{1}{2} < d < \frac{1}{2}$ the fractional difference operator is defined as.

$$(1 - L)^d = \sum_{j=0}^{\infty} \pi_j L^j \quad (\text{A-1})$$

with $\pi_0 = 1$ and

$$\pi_j = \frac{-d\Gamma(j-d)}{\Gamma(1-d)\Gamma(j+1)}, \quad j > 0.$$

For $\frac{1}{2} < d < \frac{3}{2}$, we interpret $(1 - L)^d$ as $(1 - L)(1 - L)^{d-1}$. In this case, as is well-known, the series x_t defined implicitly by

$$(1 - L)^d x_t = u_t \sim I(0) \quad (\text{A-2})$$

is nonstationary, with asymptotically infinite variance. To specify its distribution, it is necessary to specify a finite starting date $t = 0$ such that

$$x_t = \sum_{s=1}^t w_s$$

where the process $w_s \sim I(d-1)$ is stationary. This straightforwardly generalizes the standard $I(1)$ model, conceived as being the simple integral of an $I(0)$ stationary process. The theory of weak convergence for functionals of processes defined by (A-2), where the limits involve the fractional Brownian motion process defined by Mandelbrot and van Ness (1968), is treated in Davidson and De Jong (2000).

Some authors have defined the nonstationary fractional process differently, assuming that

$$(1 - L)^d x_t = u_t 1(t \geq 1) \quad (\text{A-3})$$

for $d > \frac{1}{2}$ where $u_t \sim I(0)$ and $1(\cdot)$ denotes the indicator function of its argument. In other words, the shocks are assumed to equal zero in the pre-observation period. This approach appears to us somewhat artificial, in comparison. It is important to note that it gives rise to a different asymptotic distribution theory, involving so-called ‘‘Type-2’’ fractional Brownian motion. (see Marinucci and Robinson 1999).

Since these models involve an infinite distributed lag, some scheme must be adopted for computing them from finite samples. One possibility is a truncation of fixed length J for every data point (which involves sacrificing the first J observations). Another is a moving truncation of length $t - 1$, so that every data point is included. The latter scheme is adopted for all the calculations in this paper.

References

- Box-Steffensmeier, J. M. and R. M. Smith, (1996) ‘‘The Dynamics of Aggregate Partisanship’’, *American Political Science Review* 90 (September), 567–580
- Box-Steffensmeier, J. M., S. de Boef and T.-M. Lin (2004) ‘‘The dynamics of the partisan gender gap’’, *American Political Science Review* 98(3) 515–528.
- 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

- Byers, D., J. Davidson, and D. A. Peel, (2000) The dynamics of aggregate political popularity: evidence from eight countries. *Electoral Studies* 19,1 49–62
- Byers, D., J. Davidson, and D. A. Peel, (2002) Modelling political popularity: a correction *Journal of the Royal Statistical Society Series A* 165, 187–189
- Clarke, H. D. and M. Lebo (2003) Fractional (Co)integration and Governing Party Support in Britain *British Journal of Political Science* 33(2) 283–301
- Davidson, J. (2002) A model of fractional cointegration, and tests for cointegration using the bootstrap. *Journal of Econometrics* 110, 187–212.
- Davidson, J. (2004a) Alternative bootstrap procedures for testing cointegration in fractionally integrated processes. Forthcoming in *Journal of Econometrics*, also at <http://people.ex.ac.uk/jehd201/>.
- Davidson, J. (2004b) Bootstrap tests for fractional cointegration: a reappraisal of the relationship between government popularity and economic performance in the UK, forthcoming in *New Trends in Macroeconomics*, eds. C. Diebolt and C. Kyrtsov, Springer Verlag.
- Davidson, J. (2005) Time Series Modelling Version 4.11, at <http://www.timeseriesmodelling.com/>
- Davidson, J. and R. M. de Jong (2000) The functional central limit theorem and convergence to stochastic integrals II: fractionally integrated processes. *Econometric Theory* 16, 5, 643–666.
- Davidson, J., Hendry, D. F., Srba, F. and Yeo, 5. (1978) Econometric modelling of the aggregate time-series relationship between consumers' expenditure and income in the United Kingdom, *Economic Journal* 88, 661–92.
- Davis, R. B. (1977) Hypothesis testing when a nuisance parameter is present only under the alternative, *Biometrika* 64, 2, 247-54.
- Diebold, F. X. and A. Inoue (2001) Long memory and regime switching. *Journal of Econometrics* 105, 131 –159
- Dolado, J. J. , J. Gonzalo, and L. Mayoral (2003) Long range dependence in Spanish political opinion poll data, *Journal of Applied Econometrics* 18 (2) pp. 137-155
- Doornik, J.A. (1999), *Object-Oriented Matrix Programming Using Ox*, 3rd ed. London: Timberlake Consultants Press and Oxford: www.nuff.ox.ac.uk/Users/Doornik
- Dueker, M. and R. Startz (1998) Maximum likelihood estimation of fractional cointegration with an application to US and Canadian bond rates. *Review of Economics and Statistics* 80(3) 420-426
- Granger, C. W. J. (1980) Long memory relationships and the aggregation of dynamic models, *Journal of Econometrics* 14, 227–238
- Granger, C. W. J. (1986) Developments in the study of cointegrated economic variables, *Oxford Bulletin of Economics and Statistics* 48, 213–228.
- Granger, C. W. J. and T.-H. Lee (1989) Multicointegration, *Advances in Econometrics* Vol. 8, 71-84, reprinted as Chapter 9 of *Long-run Economic Relationships* (Engle and Granger, eds.) Oxford University Press 1991.
- Johansen, S. (1988) Statistical Analysis of Cointegration Vectors, *Journal of Economic Dynamics and Control* 12, 231–54.
- Johansen, S. (1991) Estimation and hypothesis testing of cointegration vectors in Gaussian vector autoregressive models, *Econometrica* 59(6), 1551–80.
- Kwiatkowski, D., Phillips, P. C. B., Schmidt, P., and Shin, Y (1992) Testing the null hypothesis of stationarity against the alternative of a unit root. *Journal of Econometrics* 54, 159–78.
- Lebo. M. L. and W. H. Moore (2003) Dynamic foreign policy behaviour, *The Journal of Conflict Resolution* 47(1) 13-32

- Mandelbrot, B. B. and J. W. van Ness (1968) Fractional Brownian motions, fractional noises and applications. *SIAM Review* 10, 4, 422–437.
- Marinucci, D. and P. M. Robinson (1999) Alternative forms of fractional Brownian motion. *Journal of Statistical Inference and Planning* 80, 111–122.
- Robinson, P. M. (1978) Statistical inference for a random coefficient autoregressive model, *Scan. J. Statist.* 5, 163–8.
- Saikkonen, P. (1991) Asymptotically efficient estimation of cointegration regressions, *Econometric Theory* 7, 1–21.
- Shin, Y. (1994) A residual-based test of the null of cointegration against the alternative of no cointegration, *Econometric Theory* 10, 91-115.