Political Analysis (2005) 13:77–94
doi:10.1093/pan/mpi004

How Smart Is my Dummy? Time Series Tests for the Influence of Politics

Tony Caporale
Department of Economics, 331 Bentley Annex,
Ohio University, Athens, OH 45701
e-mail: caporale@ohio.edu (corresponding author)

Kevin Grier
Department of Economics, 335 Hester Hall,
University of Oklahoma, Norman, OK 73019
e-mail: angus@ou.edu

Of necessity, many tests for political influence on policies or outcomes involve the use of dummy variables. However, it is often the case that the hypothesis against which the political dummies are tested is the null hypothesis that the intercept is otherwise constant throughout the sample. This simple null can cause inference problems if there are (nonpolitical) intercept shifts in the data and the political dummies are correlated with these unmodeled shifts. Here we present a method for more rigorously testing the significance of political dummy variables in single equation models estimated with time series data. Our method is based on recent work on detecting multiple regime shifts by Bai and Perron. The article illustrates the potential problem caused by an overly simple null hypothesis, exposits the Bai and Perron model, gives a proposed methodology for testing the significance of political dummy variables, and illustrates the method with two examples.

> Before the curse of statistics fell upon mankind we lived a happy, innocent life
> —Hilaire Belloc, *On Statistics*

1 Introduction

The interaction between politics and economic policy making is an important and fascinating study. Much has been learned from the many papers using political variables to help predict the time path of economic variables (and vice versa). However, there is a large potential statistical problem in much of the literature that is seldom discussed, namely the frequent absence of a well-defined null hypothesis to the posited pattern of political influence. Of necessity, political information often enters time series models as a group of dummy variables. Often the null hypothesis is simply that the chosen political dummies are insignificant and the alternative is that politics "matters." Note that the null hypothesis implicitly embraces a constant intercept throughout the sample.

Authors' note: An earlier version of this article was presented at the 2004 American Political Science Association (APSA) meetings. The authors wish to thank John Londregan, George Krause, and three anonymous referees for their helpful comments, criticisms, and suggestions.

However, there may be significant intercept shifts in the sample that are not caused by the political factors under investigation. In this case, if the political events are correlated with the excluded true break points, they will tend to be statistically significant even though they are, in reality, not. In some sense this is just a simple omitted variable story, but in another sense it is a potential challenge to any set of results concluding that politics matters by comparing the statistical significance of political change dummies to that of a fixed intercept.

This is an important issue in political science research, as many political phenomena such as partisanship, divided government, autocracy, or coalition governments are represented empirically by dummy variables. If, in the regression under study, the dummy variable is significant, the researcher generally concludes that the political phenomenon represented is statistically important. The test is generally very stark: either the intercept shifts at the chosen point (or points) or it does not shift at all. However, life is probably not that simple. There may be other factors that in reality shift the intercept while the political factors inherently do not. But if we omit the other factors and include the political ones, we may well uncover a spuriously significant effect as the coefficient on the political variable will be biased away from zero.

Deriving the exact bias that arises when the true intercept shift is excluded in a model with multiple regressors is complicated, but the simple case of one regressor is straightforward. Suppose, for example, that the true model is given by a constant and a particular intercept shift represented by the dummy variable D*, as shown in Eq. (1):

$$Y_t = \alpha_o + \beta_1[D_t^*] + \epsilon_i \tag{1}$$

But instead, the researcher estimates an incorrect model consisting of an intercept and an incorrect intercept shift represented by the dummy variable DP as shown in Eq. (2):

$$Y_t = \alpha_o + \beta_1[DP_t] + \epsilon_i \tag{2}$$

The null hypothesis embodied in Eq. (2), that β_1 equals zero, implies that the null model is a fixed intercept with no other shift allowed. Under this incorrect null, the coefficient β_1 in Eq. (2), whose true value is zero, will have a nonzero expected value that is given by Eq. (3):

$$E[\hat{\beta}_1^{ols}] = \beta_1 \rho_{D^*,DP}[\sigma_{DP}/\sigma_{D^*}] \tag{3}$$

Equation (3) shows that the size and sign of the estimated coefficient will depend on the true coefficient on the correct variable, the correlation between the spurious and correct variables, and the relative volatilities of the variables.

If the researcher had instead tested the significance of dummy variable DP in a model that included D* [i.e., had estimated Eq. (4)], the expected value of the

$$Y_t = \alpha_o + \alpha_1[D_t^*] + \beta_1[DP_t] + \epsilon_t \tag{4}$$

coefficient on DP would have been zero.

We have concocted a simple empirical example of this phenomenon. We construct an artificial data series with 160 quarterly observations according to the following equation:

$$y_t = 6.5 + 2^* \text{Shift}_t + e_t \tag{5}$$

Shift is a dummy variable that raises the mean for 40 observations (i.e. 10 years) in the middle of the sample and the e(t)'s are independent draws from a normal distribution with mean zero and variance 0.5. The series is displayed in Fig. 1.

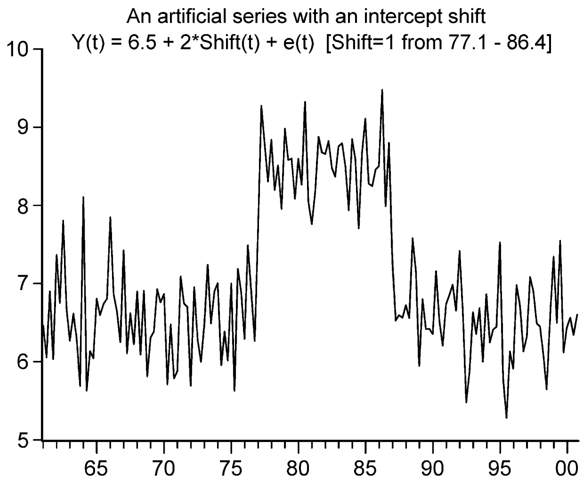

Fig. 1 An artificial series with an intercept shift Y(t) = 6.5 + 2*Shift(t) + e(t) [Shift=1 from 77.1 - 86.4].

Figure 2 illustrates our point. Suppose we have a political shift that is represented by the dummy variable shown in the figure. This spurious dummy begins three years earlier and ends four years later than the true shift. It is especially hard to see how this dummy could be said to cause the observed regime shift, since any argument involving lagged effects or anticipated effects may work at one end but not at the other. Nevertheless, under the incorrect null of no other possible shifts, this spurious dummy will have a large and significant coefficient (its linear correlation with the true shift is 0.67).

To further illustrate, we use the dummy shown in Fig. 2 along with four other dummy variables that are correlated to different degrees with the true intercept shift. These correlations range from 0.33 to 0.73 as reported in column one of Table 1. Columns 2 and 3 of Table 1 show the size and significance of these spurious dummy variables when they are estimated via ordinary least squares (OLS) under the incorrect null of no other possible

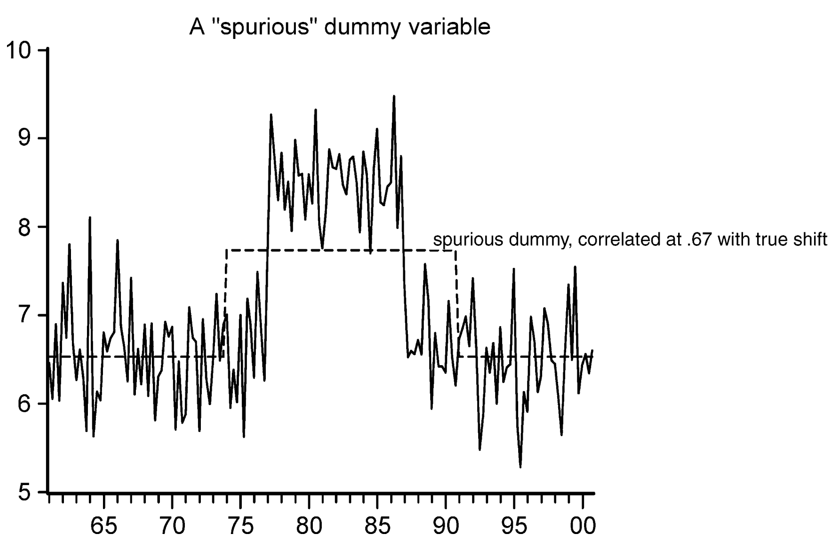

Fig. 2 A "spurious" dummy variable.

Table 1 An example of the effect of an incorrect null hypothesis on significance tests of spurious dummy variables

Correlation of Bogus Dummy with True Dummy	*OLS Coefficient on Bogus Dummy*	*Classical & HAC t stats on Bogus Dummy*	*AR(1) Corrected Coefficient & T stat on Bogus Dummy*	*Coefficient & T Stat on Bogus Dummy With True Dummy Included*
0.33	0.77	4.55 / 2.26	0.44 (1.32)	0.13 (1.3)
0.42	0.89	5.57 / 2.73	0.55 (1.76)	0.12 (1.1)
0.55	0.92	6.70 / 3.79	0.89 (3.40)	0.01 (0.0)
0.67	1.21	9.57 / 4.90	1.06 (4.93)	0.10 (0.95)
0.73	1.54	11.52 / 6.58	1.48 (6.45)	0.22 (1.63)

Note. In column 3, the first number is the classical *t* statistic; the second number is the Newey-West *t*statistic that is robust to general heteroskedasticity and autocorrelation. In columns 4 and 5 the numbers in parentheses are classical *t* statistics. All estimations for this table were done using EVIEWS3.1.

shift, while column 4 shows the size and significance of the same dummy variables when estimated using a simple correction for first-order autocorrelation in the errors under the incorrect null. Finally, column 5 of the table shows the size and significance of these spurious dummies under the correct null, which is the null that allows for the true intercept shift. The results for the particular dummy graphed in Fig. 2 appear in the fourth row of Table 1.

We can see that even when correcting the standard errors for general heteroskedasticity and autocorrelation in the errors, these five spurious dummies are positive and significant at the 0.05 level or better, with the size of their coefficients and *t* statistics being positive functions of their correlation with the true, excluded dummy. When estimated with a correction for first-order serial correlation, the last three dummies (those most highly correlated with the true dummy) are positive and significant at the 0.01 level, the second dummy is positive and significant at the 0.10 level, and the first dummy is not statistically significant. Finally, as seen in column 5, when the true dummy is included in the regression (i.e., the correct null hypothesis is imposed), the spurious dummies all have much smaller and insignificant coefficients.

The foregoing discussion is just a simple illustration of our overall point: in order to have confidence in the reported significance of political dummy variables, there needs to be a realistic null hypothesis under consideration when significance tests are performed.

In the rest of this article we outline a method to provide a more stringent null hypothesis from which to test the importance of politics. The first step is to determine the number, location, and confidence intervals for intercept shifts in the sample via time series techniques. In particular, we employ the methods recently developed by Bai and Perron (1998, 2000, 2003).

Given these statistically optimal break points, the significance of political dummy variables can then be considered in two separate stages. If the political dummies imply the same number of intercept shifts that are found by the time series methods and these political shifts fall inside the confidence intervals of the time series shifts, then there is an extremely strong case for the argument that politics fundamentally matters. If some of the politically derived shifts match up with time series shifts, the case is weakened but potentially sustainable.

If none of the political shift points line up with the time series break dates, then it is difficult to make the case that political factors cause major movements in the variable under study.

However, they may still matter in the following sense. One can take the time series break points as given and test to see whether the political dummies have any incremental explanatory power. That is to say, even though politics may not be causing large changes in the behavior of the variable, they do cause statistically significant changes, even when the larger changes are taken into account. Even this second type of demonstration would be a much more compelling argument in favor of the importance of politics than the common practice of simply testing political dummies against an otherwise fixed intercept.

In what follows we present a method for statistically determining the optimal break points based on the work of Bai and Perron, describe our suggested methodology for testing the significance of political dummy variables in more detail, and then apply this method to an investigation of the determinants of shifts in U.S. and U.K. monetary policy as measured by shifts in the real interest rate.

2 Multiple Structural Change Models

In a recent series of papers Bai and Perron (1998, 2000, 2003) developed methods to test for and estimate multiple structural changes in stationary time series.[1] The Bai and Perron (BP) methodology has the advantage of allowing the break points to be determined endogenously rather than being chosen a priori by the statistical analyst. In addition, unlike the Chow (1960) test, the CUSUM procedure of Brown et al. (1975), or the work of Andrews (1993), the BP methodology allows for possible multiple breaks to be uncovered. Finally, BP methods allow testing for significant breaks in subsets of parameters (partial multiple break models) rather than testing for the stability of all coefficients jointly.

Before we present a more formal discussion of the BP model, it may be useful to discuss at an intuitive level the logic behind the tests detailed below. As is well known, a linear time series model with K coefficients can be tested for parameter instability at a known break point by comparing the sum of squared errors (ESS) of the model estimated on the full sample (the restricted model) with the sum of the separate ESSs in the two subsamples (the unrestricted model). If the difference between the two is large enough, the hypothesis that the parameters are time invariant will be rejected using a standard F test.

If the break point date is not known with certainty, a series of F tests for each of the possible breaks can be calculated, and the break point associated with the largest F statistic would be the logical candidate for the location of a break, if the test were statistically significant. Clearly, though, distribution of this maximal F (labeled SupF for supremum of the possible F's) is not the same as a single F statistic, as we have grossly violated classical hypothesis testing assumptions by using the same sample to both find the break and then test its significance. In cases like this, the appropriate critical values have to be determined by simulation exercises (i.e., Monte Carlo analyses).

The BP methodology can be seen as a logical extension of the prior break point research. It deals with the problems raised by data series that may have more than one break and develops a way to optimally find the number as well as the location of the structural shifts in the parameters. The method finds the best one-break model via the SupF method, then the best two-break model by the same method. At this point we can test whether the optimal one-break model fits the data better than the optimal two-break model using a different nonstandard F test that compares the fit of the two empirically derived

[1]The endogenous determination of break points has been an important research topic in empirical macroeconomics since Perron's (1989) demonstration that standard Dickey-Fuller unit root tests are sensitive to allowing for intercept and/or trend shifts (see Zivot and Andrews 1992 and Perron 1997).

models. This test, which BP would call SupF (2/1), also requires its critical values to be determined experimentally. This process can continue up to the largest number of breaks the researcher is willing to consider and will produce an estimate of both the number of structural breaks and their locations in time.

Now we will more formally present the Bai and Perron (1998) framework using its most simple (intercept shift only) version. However, it is important to note that the procedure is in no way limited to this simple case. There can be other variables in the model, and their coefficients can either be subject to shifts or remain constant over the sample depending on the choice of the investigator. We present this simple case for ease of exposition using the following multiple linear regression with m break points (m+ 1 regimes):

$$\begin{aligned} y_t &= \beta_1 Z_{t1} + E_t; \qquad t = 1, 2, \ldots, T_1 \\ y_t &= \beta_2 Z_{t2} + E_t; \qquad t = T_1 + 1, \ldots, T_2 \\ &\vdots \\ &\vdots \\ y_t &= \beta_j Z_{tm+1} + E_t; \qquad t = T_{m+1}, \ldots, T \end{aligned} \tag{6}$$

According to this specification, y_t is the observed dependent variable at time t, Z_t is a column of ones, β_j (j = 1, 2,. . .m+ 1) are coefficients (the value of the constant for each regime), and E_t is the disturbance term. The break points (T_1,. . . T_m) are explicitly treated as unknown. We seek to estimate the unknown regression coefficients together with the break points. This allows us to test for structural changes in the mean of the series.[2]

Bai and Perron (1998) present test statistics to detect multiple breaks. For each partition m, (T_1,. . . Tm), the associated least squares estimates of β_j are obtained by minimizing the sum of squared residuals:

$$\sum_{i=1}^{m+1} \sum_{t}^{Ti} (Y_T - Z_T' \beta_j)^2 \tag{7}$$

This provides us with $\hat{\beta}_j(T_1 \ldots T_M)$, which are the estimates associated with the m partition ($T_1 \ldots T_M$). Substituting them into the objective function (7) yields the estimated break points

$$(\hat{T}_1, \ldots \hat{T}_M) = \min{}_{T_1 \ldots T_M} S_T(T_1, \ldots T_M) \tag{8}$$

where S_t ($T_1 \ldots T_M$) is the sum of squared residuals. Using the estimated break points ($\hat{T}_1$,. . .$\hat{T}_M$) the parameter estimates found are $\hat{\beta}_j(\hat{T}_1, \ldots \hat{T}_M)$.

Bai and Perron first suggest a SupFt(L) F statistic to test the null of no structural breaks (L = 0) versus the alternative hypothesis that there are m = k breaks. This tests $\beta_1 = \beta_2 = \ldots \beta_{m+1}$. The procedure searches all possible break dates and minimizes the difference between the restricted and unrestricted sum of squares over all the potential breaks.[3]

[2]The Bai-Perron (1998) procedure corrects for serial correlation and different variances across segments by incorporating Andrews (1991) robust standard errors. All break models employed in this paper utilize this correction.

[3]The break values can depend on the imposition of the minimal length of a segment (h). This is determined by the value of the trimming parameter (E) that must be specified to estimate the model. Since E = h/T, a lower value of E implies a smaller minimum regime size. Bai and Perron (1998) provide recommendations for E based on sample size and the maximum number of break points allowed.

Bai and Perron (1998, 2003) next propose two tests of the null hypothesis of no breaks against at least 1 through M breaks. These are called double maximum tests. The first, a Udmax statistic, is the maximum value of the SupFt(L), where L is an upper bound on the possible number of breaks. The second, a Wdmax test, weights the individual statistics such that the marginal p values are equal across values of m. This implies weights that depend on the significance levels of the tests. The null hypothesis of both tests is no structural breaks against an unknown number of breaks given some specific upper bound on the possible number of break points.

If the null of no break is rejected by the double maximum tests, Bai and Perron next suggest a sequential SupFt(L+1/L) procedure to determine the number of structural breaks. The statistic tests the null of L breaks against the alternative of L+1 breaks. Rejection in favor of a model with L+1 breaks occurs if the overall minimum value of the sum of squared residuals is sufficiently smaller than the sum of squared residuals from the one-break model. The number of break dates selected is the number associated with the overall minimum error sum of squares.[4]

Finally, estimates of the break dates need not be the global minimizers of the sum of squared residuals. A sequential procedure can also be used to select the number of breaks in which, if an initial break is found [based on the initial SupFt(1) test], the sample is then divided into subgroups at the break point, and the same parameter constancy test is then performed on the subsamples. The partitioning of the subsamples continues until the parameter constancy test fails to reject the null. Bai and Perron (1998, 2003) are able to develop a method to compute confidence intervals for the sequential break points by employing a novel asymptotic theory that assumes the magnitudes of the breaks decline as the sample size increases.

A useful check on the number of breaks found using the sequential method is supplied by Bai's (1997) repartition estimation procedure. Starting with T-consistent estimates of k_i from the sequential procedure ($\hat{k}_i, i = 1, 2.$), k_1^0 is reestimated using the subsample $[1, \hat{k}_2]$ and k_2^0 is reestimated using $[\hat{k}_1, T]$ Cases in which these estimators (call them $\hat{k}_1^*$ and $\hat{k}_2^*$) reveal the same number and location of breakdates provides us with additional confidence in our results based on the sequential procedure.[5]

In sum, the BP methods can be used to test for the existence of structural breaks, along with the number and location of those breaks. Confidence intervals on the break dates can also be constructed. We argue that this method imposes a needed reality check on the true significance of political dummy variables that are often included in regressions where the intercept is otherwise constant.

3 A Proposed Methodology for Evaluating the Significance of Political Dummy Variables in Time Series Equations

In this section, we outline the steps involved for testing the significance of political dummy variables in time series equations using the BP methods described above. Put briefly, the proposed methodology goes as follows:

(1) Choose the maximum number of break points allowed and the error structure to be assumed in the tests;

[4] Critical values for these tests are found in Bai and Perron (1998, 2003).

[5] Two final methods, the Bayesian information criteria (BIC) and Schwartz's criteria (LWZ), have also been proposed as additional ways of determining break dates. However, Bai and Perron (2003), using Monte Carlo experiments, show that the sequential procedure works better than these alternatives, and thus we do not explore them here.

(2) Estimate the number, location and confidence intervals for the structural breaks using the BP methods;
(3) Check to see if the political dummy variables fall within the confidence intervals for the empirically optimal breaks. If so, declare victory as the political variables are closely associated with major shifts in the series. If not:
(4) Check to see if, given the empirically optimal break points, the political dummies still have incremental explanatory power via non-nested hypothesis tests.

An important preliminary step is to determine whether the variables under consideration are stationary.[6] Bai and Perron (2000, p. 10) explain that their methodology "precludes intergrated variables (with an auto-regressive unit root) but permits trending regressors." This issue is even less straightforward than usual here because the possibility of regime shifts makes standard unit root tests potentially biased. In the interests of continuity and brevity, we relegate a full discussion of this issue to the appendix and proceed to elucidate the four-point process outlined above.

(1) Given a resolution to the stationarity question, we next turn to specifying the model. An initial trimming percentage (ε) must be specified in order to ensure a reasonable amount of degrees of freedom to calculate an initial error sum of squares. For example, if $\varepsilon = .15$, then breaks will be considered for the middle 70% of the sample. Bai and Perron's Gauss break point program allows for $\varepsilon = .05, .10, .15, .20, .25$.

The trimming specification determines the maximum possible number of breaks as well as the minimum regime size. For instance, when $\varepsilon = .10$ the maximum number of breaks is eight, since allowing nine breaks forces the break dates to be exactly at 10% intervals (.1, .2,. . .9). When $\varepsilon = .15$ the maximum (M) number of breaks is 5; M = 3 for $\varepsilon = .20$ and M = 2 for .25. Therefore, for a series with sample = 100 quarters and $\varepsilon = .15$, there is a maximum of five breaks (six regimes) where each regime has a minimum length of 15 quarters.

Next, assumed error structure inside each segment and across segments must be specified and modeled. Bai and Perron consider the following possibilities:

A. No serial correlation and constant variances of the errors within and between segments.
B. No serial correlation and different variances of the errors in between segments.
C. Serial correlation and constant error variances within and across segments.
D. Serial correlation in the errors and nonconstant error variances within and between segments.

For case A, the critical F values, point estimates as well as confidence intervals of the break points are generated using classical OLS assumptions. Cases B and D require the researcher to use the Andrews (1991) HAC standard error option to estimate the model as well as choose an option (for case D) that allows the variance of the residuals to be different across all segments. Case C requires a prewhitening of the data prior to estimation of the structural break model.

As a practical matter, Bai and Perron have created free GAUSS code that implements any of these cases. They recommend choosing the most general error structure (case D),

[6]Variables that have a unit root (or stochastic trend) have population parameters (mean or variance) that are time dependent. Since it is well known that correlations between independent nonstationary time series can often be spurious, variables must be rendered stationary prior to being used for parameter estimation and hypothesis testing.

since consistent estimates of the break points are still assured even if the corrections are implemented on a case A dataset.

(2) Bai and Perron (1998, 2003) argue that the global and sequential SupFt(L+1/L) tests provide the most reliable estimates of the number and location of the break dates. Clearly the strongest evidence would involve both methods yielding the same answers. In cases in which they disagree, Bai and Perron (1998) suggest that the global procedure should be used for any model with more than one significant break.[7]

Beyond determining the number and location of the breaks, 90, 95, and 99% confidence intervals for the break dates are provided. The derivations of these values are explained in Bai and Perron (2000, pp. 11–13) and rest on the use of a novel asymptotic framework in which the magnitudes of the shifts converge to zero as the sample size increases.

(3) Given these confidence intervals, we can examine how well the proposed political dummy variables (which are intercept shifts) correlate with the break points estimated via BP methods. If the number of regimes implied by the political model is equal to the number implied by the time series model, and each political break point falls inside the confidence interval of a BP break, that will be very strong evidence in favor of the primacy of the political effects. They are responsible for all the major break points in the sample. If there is a partial match between political breaks and BP breaks, the importance of politics can still be argued, but the matter becomes open to interpretation.

(4) Even if the political breaks are not closely related to the BP regime shifts, they may still be statistically significant and important variables. One can allow for this possibility by testing whether, taking the time series shift points as given, the political dummy variables have any incremental explanatory power.

Perhaps the most straightforward way to accomplish this is via Davidson and McKinnon's J-test methodology for non-nested hypothesis testing. This allows us to test the hypothesis that model X rejects (or dominates) model Y by including the fitted value of model X as an additional variable in the model Y regression. If model X's predicted values are insignificant, the validity of model Y is not rejected. Alternatively, if model X's predicted value is significant, we conclude that model Y is rejected by model X. The same procedure can then be used to test model X against Y. The strongest evidence concerning the superiority of a model (say X) would be that it dominates model Y and fails to be dominated by model Y. In our case, we would be testing whether the fitted values from a political intercept shift model had any incremental explanatory power in the BP intercept shifting model.[8] The following section presents two real-world applications of our proposed method.

4 Application: Testing for Politically Related Shifts in U.S. and U.K. Real Interest Rates

The effect of political changes on monetary policy is an important and long-studied subject that of necessity often uses dummy variables to test for political effects. Papers using this method routinely appear in top journals in both economics and political science.[9] Of

[7] A divergence can occur in the case in which the SupFt(1) test is insignificant but the SupFt(2) test rejects zero breaks in favor of 2. The sequential procedure will then stop at zero breaks whereas the global SupFt(2/1) test may suggest a two-break model.

[8] One could also consider testing this hypothesis by including both the empirical break point dummies and the political break point dummies in the regression of interest and using an F test of the hypothesis that the political dummies all had coefficients of zero.

[9] A few examples include Hibbs (1977), Beck (1982), Alesina and Sachs (1988), Hakes (1990), Grier (1991, 1996), and Krause (1994).

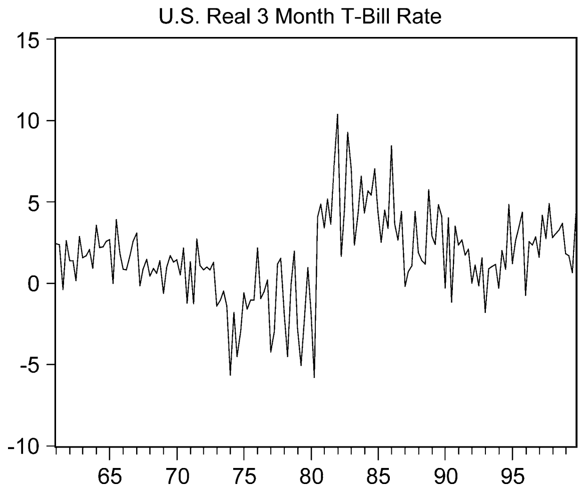

Fig. 3 U.S. Real 3 Month T-Bill Rate.

course, there are many (all imperfect) ways to measure monetary policy. Given the decline in correlation between monetary aggregates and economic outcomes, a consensus has emerged in favor of using a short-term interest rate as the policy measure. Yet nominal interest rates can be misleading policy indicators without controlling for inflation. That is to say, an 8% interest rate in a zero inflation environment is indicative of restrictive policy, but that same rate in a period of 10% inflation is not at all restrictive.[10]

Caporale and Grier (2000) use a methodology similar to the one exposited above and show that, in the U.S. case, large shifts in the real interest rate are closely related to changes in party control of either the executive or one of the legislative branches of the federal government, and not at all related to changes in the chairmanship of the Federal Reserve.

4.1 *U.S. Real Rate Break Points*

We begin by reestimating their model using a longer sample for the United States. We use a quarterly sample of 156 observations from 1961.1 through 1999.4. The real interest rate is taken from the St. Louis Federal Reserve database and is defined as the three-month treasury bill rate minus the inflation rate calculated using the consumer price index. Figure 3 displays the data.

We consider two specific political-economic hypotheses: (1) political change affects monetary policy, and (2) bureaucratic change affects monetary policy. Table 2 describes the changes in party control of branches of the federal government and changes in the chairmanship of the Federal Reserve during the sample. The first two columns of Table 3 report simple dummy variable regressions showing whether either set of political variables is statistically significant in isolation. Specifically, political change dummies account for about 47% of the variation in the real rate and the Federal Reserve chair dummies for about 38%.[11]

[10]For a further discussion, see Caporale and Grier (1998).

[11]Note that while these regression are extremely parsimonious, the standard errors of the coefficients are estimated using the Newey-West formula, making them consistent in the face of arbitrary types of autocorrelation and heteroskedasticity.

Table 2 U.S. political and bureaucratic regimes 61–99

Political		*Federal Reserve Chairs*	
Clinton	93.1–99.4	Greenspan	87.4–99.4
Reagan-Bush	81.1–92.4	Volcker	79.4–87.3
Carter	77.1–80.4	Miller	78.2–87.2
Nixon-Ford	69.1–76.4	Burns	70.1–78.1
Kennedy-Johnson	61.1–68.4	Martin	61.1–77.4
Republican Congress	81.1–86.4 and 95.1–99.4		

Now we consider finding the optimal number and location of break points using the BP methods outlined above. As shown in Table 4, the two general tests for the presence of structural breaks, the UDmax and Wdmax tests, are both significant at the 0.01 level. We consider finding the optimal number of breaks by using both the SupF $(L+1 \mid L)$ tests and the sequential procedure. In both cases, four breaks are chosen with dates of 1967.2, 1973.1, 1980.3, and 1986.2.[12] These break points are fairly tightly estimated; the 95% confidence intervals are 1966.1–1969.2, 1971.4–1973.2, 1980.1–1981.1, and 1985.1–1987.3. These results are reported in Table 2, and column 3 of Table 3 shows that the BP dates account for around 58% of the variation in the real rate.

As can be seen by comparing the political and bureaucratic change dates in Table 2 with the confidence intervals for the empirically estimated break dates in Table 4, the BP estimated breaks capture the Democrat to Republican presidential change in 1969.1, the Democrat to Republican presidential and Senate change in 1981.1, and the Republican to Democrat Senate change in 1987.1. The Republican to Democrat presidential change in 1977.1 does not correlate with a structural shift, nor do the Republican to Democrat presidential change in 1993.1 or the Democrat to Republican House and Senate change in 1995.1.[13]

By contrast, each change of the chairmanship of the Federal Reserve occurred outside the confidence intervals for the structural breaks. This is a striking result, as it shows that despite popular belief, large changes in monetary policy are not largely determined by changes in Federal Reserve leadership.

Given the coincidence of the BP break points and political changes, we believe that there is a significant effect of large political changes on the real interest rate. In this subsection, we go on to consider whether, given the null hypothesis of break points only at the dates uncovered by the Bai and Perron procedures, there is any additional evidence of political influence on the U.S. real interest rate. That is to say, we assume the BP break points are not political and test to see if political changes achieve statistical significance taking the BP dates as given. We accomplish this by means of non-nested hypothesis tests.[14]

Columns 1 through 3 of Table 3 report OLS regressions of the BP break dates, the party change break dates, and the Federal Reserve chair change break dates on the real interest rate. In column 4 the predicted values of the political change regression are used as an additional regressor in the BP break date regression. As can be seen, the resulting

[12] These are the exact same dates reported in Caporale and Grier (2000). Also note that for both our U.S. and U.K. real rate results, the repartition method produced the identical number, location, and confidence intervals for the break dates as the sequential procedure.

[13] However, these last two omissions may not be surprising because they happen very close to each other and intuitively we might expect the two changes to offset each other.

[14] Our test can be recognized as one-half of the so-called J-test methodology of Davidson and Mackinnon (1981).

Table 3 Alternative mean shifting models of U.S. real rates

	Political Model	*Federal Reserve Chair Model*	*Structural Break Model*	*Structural Break + Political Fit*	*Structural Break + Federal Reserve Fit*
Constant	1.64 (9.10)	1.55 (8.08)	1.88 (11.62)	1.36 (5.14)	1.94 (6.58)
Clinton	−1.29 (−2.46)				
Reagan-Bush	0.60 (1.44)				
Carter	−2.69 (−4.10)				
Nixon-Ford	−1.90 (−3.32)				
Rep Congress	2.61 (5.67)				
Burns		−2.06 (−3.72)			
Miller		−3.65 (−6.77)			
Volcker		2.38 (3.01)			
Greenspan		0.65 (1.82)			
BP Regime 2			−1.03 (4.99)	−0.61 (−1.95)	−1.07 (−4.06)
BP Regime 3			−2.45 (6.60)	−2.16 (−4.93)	−2.49 (−6.17)
BP Regime 4			6.84 (15.06)	5.29 (6.79)	7.03 (8.96)
BP Regime 5			−3.04 (−72.9)	−2.42 (−6.97)	−3.10 (−7.18)
Pol fit				0.31 (2.51)	
Fed fit					−0.04 (−0.27)
Adjusted R^2	.47	.38	.58	.59	.58

Note. All regressions are estimated using the Newey-West HAC corrected standard errors with lag truncation = 4. *T* statistics in parentheses.

coefficient is positive and significant at the 0.05 level, indicating that even with this stringent null hypothesis, party change is a significant determinant of the real interest rate. By contrast, the predicted values from the Federal Reserve chair regression are completely insignificant when added as an ancillary regressor to the BP break date regression, as shown in column 5.

In sum, then, the evolution of the U.S. real interest rate is significantly influenced by changes in party control of branches of the federal government, but not by changes in the chairmanship of the Federal Reserve. The structural break points in the series are reasonably closely correlated with political changes, and additional significant information is carried in the political change dates beyond the information in the optimal break points. Neither of these results obtain for the Federal Reserve chair change dates.

4.2 *The U.K. Real Rate Break Points*

We now turn our attention to the U.K. real interest rate. We use the same 1961.1–1999.4 quarterly sample and again define the real interest rate as the three-month treasury bill rate minus inflation calculated using the CPI. In this case the raw data come from the International Monetary Fund's IFS CD-ROM. The U.K. real rate variable is displayed in Fig. 4, while Table 5 lists the dates of changes in the political party controlling the government and changes in the person holding the governorship of the Bank of England. Columns 1 and 2 of Table 6 report regressions of the U.K. real rate on the set of dummy variables obtainable from these political (column 1) and bureaucratic (column 2) changes. In each case the dummy variables are significant as a group at the 0.01 level. The political (bureaucratic) change dummies account for 32% (24%) of the variation in the U.K. real rate.

As shown in Table 7, when we test for structural breaks in the mean using the BP UDmax and Wdmax tests, we reject the null hypothesis of no breaks at the 0.01 level. We

Table 4 Pure structural break model results: U.S. real rate

Sup Ft(1)	Sup Ft(2)	Sup Ft(3)	Sup Ft(4)	Sup Ft(5)
26.26*	29.43*	39.12*	37.26*	31.49*
Sup Ft(2/1)	Sup Ft(3/2)	Sup Ft(4/3)	Sup Ft(5/4)	
37.35*	37.35*	23.62*	1.57	
Ud_{max}	Wd_{max} (10%)	Wd_{max} (5%)	Wd_{max} (1%)	
39.12*	63.89**	69.11**	78.83*	
Number of breaks selected				
Sequential procedure		4		
Repartition procedure		4		
Break point dates and 95% confidence interval				
		$\hat{T}_1$	67.2	(66.1–69.2)
		$\hat{T}_2$	73.1	(71.4–73.2)
		$\hat{T}_3$	80.3	(80.1–81.1)
		$\hat{T}_4$	86.2	(85.1–87.3)

$^*p < .01$, $^{**}p < .05$, $^{***}p < .10$

again consider finding the optimal number of breaks by using both the SupF (L+1 | L) tests and the sequential procedure. In both cases, three breaks are chosen, with dates of 1970.4, 1980.3, and 1993.3. The 1980.3 break point has a tightly estimated confidence interval of 1980.1–1982.4. The other two breaks are not as tightly estimated. Their confidence intervals are 1964.2–1971.2 and 1991.2–1999.1.[15]

We can see that the first U.K. break point corresponds with the change in government from Liberal to Conservative in 1970.2, and the last U.K. break point corresponds with the change in government from Conservative to Liberal in 1997.1, though this is due mainly to the imprecise estimate of the BP break date. Interestingly, the famous Liberal to Conservative government switch in 1979.2 is close to but outside of the 95% confidence interval for the second U.K. break point by three quarters. The Conservative to Liberal government change in 1964.4 is not related to any BP break date.

Turning to changes in the director of the Bank of England, the change from Cromer to O'Brien in 1966.3 falls in the confidence interval for the first BP break point, and the change from Leigh to George in 1993.3 coincides exactly with the second BP break point. The other two changes in the director of the Bank of England (O'Brien to Richardson in 1973.3 and Richardson to Leigh in 1983.3) are unrelated to any BP break dates.

The U.K. case is less clear cut than the U.S. case. There are five political regime shifts, two of which are in BP shift confidence intervals, but in neither of these cases is the political shift "close" to the exact BP break point. Further, the most famous political shift in the sample, from Callahan to Thatcher, is not related to any BP shift. There are four Bank of England director shifts in the data; one of them coincides directly with a BP shift point and another falls inside a BP shift confidence interval but again is not very close to the estimated break point.

[15]Note that one of these breaks corresponds exactly to a U.S. break, namely the one at 1980.3. The other two estimated U.K. break dates do not fall inside the confidence intervals for the U.S. breaks. Thus there is a significant amount of independent variation in the U.K. series, and we are probably justified in treating it as a separate case. Using some relatively strong assumptions (relative purchasing power parity, the international Fisher effect, and uncovered interest parity), one can derive a real interest parity condition for international interest rates. This condition implies that single country real rates do not move independently of the "world" real rate. Real interest parity is seldom found in the data.

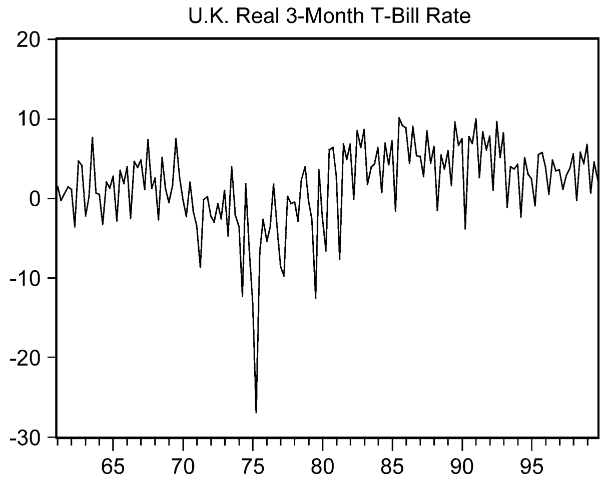

Fig. 4 U.K. Real 3-Month T-Bill Rate.

Neither set of changes relates to the real rate as well as does political change in the United States, nor does one type of change dominate the other as was the case in the United States.

We now consider whether, taking the BP break points as given, either party government change or Bank of England director change has any additional, independent explanatory power for the U.K. real rate. Columns 4 and 5 of Table 6 show that neither the predictions of the party change model nor the Bank of England change model have any significant explanatory power above and beyond the BP break points. That is to say, the political and bureaucratic variables that are significant when tested against the null hypothesis of an otherwise fixed intercept are insignificant when tested against a null that includes the statistically optimal intercept shifts.

4.3 *Discussion*

Of the four sets of dummy variables studied (political change U.S., political change U.K., Federal Reserve chair change U.S., and Bank of England director change U.K.) all are significant against the null of a fixed intercept. None of these sets of dummy variables match up precisely to the statistically optimal intercept shifts uncovered by the BP methods used here, though the political change U.S. dummies are reasonably close.

Further, only the political change U.S. dummies are significant when tested against a null that includes the BP optimal break points. These examples illustrate both the use of the BP methodology and the problems that can arise when testing for the influence of dummy variables against an overly simple null hypothesis.

Table 5 U.K. political and bureaucratic regimes 61–99

Prime Ministers with Same Party		*Bank of England Directors*	
Blair$^{(L)}$	97.2–99.4	George	93.3–99.4
Thatcher-Major$^{(C)}$	79.2–97.1	Leigh	83.3–93.2
Wilson-Callahan$^{(L)}$	74.1–79.1	Richardson	73.3–83.2
Heath$^{(C)}$	70.2–73.4	O'Brien	66.3–73.2
Wilson$^{(L)}$	64.4–70.1	Cromer	61.2–66.2
MacMillan-Douglas Home$^{(C)}$	61.1–64.3		

Table 6 Alternative mean shifting models of U.K. real interest rates 1961.1–1999.4

	Political Model	*Central Bank Chair Model*	*Structural Break Model*	*Structural Break + Political Fit*	*Structural Break + Central Bank Fit*
Constant	1.03 (2.29)	1.07 (3.21)	1.63 (4.4)	1.32 (2.21)	1.48 (3.64)
Blair	2.51 (3.91)				
Thacker-Major	3.04 (3.96)				
Wilson-Callaghan	−5.67 (−2.85)				
Heath	−2.64 (−3.91)				
Wilson	1.18 (1.92)				
George		2.19 (5.02)			
Leigh		4.29 (8.06)			
Richardson		−2.68 (−1.56)			
O'Brien		−0.55 (−0.51)			
Regime 2			−5.28 (−4.42)	−4.44 (−4.20)	−4.92 (−4.32)
Regime 3			8.88 (7.35)	7.53 (4.62)	7.97 (6.42)
Regime 4			−2.13 (−3.92)	−2.09 (−3.73)	−2.06 (−3.88)
PM fit				0.20 (0.64)	
Bank fit					0.19 (1.26)
Adjusted R^2	.32	.24	.41	.41	.41

Note. All regressions are estimated using the Newey-West HAC corrected standard errors with lag truncation = 4. *t* statistics in parentheses.

These examples also allow us to point out some of the limitations of the BP methodology. Since the BP procedure can only detect statistical break points in the data, it cannot tell us why the shift is occurring (nor does it purport to). Clearly no empirical technique is a substitute for a well-specified theory.

To further illustrate the importance of theory, note that our empirical hypotheses in this article are tested with important auxiliary assumptions, the most relevant being that political or bureaucratic regime changes should have a contemporaneous effect on the variable under study. This allows us to argue that the Thatcher regime switch in 79.2 did not have a significant effect on U.K. real interest rates, since the break point in the series that takes place in 80.3 has a confidence interval that only reaches back to 80.1. Certainly a model that allows for a delayed effect of the policy change may well argue the opposite.

If a researcher has a well-specified theory that includes lead or lag effects of policy changes, there is nothing that precludes that theory from being tested using BP techniques. Clearly for some economic variables (like real output), longer lags may be easily justified in picking up the effects of the policy regime changes, while for others (like interest rates), a longer lag structure would be harder to justify. The danger with overreliance on (especially overly loose) lead or lagged effects is that it can lead to any result "confirming" a theory.

Our tests also assume that the real interest rate is constant inside each regime and subject to infrequent regime shifts. This may seem like a stark model, but it has been employed several times in the literature.[16] Other theories may imply a range of slope coefficients that shift over time or an intercept that shifts along with slope coefficients that do not. The BP methodology will accomodate these permutations, but the substantive

[16]For example, Garcia and Perron (1996), Caporale and Grier (2000), Bai and Perron (2003).

Table 7 Pure structural break model: U.K. real rate

Sup Ft(1)	Sup Ft(2)	SupFt(3)	Sup Ft(4)	Sup Ft(5)
35.86*	31.58*	23.76*	20.06*	16.47*
Sup Ft(2/1)	Sup Ft(3/2)	Sup Ft(4/3)	Sup Ft(5/4)	
16.61*	15.93*	0.72	0.00	
Ud_{max}	Wd_{max} (10%)	Wd_{max} (5%)	Wd_{max} (1%)	
35.86*	35.87*	37.53**	41.47**	
Number of breaks selected				
Sequential procedure		3		
Repartition procedure		3		
Break point dates and 95% confidence interval				
		$\hat{T}_1$	70.4	(64.2–71.2)
		$\hat{T}_2$	80.3	(80.1–82.4)
		$\hat{T}_3$	93.3	(91.2–99.1)

*$p < .01$, **$p < .05$, ***$p < .10$

answer obtained clearly can depend on the specification of the model. That is, if the reader is not happy with the form of the regression we used, that is an issue to take up with us and not with the BP method.

5 Conclusion

In this article we proposed using the statistical methods of Bai and Perron to create more stringent tests for the influence of political dummy variables on time series data. We show that the correlation between central bankers and monetary policy is more apparent than real in the United States and the United Kingdom and that political change is an important determinant of monetary policy (as measured by short-term real interest rates) in the United States.

Our examples involve time series models of monetary policy with intercept shifts only. However, the technique applies more broadly. The methods can be used to uncover the number and location of break points in slope coefficients in a model as well as just intercept shifts. We are currently working on extending our tests for political influence on monetary policy in this direction.

Finally, to close as we began, both the problem described and the method presented here have applications far beyond the study of monetary policy, since political dummy variables are widely used in empirical research in the social sciences on topics like government spending, taxation, tariff rates, exchange rates, and probably many others.

Appendix: Testing for Stationarity in Series with Possible Regime Shifts

Although the time series literature now contains many different tests for unit roots, the most commonly used procedure is the augmented Dickey Fuller (ADF) test. It tests for an auto-regressive unit root using the following equation:

$$\Delta Y_t = B_0 + \delta Y_{t-1} + \psi_1 \Delta Y_{t-1} + \ldots \psi_p \Delta Y_{t-p} + U_t, \tag{A1}$$

where the null hypothesis that the series has a unit root ($\delta = 0$) is tested against the alternative of stationarity ($\delta < 0$). The number of lags in the ADF regression must be determined by some criterion. A common choice is to use the number of lags that minimize the value of the Akaike information criteria (AIC).

The ADF test can also be modified by allowing for an alternative hypothesis in which the series contains a deterministic trend. The appropriate ADF regression becomes

$$\Delta Y_t = B_0 + \alpha T + \delta Y_{t-1} + \psi_1 \Delta Y_{t-1} + \ldots \psi_p \Delta Y_{t-p} + U_t \tag{A2}$$

and the computed ADF statistic is the OLS t-statistic testing $\delta = 0$ against $\delta < 0$.

However, for present purposes, standard ADF tests are problematic. Zivot and Andrews (1992) explain that these tests can be misleading (fail to reject a false null) if the time series undergo discrete structural changes, which is exactly the situation we are studying. They demonstrate that adding level and/or trend shifts to ADF equations lowers the chances of falsely concluding that a break or trend break stationary series is nonstationary.[17]

Their procedure involves altering the standard ADF test by estimating

$$\Delta Y_t = B_0 + \alpha T + \varphi DU_t(\lambda) + \delta Y_{t-1} + \psi_1 \Delta Y_{t-1} + \ldots \psi_p \Delta Y_{t-p} + U_t \tag{A3}$$

where (λ) indicates the potential break point in the time series and DU_t denotes a level shift that equals one at and after the break point and zero before. The unit root test is then the *t* statistic evaluating the null of $\delta = 0$ (nonstationary) against $\delta < 0$ (break stationary) using new critical values supplied by Zivot and Andrews. They propose estimating the model for every possible (λ) and choose the one that minimizes the *t* statistic on the coefficient δ.[18]

References

Alesina, Alberto, and Jeffrey Sachs. 1988. "Political Parties and the Business Cycle in the United States." *Journal of Money Credit and Banking* 20:63–82.

Andrews, Donald W. K. 1991. "Heteroskedasticity and Autocorrelation Consistent Covariance Matrix Estimation." *Econometrica* 59(3):817–858.

Andrews, Donald W. K. 1993. "Tests for Parameter Instability and Structural Change with Unknown Change Point." *Econometrica* 61(4):821–856.

Bai, Jushan. 1997. "Estimating Multiple Breaks One at a Time." *Econometric Theory* 13:315–352.

Bai, Jushan, and Pierre Perron. 1998. "Estimating and Testing Linear Models with Multiple Structural Changes." *Econometrica* 66:47–78.

Bai, Jushan, and Pierre Perron. 2000. "Multiple Structural Changes: A Simulation Analysis." Boston University Working Paper.

Bai, Jushan, and Pierre Perron. 2003. "Computation and Analysis of Multiple Structural Change Models." *Journal of Applied Econometrics* 18:1–22.

Beck, Nathaniel. 1982. "Presidential Influence on the Federal Reserve." *American Journal of Political Science* 26:415–445.

Brown, R. L., J. Durbin, and J. M. Evans. 1975. "Techniques for Testing the Constancy of Regression Relationships over Time." *Journal of the Royal Statistical Society, Series B* 37:149–192.

Caporale, Tony, and Kevin Grier. 1998. "A Political Model of Monetary Policy with Application to the Real Fed Funds Rate." *Journal of Law and Economics* 41:409–428.

Caporale, Tony, and Kevin B. Grier. 2000. "Political Regime Change and the Real Interest Rate." *Journal of Money, Credit, and Banking* 32:320–334.

Chow, Gregory. 1960. "Testing the Equality between Sets of Coefficients in Two Linear Regressions." *Econometrica* 28:591–605.

[17] Both of the real interest rates that we investigate in this paper were overwhelmingly found to be break point stationary using the ZA procedure. This finding is consistent with many recent studies on real rates (see Garcia and Perron 1996 and Caporale and Grier 2000).

[18] The test can also be performed by interacting the trend with the break point dummy and by allowing both an intercept and a trend shift. Zivot and Andrews (1992) provide critical values for these alterative tests as well.

Davidson, Russell, and James G. MacKinnon. 1981. “Several Tests for Model Specification in the Presence of Alternative Hypotheses.” *Econometrica* 49:781–793.

Garcia, Rene, and Pierre Perron. 1996. “An Analysis of the Real Interest Rate under Regime Shifts.” *Review of Economics and Statistics* 79:327–337.

Grier, Kevin. 1991. “Congressional Influence on U.S. Monetary Policy: An Empirical Test.” *Journal of Monetary Economics,* October:201–220.

Grier, Kevin. 1996. “Congressional Influence on U.S. Monetary Policy Revisited.” *Journal of Monetary Economics,* December:571–580.

Hakes, David. 1990. “The Objectives and Priorities of Monetary Policy under Different Federal Reserve Chairmen.” *Journal of Money, Credit and Banking* 22:327–337.

Hibbs, Douglas. 1977. “Political Parties and Macroeconomic Policy.” *American Political Science Review* 71:1467–1487.

Krause, George. 1994. “Federal Reserve Policy Decision Making: Political and Bureaucratic Influence.” *American Journal of Political Science* 38:124–44.

Perron, Pierre. 1989. “The Great Crash, the Oil Price Shock and the Unit Root Hypothesis.” *Econometrica* 57:1361–1401.

Perron, Pierre. 1997. “Further Evidence from Breaking Trend Functions in Macroeconomic Variables.” *Journal of Econometrics* 80:355–385.

Zivot, Eric, and Donald W. K. Andrews. 1992. “Further Evidence on the Great Crash, the Oil-Price Stock, and the Unit-Root Hypothesis.” *Journal of Business and Economic Statistics* 10(3):251–270.