Bangor Business School Working Paper

BBSWP/10/007

THE INTERTEMPORAL RELATION BETWEEN GOVERNMENT REVENUE AND EXPENDITURE IN THE UNITED KINGDOM, 1750–2004

By

Lusine Lusinyan
European Department
International Monetary Fund
Washington DC, USA

John Thornton
Division of Financial Studies
Bangor Business School

March, 2010

Bangor Business School
Bangor University
Hen Goleg
College Road
Bangor
Gwynedd LL57 2DG
United Kingdom
Tel: +44 (0) 1248 382277
E-mail: business@bangor.ac.uk

The intertemporal relation between government revenue and expenditure in the United Kingdom, 1750–2004

Abstract

We examine the intertemporal relation between government revenue and expenditure in the UK during 1750–2004. We pay particular attention to long-run trends by applying a battery of unit root and cointegration techniques to the data, and we use a modified Granger-causality test on data spans organized around structural breaks in the series. The results suggest that, allowing for structural breaks, UK real revenue and spending are $I(1)$ series and cointegrated and that Granger-causality runs from government spending to revenue. As such, the 'spend-tax' hypothesis appears to best characterize the long-run intertemporal relation between government revenue and spending in the UK.

Keywords Government revenue and expenditure · Unit roots · Cointegration · Causality · Structural breaks

JEL Classification E62 · H61 · H62

The intertemporal relation between government revenue and expenditure in the United Kingdom, 1750–2004

1 Introduction

There is a substantial empirical literature on the intertemporal relation between government expenditure and revenue in the generation of fiscal deficits.[1] Much of it has focused on whether a 'causal' relation exists between revenue and spending and the resulting implication for fiscal consolidation strategies. There are four main hypotheses in this regard. Unidirectional causality from taxation to expenditure ('tax-spend') is suggested by Friedman (1978, 2003), who argues that tax cuts generate politically intolerable fiscal deficits that eventually force spending cuts, and that a strategy of raising taxes to reduce fiscal deficits will likely fail because the additional revenues will result in an increase in government expenditure. This pattern of causality is also consistent with Buchanan and Wagner (1977), who argue that tax cuts reduce the perceived cost of government programs, leading to a greater demand for such programs, more government spending and larger fiscal deficits. Unidirectional causality from spending to taxation ('spend-tax') is consistent with Peacock and Wiseman's (1961, 1979) 'displacement effect', which stresses the propensity for government expenditures to increase permanently in the face of temporary developments (such as major wars); and Barro's (1979) Ricardian equivalence proposition that government borrowing (to finance spending) today results in an increased future tax liability. In these contexts, fiscal consolidation should focus on cutting spending. Bidirectional causality is suggested by the 'fiscal synchronization hypothesis' associated with Musgrave (1966) and Meltzer and Richard (1981), who argue that the public chooses simultaneously an optimal package of spending programs and the taxes

[1] Payne (2003) provides a survey of recent studies on the issue.

necessary to fund the spending. Finally, Wildavsky (1988) and Baghestani and McNown (1994) have advanced an 'institutional separation hypothesis' under which decisions on taxation are taken independently from the allocation of government expenditure, such that no causal relation between revenue and spending is to be expected. In the case of bidirectional causality or no causality, whether fiscal consolidation is carried out through tax increases or cuts in spending would not affect the likelihood of the consolidation being successful.

A notable feature of the many associated empirical studies is that they often draw quite different conclusions as to the intertemporal relation between government revenue and spending for the same country. In the United Kingdom (UK) context, for example, support for a 'tax-spend' hypothesis is reported by Koren and Stiassny (1998) during 1956–92, and Chang et al. (2002) for 1951–96; support for a 'spend-tax' tax hypothesis is reported by Joulfaian and Mookerjee (1991) for 1961–86; and support for a 'fiscal synchronization' hypothesis is reported by Hasan and Lincoln (1997) for 1961–93 (quarterly data), and Owoye (1995) for 1961–90. Finally, Ram (1988) finds support for the tax-spend hypothesis for the period 1958–85 using constant price fiscal variables, but support for the fiscal synchronization hypothesis when the variables are expressed in current prices. The differences in these results could be attributed to the differences in the sample periods, model specifications, and the choice of econometric methodology. But the reliability of the test results can also be affected by the empirical strategies employed that are based on relatively short data spans and Granger-causality tests or impulse response functions using mainly error-correction and cointegrated vector autoregressive (VAR) models. In particular, the validity of causality test results can be questioned when these are formulated on the basis of the outcomes of unit root and cointegration tests that are known to suffer from size and power problems in small samples and are suspect if there is reason to believe that the country has experienced a structural break in its fiscal policies over the sample period (Perron 2006).

In this paper, we contribute to the empirical literature on the intertemporal relation between revenue and spending in the UK in several respects. First, we focus on the historical experience using a data span of more than 200 years, which substantially increases the number of observations used compared to the studies noted above. Second, we take as our point of departure the framework of the government intertemporal budget constraint and fiscal policy sustainability, the most recent literature on which has emphasized flexible testing strategies based on cointegration.[2] Thus, we pay particular attention to long-run trends in the data by applying a battery of unit root tests and cointegration techniques, including tests that allow for structural breaks in the data. The analysis of structural breaks in the context of nonstationarity and cointegration tests provides us with a basis for investigating revenue-expenditure causalities over specific sub-samples where no such breaks are identified.[3] Finally, we report results from a modified Granger-causality test suggested by Dolado and Lütkepohl (1996) applied to data spans organized around structural breaks in the series. Our results suggest that, among the hypotheses discussed above, the 'spend-tax' hypothesis appears to be the one that would most accurately characterize the long-run intertemporal relation between government revenue and expenditure in the UK.

The next section of the paper provides a brief description of long-run trends in government revenue and expenditure in the UK during 1750–2004; the empirical

[2] In particular, as shown by Quintos (1995), given that government revenues and expenditures are nonstationary, sustainability requires these variables to be cointegrated with a unit slope on expenditures.

[3] Even though we report Granger-type causality results only for the sub-samples that are free of structural breaks, our results cover a longer time span and are based on more observations than the other empirical studies of the UK experience. For example, our shortest sub-sample is 1951–2004, which still has more observations that the samples analyzed in the other studies with the exception of Hasan and Lincoln (1997) who use quarterly data for 1961–93.

methodology and results are presented in Section 3; and concluding observations are summarized in Section 4.

2. The data and some stylized facts

The choice of specific definitions of fiscal variables varies widely in the empirical literature on the relation between government spending and revenue, and the appropriateness depends largely on the objectives of the study (Baghestani and McNown 1994). The studies whose main focus has been on revenue-expenditure causality have used real or nominal series as well as GDP ratios expressed in logarithms; and the VAR models employed in these studies have often included additional macroeconomic control variables (such as GDP growth and inflation). In contrast, studies that have examined the revenue-spending relationship from the viewpoint of fiscal policy sustainability, for which the government's intertemporal budget constraint has served as an analytical framework, have used mainly levels of real fiscal variables (e.g., Ahmed and Rogers 1995; Quintos 1995). As our point of departure is also the government's intertemporal budget constraint framework, we also employ levels of real fiscal variables in our study.

Annual data for the UK government real revenue, real spending, and the fiscal balance for the period 1750–2004 are shown in levels and first differences in Figure 1 (panels A and B, respectively) and natural logarithms of levels and the first differences in Figure 2 (also panes A and B, respectively).[4] Barro (1987) and Clark and Dilnot (2002) have stressed the short-run dominant effect of wars on the UK public finances in the 18th and 19th centuries and the first half of the 20th

[4] The series are from Mitchell (1988) and the Annual Abstract of Statistics published by the UK National Statistics Office (NSO), and are deflated using the consumer price index compiled by O'Donoghue et al. (2004). The data are calendar year; the fiscal data are for the central government and are total revenue and expenditure (including transfers). While the information on the details about the data used in other similar studies is not readily available, some studies have used the Government Finance Statistics or the International Financial Statistics of the IMF, or the OECD Economic Indicators as the sources for the data on central government expenditure and revenue

century, whereby wars were typically associated with large increases in public spending and relatively smaller increases in taxation. As noted above, Peacock and Wiseman (1961, 1979) have stressed the enduring effect of the size of government in the UK data, with public spending falling back after the wars but to levels much higher than those that had preceded the war, referring to this as the 'displacement' effect. Clark and Dilnot (2002) suggest that the 'displacement effect' may occur because wars ease the constraint on politicians' ability to raise taxes as much as they would like and provide scope for higher levels of non-military spending once war is over. After 1945, the main features of the data on the spending side are the temporary boost from the Korean War in the early 1950s, and two periods of rapid growth in the first half of the 1970s and in the early 1990s, which mainly reflected higher cyclical spending associated with downturns in the economy. In the second half of the 1990s, spending declined as the economy strengthened but picked up again during the last years of the sample period because of a structural increase in social benefits and pension payments (Emmersen et al. 2003).

The main features on the revenue side are the sharp increase in the tax burden in the mid-1960s to 1970 and the decline in revenues in the late 1980s to the early 1990s, followed by a rapid increase thereafter.[5] Reflecting the above developments, the fiscal position shifted from being generally in deficit in the early part of the sample period to being mainly balanced or in surplus in the period between the Napoleonic War and World War I, and persistently in deficit after the early 1950s. Thus, UK government revenue and spending generally display a high degree of co-movement over the long run, which is suggestive of cointegration, but Figures 1 and 2 make clear that there have been several periods in which the series have diverged markedly, indicating the need to take account of possible structural

[5] Clark and Dilnot (2002) note that between 1964 and 1970, the increase in UK government revenue was the largest significant and sustained increases in the tax burden in the 20th century that was not associated with a war or pre-war military build-up.

changes in fiscal policy. Indeed, the data in levels, both with or without logarithmic transformation, show the episodes, mainly coinciding with major wars, where a change in the level of the series and/or in trend could have potentially taken place.

4. Empirical methodology and results

In this section, we focus on the time series properties of UK government real revenue and expenditure and present the results of unit root and cointegration tests, including from tests that allow for structural break(s) in the series.

3.1 Unit Root Tests

Three sets of unit root tests are used to examine the stationarity properties of the levels and first difference of the government revenue and spending series. The first set comprises tests that are relatively common in the literature but have been criticized for their bias towards non-rejection of the null hypothesis of a unit root against the alternative of (trend) stationarity in the presence of structural breaks and low power for near-integrated processes. These are the Augmented Dickey-Fuller (ADF) test developed by Dickey and Fuller (1979) and Said and Dickey (1984); the DF-GLS test developed by Elliot et al. (1996), which is a modified Dickey-Fuller test that has improved power in small samples; and the Phillips and Perron (1988) test.

The second set of tests allows for endogenously determined structural breaks in the series and comprises the Zivot and Andrews (1992) and Lee and Strazicich (2003, 2004) tests. With these tests, the structural breakpoint is determined by utilizing a grid search over a range of possible breakpoints and choosing the year when the unit root t statistic is minimized. The former test has been criticized because it does not allow for breaks under the null hypothesis of a unit root, which may bias the test and lead to size distortions in the presence of a unit root with a break and/or loss of power (Nunes et al. 1997, Glynn et al. 2007, and Perron 2006). However, the minimum Lagrange Multiplier (LM) unit root test proposed by Lee and Strazicich (2003, 2004) for one and two breaks allows for breaks under both the null and alternative hypotheses and avoids the problems of bias and spurious rejections of the null hypothesis, with the alternative hypothesis unambiguously implying trend-stationarity.[6] We consider two types of breaks for the Lee-Strazicich test—a level shift and a shift in both level and trend, which corresponds to Perron's (1989) "crash" Model A (change in level/intercept) and Model C with a simultaneous change in level and growth path, respectively.[7] The data suggest shifts around the World War I and II (and also around the turn of the 19th century for the log-series); in particular a possible level-trend shift for the series in levels and a level shift for the log-levels. We follow the literature and report the results for all four specifications (two models for one and two-break tests each).[7]

The final set of tests allows for an exogenously determined break point, and comprises the test proposed by Saikkonen and Lütkepohl (2001, 2002) and Lanne et al. (2002). This test is based on estimating the deterministic term first under the unit root null hypothesis and then performing an ADF-type test on the adjusted series, including terms to correct for estimation errors in the parameters of the deterministic part. We report the test results when the beginning of World War II

[6] In addition, Zivot-Andrews and other similar ADF-type endogenous break tests tend to select the break point incorrectly (one time period before the true break), where the bias and spurious rejections are the greatest. In contrast, the break point(s) tend to be determined correctly at the true break when using the LM test, with the accuracy increasing with the magnitude of the break(s). But even when the size of the break(s) is small and the break point cannot be accurately estimated, the LM test does not suffer a significant loss of power in this case as this is similar to having no break (Lee and Strazicich 2003, 2004).

[7] Lee and Strazicich (2003, 2004) do not examine Perron's (1989) "changing growth" Model B as it is commonly held that most economic time series can be adequately described by model A or C.

is taken (1939) as a break point (though the results remain largely robust to other choices for a break point).[8]

The results of these different unit root tests for the real revenue and real expenditure series are reported in Table 1 for the levels and first differences of the data and in Table 2 for the natural logarithms of the series. Most of the tests, with and without breaks, suggest that the series in levels behave like unit root processes, while for the first differences the unit root can be strongly rejected; support for the presence of a unit root is particularly strong for the revenue series.[9] The tables show that the most frequent break points determined by the unit root tests are associated with war events, including the Anglo-Boer War (1899–1902), World War I (1914-18), and World War II (1939-45). We conclude from these results that government revenue and expenditure are integrated I(I) provided that structural breaks in the series are taken into account (particularly level shift(s)), and that cointegration analysis of the revenue and expenditure series is appropriate.[10]

3.2 Cointegration of revenue and expenditure

An assessment of the cointegration of UK real revenue and expenditure is needed to determine the appropriate formulation of Granger-type causality tests

---

[8] Unit root tests with more than two breaks are often the extensions of the tests that do not allow for break(s) under the null hypothesis of unit root, while such extensions for the Lee-Strazicich type tests are, to our knowledge, not available. Hence, note that some loss of power can be expected from ignoring more than two breaks in the one- or two-break test.

[9] This is in line with the findings of Ahmed and Rogers (1995), who use only the Phillips and Perron (1988) and Perron (1989) tests, the latter of which is based on the assumption of a known single structural break (including for the year 1939, for which, however, Ahmed and Rogers reject the unit root null for the tax revenue series).

[10] We also carried unit root tests of the revenue and expenditure series over the different sub-samples but do not report them here because our main focus is the causal relation between revenue and expenditure, and because the validity of the results from the modified Granger-causality test suggested by Dolado and Lütkepohl (1996) does not depend on knowing the order of integration of the variables and the test does not assume that the cointegration structure of the system under investigation is known. The unit root test results for the sub-samples suggest a likely I(1) behavior for both revenue and expenditure, though there is more evidence of trend stationarity in the post-World War II period when allowing for the presence of structural breaks. These results are available on request.

(Engle and Granger 1987). We report the results from several types of test in Table 3 (levels of the series) and Table 4 (logarithms of the series). The first type of cointegration tests is the residual-based test of Gregory and Hansen (1996a,b) with one endogenously determined break. Tests are carried out for four cases of a structural change in the cointegrating relation, including a level shift, a level shift with trend, a level shift with a change in the slope coefficient or a regime shift (with no trend), and a regime shift with a trend shift. Three test statistics, based on ADF and the Phillips Z(t) and Z(a) tests, are reported for the null of no-cointegration against the alternative of cointegration with a break. The results in panel I of Tables 3 and Table 4 suggest that the series are cointegrated; however, as in the case for the unit root tests discussed above, the value of the break point associated with the minimal value of a given statistic is not, in general, a consistent estimate of the break date if a change is present. Accordingly, in panels II–IV of Tables 3 and 4 we report the results from the tests that are based on a multivariate approach where revenue and spending are modeled by a VAR process. In panel II, we report results from the Saikkonen and Lütkepohl (2000a,b,c) cointegration tests, which involve a GLS-type detrending procedures for estimating the trends component of the series for the cases where: (i) there is a linear trend in both the variables and cointegrating relation and (ii) there is a trend only in data and not in the cointegrating relation. For each case, specifications with and without level shifts have been considered but only the results without level shifts are presented.

The last two panels in Tables 3 and 4 present the results from the cointegration analysis based on a rank test developed by Johansen et al. (2000), in which government spending and revenue are modeled by a VAR process with one or two breaks in the deterministic component and known break points. When the structural breaks are chosen around the major World Wars, the test results for the VAR in levels appear to suggest that the system is stationary, hence instead of a cointegration analysis, standard inference would apply (Table 3). However, the limit distribution of the Johansen et al.(2000) test depends on exact specification of the deterministic terms and on the true break dates and, as argued by Lütkepohl et al. (2003), for cases of a VAR process with level shift, size properties of the Johansen et al.-type tests could be unsatisfactory. We also performed the trace tests developed by Johansen (1988, 1991, 1995) with dummy variables to correct for large outliers (mainly in the expenditure series). The model in levels passes only partly the misspecification tests (including the multivariate and univariate residual-based tests) and does not give robust support for cointegration.[11] In general, more robust and stronger cointegration results for all specifications are obtained when the logarithms of the revenue and spending series are used (Table 4).[1212]

As discussed earlier, the behavior of the fiscal balance appears to have changed since the early 1950s, which is suggestive of a change in the long-run relation between government revenue and spending. Indeed, when, as a break point, the post-1950 years are chosen, we find evidence of cointegration particularly when allowing for joint break(s) in level(s) and trend. In turn, pattern that emerged of larger and more slowly reversing fiscal deficits from the 1950s would be consistent with Peacock and Wiseman's 'displacement effect' with regard to government expenditure, the political economy difficulties in maintaining the associated higher levels of taxation that this implied, and the deepening of UK government debt markets which facilitated the financing of fiscal deficits. Not surprisingly, the tests suggest a presence of strong cointegrating relation between spending and revenue in the sub-samples when the war-related breaks are excluded.

---

[11] The results from the Johansen (1988, 1991, 1995) trace tests with dummy variables to correct for large outliers are available on request.

[12] This is likely to reflect a transformation of the VAR system into a more suitable model for the Johansen-type tests, i.e. closer to a Gaussian model.

3.3 Revenue and expenditure 'causalities'

The findings from the unit root and cointegration analyses underscore the importance of taking due account of structural breaks when analyzing the long-run relation between revenue and expenditure in the UK. Similarly, structural stability is an important condition for standard Granger-causality tests. Lütkepohl (1991), for example, shows that the Granger-causality tests may over-reject the true null when mean shifts are ignored, although the problems may be avoided if the number and the dating of the breaks are known. Baldé and Rodríguez (2005) argue that the presence of additive outliers contaminates the exact size of the Granger-causality statistic and propose adjusting for additive outliers (using dummy variables) when conducting hypothesis testing in a VAR. In this section, we focus on the causality tests for the sub-samples that have been partitioned based on the break points detected and used in the previous sections.[13] In particular, using the break points around the major wars we analyze revenue-spending causalities over the pre-World War I (1750–1913) and post-World War II periods (1947–2004) for the series in levels and logarithms.

We carry out tests for modified Granger-causality using the methodology suggested by Dolado and Lütkepohl (1996). They proposed a simple method which guarantees standard χ^2 asymptotics for Wald tests performed on the coefficients of cointegrated VAR processes with $I(1)$ variables if at least one coefficient matrix is unrestricted under the null hypothesis. If all the matrices are restricted, adding one extra lag to the process and concentrating on the original set of coefficients result in Wald tests with standard asymptotic distributions. In addition, despite a potential loss in efficiency when using the modified procedure, the test performs well if the true order of the process is not known, as is often be the case in

[13] It is well known that in a bivariate cointegrated system there must be Granger-causality in at least one direction (Engle and Granger 1987).

empirical studies. Bauer and Maynard (2006) extend this methodology and show that it can provide Granger-causality tests that accommodate stationary, nonstationary, near-stationary, long-memory, and un-modelled structural break processes within the context of a single $\chi 2$ null limiting distribution.

The Dolado and Lütkepohl (1996) test can be described as follows: let $y_t = (REV_t, EXP_t)'$ be a VAR(k) cointegrating process for real revenue and expenditure with vector error correction model (VECM) representation given by:

$$\Delta y_t = \Pi y_{t-1} + \sum_{i=1}^{k-1} \Gamma_i \Delta y_{t-i} + \Phi D_t + \varepsilon_t \qquad t = 1, \ldots, T \tag{1}$$

where Π (with rank r) and Γ are coefficient matrices, D_t are the deterministic terms, and ε_t is the error term. For testing whether one variable of the system is Granger-noncausal for the other, the levels VAR form of the model is considered but the test is based on a model with $k+1$ lags of the endogenous variables

$$y_t = \sum_{i=1}^{k+1} \begin{bmatrix} \alpha_{11,i} & \alpha_{12,i} \\ \alpha_{21,i} & \alpha_{22,i} \end{bmatrix} y_{t-i} + \Phi D_t + \begin{bmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \end{bmatrix} \tag{2}$$

The null hypothesis that REV_t (EXP_t) is Granger-noncausal for EXP_t (REV_t) is tested by checking the hypothesis $\alpha_{21,i} = 0$ ($\alpha_{12,i} = 0$), $i = 1, \ldots, k$. In this setup, one can also test for so-called instantaneous causality, which is characterized by nonzero correlation of ε_{1t} and ε_{2t}, and for which the null hypothesis $\mathrm{E}(\varepsilon_{1t}\varepsilon_{2t}') = 0$ is tested against the alternative of a nonzero covariance between the two error vectors.

The Granger-causality test results are reported in Table 5, where panels I and II correspond to the VAR model based on the levels and log-levels of the revenue and spending series, respectively. These tests are based on the cointegrated VAR models specified for each sub-sample, including by: (i) establishing the significance of the deterministic terms and whether these are restricted or unrestricted; and (ii) determining the lag order on the basis of the information criteria and checking this with a residual-based analysis focusing on the tests for autocorrelation as well as with the significance of lagged covariates. Despite some differences between the results from the two types of models, the null-hypothesis of non-causality is rejected almost in all cases for the expenditure series while the opposite is true for the revenue series. The results thus indicate the presence of unidirectional causality from spending to taxation. As such, of the four main hypotheses of around which the empirical literature is organized, the 'spend-tax' hypothesis associated with Peacock and Wiseman's (1961, 1979) 'displacement effect' and Barro's (1979) Ricardian equivalence proposition appears to be the one that would most accurately characterize the intertemporal relation between government revenue and expenditure in the UK.

4. Conclusions

In this paper, we examined the intertemporal relation between government revenue and expenditure in the UK during 1750–2005. We have built on the previous studies of the issue by focusing on the UK's long-run experience with data spanning more than 200 years, paying particular attention to long-run trends in the data by applying a battery of unit root and cointegration techniques, including tests that allow for structural breaks, and by applying a modified Granger-causality test to data spans organized around structural breaks in the series. Most of the unit root and cointegration tests, both with and without breaks, and when applied to the data in levels or natural logarithms of the levels, suggest that allowing for structural breaks, UK real revenue and spending during the period were $I(1)$ series and cointegrated. In addition, the results from the modified Granger-causality tests generally support a pattern of causality running from government spending to government revenue. This is consistent with the long run intertemporal relation between government revenue and expenditure in the UK being characterized by the 'spend-tax' hypothesis associated with Peacock and Wiseman's 'displacement effect' and Barro's Ricardian equivalence proposition.

References

Ahmed, S., Rogers. J.H.,1995. Government budget deficits and trade deficits: Are present value constraints satisfied in long-term data? Journal of Monetary Economics 36, 351–374

Baghestani H., McNown, R., 1994. Do revenues or expenditures respond to budgetary disequilibria? Southern Economic Journal 60, 24–31.

Baldé, T.A., Rodríguez, G., 2005. Finite sample effects of additive outliers on the Granger-Causality test with an application to money growth and inflation in Peru. Applied Economics Letters 12, 841-844

Barro, R. J., 1987. Government spending, interest rates, prices, and budget deficits in the United Kingdom, 1701–1918. Journal of Monetary Economics 20, 221–247

Barro, R. J., 1979. On the determination of government debt. Journal of Political Economy 81, 940–971

Bauer, D., Maynard, A., 2006. Robust Granger causality tests in the VARX framework. Unpublished manuscript. Available at: www.amaynard.ca/papers/grangint.pdf

Buchanan, J.M., Wagner, R.E., 1977. Democracy in deficit: the political legacy of Lord Keynes (New York: Academic Press).

Chang, T., Liu, W.R., Caudill, S.B., 2002. Tax-and-spend, spend-and-tax, or fiscal synchronization: new evidence for ten countries. Applied Economics 34, 1553–1561

Clark, T., Dilnot, A., 2002. Long-term trends in British taxation and spending. London: Institute of Fiscal Studies, Briefing Note 25

Dickey, D. A., Fuller, W.A., 1979. Distribution of the estimators for autoregressive time series with a unit root. Journal of the American Statistical Association 74, 427–431

Dolado J.J., Lütkepohl, H., 1996. Making Wald tests work for cointegrated VAR systems. Econometric Reviews 154, 369–386

Elliot, G., Rothenberg, T.J., Stock, J.H.,. 1996. Efficient tests for an autoregressive unit root. Econometrica 4, 813–836

Emmersen, C., Frayne, C., Love, S., 2003. A survey of public spending in the UK. London: Institute for Fiscal Studies, Briefing Note Number 43.

Engle, R.F., Granger, C.W.J., 1987. Cointegration and error correction: representation, estimation and testing. Econometrica 55, 251–76.

Friedman, M., 1978. The limitations of tax limitation. Policy Review, Summer, 7–14.

Friedman M., 2003. What every American wants. The Wall Street Journal. Available at: http://www.opinionjournal.com/editorial/feature.html?id=110002933.

Glynn, J., Perera., 2007. Unit root tests and structural breaks: a survey with applications. Journal of Quantitative Methods for Economics and Business Administration 3, 63–79.

Gregory, A. W., Hansen, B.E., 1996a. Residual based tests for cointegration with regime shifts. Journal of Econometrics 70, 251–76

Gregory, A. W., Hansen, B.E.,1996b. Test for cointegration in models with regime and trend shifts. Oxford Bulletin of Economics and Statistics 58, 555–560

Hasan, M., Lincoln, I., 1997. Tax then spend or spend then tax? Experience in the UK, 1961-1993. Applied Economics Letters 4, 237–239

Johansen, S., 1988. Statistical analysis of cointegrating vectors. Journal of Economic Dynamics and Control 12, 231–254

Johansen, S., 1991. Estimation and hypothesis testing of cointegrating vectors in Gaussian vector autoregressive models. Econometrica 59, 1551–1580

Johansen, S., 1995. Likelihood-based inference in cointegrated vector autoregressive models (Oxford: Oxford University Press).

Johansen S., Mosconi, R., Nielsen, B., 2000. Cointegration analysis in the presence of structural breaks in the deterministic trend. Econometrics Journal 3, 216–249

Joulfaian, D., Mookerjee, R., 1991. Dynamics of government revenues and expenditures in industrial economies. Applied Economics 23, 1839–1844

Koren, S., Stiassny, A., 1998. Tax and spend, or spend and tax? An international study. Journal of Policy Modeling 20, 127–149

Lanne, M., Lütkepohl, H., Saikkonen. P., 2002. Comparison of unit root tests for time series with level shifts. Journal of Time Series Analysis 23, 667–685

Lee, J., Strazicich, M.C., 2003. Minimum Lagrange Multiplier unit root test with structural breaks. Review of Economics and Statistics 85, 1082–1089

Lee, J., Strazicich, M.C., 2004. Minimum LM unit root test with one structural break. Appalachian State University Working Paper No. 04-17, Department of Economics. Available at: http://econ.appstate.edu/RePEc/pdf/wp0417.pdf

Lütkepohl, H., 1991. Introduction to multiple time series analysis (Berlin: Springer-Verlag).

Lütkepohl, H., Saikkonen, P., Trenkler, C., 2003. Comparison of tests for the cointegrating rank of a VAR process with a structural shift. Journal of Econometrics 112, 201–229

Meltzer, A.H., Richard, S.F., 1981. A rational theory of the size of government. Journal of Political Economy 89, 914–927

Mitchell, B.R., 1988. British historical statistics (Cambridge: Cambridge University Press).

Musgrave, R., 1966. Principles of budget determination, in H.A. Cameron and W. Henderson, eds., Public Finance: Selected Readings (New York: Random House), 15–27

Nunes, L.C., Newbold P., Kaun, C-M.,1997. Testing for unit roots with breaks: evidence on the great crash and the unit root hypothesis reconsidered. Oxford Bulletin of Economics and Statistics 59, 435–448

O'Donoghue, J., Goulding, L., Allen, G., 2004. Consumer price inflation since 1750. Economic Trends 604, 38–45

Owoye, O., 1995. The causal relationship between taxes and expenditures in the G7 countries: cointegration and error correction models. Applied Economics Letters 2, 19–22

Payne, J.E., 2003. A survey of the international empirical evidence on the tax-spend debate. Public Finance Review 31, 302–324

Peacock, A.T., Wiseman, J., 1961. The growth of public expenditure in the United Kingdom (Princeton: Princeton University Press).

Peacock, A.T., Wiseman, J., 1979. Approaches to the analysis of government expenditure growth. Public Finance Quarterly 7, 3–23

Perron, P., 1989. The great crash, the oil price shock and the unit root hypothesis. Econometrica 57, 1361–1401

Perron, P., 2006. Dealing with structural breaks, in T. C. Mills and K. Patterson, eds., Handbook of Econometrics, Vol 1 (London: Palgrave Macmillan), 278–352

Phillips, P. C. B., Perron, P., 1988. Testing for a unit root in time series regression. Biometrika 75, 335–346

Quintos, C.E., 1995. Sustainability of the deficit process with structural shifts. Journal of Business and Economic Statistics 13, 409–417

Ram, R., 1988. A multicountry perspective on causality between government revenue and government expenditure. Public Finance 43, 261–270

Said, S.E., Dickey, D.A., 1984. Testing for unit roots in autoregressive-moving average models of unknown order. Biometrika 71, 2599–2607

Saikkonen, P., Lütkepohl, H., 2000a. Testing for the cointegrating rank of a VAR process with an intercept. Econometric Theory 16, 373–406

Saikkonen, P., Lütkepohl, H., 2000b. Testing for the cointegrating rank of a VAR process with structural shifts. Journal of Business and Economic Statistics 18, 451–464

Saikkonen, P., Lütkepohl, H., 2000c. Trend adjustment prior to testing for the cointegrating rank of a vector autoregressive process. Journal of Time Series Analysis 21, 435–456

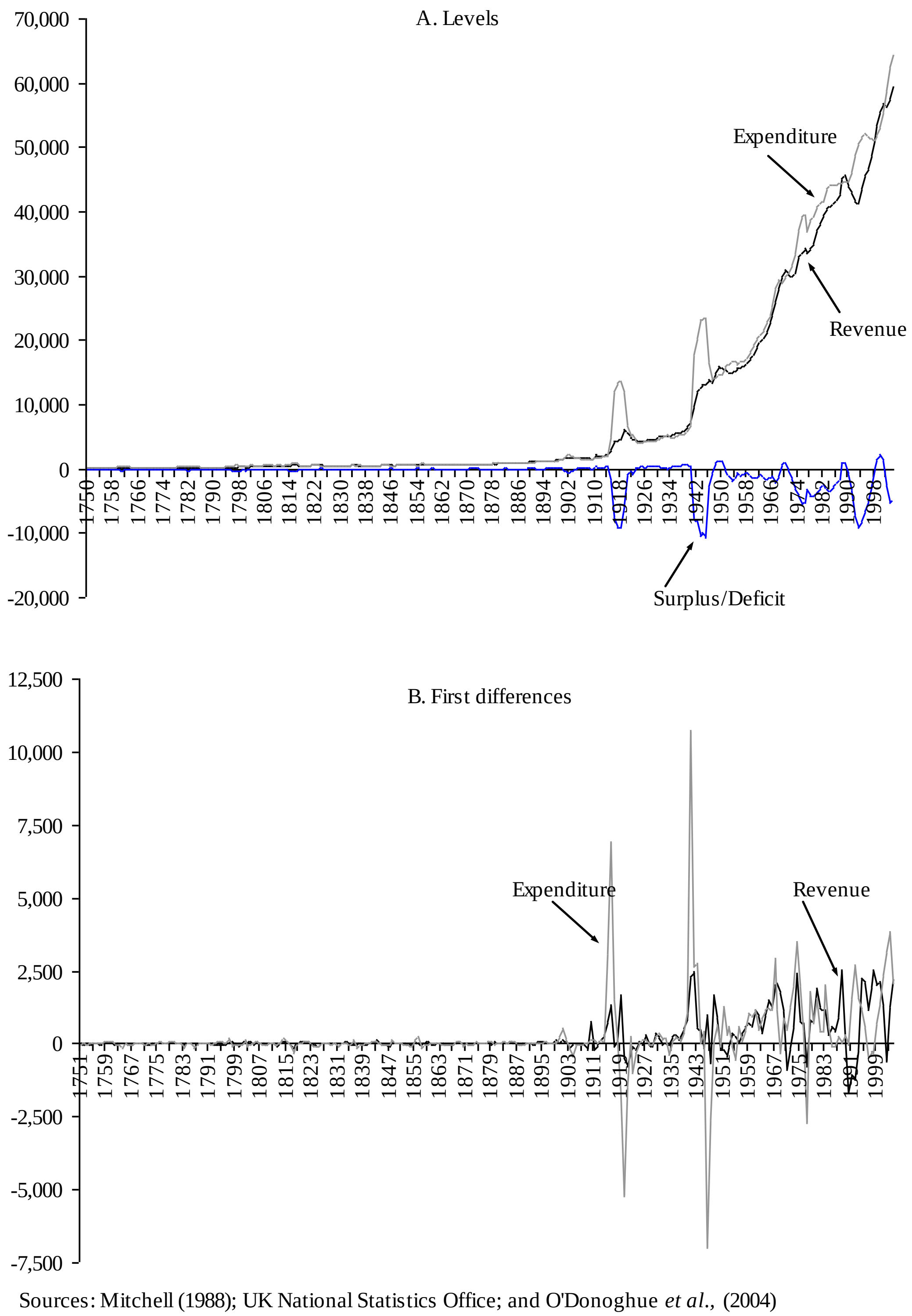
Saikkonen, P., Lütkepohl., 2001. Testing for unit roots in time series with level shifts. Allgemeines Statistisches Archiv 85,1–25

Saikkonen, P., Lütkepohl., 2002. Testing for a unit root in a time series with level shift at unknown time. Econometric Theory 18, 313–348

Wildavsky, A., 1988. The new politics of the budgetary process (Glenview, IL: Scott, Foresman).

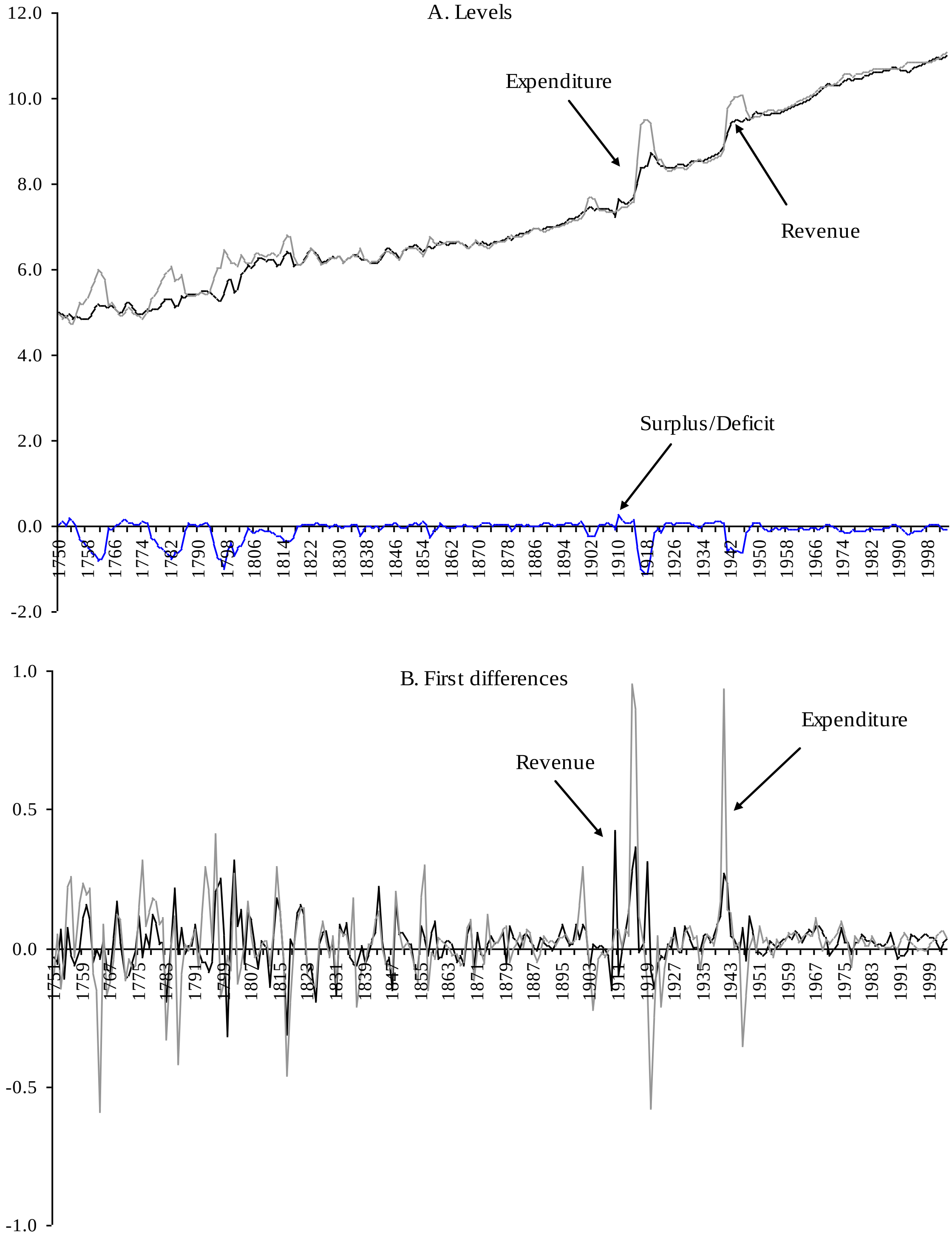
Zivot, E., Andrews, D.W.K., 1992. Further evidence of the great crash, the oil price shock and the unit root hypothesis. Journal of Business and Economics Statistics 10, 251–70

Figure 1. Government Revenue and Expenditure in 1974 Prices, 1750–2004
(millions of pounds)



Sources: Mitchell (1988); UK National Statistics Office; and O'Donoghue *et al.*, (2004)

Figure 2. Government Revenue and Expenditure in 1974 Prices, 1750–2004
(millions of pounds, logarithmic scale)



Sources: Mitchell (1988); UK National Statistics Office; and O'Donoghue et al., (2004)

Table 1. Univariate unit root tests for series in levels, 1750–2004

	Revenue		Expenditure	
	Level	Change	Level	Change
(i) No breaks (lags in parenthesis)				
Dickey and Fuller (1979), ADF test	3.56 (10)	-2.28 (8)	1.97 (5)	-6.22 (5)*
Elliott *et al* . (1996), DF-GLS test	0.42 (9)	-1.92 (8)+	0.26 (11)	-3.23 (10)*
Phillips and Perron (1988) test				
Z(rho)	4.26 (4)	-96.05 (4)*	2.40 (4)	-138.76 (4)*
Z(t)	3.62 (4)	-7.51 (4)*	1.04 (4)	-9.69 (4)*
(ii) One break (lags in parenthesis, breakpoint below)				
Zivot and Andrews (1992) with a break in:				
Intercept	0.15 (2)	-9.86 (0)*	-1.49 (3)	-9.18 (3)*
	1961	1938	1963	1939
Trend	-3.62 (2)	-9.94 (0)*	-5.19 (3)*	-9.07 (3)*
	1933	1907	1933	1923
Intercept and trend	-3.48 (2)	-10.00 (0)*	-5.10 (3)*	-9.18 (3)*
	1920	1937	1932	1938
Lee and Strazicich (2004), with a change in:				
Intercept	-0.79 (7)	-3.63 (6)*	-1.37 (6)	-7.09 (5)*
	1946	1977	1941	1971
Intercept and trend	-4.32 (7)+	-8.49 (7)*	-4.87 (6)*	-10.45 (4)*
	1926	1918	1920	1900
Saikkonen and Lütkepohl (2001,2002) and Lanne, Lütkepohl, and Saikkonen (2002), with a break given by:				
Shift dummy	-0.99 (10)	-6.04 (7)*	-0.80 (5)	-9.15 (4)*
	1939	1939	1939	1939
Rational shift dummy	-0.66 (11)	-6.07 (7)*	0.24 (5)	-4.77 (4)*
	1939	1939	1939	1939
(iii) Two breaks (lags in parenthesis, breakpoints below)				
Lee and Strazicich (2003), with changes in:				
Intercept	-0.92 (9)	-4.11 (6)*	-1.67 (6)	-7.20 (5)*
	1942; 1946	1968; 1973	1916; 1941	1856; 1958
Intercept and trend	-7.29 (6)*	-9.09 (7)*	-7.06 (6)*	-12.96 (4)*
	1901; 1947	1918; 1974	1930; 1945	1939; 1944

* and + indicate statistical significance at the 5 (at least) and 10 percent levels.

Table 2. Univariate unit root tests for natural logarithms of constant price series, 1750–2004

	Revenue		Expenditure	
	Level	Change	Level	Change
(i) No breaks (lags in parenthesis)				
Dickey and Fuller (1979), ADF test	-1.68 (2)	-12.08 (1)*	-3.15 (1)$^{+}$	-7.44 (10)*
Elliott *et al*. (1996), DF-GLS test	-0.82 (10)	-1.63 (15)$^{+}$	-1.17 (11)	-1.80 (13)$^{+}$
Phillips and Perron (1988) test				
Z(rho)	-5.40 (4)	-189.43 (4)*	-14.91 (4)	-159.46 (4)*
Z(t)	-1.78 (4)	-13.66 (4)*	-2.74 (4)	-11.03 (4)*
(ii) One break (lags in parenthesis, breakpoint below)				
Zivot and Andrews (1992) with a break in:				
Intercept	-3.63 (3)	-10.89 (2)*	-6.38 (1)*	-8.81 (3)*
	1910	1909	1915	1909
Trend	-3.29 (3)	-10.71 (2)*	-5.78 (1)*	-8.74 (3)*
	1877	1942	1881	1817
Intercept and trend	-3.88 (3)	-11.11 (2)*	-6.73 (1)*	-8.84 (3)*
	1910	1909	1915	1912
Lee and Strazicich (2004), with a change in:				
Intercept	-1.76 (6)	-12.10 (1)*	-3.30 (4)$^{+}$	-11.10 (0)*
	1799	1821	1940	1917
Intercept and trend	-4.11 (1)	-13.44 (0)*	-5.76 (1)*	-11.14 (0)*
	1903	1815	1893	1918
Saikkonen and Lütkepohl (2001,2002) and Lanne, Lütkepohl, and Saikkonen (2002), with a break given by:				
Shift dummy	-1.44 (2)	-11.34 (1)*	-3.17 (1)$^{+}$	-11.04 (0)*
	1939	1939	1939	1939
Rational shift dummy	-1.95 (2)	-7.77 (1)*	-4.49 (1)*	-4.04 (4)*
	1939	1939	1939	1939
(iii) Two breaks (lags in parenthesis, breakpoints below)				
Lee and Strazicich (2003), with changes in:				
Intercept	-2.18 (1)	-13.41 (0)*	-4.25 (1)*	-11.95 (0)*
	1799; 1909	1797; 1917	1785; 1940	1915; 1923
Intercept and trend	-5.06 (1)	-13.47 (0)*	-6.11 (1)*	-12.12 (0)*
	1802; 1909	1815; 1950	1822; 1903	1914; 1921

* and $^{+}$ indicate statistical significance at the 5 (at least) and 10 percent levels.

Table 3. Cointegration test results for constant price series, 1750–2004

Cointegration test:				Cointegration ?
I. Residual-based test, with one break (lags in parenthesis, breakpoint below)				
Gregory and Hansen (1996a,b), with a break in:	ADF*	Z(t)	Z(a)	
Level shift	-6.38 (6)*	-6.06*	-69.74*	Yes
	1908	1945	1945	
Level shift with trend	-6.38 (6)*	-6.11*	-70.61*	Yes
	1908	1945	1945	
Regime shift	-9.88 (2)*	-7.19*	-92.83*	Yes
	1950	1945	1945	
Regime and trend shifts	-7.84 (2)*	-6.11*	-71.20*	Yes
	1961	1947	1947	
II. VAR models with outliers/level shifts				
Saikkonen and Lütkepohl (2000a,b,c) (without dummies):	H_0: r=0	H_0: r≤1	Lag length[1]	
Trend in data and cointegrating relation	78.53*	0.17	3	Yes
Trend in data; no trend in cointegrating relation	95.83*	-	3	Yes [2]
III. VAR models with one break				
Johansen et al. (2000) with given break in:	H_0: r=0	H_0: r≤1	Lag length[1]	
Levels (breakpoints below)				
1939	139.44*	40.43*	3	No
1914	132.95*	34.14*	3	No
1951	60.87*	29.49*	8	No
Levels and trends (breakpoints below)				
1939	108.64*	41.14*	7	No
1914	133.97*	35.35*	3	No
1951	44.77*	5.46	8	Yes
IV. VAR models with two breaks				
Johansen *et al*. (2000) with given break in:	H_0: r=0	H_0: r≤1	Lag length[1]	
Levels (breakpoints below)				
1914; 1939	127.02*	22.04*	9	No
1917; 1939	85.33*	9.37	9	Yes
1818; 1951	64.36*	30.63*	8	No
Levels and trends (breakpoints below)				
1914; 1939	131.77*	60.19*	10	No
1917; 1940	93.87*	6.58	7	Yes
1818; 1951	47.65*	8.34	8	Yes

* and + indicate statistical significance at the 5 (at least) and 10 percent levels.

1. As common practice, the lag-length is determined by the Hannan-Quinn information criterion (Johansen et al., 2000).

2. Yes, if the stationary altenative is excluded.

Table 4. Cointegration test results for series in natural logarithms of levels, 1750–2004

Cointegration test:				Cointegration ?
I. Residual-based test, with one break (lags in parenthesis, breakpoint below)				
Gregory and Hansen (1996a,b), with a break in:	ADF*	Z(t)	Z(a)	
Level shift	-7.67 (6)*	-6.03 (6)*	-67.93 (6)*	Yes
	1811	1814	1814	
Level shift with trend	-5.26 (6)*	-5.40 (6)*	-51.16 (6)*	Yes
	1811	1800	1800	
Regime shift	-7.29 (2)*	-6.38 (2)*	-74.16 (2)*	Yes
	1802	1800	1800	
Regime and trend shifts	-6.96 (1)*	-6.36 (1)*	-72.85 (1)*	Yes
	1801	1800	1800	
II. VAR models with outliers/level shifts				
Saikkonen and Lütkepohl (2000a,b,c) (without dummies):				
Trend in data and cointegrating relation	39.27*	1.25	3	Yes
Trend in data; no trend in cointegrating relation	37.27*	-	3	Yes [2]
III. VAR models with one break				
Johansen et al. (2000) with given break in:	H_0: r=0	H_0: r≤1	Lag length[1]	
Levels (breakpoints below)				
1939	64.17*	7.30	3	Yes
1914	65.83*	5.27	3	Yes
1951	60.06*	4.68	3	Yes
Levels and trends (breakpoints below)				
1939	64.07*	7.35	3	Yes
1914	72.78*	12.08	3	Yes
1951	60.51*	4.80	3	Yes
IV. VAR models with two breaks				
Johansen *et al* . (2000) with given break in:	H_0: r=0	H_0: r≤1	Lag length[1]	
Levels (breakpoints below)				
1914; 1939	75.71*	13.38	3	Yes
1917; 1939	71.66*	14.98	2	Yes
1818; 1951	47.88*	9.10	2	Yes
Levels and trends (breakpoints below)				
1914; 1939	77.38*	16.07	3	Yes
1917; 1940	61.41*	5.57	4	Yes
1818; 1951	81.10*	10.07	7	Yes

* and $^{+}$ indicate statistical significance at the 5 (at least) and 10 percent levels.
1. As common practice, the lag-length is determined by the Hannan-Quinn information criterion (Johansen et al., 2000).
2. Yes, if the stationary altenative is excluded.

Table 5. Granger-causality tests for revenue and expenditure

	H_0: Revenue does not Granger-cause expenditure	H_0: Expenditure does not Granger-cause revenue
I. Series in levels		
(i) Pre-World War I: 1750–1913		
Test statistic	3.66	6.99
p -value	0.00	0.00
(ii) Post-World War II: 1947–2004		
Test statistic	0.04	2.48
p -value	0.96	0.09
(iii) Post-World War II: 1951–2004		
Test statistic	0.07	2.06
p -value	0.93	0.13
II. Series in natural logarithm		
(i) Pre-World War I: 1750–1913		
Test statistic	0.01	2.16
p -value	0.93	0.14
(ii) Pre-World War I: 1820–1913		
Test statistic	2.15	2.78
p -value	0.12	0.06
(iii) Post-World War II: 1947–2004		
Test statistic	0.13	3.64
p -value	0.88	0.03
(iv) Full sample: 1750–2004		
Test statistic	0.85	10.07
p -value	0.47	0.00