Conflict Management and Peace Science, 24:55–64, 2007

ISSN: 0738-8942 print / 1549-9219 online
DOI: 10.1080/07388940601102845

Data, Models, Coefficients: The Case of United States Military Expenditure

JURGEN BRAUER

James M. Hull College of Business
Augusta State University
Augusta, Georgia, USA

This article is an exercise in economic methodology. It replicates two published models of the effect of military expenditure on the United States economy but, in order to study variations in the relevant estimated parameters, applies two different military expenditure data sets to the models (budget vs. *National Income and Product Accounts [NIPA] data). In an extension, the article examines coefficient stability when the economically preferred NIPA data are applied across varying time-periods. Two major findings are that economic models should avoid the use of budget data and that even when the preferred NIPA data are used, estimated parameters are highly unstable across time.*

Keywords defense economics, military expenditure, United States, data sources, methodology, replication.

The purpose of this article is to demonstrate the obvious: that use of the "wrong" data leads one to obtain improper results. Even if one employs the "right" (or at least "better") data, one can obtain results that are inconsistent when the sample is varied. To demonstrate the obvious is necessary because of the continuing practice in work submitted for publication and in published work to neglect investigating the sensitivity of empirical results to changes in the underlying data and to variation in sample size, composition, or time period. This article may thus be viewed as a case study complementing Smith (1998), who "argues that quantitative peace research could be improved if authors put more emphasis on the substantive issues and less on the mechanical application of rule-based statistical techniques." It is part of a tradition of a small body of methodological work investigating the (im)proper use of models and data in defense and peace economics.[1]

The article replicates two published models of the effect of military expenditure on the economy of the United States. To study variations in the relevant estimated parameters, it applies two different military expenditure data sets to the models (budget-based vs. National Income and Product Accounts [NIPA]-based data). In an extension, the article examines coefficient stability when the economically preferred NIPA-based data are applied across varying time-periods. Two findings are (a) that economic models should avoid the use of

Address correspondence to Jurgen Brauer, Professor of Economics, James M. Hull College of Business, Augusta State University, Augusta, GA 30904, USA. E-mail: jbrauer@aug.edu

[1]In addition to Smith (1998), see for example, Brzoska (1982), Scheetz (1985), Brzoska (1995), Brauer (2002), and Dunne, Smith, Willenbockel (2005). The classic reference for replication in economics is Dewald, Thursby, and Anderson (1986). Coefficient instability is amply demonstrated in Stock and Watson's (1996) experimental piece. An application to the estimation of production functions, just to choose one non-defense example, can be found, e.g., in Felipe and Adams (2005, esp. p. 432). Data issues are discussed, e.g., in Kennedy (2003) and in a number of symposium papers in the *Journal of Economic Perspectives* (Vol. 15, No. 4, Fall 2001).

budget-based military expenditure data and (b) that even when the preferred NIPA data are used, estimated parameters are highly unstable across sample time periods.

United States Military Expenditure: The Data

In the United States, there is one ultimate source for military expenditure data. That source is the United States Budget or, more precisely, the *Historical Tables*, a document supplementary to the fiscal year budget request by the administration to Congress and issued annually to Congress by the Office of Management and Budget (OMB) from within the president's office. The *Historical Tables* capture outlays for prior fiscal years, not presidential budget requests, nor Congressional appropriations. Outlays data are revised when supplementary budgets are passed, funds appropriated, and monies expended.

The outlays data are *compiled* in two different ways. One compilation, the one that NATO, SIPRI, the Bureau of Verification and Compliance (BVC),[2] and therefore most researchers looking for cross-national data rely upon, picks certain budget line items that fit NATO's military expenditure definition.[3] Thus, the three major international sources for cross-country military expenditure data all report the same figures for United States military expenditure.[4]

OMB's *Historical Tables* distinguish between line item 051, which includes all Department of Defense (DoD) outlays, and line item 050, which includes all national defense outlays.[5] Since the latter include military expenditure in agencies other than the DoD—for instance, military-nuclear activities budgeted within the Department of Energy—the numbers for line item 050 are larger than the DoD numbers in line item 051. Except for some minor adjustments, this more comprehensive budget line item is essentially the number that NATO uses for U.S. military expenditure, so that the differences between line item 050 and NATO's numbers are insubstantial (see Figure 1). For 1996, for instance, NATO reports

[2]Bureau of Verification and Compliance, see BVC (2002). This is the former U.S. Arms Control and Disarmament Agency, ACDA.

[3]NATO's definition is *not* available on its web site [accessed 4 April 2006] but is cited in BVC (2002, p. 194) as follows: "In this definition, (a) civilian-type expenditures of the defense ministry are excluded and military-type expenditures of other ministries are included; (b) grant military assistance is included in the expenditures of the donor country; and (c) purchases of military equipment for credit are included at the time the debt is incurred, not at the time of payment." NATO revised its definition in 2004, to exclude a category called "Other Forces," but reports that not all members as yet comply with the new definition (see http://www.nato.int/docu/pr/2005/p050609e.htm [accessed 4 April 2006].

SIPRI takes NATO figures for NATO states. SIPRI's definition is: "Where possible, SIPRI military expenditure data include all current and capital expenditure on: (a) the armed forces, including peacekeeping forces; (b) defence ministries and other government agencies engaged in defence projects; (c) paramilitary forces, when judged to be trained and equipped for military operations; and (d) military space activities. Such expenditures should include: (a) military and civil personnel, including retirement pensions of military personnel and social services for personnel; (b) operations and maintenance; (c) procurement; (d) military research and development; and (e) military aid (in the military expenditure of the donor country). Civil defence and current expenditures on previous military activities, such as veterans' benefits, demobilization, conversion and weapon destruction are excluded." This is taken from http://www.sipri.org/contents/milap/milex/mex_sources.html [accessed 4 April 2006].

[4]Any data differences are due solely to reporting lags but, eventually, all three sources report the same military expenditure number for the United States. The reporting sequence is *Historical Tables*, then NATO, then SIPRI, then BVC. (Actually, BVC has not produced a report since 2002, the last data point being for 1999.)

[5]The easiest way to access time series for these numbers is via the statistical appendix to the annually issued *Economic Report of the President*, available online at http://www.gpoaccess.gov/eop/index.html [accessed 5 April 2006].

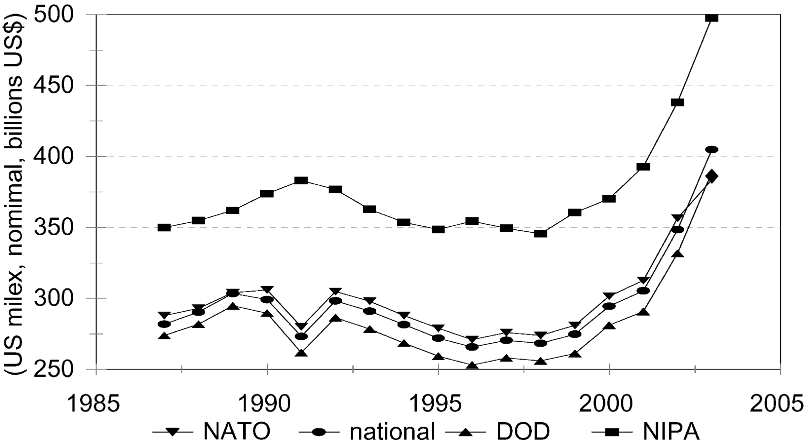

FIGURE 1 U.S. military expenditure, 1987–2003, according to NATO/SIPRI/BVC, the *Historical Tables* (total national defense outlays and DoD; budget line items 050 and 051, respectively), and NIPA measures; nominal billions of US dollars. NATO 2003 is an estimate.

U.S. military expenditure of $271 billion (nominal), whereas the *Historical Tables* report $266 billion for national defense outlays (line item 050).

The other compilation of U.S. budget data is contained in the National Income and Product Accounts (NIPA), produced by the Bureau of Economic Analysis (BEA) in the U.S. Department of Commerce. Here, a truly astonishing difference arises (Figure 1). For 1996, NIPA counts U.S. military expenditure at $355 billion (nominal), a difference of $84 billion, or 31%, over the NATO numbers. For a more recent year, fiscal year (or FY) 2005, NIPA's national defense outlays are $581.9 billion, which breaks down into $511.4 billion for consumption expenditures and $70.4 billion for gross investment in structures and equipment. This contrasts to $495.3 billion for budget line item 050 national defense outlays, an $86.6 billion difference (Mandel & Roy, 2006, p. 17, Table 6). Apart from some minor items, the main difference arises primarily from NIPA adjustments concerning contributions to retirement funds and from calculations of consumption of general government fixed capital related to the defense function and national defense gross investment.[6]

Budgets are constructed on a cash basis for administrative purposes and as such have their uses. But budgets do not measure the economic resources devoted to a state's defense function. The NIPA exercise is a step in that direction. All federal receipts and expenditures are translated into NIPA categories and then adjusted to fit the national income accounting framework which, in turn, is used for macroeconomic analysis of proposed changes in federal government activity. As such, the longstanding NIPA effort appears not so different from the relatively new "Resource Accounting and Budgeting" effort in the United Kingdom, which now also includes depreciation and capital costs of the defense sector and has led to vast "increases" in U.K. defense spending, increases that always existed but were hidden by the way the accounting was done.[7]

[6]In addition to the Mandel and Roy (2006) piece and, in exhaustive detail, BEA (2005), a short explanation of differences between budget and NIPA data may be found in *Budget of the United States Government, Analytical Perspectives, Fiscal Year 2004* (Washington, DC: Government Printing Office, February 2003, pp. 383–387), see http://www.gpoaccess.gov/usbudget/fy04/pdf/spec.pdf [accessed 4 April 2006].

[7]I am thankful to an anonymous reviewer to point me to this U.K. defense statistics web site: http://www.dasa.mod.uk/natstats/ukds/2005/c1/table11.html [accessed 5 April 2006].

For the United States, the NIPA numbers are the most comprehensive and conceptually complete national defense outlays data we have, since they are expressly based on economists' national income accounting framework rather than on politicians' need to review and pass budget requests. NIPA data are therefore the preferred data to use in economic analysis.[8] Since NATO's (and SIPRI's) numbers are essentially equivalent to budget line item 050, I refer to these as the "budget" or *Historical Tables* data. The NIPA numbers are referred to as "NIPA" data.

Analytic Considerations: The Data

This section presents two examples of how the use of different data sets affects model estimates. The first follows a model published in Ramanathan (2002, pp. 464–471), the second a model published by Atesoglu (2002). It is important to appreciate that the models are not criticized. Instead, the exercise investigates what happens to the parameter estimates in these models when two different data sets—both purporting to measure United States military expenditure—are used.

Ramanathan and Blackburn's Model

Ramanathan and Blackburn argue that even though total government and military expenditure diverge in the short run, a stable relation exists between the two in the long run. This can be derived from the following model. Suppose government maximizes a Cobb-Douglas utility function $U(M,N) = M^{\alpha}N^{1-\alpha}$, where M and N are military and nonmilitary (civilian) government expenditure and α is a fixed parameter. Subject to a conventional budget constraint, $p_M M + p_N N = G$, it can be shown that utility maximizing military expenditure is a function of G, namely $M = \lambda\ G$, where λ reflects the long-run relation between M and G. Short-run deviations are possible, and the return path to long-run equilibrium can be captured with an error-correction model of the form

$$\Delta m_t = \alpha + \beta \Delta g_t + \gamma(g_{t-1} - m_{t-1}) + u_t \tag{1}$$

where lower-case letters indicate the log-form of M and G, respectively. Thus, Δm_t is the percentage change in military expenditure from period $t-1$ to period t. β captures the contemporaneous effect of Δg_t on Δm_t, while γ reflects the long-run adjustment of the prior-period deviation of military from total government expenditure on Δm_t.

Using Ramanathan and Blackburn's own budget-based DoD data (line item 051, for 1940 to 1987), a replication using E-Views (version 4) arrives at broadly consistent results.[9] Since NIPA data are not available prior to 1962, the model was re-run with data for 1962 to 2002. Running equation (1) with the budget data (line item 050) yields $\beta = 2.205$ ($p = 0.0000$) and $\gamma = 0.045$ ($p = 0.1806$), whereas with the NIPA data, $\beta = 2.122$

[8]Even so, the NIPA defense function numbers do not include the budget of the Department of Veterans Affairs. This includes health, burial, survivor, and other benefits but, with very minor exceptions, not military retirement benefits (contributions to which are part of the DoD budget). For a brochure explaining Veterans Affairs benefits, see http://www1.va.gov/opa/vadocs/fedben_pt1.pdf [accessed 4 April 2006]. Nor do the NIPA data include an allocation of federal government interest payments on the federal debt to the defense function (BEA, 1988, p. 4; BEA, 2005, p. IV-4). Since both categories reflect legacy costs (cost of past defense-readiness coming due in the current fiscal or calendar year), both should be included. However, in this article the empirical analysis is strictly confined to the budget and NIPA data as officially reported.

[9]See the table below (R/B results from Ramanathan, 2002, p. 466). Note that Ramanathan and Blackburn use Department of Defense outlays (line item 051 data)

($p = 0.0000$) and $\gamma = 0.052$ ($p = 0.0329$).[10] For the run with budget data, the adjustment parameter, γ, is not statistically significant, but for the run with NIPA data it is, as indeed it should if the long-run adjustment the theory predicts exists. The sole data difference lies in the measurement of military expenditure (budget vs. NIPA). All other data are identical.

Following a suggestion by Ramanathan and Blackburn, amend equation (1) to include dummy variables for the height of the Vietnam war years (1965–1969 = 1), the Carter/Reagan military build-up (1976–1987 = 1), and for the one post-9/11 year in the data set (2002 = 1). Allowing for interaction effects of the long-run adjustment parameter, γ, with short-run special-tension circumstances, one would expect the relevant coefficients to be positive. In both runs, the relevant coefficients are positive, as expected, and are even of similar size (see γ_B and γ_N in Table 1). In the NIPA formulation, all coefficients, including those for the interaction terms, are statistically significant (at 4% or better). But in the budget data formulation, γ_B and the interaction effects on Vietnam and post-9/11 are not (at 11% and worse).[11] Evidently, using budget as opposed to NIPA data results in substantively different conclusions.

Encouragingly, the results from the NIPA run make economic and political sense. Using the budget data, the results suggest that the long-run adjustment coefficient is, statistically speaking, zero and that neither the Vietnam war nor 9/11 had a statistically significant effect on military expenditure. In contrast, using the NIPA data, all variables are of the proper sign and statistically significant, just as one would expect.

Atesoglu's Model

Rather than employing a budget-constraint model, Atesoglu (2002) recently published an interesting Keynesian macroeconomics model, based on Romer (2000) and Taylor (2000, 2001). Let

$$QL_t = CL_t + IL_t + XL_t + ML_t + GL_t \tag{2}$$

$$CL_t = a + b(QL_t - TL_t) \tag{3}$$

$$TL_t = n + gQL_t \tag{4}$$

as reported in the *Historical Tables* (Ramanathan, 2002, p. 669), and so does the replication.

	α	Δg_t	$(g_{t-1} - m_{t-1})$
R/B			
Coefficients	−0.172	+1.579	+0.169
t-stats	−2.8	12.9	2.8
Adj R^2	0.792		
SE regression	0.1741		
Replication			
Coefficients	−0.167	+1.597	+0.129
t-stats	−2.80	13.82	2.17
Adj R^2	0.821		
SE regression	0.167		

[10] To remain faithful to R/B's original intent, equation (1) was also run with budget line item 051 data. For 1962 to 2002, there are no essential differences in the estimation results between the use of 051 and 050 data. The differences reported in the text come about only as compared to using NIPA data.

[11] Both runs contain some residual serial autocorrelation that would ordinarily require statistical "cleaning," but the point of the exercise here regards only the differences obtained in running the same model with different data.

TABLE 1 Regressions with budget vs. NIPA military expenditure

Dependent variable: Δm_t (budget line item 050 data)
Sample (adjusted): 1963 2002
Included observations: 40 after adjusting endpoints

Variable	Coefficient	SE	t-Statistic	Probability
Constant	−0.119516	0.052288	−2.285703	0.0286
Δg_t	1.666627	0.395125	4.217974	0.0002
γ_B	0.047420	0.037284	1.271885	0.2120
γ_B*VWAR	0.040296	0.029359	1.372515	0.1789
γ_B*CARTREAG	0.036376	0.012699	2.864510	0.0071
γ_B*T911	0.045199	0.027704	1.631496	0.1120
R^2	0.635367	Mean dependent var		0.001902
Adjusted R^2	0.581744	S.D. dependent var		0.070814
S.E. of regression	0.045797	Akaike info criterion		−3.191703
Sum squared resid	0.071311	Schwarz criterion		−2.938371
Log likelihood	69.83405	F-statistic		11.84887
Durbin-Watson stat	1.790679	Prob(F-statistic)		0.000001

Dependent variable: Δm_t (NIPA data)
Sample (adjusted): 1963 2002
Included observations: 40 after adjusting endpoints

Variable	Coefficient	SE	t-Statistic	Probability
C	−0.111170	0.026397	−4.211524	0.0002
Δg_t	1.708453	0.180517	9.464209	0.0000
γ_N	0.052007	0.021703	2.396286	0.0222
γ_N*VWAR	0.034140	0.015853	2.153518	0.0385
γ_N*CARTREAG	0.038383	0.007002	5.481919	0.0000
γ_N*T911	0.034287	0.015359	2.232395	0.0323
R^2	0.871536	Mean dependent var		0.003734
Adjusted R^2	0.852644	S.D. dependent var		0.056329
S.E. of regression	0.021623	Akaike info criterion		−4.692643
Sum squared resid	0.015897	Schwarz criterion		−4.439311
Log likelihood	99.85285	F-statistic		46.13298
Durbin-Watson stat	1.487317	Prob(F-statistic)		0.000000

Notes: Δm_t: differenced log of real budget/NIPA military expenditure; Δg_t: differenced log of real budget/NIPA total government expenditure; γ_B and γ_N: the gamma function in model (4)—$\gamma\ (g_{t-1} - m_{t-1})$—where g is total government and m is the government's military expenditure; the dummy variables are described in the text. SE: standard error.

$$IL_t = e - fR_t \tag{5}$$

$$XL_t = z - mQL_t - nR_t \tag{6}$$

where equation (2) is the national income and expenditure identity with output equal to consumption, investment, net exports, and military and nonmilitary government expenditure (all in natural log terms, L). Consumption, in equation (3), is a function of output minus taxes. The tax take, equation (4), consists of lump-sum and income taxes. Investment, equation (5), is a function of the real interest rate, R (non-log), and net exports, in equation

(6), are a negative function of national output and also a negative function of the real interest rate (i.e., higher interest rates are modeled to reduce exports and attract imports via a more highly valued currency).

By substitution, equation (7) can be derived

$$QL_t = \alpha + \beta ML_t + \delta GL_t - \lambda R_t + u_t \tag{7}$$

where QL is aggregate real output, ML is real military expenditure, GL is real nonmilitary expenditure (all in logs), R is the real interest rate, and u is an error term (and $\beta = \delta$).

Atesoglu (2002) correctly employs quarterly United States military expenditure data as reported by NIPA, 1947:Q2 to 2000:Q2.[12] Replicating Atesoglu's estimation, one arrives at similar baseline results and diagnostics.[13] Since Atesoglu finds his variables to be I(1), the OLS results could be spurious,[14] and he thus employs the Johansen cointegration routine to generate estimates for the parameters of the long-run cointegration relation between QL and ML, GL, and R. The results were replicated, again with close agreement of coefficients and diagnostics.[15] The vector error correction (VEC) model also resulted in broadly consistent short-run adjustment coefficients and diagnostics between Atesoglu and the replication.

[12]In a model derived from national income accounting, obviously national income accounts data should be used.

[13]For raw data, Atesoglu used the FRED data base at the St. Louis Federal Reserve. By the time of the replication, this had become FRED II, with corrected and updated data and a changed base year. Estimating equation (7) with OLS using E-Views (version 2), Atesoglu obtained the following results:

Atesoglu	α	βML	δ GL	λR
Coefficients	+3.348	+0.233	+0.813	−0.011
SE-coeffs	0.164	0.035	0.024	0.005

The replication with E-Views (version 4), FRED II data, and the same time frame, yields almost identical results:

Replication	α	βML	δ GL	λR
Coefficients	+3.167	+0.265	+0.815	−0.013
SE-coeffs	0.167	0.035	0.025	0.005

[14]Running the augmented Dickey-Fuller (ADF) unit root test, with intercept and trend, and the two significant lags, the log-values of the FRED II (replication) data set turn out not to contain a unit root (ADF test statistic: −3.965, as against the MacKinnon critical values of −3.43 at the 5% level and -4.00 at the 1% level). When one uses FRED II data and restricts the sample to match Atesoglu's, i.e., without the post-9/11 upswing in U.S. military expenditure, the ADF-test (level, intercept and trend, two significant lags) still does not show a unit root (ADF test statistic: −3.712, as against the same MacKinnon critical values). However, despite detrending in the unit root test, in both cases *rho* is close to unity at 0.96, and it might therefore be wise to convert this data series to percentage changes—d[ln(ML)]—to remove what may be a unit root in a practical if not statistical sense and run OLS and/or to use ln(ML) in a cointegration equation.

[15]Atesoglu used E-Views (version 2); the replication uses E-Views (version 4). Replication with FRED II data, but otherwise identical procedures and specifications.

	α	βML	δGL	λR
Atesoglu				
Coefficients	+1.237	+0.572	+1.106	−0.104
SE-coeffs	1.418	0.265	0.127	0.034
Replication				
Coefficients	+0.691	+0.669	+1.135	−0.118
SE-coeffs	1.780	0.345	0.147	0.030

With this baseline replication, the remainder of the article now addresses two questions: first, what happens to the parameter estimates in Atesoglu's model [equation (7)] if one (wrongly) uses budget rather than NIPA data and, second, what happens to the parameter estimates if one (correctly) uses NIPA data but changes the time period for which the parameters are estimated.

Regarding the first question, comparable inflation-adjusted annual budget and NIPA data are available for 1962 to 2002. In levels, the NIPA series has a unit root (level, intercept, 2 lags) which is removed for the first difference. The budget series likewise has a unit root (level, intercept, 1 lag), also removed for the first difference. Both series are integrated of order one, I(1), and the first differences are integrated of order zero, I(0). The *log* of the NIPA series shows no unit root (*rho* = 0.91; level, intercept, 1 lag) at the 10% level, but the log of the budget series does show a unit root (*rho* = 0.86; level, intercept, 1 lag) at the 10% level. One can difference the data and decide to work either with $\Delta y_t = y_t - y_{t-1}$, where y_t refers to the level data, or with $\Delta ly_t = ly_t - ly_{t-1}$, where ly_t refers to the log data.[16]

Suppose one decides to work with differenced-log data. An OLS estimation of equation (7) using the budget data results in an elasticity of output with respect to military expenditure of +0.0167 (growth in military expenditure has a positive effect on growth in national output). Using NIPA data, the elasticity is −0.0197. The sign of the crucial parameter estimate has been reversed. The budget data suggest a growth-enhancing effect, the NIPA data suggest a growth-dampening effect. Now employ OLS to run the log-data rather than the differenced-log data. The coefficient for the log of real military expenditure measured by budget data on real output is −0.036, whereas it is −0.025 when measured by NIPA data. Running this as a Johansen cointegration routine, because of the possibility of unit roots, the long-run estimate with budget data is −0.204 (using 4 lags; estimating an intercept and no trend), whereas it is +0.034 (also using 4 lags; intercept, no trend) with NIPA data. Again, the sign is reversed.

For the Ramanathan/Blackburn and for the Atesoglu models, use of budget versus NIPA data results in substantively different conclusions, even as both data sources purport to measure the same thing, United States military expenditure.

Analytic Considerations: The Timespan

Reconsider Atesoglu's original model (with NIPA data and use of the Johansen cointegration routine) but change the time period for which the estimation is done. For the following exercise run, first, 33 "rolling 20-year" estimations with quarterly data (1952:Q1–1971:Q4, 1953:Q1–1972:Q4, ..., 1984:Q1–2003:Q4); second, 23 "rolling 30-year" estimations with quarterly data (1952:Q1–1981:Q4, 1953:Q1–1982:Q4, ..., 1974:Q1–2003:Q4); third, 13 "rolling 40-year" estimations with quarterly data (1952:Q1–1991:Q4, 1954:Q1–1993:Q4, ..., 1964:Q1–2003:Q4); and fourth, 13 "rolling 40-year" estimations with annual data (1952–1991, 1953–1992, ..., 1964–2003). In each case, the Johansen cointegration technique is applied "mechanically" and only the estimated parameter for the long-run relation in equation (7) is recorded. For the runs with quarterly data, a lag-length of 16 is used; for the annual runs, a lag-length of 4 is employed. For the 33 20-year runs with quarterly data, one obtains 12 sign reversals (i.e., in 36% of the cases). The ratio of the largest to the smallest value (in absolute terms) is 7.769/0.006 = 1,294.8. For the twenty-three 30-year runs with quarterly data, there are 9 sign reversals (39%). The ratio of largest to smallest parameter value is 1,412.8. For the 13 40-year runs with quarterly data, two sign reversals

[16] Differenced-logs, the second expression, are approximations to growth rates. But differencing data—level or log—to achieve I(0) "limits the scope of the questions that we can answer" (Wooldridge, 2003, p. 615).

occur (15%). The ratio of largest to smallest value is 40.5. And for the thirteen 40-year runs with annual data, there are three sign reversals (23%), and the ratio of largest to smallest parameter value is 216.9. In a word, one finds frequent sign reversals and considerable instability in the estimated parameter values.

Even though the Keynesian model in equation (7) has good theoretical foundations and uses the preferred data—one reason why it was chosen for replication—it is by no means clear how to implement it empirically, i.e., over which time period to run it or whether to choose quarterly or annual data. Because of the statistical properties of the cointegration technique, it is generally preferable to work with as many observations and degrees of freedom as possible. This would speak in favor of quarterly data, and is possible to do that for the United States but not for most other countries. But why choose 40, or 30, or 20, as this exercise has done, or any other particular number of years? This, it seems to me, is a major methodological issue that needs debate since, as shown in this article, changing the timespan over which an estimation is run, can lead to dramatic sign reversals and drastic differences in the magnitude of the estimates even if they are of the same sign. At issue is what reliable statements we can or cannot make about the effect of military expenditure on an economy.

Conclusions

First, those interested in studying economy-wide effects of military expenditure should discontinue use of the budget line item 050 national defense outlays data (i.e., discontinue the use of NATO/SIPRI data for the United States). This number does not capture U.S. military expenditure. Instead, use of NIPA data is recommended. Even these, however, do not measure the full resource cost that the government defense function lays on the country. Second, a replication of Ramanathan and Blackburn's model yields statistically insignificant coefficients when run with budget data but results in statistically significant coefficients when run with NIPA data. Running an expanded equation to capture special-tension circumstances results in nonsensical coefficients with budget data but yields the expected and statistically significant coefficients when run with NIPA data. Third, a replication of Atesoglu's Keynesian model resulted in reversed signs for the relevant estimated parameter when run with budget and NIPA data. Moreover, when Atesoglu's model (with NIPA data) is run over different time periods, massive coefficient instability and frequent sign reversals are observed. The results also differ in magnitude and sign, depending on whether quarterly or annual NIPA data are employed. Atesoglu cannot be faulted since he properly used quarterly NIPA numbers, as do most analysts when dealing specifically with the United States. Rather, the point of this particular exercise is to examine the validity of results obtained for states for which quarterly NIPA-equivalent data are not available, cases in which researchers then routinely access the annual NATO/SIPRI, non-NIPA equivalent data. Were the proper data available, would the coefficients also change in magnitude and sign as for the case of the United States?

Ultimately, economists are interested in making reliable statements about the effect of military expenditure on an economy. From the work presented here, it would seem that our empirical work may need more validation and hedging than it ordinarily receives.

Acknowledgments

I thank Prof. Atesoglu for his kind assistance in replicating and interpreting the results of his original model. For comments, I also thank participants at the International Conference on Economics and Security, Technical Institute of Larissa, Greece, June 2004, participants at the Eastern Economic Association, New York, March 2005, three anonymous reviewers, and the editor of this journal. Any remaining shortcomings are solely the author's responsibility.

References

Atesoglu, S. H. 2002. Defense spending promotes aggregate output in the United States—Evidence from cointegration analysis. *Defence and Peace Economics* 13(1): 55–60.

[BEA] 1988. *Government transactions*. Methodology Paper Series MP-5. U.S. National Income and Product Accounts, Bureau of Economic Analysis, U.S. Department of Commerce. Washington, DC: Government Printing Office. 1988. Available at http: //www.bea.gov/bea/ARTICLES/NATIONAL/NIPA/Methpap/methpap5.pdf [accessed 4 April 2006].

[BEA] 2005. *Government transactions*. Methodology Paper Series MP-5. U.S. National Income and Product Accounts, Bureau of Economic Analysis, U.S. Department of Commerce. Washington, DC: Government Printing Office. September. Available at http: //www.bea.gov/bea/mp5.pdf [accessed 4 April 2006].

Brauer, J. 2002. Survey and review of the defense economics literature on Greece and Turkey: What have we learned? *Defence and Peace Economics* 13(2): 85–107.

Brzoska, M. 1982. Arms transfer data sources. *Journal of Conflict Resolution* 26(1): 77–108.

Brzoska, M. 1995. World military expenditures. In *Handbook of defense economics* Vol. 1, eds., K. Hartley & T. Sandler, 46–67. Amsterdam: Elsevier.

[BVC] 2002. *World Military Expenditures and Arms Transfers 1999–2000*. Bureau of Verification and Compliance, U.S. Department of State. Washington, DC: Government Printing Office.

Dewald, W. G., J. G. Thursby, and R. G. Anderson. 1986. Replication in empirical economics: The *Journal of Money, Credit and Banking Project*. *American Economic Review* 76(4): 587–603.

Dunne, J. P., R. P. Smith, and D. Willenbockel. 2005. Models of military expenditure and growth: A critical review. *Defence and Peace Economics* 16(6): 449–461.

Felipe, J., and F. G. Adams. 2005. A theory of production. The estimation of the Cobb-Douglas function: A retrospective view. *Eastern Economic Journal* 31(3): 427–445.

Kennedy, P. 2003. *A guide to econometrics*. 5th ed. Cambridge, MA: The MIT Press.

Mandel, B. A., and M. L. Roy. 2006. Federal budget estimates for fiscal year 2007. *Survey of Current Business* 86(3): 12–22.

Ramanathan, R. 2002. *Introductory econometrics with applications*. 5th edition. Mason, OH: South-Western.

Romer, D. 2000. Keynesian macroeconomics with the LM curve. *Journal of Economic Perspectives* 14(2): 149–169.

Scheetz, T. 1985. Gastos militares en Chile, Peru y la Argentina. *Desarrollo Económico* 25(99): 315–327.

Smith, R. P. 1998. Quantitative methods in peace research. *Journal of Peace Research* 35(4): 419–427.

Stock, J. H., and M. W. Watson. 1996. Evidence on structural instability in macroeconomic time series relations. *Journal of Business and Economic Statistics* 14(1): 11–30.

Taylor, J. 2000. Teaching modern macroeconomics at the principles level. *American Economic Review* 90(2): 90–94.

Taylor, J. 2001. *Principles of Macroeconomics*. 3rd ed. Boston: Houghton-Mifflin.

Wooldridge, J. 2003. Introductory econometrics: A modern approach. Cincinnati, OH: Thomson.