

Military expenditure data: theoretical and empirical considerations

Ron P. Smith¹

Department of Economics Mathematics and Statistics, Birkbeck University of London.

This paper discusses some of the methodological issues involved in analysing military expenditure data, with particular reference to the extended SIPRI dataset. The discussion is organised under the headings of validity, what is the appropriate concept to measure? reliability, how well is it being measured? and comparability, is the same thing being measured over time and space? The paper then considers some of the econometric issues involved in the use of such data.

Keywords: Military Expenditures, Data and Measurement, Panel data

JEL codes: C18, C82, H56

1. Introduction

This paper discusses some of the methodological issues involved in analysing military expenditure data, with particular reference to the extended SIPRI dataset. SIPRI military expenditure data, which previously started in 1988 now extends back to 1945. The other main source of data on military expenditures, the US Department of State WMEAT(2015) only gives eleven years of data, 2002-2012.

As is pointed out in SIPRI (2015, p400) there are three main types of issue with the data: validity, what is the appropriate concept to measure? reliability, how well is it being measured? and comparability, is the same thing being measured over time and space? This paper considers each of these in turn and then looks at the econometric issues in the use of such data. Little in this paper is new. The difficulties with measuring military expenditure are well known, Brzoska (1996) begins his survey by saying “Military Expenditure is difficult to define”. But since the points are not always appreciated, they may bear repetition. Many of

¹ Department of Economics, Mathematics and Statistics, Birkbeck, University of London, Malet Street, London WC1E7HX, UK. Tel. +44 2076316413; Fax: +44 2076316416. Email r.smith@bbk.ac.uk

these points are discussed in the SIPRI Sources and Methods section of the military expenditure chapter, SIPRI (2015, p400-4002) also available on the SIPRI webpages. Omitoogun and Skons (2006) discuss the origins of the SIPRI Military Expenditure Project. They note that the rationale for establishing SIPRI, and by implication the Military Expenditure Project was to produce a “factual and balanced account of a controversial subject - the arms race and attempts to stop it”. It is somewhat ironic that many of the most devoted users of SIPRI data are in defence establishments.

2. Validity

The validity of the measure, how closely it relates to the relevant concept, depends on the purpose for which the concept is being used and military expenditure measures are used for many purposes. As Dunne (1999, p377) points out “Military statistics, like other statistics, are not objective facts; they are constructed using theory and reflect social relations. They are also estimated using available information (there are no true statistics hidden in the military establishment waiting to be discovered).”

Similarly WMEAT (2015p1) begins “Readers should note that WMEAT figures, especially for armed forces personnel, military expenditures and arms transfers, are neither so accurate nor so reliable as uniform presentation in statistical tables might seem to imply, due to incompleteness, ambiguity, or total absence of data for some countries either in those parameters or in parameters, such as GDP price deflators or rates used to convert national-currency-denominated values to US dollars.”

In the country itself, it is the government that constructs the military expenditure data, primarily for purposes of budgetary control, fiscal policy and national defence planning. The measures may be used to assess the opportunity costs of military expenditure, the other uses to which the resources could be put, and thus contribute to arguments for reduced military spending or disarmament. Abroad, other countries may use the figures to assess threats from enemies or burden sharing by allies. Finally, the data are widely used for qualitative and quantitative academic research. These different users have different demands of the data and thus different concepts of validity.

SIPRI publishes three main measures of military expenditure: in local currency at current prices; in constant US dollars; and as a share of gross domestic product, GDP. For national budgeting purposes, the local currency, current price measure is the most relevant. For comparisons of the levels of military expenditure across time, one needs to remove the effects of inflation, for comparisons of the levels of military expenditure across countries one needs to express the figures in a common currency. The constant US dollar figure does both. We discuss the issues involved in comparisons further below.

The share of military expenditure in GDP indicates the priority attached to the military, how much of available output is devoted to defence, and is sometimes referred to as the military burden. The share, being a pure number, is comparable over time and space, without the complications of having to choose a price index or exchange rate, but it reveals nothing about the level of military expenditure. The share may be taken as an indication of commitment or intent as suggested by the importance attached by NATO to the agreement at the 2014 Cardiff summit to spend at least 2% of GDP on defence.² GDP is itself a problematic concept and revisions to the GDP figures can have a substantial impact on the estimated share.

SIPRI also provide, in a worksheet, military expenditure per capita in current US dollars and military expenditure as a share of general government expenditure. The share of government expenditure may be relevant for budgeting purposes; but given differences in the size and definition of government expenditure, over time and between countries, the share of government expenditure is less comparable than the share of GDP. There are many other potentially interesting ratios such as military expenditure per member of the armed forces, which gives some indication of the capital intensity of the provision.

There is an issue as to what price index should be used to deflate the current price military expenditure series. SIPRI uses a Consumer Price Index, CPI, which is consistent with an opportunity cost interpretation of the measure: the real amount of consumer goods given up to finance the military. To get a measure of the real output of defence would require a defence specific deflator. SIPRI comment that military specific deflators are not available for most countries. Using military specific deflators for some countries and CPIs for others

² Despite this agreement, only 4 of the 28 countries met the target in 2015: US, Poland, Greece UK, and Estonia.

would reduce international comparability. Even where they are available, military specific deflators are controversial, since the price and quantity of defence are not well defined. A series of articles on the measurement of defence inflation appear in a special issue of *Defence and Peace Economics* 27(2) April 2016.

As SIPRI emphasise, military expenditure is an input measure, the flow of resources devoted to the military over a period; not an output measure of the stock of armed forces or their military effectiveness. The forces obtained for the expenditure will depend on their price, the wages of troops and the cost of weapons. The wages will depend on such factors as the level of income of the country and whether there is conscription. The cost of weapons will depend on the efficiency of the procurement process and of the arms industry. The military effectiveness of those forces will depend on training, logistics, maintenance of the existing equipment, morale and leadership, strategy and tactics.

The correlation between military expenditure and military forces may be very low. The issues are illustrated by the 1976 revision of estimated Soviet military spending by the US Central Intelligence Agency, CIA. The incident is described by one of the analysts involved in Noren (2003). The CIA calculated Soviet military spending by a building block method. It first estimated the forces: numbers of troops, tanks, ships and missiles. Reasonable estimates of these could be obtained from technical intelligence sources such as satellite photographs. It then estimated what it would have cost the USA to purchase these forces, giving a dollar figure. This was then multiplied by an estimated rouble-dollar exchange rate, to get a rouble figure, which could then be expressed as a share of CIA estimates of Soviet GNP. In 1975, the CIA got "an unusually large body of new information" that suggested the Soviet military industry was much less efficient than it had previously thought and altered the exchange rate to reflect this. This raised the estimated share of military expenditure in GNP from less than 8% to 11-13%. Noren (2003) emphasises that the revision did not affect CIA's appraisal of the size or capabilities of Soviet military forces or change appreciably the estimated dollar cost of reproducing Soviet defence programs. The revision reflected the fact that the CIA thought that it was costing the Soviets far more resources to produce those forces. Although the revision reflected a judgement that the Soviets were less efficient than had been previously thought, it was interpreted as showing that the Soviets were more threatening than had been previously thought. Subsequently some have argued that the Soviet economy was

even less efficient than the CIA thought, and the actual share of military expenditure could have been over 20%.

The dangers of using measures of military expenditure as a measure of strength is also indicated by the fact that, contrary to Voltaire's judgement that God was on the side of the big battalions, the weak often win wars, Arreguin-Toft (2005). They can do this through asymmetric warfare, fighting in ways that the strong do not expect. A recent example of the difference between expenditure and effectiveness is the cost of preparing the 9-11 attack by the four planes. The 9-11 Commission (2004) estimate that the cost was between \$400,000 and \$500,000, a tiny expenditure compared either to its effect or to typical military spending numbers.

3. Reliability

Since different users will require different concepts, the issue of reliability, how well the concept is measured? cannot be completely separated from the issue of validity, what is the relevant concept? In most countries the definition of military expenditure is determined by institutional factors: what is included in the budget of the Ministry of Defence; rather than functional factors: what contributes to military capability. Thus it can be rather arbitrary whether functions like civilian search and rescue or disaster relief are included in the defence budget, depending on whether they happen to be done by the military or not. SIPRI has a standard definition which they try to apply but note that "While this definition serves as a guideline, in practice it is often difficult to adhere to due to data limitations."

Even among countries that provide a lot of data, reliability is reduced by a lack of uniformity in definitions of military expenditure. There are also problems of valuation, such as conscripts paid below market wages, or direct expropriation of resources without payment. Definitional differences include the line drawn between military and internal security functions including paramilitaries like the French Gendarmerie and the extent to which intelligence functions are included in the defence budget. In order to meet its commitment to NATO to spend 2% of GDP on defence, in 2015 the UK redefined the spending it declared to NATO, including elements of intelligence and security services that had not previously been included.

New approaches to financing, such as public private partnerships, where the private partner makes the investment and the public partner pays for the services over the years, complicate the calculations. The accounting may be on a cash basis, when the money is spent, or an accruals basis, when the liability is incurred. Some governments are following private sector accounting, and reporting a balance sheet of assets with associated charges for capital employed and depreciation. Pensions are a cost of doing business and should be included in the defence budget but institutional systems differ. Some include payments into a pension fund for serving troops, some include payment of pensions to past veterans.

Some countries, particularly poorer ones, refuse to publish any numbers or provide little detail, beyond a single undefined number. This may be the amount budgeted, rather than the amount spent, and may exclude categories like arms imports or capital expenditure. There is also deliberate manipulation of the data, including financing items off-budget. Some uses of the data can be counter-productive. As Omitoogun and Skons (2006) note, some foreign aid donors made lower military expenditure a condition for giving aid. This provides a direct incentive for countries to under-estimate military expenditure.

Ferguson and Perlo-Freeman (2015) find that countries that are less free (based on the Freedom House measure) provide fewer years of data, as do more fragile states (based on the World Bank State Fragility Index). It does not seem to be available resources that determine the availability of data, since after controlling for freedom and fragility, income is insignificant. The low transparency seems a matter of choice for low freedom states and a matter of capability for fragile states: they are unable to produce high quality numbers. Many other factors must matter since freedom and fragility only explain 29% of the variation in the availability of data.

The quality of military expenditure numbers deteriorates during conflict and the definitional difficulties can increase, since most conflicts now are intra-state rather than inter-state. This makes it harder to distinguish military functions from police or internal security functions.

4. Comparability

There are trade-offs between comparability over time, inter-temporal consistency, and comparability over space, international consistency between countries. As noted above, to

compare levels of military expenditure across time, one needs to remove the effects of inflation to get a real, constant price, series. To compare military expenditures across countries one needs to express the figures in a common currency, typically dollars. There are a number of ways one could get a constant dollar figure. WMEAT(2015) gives estimates based on five different methods of converting to US dollars and notes that the results yielded by these methods may differ substantially, both over time and across countries. It discusses the advantages and disadvantages of each. For instance, one could either (a) deflate the local currency series by a national price index and convert at the exchange rate of a base year or (b) convert the local currency current price series to dollars using the exchange rates of each year and then deflate by a US price index. SIPRI uses procedure (a) since it maintains the structure of the time series of real expenditure for the country, undistorted by exchange rate variations. But the cross-section pattern can be very sensitive to the base year chosen for the exchange rate and relative military expenditures between countries can change a lot when the base year is changed. Comparisons of the relative military expenditure of two countries far from the base year can also be misleading. SIPRI uses market exchange rates, rather than purchasing power parity rates, for reasons explained in SIPRI (2015, 401-2). WMEAT (2015, p29-34) discusses some of the considerations and describe their method 5, which attempts to construct a defence specific PPP measure. This uses the ratio of military spending per member of the armed forces to GDP per member of the labour force (as measures of capital intensity) to assign the defence sector dollar PPP value of military spending to a point on the interval between the market exchange rate and PPP values.

Revisions are inevitable as new data on both military expenditures and economic variables like GDP become available, so there can be big differences in series in different Yearbooks. Revisions can also be caused by rebasing of price indices or exchange rates. Brauer (2002), reviewing the literature on Greek Turkish arms races, noted that there were some “astonishing differences” in the time series graphs in two papers, despite being for the same time period and from the same source, SIPRI, though they used different yearbooks and different exchange rate bases. Definitional changes raise difficult issues. Although the NATO definition of military expenditure is similar to that used by SIPRI; in 2004 NATO changed its definition to exclude paramilitary forces. Had, SIPRI followed NATO, it would cause a large change in the estimates for countries that have paramilitaries, like France. WMEAT (2015, p8) followed NATO in changing the definition and notes “The data discontinuities resulting from the applications of the new definitions are large for some

countries.” SIPRI chose to maintain comparability over time, at the cost of comparability across countries. SIPRI (2015, p400) say “priority is given to choosing a uniform time series for each country, in order to achieve consistency over time, rather than to adjusting the figures for individual years according to a common definition.” The detailed data required to adjust for classification differences over time are rarely available.

When faced with two overlapping estimates on a different basis, it is common to project the later series backwards by “splicing” them using the ratio during the overlap period. This preserves the growth rates of the two series, but it imposes the level difference at the time of overlap on the whole of the earlier spliced series. In addition, such splicing may not preserve accounting identities, e.g. making components of military expenditure sum to the total, though, since SIPRI does not currently report components, this is not yet an issue.

5. Econometric issues

There is now data for $t=1,2,\dots,T$ for countries $i=1,2,\dots,N$ where the maximum values of both $T=65$ and $N=172$ are large. This allows considerable flexibility in how one models the data.

The first econometric issue in studies using military expenditure data is whether one is using the data in time-series, cross-section or in panel. The main benefit of the extended series is for time series analysis, since the large T allows for a more flexible treatment of the dynamics. This is important, since both the level and share seem to be integrated of order one, they need to be differenced once to make them stationary, and allowing for lags is important to reproduce the time-series properties of the data, including any cointegration among the levels of the variables. The large T of the extended SIPRI data set allows one to model not just the dynamics, but also to allow for heterogeneity and cross-section dependence in panels and to test for structural breaks, e.g. at the end of the Cold War.

In cross-section, a central issue tends to be functional form, for instance the effect of military expenditure on growth may be different at different levels of military expenditure or different levels of threat. While one needs data in comparable units, such as constant US dollars, for cross-section analysis, the treatment of prices and exchange rates in the conversion may add more noise into the relationships. In spatial models, which can be estimated from both cross-section and panel data, one is interested in measuring the effect of neighbours behaviour.

Therefore it is important to have a full list of neighbours available, and this may restrict the time period for which data for all the neighbours are available.

In panels if one is using balanced data, with the same time period for each country, there is a clear trade-off: the larger T the smaller N; fewer countries have data for long time periods. The set of missing countries is not random. If one wants a long run of data back to 1945, this excludes both many poor colonies, that only became independent later, and all the former Soviet republics, which did not exist before the end of 1991. Countries with much missing data are also not a randomly selected group of countries; as noted above, fragile and less free states are less likely to provide data. Thus with an unbalanced panel, changes over time may reflect the mix of countries in the sample rather than any real changes in the relationship.

A second econometric issue is whether military expenditure is the dependent or an independent variable in an equation. For instance in the arms race or demand for military expenditure literature, military expenditure is the dependent variable and GDP is typically treated as an exogenous independent variable. In the military expenditure and growth literature, military expenditures is typically treated as an independent variable and GDP is treated as the dependent variable. This simultaneity, both GDP and military expenditure being treated as endogenous in different literatures, raises delicate issues of identification. In particular, to identify both relationships in a system that determines both GDP and military expenditure one needs some exogenous measure of the threat that moves military expenditure but not GDP and some exogenous supply variable that moves GDP but not military expenditure. The relevant theory might also suggest using different measures of military expenditure depending on whether it is a dependent or independent variable.

A third issue is whether the level or share is the appropriate measure. Brauer (2002) has a nice discussion of this choice. While the choice depends on the purpose of the exercise and the question being asked, one can often include both the logarithm of the level of military expenditure and the logarithm of GDP, in common units, and test for the restriction that it is the share, the difference between log military expenditure and GDP, that matters. An income elasticity of demand for military expenditure close to one often seems reasonable. In models explaining growth, one tends to include the share of investment in GDP as a determinant, so by analogy the share of military expenditure may be an appropriate determinant.

A fourth issue is the treatment of uncertainty. SIPRI signal degrees of uncertainty with round or square brackets and this information could be used to examine the effects of measurement error and the sensitivity of the conclusions to the use of different quality measures of military expenditure. But as noted above, the quality of military expenditure data is correlated with other things such as state capacity, and this introduces further complications. A number of studies have compared SIPRI measures with other measures, such as WMEAT, and used differences between them as another indicator of uncertainty. But if they both use the same single source, they may agree on a poor quality number.

6. Conclusion

The extended SIPRI data set allows a range of new questions, particularly in the time series dimension, to be answered with a high quality data set. But the limitations of measurement and the concerns of validity, reliability and comparability must be appreciated. Feeding the data into the computer is not enough, sensitivity to the question being asked, to the qualitative historical features of the countries concerned, and the econometric issues, particularly identification, is also required.

ACKNOWLEDGEMENTS

Prepared for the special issue of *Defence and Peace Economics*, on the Extended SIPRI military expenditure dataset. This paper is part of a larger SIPRI project supported by Sweden's Riksbanken Jubileumsfond. Sam Perlo-Freeman, Christos Kollias, Jurgen Brauer Paul Dunne and an anonymous referee provided useful comments on earlier drafts.

DISCLOSURE STATEMENT

No potential conflict of interest was reported by the author.

References

- Arreguin-Toft, I (2005) *How the weak win wars: a theory of asymmetric conflict*, Cambridge University Press.
- 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), 85-108.
- Brzoska, M. (1995) World Military Expenditures, p45-67 (chapter 3) of *Handbook of Defence Economics, Vol. 1*, K. Hartley and T. Sandler, (eds) Elsevier.

Dunne, Paul (1999) The Statistics of Militarism in *Statistics in Society: the arithmetic of politics*, Daniel Dorling and Stephen Simpson (eds), 376-83, Arnold Applications of Statistics Series, London.

Ferguson, N. T. N. and S. Perlo-Freeman (2015) The availability of military expenditure data, Chapter 9 III, p360-368, *SIPRI Yearbook 2015: Armaments, Disarmament and International Security*. Oxford University Press.

Noren, J. (2003) CIA's Analysis of the Soviet Economy, Chapter 2 of *Watching the Bear, Essays on the CIA's Analysis of the Soviet Union*, edited by Gerald K Haines and Robert Legget Center for the Study of Intelligence monograph,..CIA.

Omitoogun, W and E Skons (2006) Military expenditure data: a 40-year overview, p269-294, (chapter 7) *SIPRI Yearbook 2006: Armaments, Disarmament and International Security*. Oxford University Press.

SIPRI, various years, *Armaments, Disarmament and International Security* Stockholm International Peace Research Institute Yearbook, Oxford University Press.

The 9-11 Commission (2004) *The National Commission on Terrorist Attacks Upon the United States*, US Government available <http://www.9-11commission.gov/>

WMEAT (2015) *World Military Expenditures and Arms Trade. Sources, data and methods*, US Department of State available at <http://www.state.gov/t/avc/rls/rpt/wmeat/>.