



Issues in Measuring Defence Output: A Review

Ron P. Smith

ESCoE Discussion Paper 2024-06

April 2024

ISSN 2515-4664

DISCUSSION PAPER

Issues in Measuring Defence Output: A Review
Ron P. Smith
ESCoE Discussion Paper No. 2024-06
April 2024

Abstract

This paper reviews the issues involved in the measuring defence output, inputs and productivity within the UK national accounts. The little literature that there is on this subject suggests that providing such measures would be a challenging task. Given the diverse audience this paper may have - in academic, national accounting and defence communities - I have included discussion of issues that may be well known among those who work in each particular area. Until 1998, measures of government output in the UK national accounts assumed the volume of inputs equals the volume of outputs, which implies that there can be no productivity growth. Subsequently the UK began to adopt measures of the volume of output for some government services and this process was accelerated by the Atkinson Report of 2005. The volume of output for many government services, like elements of health and education, was measured by the number of activities performed. For defence, there is a problem of comparability because the nature of the activities, capabilities and objectives of defence change over time, and for good reasons, as threats, technology and strategy evolve. These issues are discussed in the paper. When these activities or capabilities are discontinued to reorient to the new context, measured output will have fallen while the military are fully occupied doing different activities. For defence, performance measures include elements like success in operations, maintaining readiness, and stopping equipment being delivered late, over budget and not meeting technical requirements. These are difficult to convert into indicators that would match national income accounting criteria.

Keywords: national accounts, defence output, defence inflation

JEL classification: C8, H56

Ron P. Smith, Birkbeck University of London
r.smith@bbk.ac.uk

Published by:
Economic Statistics Centre of Excellence
King's College London
Strand
London
WC2R 2LS
United Kingdom
www.escoe.ac.uk

ESCoE Discussion Papers describe research in progress by the author(s) and are published to elicit comments and to further debate. Any views expressed are solely those of the author(s) and so cannot be taken to represent those of the Economic Statistics Centre of Excellence (ESCoE), its partner institutions or the Office for National Statistics (ONS).

Issues in Measuring Defence Output: A Review

Ron P. Smith, Birkbeck University of London

April 2024

This paper reviews the issues involved in the measuring defence output, inputs and productivity within the UK national accounts. The little literature that there is on this subject suggests that providing such measures would be a challenging task. Given the diverse audience this paper may have - in academic, national accounting and defence communities - I have included discussion of issues that may be well known among those who work in each particular area. Until 1998, measures of government output in the UK national accounts assumed the volume of inputs equals the volume of outputs, which implies that there can be no productivity growth. Subsequently the UK began to adopt measures of the volume of output for some government services and this process was accelerated by the Atkinson Report of 2005. The volume of output for many government services, like elements of health and education, was measured by the number of activities performed. For defence, there is a problem of comparability because the nature of the activities, capabilities and objectives of defence change over time, and for good reasons, as threats, technology and strategy evolve. These issues are discussed in the paper. When these activities or capabilities are discontinued to reorient to the new context, measured output will have fallen while the military are fully occupied doing different activities. For defence, performance measures include elements like success in operations, maintaining readiness, and stopping equipment being delivered late, over budget and not meeting technical requirements. These are difficult to convert into indicators that would match national income accounting criteria.

This research has been funded by the Office for National Statistics as part of the research programme of the Economic Statistics Centre of Excellence (ESCoE). I am grateful for comments on earlier versions from colleagues at ESCoE, ONS and MOD as well as Paul Dunne, David Kirkpatrick and Keith Hartley. The views in this paper are my own and should not be attributed to ESCoE, ONS or MOD.

Keywords: national accounts, defence output, defence inflation.

JEL codes: C8, H56

1. Context

This paper reviews the issues involved in the measuring defence output, inputs and productivity within the UK national accounts. The little literature that there is on this subject suggests that providing such measures would be a challenging task. Given the diverse audience this paper may have - in academic, national accounting and defence communities - I

have included discussion of issues that may be well known among those who work in each particular area.

Until 1998, measures of government output in the UK national accounts assumed the volume of inputs equals the volume of outputs, which implies that there can be no productivity growth. Subsequently the UK began to adopt measures of the volume of output for some government services and this process was accelerated by the Atkinson Report of 2005.¹ The volume of output for many government services, like elements of health and education, was measured by the number of activities performed.

Following the Atkinson Report, a November 2008 Scoping Paper² reported on the joint work by ONS (UKCeMGA) and MOD to consider improved measures of defence in the National Accounts. The Scoping paper pointed out many difficulties involved in measuring defence output and inputs, discussed further below. Section 6 of the Scoping paper was titled “Next Steps”. With respect to inputs there were some specific suggestions as to how their measurement could be improved. With respect to output it said that ONS would seek to agree a work plan with MOD. It is not clear how these initiatives developed. In the event the inputs equals outputs approach was maintained for defence.

In 2023 a Public Services Productivity Review was commissioned and the question of how to measure defence output and productivity was reopened. This paper is a contribution to that discussion. While I refer to defence without qualification, under the classifications of the functions of government, COFOG, developed by the OECD, the defence category does not exactly match the UK defence budget, for instance excluding military health and education services. The COFOG second level headings are military defence; civil defence; foreign military aid; research and development defence; and defence not elsewhere classified.

Elsewhere³ I discuss the more general methodological issues involved in measuring military expenditure under the headings of validity: what is the appropriate concept to measure? reliability: how accurately is that concept measured? And comparability: is the same thing being measured over time and space? For national accounts purposes comparability over time

¹ Atkinson, A.B. (2005) Atkinson Review: Final Report: Measurement of Government Output and Productivity for the National Accounts. Basingstoke: Palgrave Macmillan.

² Scoping Paper on Possible Improvements to Measurement of Defence in the UK National Accounts, UK Centre for the Measurement of Government Activity, ONS 19 November 2008. See also “Measuring Defence” Mavis Anagboso and Alison Spence, Economic & Labour Market Review, vol3 3, No.1, January 2009, ONS.

³ R.P. Smith (2017) “Military Expenditure Data: Theoretical and Empirical Considerations” *Defence and Peace Economics*, 28:4, 422-428.

is a major issue for defence. What constitutes a defence output changes over time as threat, technology and strategy evolve. Similarly, those activities the defence budget includes or excludes changes over time.

There are international differences in how defence expenditure is measured and there have been efforts to construct comparable series for international comparisons including using defence specific purchasing power measures. But, as far as I am aware, no country tries to measure defence output directly in its national accounts.⁴ There are two questions to address: what indicators to choose? And how to combine those indicators into a single measure? The choice of indicators could be politically very sensitive and raise difficult questions for national statisticians, such as what the 2003 invasion of Iraq or the two failures of UK Trident missile tests implied for defence output. There is also a danger that the use of particular output indicators may lead to perverse incentives or misleading signals, examples of which will be provided below.

Section 2 has a brief review of the academic treatment of the output of defence. Unfortunately, this literature provides little operational advice for national accountants. The measurement of government output for national accounts purposes is clearly related to the measurement of performance for the management of public services, but there are differences. Section 3 examines the relationship between these two sets of concepts. Section 4 reviews the 2008 Scoping Paper which provided a framework for the analysis of alternative measures. Its preferred measure of output was one based on capability. Capability is considered in section 5. Capabilities are defined relative to objectives and section 6 looks at how UK defence objectives have changed over time. Section 7 reviews a range of other suggested approaches. The discussion so far has focussed on output measures, section 8 looks at input measures. Any measures will rely on disaggregated data and section 9 reviews the data available. Section 10 contains some concluding comments. The Atkinson Principles for the Measurement of Government Output and Productivity are provided in an Appendix.

2. Academic approaches to defence output

⁴ The attempts to measure defence output in a number of countries, together with the more general issues, are discussed by Keith Hartley and Binyam Solomon in "Measuring Defence Output: An Economics Perspective" Chapter 3 of *Military Cost-Benefit Analysis*, Edited By Francois Melese, Anke Richter, Binyam Solomon, Routledge 2015.

There is a large academic literature on the determination of defence budgets and within that the most common concept of output is an unobserved variable: security. A common framework in the demand for military expenditure literature is to posit that a country has a welfare function, which depends on security and consumption. Security, which corresponds to the output of defence, depends positively on the country's own military expenditure and the support provided by allies and negatively on the threat represented by the military expenditures of its adversaries. Then, given a budget constraint, such as income equals consumption plus military expenditure, and the relative prices of military and consumption goods, one can derive a welfare maximising defence budget conditional on prices, income and the military expenditures of others.⁵ In principle, one might be able to derive a measure of security from these estimates, but few of these authors seem to have regarded their estimates as sufficiently robust to attempt this.

Empirical implementation of estimation of a true demand function for military expenditures requires price data. However, relatively few studies have trusted the available measures of defence prices and most omit price. How to measure the threat has also been an issue. For instance, during the Cold War Soviet military expenditure was used in a number of papers. But it was unclear, perhaps even to the Soviets, how much the USSR spent on the military. More amorphous threats like terrorist attacks are even more difficult to quantify. The literature on the demand for military expenditure is a subset of a much larger arms race literature. Even where there are what in qualitative terms appear to be clearly identifiable arms races, such as India and Pakistan or Greece and Turkey, it has been difficult to find stable functions linking the military expenditures of the two countries.⁶ Similarly, while it was widely believed that there was an action-reaction process between the nuclear arsenals of the US and USSR it was difficult to find quantitative evidence of it.

Alliances play a central role in this literature and within NATO there has been a long-running debate about burden sharing and the extent to which the smaller countries have been able to free-ride on the defence expenditure of the other countries.⁷ UK and US shares of GDP spent

⁵ This approach is surveyed in Ron P. Smith, *The Demand for Military Expenditure. Handbook of Defense Economics, Vol 1*, edited by K Hartley and T Sandler, Elsevier Science, p70-87, 1995.

⁶ J.P. Dunne and R.P. Smith, *The Econometrics of Military Arms Races* Ch 28 p913-940 of *Handbook of Defense Economics, Vol 2: Defence in a Globalised World*, ed T. Sandler and K. Hartley, Elsevier 2007.

⁷ A recent contribution is Wukki Kim & Todd Sandler (2024) *NATO Security Burden Sharing, 1991–2020*, *Defence and Peace Economics*, 35:3, 265-280, DOI: 10.1080/10242694.2023.2230408

on defence are positively correlated though presumably this is because they are both responding to the same threat rather than because the UK regards the US as an antagonist.

Parts of the academic literature tends to treat the defence budget as a homogeneous whole, but other parts disaggregate it into personnel, equipment and other elements. Questions addressed include substitution of labour for equipment and volunteers for conscripts.⁸

3. National Accounts and Defence Management Concepts

Principle A of the Atkinson Report was that Government non-market output should, as far as possible, be measured parallel to that adopted for market output. Defence is an extreme example of a non-market activity. There are no prices, it is provided free to the whole population, rather than to individuals as are health and education. Although not all agree, defence is commonly given as the classic example of a public good that is non-excludable and non-rival. In economics jargon, the optimal provision of a public good occurs when the sum of the marginal benefits (the marginal rates of substitution between public and private goods) over consumers equals the marginal cost of production (the marginal rate of transformation) or where the marginal security benefit equals the opportunity cost, the other goods and services foregone. The theory of public goods is of little help in determining the quantity of defence output. Rather it might suggest that the framework appropriate for private goods, where expenditure equals quantity times price, is inappropriate for defence.

Within the framework of expenditure equals quantity times price, the national accounts measure quantity, the volume of output, by constant price GDP and prices by the GDP deflator. Given that one observes defence expenditure, but not quantities and prices, one could either take expenditure, try to measure output and derive price as the ratio.

Alternatively, one could take expenditure, try to measure prices and derive output as the ratio. There are a range of linked relationships. The rate of growth of expenditure is equal to the rate of growth of output plus the rate of growth of prices, inflation. The rate of growth of productivity is the rate of growth of output less the rate of growth of inputs. For labour productivity the input is just labour, the number employed, for total factor productivity the

⁸ Bove, V., and E. Cavatorta. 2012. "From Conscription to Volunteers: Budget Shares in NATO Defence Spending." *Defence and Peace Economics* [23](#) (3): 273–288. doi:10.1080/10242694.2011.563973.

input is a weighted average of all the factors of production, such as labour, capital, intermediate inputs and research and development, R&D, to improve products or processes.

There is a potential confusion in that the measure of inflation that is relevant to the national accounts is different from the measures of inflation that are relevant for defence management. Within defence there are three measures of inflation that matter: inflation between generations of equipment, how much more new equipment costs e.g. an F35 costs relative to a Typhoon/Eurofighter; inflation on projects, the cost escalation on a particular procurement like the Ajax armoured vehicle; and inflation of the prices of inputs. None of these concepts of inflation are directly related to the national accounts measurement of defence output and productivity.⁹

There is considerable related work within the MOD that is distinct from the national accounting issues. Measures of defence inflation are regularly calculated from input cost growth numbers and these were published in the past¹⁰ and although publication ceased in 2017, they are available on a restricted basis. There has also been work to identify the drivers of defence cost inflation¹¹ and to identify appropriate indexes for projects which have an inflation adjustment on fixed price contracts. During low inflation times the MOD moved to fixed price contracts which are not indexed, but in high inflation times these can involve a large risk premium and following the increase in inflation MOD is now moving back to indexed contracts. There has been considerable work within MOD on the choice of appropriate indexes, which potentially could be used in improving input price indices.

Just as the concepts of inflation used in MOD differ from those used in national accounts, the concepts of output, productivity and efficiency used in defence management differ from those used in national accounts. For instance in oral evidence to the HCDC Chief of the General Staff, General Sir Patrick Sanders, said: “I think—I know—that the British Army is the most productive army in Europe. We are delivering at 130% of our capacity. We are delivering the same outputs at the same levels of deployment that we did from an Army that was 20,000

⁹ The general issue of defence inflation is discussed by MOD economists in N. Davies, T. Turner, A. Gibbons, S. Davies, D. Jones and N. Bennett, “Helping secure ‘the biggest bang for the taxpayers’ buck’ defence resource management in the United Kingdom, Ch.17 of *Handbook on the Economics of Conflict*, edited by Derek L. Braddon and Keith Hartley, Edward Elgar 2011.

¹⁰ https://assets.publishing.service.gov.uk/media/5a7f84f740f0b62305b87a39/20170126-Defence_Inflation_Statistical_Notice_1516-OS.pdf. The methods are described in “Estimating Defence Inflation” D.L. Jones and N.J. Woodhill, Defence Statistics Bulletin No. 10 March 2010.

¹¹ Evidence Summary, The Drivers of Defence Cost Inflation, MOD, February 2022.

bigger than it is now. This is an extraordinarily productive Army.”¹² Operationalising those claims might be difficult and sensitive issues might arise if military and statistical judgements of defence productivity conflicted.

Edward Luttwak in a chapter entitled *The Materialist Bias: Why we need more “fraud waste and mismanagement”* has argued that the notions of output and efficiency used by economists and accountants conflict with sound military practice. “The trouble is that the outputs that count in war are very particular and very different from the outputs that count in peacetime, and when civilian notions of efficiency are applied, the difference is routinely overlooked.”¹³

Measures of public sector output are inherently political, both in terms of party politics and bureaucratic politics. They will be interpreted or misinterpreted to make political points. The measures are also political in that incentive effects mean that Goodhart’s Law operates, when a measure becomes a target it ceases to be a good measure, and anything that can be measured will be gamed. That was a particular problem when public expenditures were measured in real terms under the PESC (Public Expenditure Survey Committee) system. This was gamed through arguments about “relative price effects”: the extent to which Departments claimed compensation for increased input costs. The PESC system of volume planning was abandoned in favour of cash planning in 1982 partly for this reason. For public expenditure planning purposes all departmental budgets are deflated by the same GDP deflator. The bureaucratic politics of output measurement are delicate because Departments have incentives to claim both its success in raising output and productivity and its need for a larger budget because of the high inflation it faced. The fact that they are inconsistent does not inhibit such claims.

There is a value chain in which money is spent paying for inputs, providing outputs and contributing to outcomes. In defence, the budget is spent to procure forces, troops and equipment; the forces provide capability, the ability to do things; these capabilities influence operational outcomes, whether you prevail in combat; which contribute to the ultimate outcome, UK security. The National Audit Office conducts regular reviews of the extent to which value for money is obtained in this process, but this is a different concept of productivity from that in the national accounts.

¹² HC Defence Committee, Oral evidence, 7 November 2023. Q223..

¹³ P133, of *The Pentagon and the Art of War*, Simon and Schuster 1985.

It can be difficult to measure output when the desirable outcome, the indicator of success, is that nothing happens. Defence is effective when threats are deterred, police are effective when there is no crime, doctors are effective when disease is prevented. In each case the effectiveness will be a function of the environment, how great are the threats from attack, crime, or disease. With preventative medical and public health services, one can estimate the effect of the treatment (e.g. heart attacks avoided). With defence, such counterfactuals (e.g. wars avoided) are more speculative.

During the Cold War, where the concern was with potential megadeaths, the threat was often quantified using the time before midnight on the clock published in the Bulletin of Atomic Scientists. This is still published and regularly reset. It reflects a subjective judgement about a global threat so may not be suitable as a UK national accounts input.

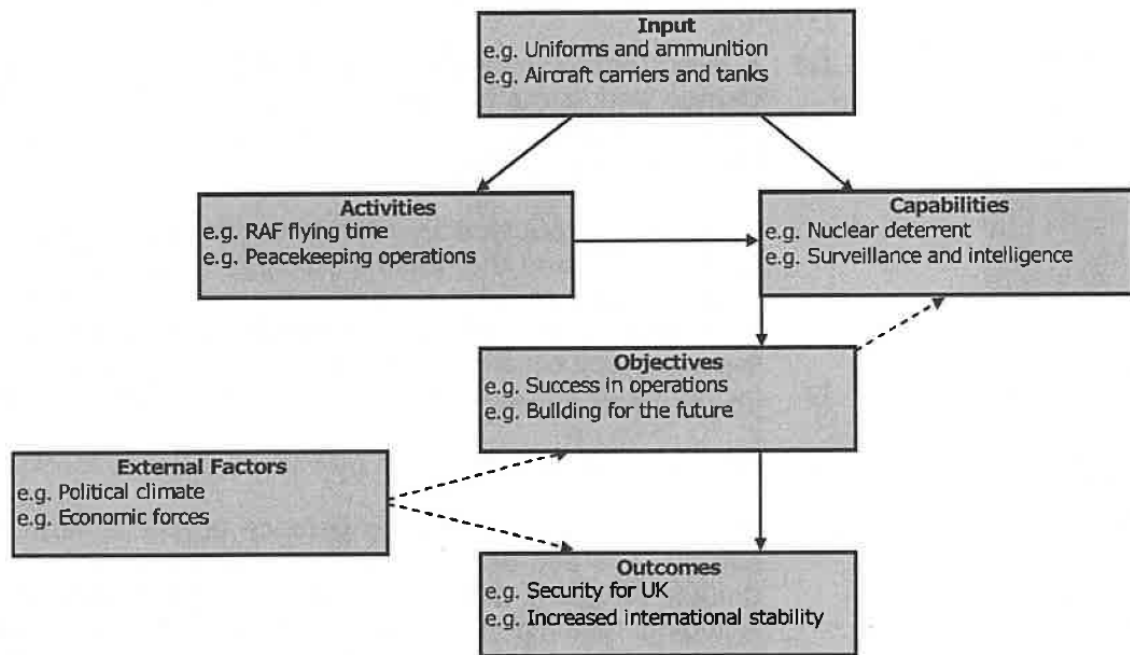
The UK share of defence in GDP tracks movements in perceived threat quite well. Using Stockholm International Peace Research Institute (SIPRI) numbers, it rose from 7.2% in 1950 to 10.9% in 1953 with the Korean War; then slowly declined to 4.6% in 1979, but rising to 5.5% in 1985 with the increased tension associated with Soviet nuclear enhancements and invasion of Afghanistan. From 1985 there was a widespread perception that the threat from the Soviet Union had decreased and the UK defence budget declined in line with that perception, stabilising around 2.5% in the middle 1990s, and subsequently drifting down to 2%. Transitory operations like the 1982 Falklands War or the 1991 Gulf War are not apparent in the movements of the share. They were “come as you are” wars largely fought using what was at hand. While the share of GDP devoted to defence might be regarded as a good indicator of perceived threat, that does not help provide an output measure.

4. The 2008 Scoping Paper

The 2008 Scoping Paper, mentioned above, discusses the measurement of both inputs and outputs. It provides a good account of the difficulties involved. As a result, the treatment of defence in the national accounts still follows the output-equals-inputs approach, because, as yet, there is no agreed definition of output. The framework used to relate inputs, outputs and outcomes is shown in the Figure below taken from the Scoping Paper. The inputs (e.g. aircraft carrier and tanks) provide resources for activities and capabilities. The activities (e.g. RAF flying time) feed into the capabilities (e.g. the nuclear deterrent). Capabilities feed into objectives, (e.g. success in operations), which feed into outcomes, (e.g. security for the UK). Both objectives and outcomes are influenced by external factors (e.g. the political climate).

Figure 5.1

Relationship between inputs, output and outcomes



The Scoping Paper proposed and discussed three potential measures of defence output – activities, capabilities and objectives, with a preference for capabilities. It said that UKCeMGA will seek to agree a work plan with MOD which will take forward jointly development of at least one of these approaches. It is not clear how much development was in fact done.

As inputs to a measure of output, one requires measures, weights or values and if possible quality adjustments. One of the Atkinson quality criteria was sustainability: measures should be available for the foreseeable future and for a reasonable number of periods in the past. Thus, one is looking for measures that are regularly reported in a consistent way as part of the normal administrative processes. Number of troops on operations or maintaining a Continuous at Sea Deterrent would be examples that meet these criteria. Again, one also needs to be careful that one does not choose a measure that may send misleading signals. Operational activity is driven by conflicts such as Iraq and Afghanistan. It might be thought undesirable to have defence output fall just because a war finished and operational activity declined.

5. Capability Measures

The Scoping Paper has a discussion of the use of training data and trained strength, 5.2.3-5.2.6 and concludes that an activity index (built up from training data) is unlikely to be a suitable measure of defence output. It preferred a capability measure, the ability to do things rather than the things actually done and suggested quality adjusted equipment and manpower measures. However, equipment counts can be misleading, since readiness can vary. For the 1991 Gulf War it was discovered that the usable number of armoured vehicles was much smaller than the number nominally on the books. Paras 5.3.15 of the scoping paper gives a number of difficulties with using an output measure based on capabilities and Para 5.4.15 points out that the way the MOD publishes readiness data is constantly evolving and as a result it may be difficult to obtain sustainable measures. The 2024 House of Commons Defence Committee Report “Ready for War?”¹⁴ discusses three concepts of readiness (operational, warfighting and strategic) and notes that it was hampered in its attempts to assess readiness by a lack of government transparency. Information on readiness that was previously in the public domain was now classified. Classified material on readiness may not be available to ONS to enable them to construct capability measures.

Data on UK armed forces equipment and formations are published regularly, though the categories used change through time with the evolution of the threat, technology and strategy. These numbers could be used to construct a measure of forces: the weighted sum of numbers of particular types of equipment or formation. The SIPRI construct measures of the volume of the arms trade in this way, using a weighted average of particular types of system. They emphasise that these weights, labelled trend indicator values, do not correspond to actual prices paid, which are often unknown.

Over time in the UK both numbers in the armed forces and number of platforms (aircraft, ships, tanks) have declined as fewer troops are recruited or platforms are procured in each generation. However, each element of the new generation may have more capability than the previous generation. There is a large literature on such comparisons with little in the way of quantitative conclusions. To construct a quantity measure one would need to be able to weight the capability of an F35 relative to a Typhoon/Eurofighter. With the reduction in numbers, capabilities can become quite lumpy because they often depend on relatively few platforms, such as two aircraft carriers or four nuclear missile submarines (SSBN). Events

¹⁴ House of Commons Defence Committee, First Report of Session 2023-24 HC 26, published 4 February 2024.

like the October 2020 burst fire main which flooded the engine room of the aircraft carrier HMS Prince of Wales and subsequent problems with its propeller shafts, both of which put it out of action for significant periods could cause large reductions on measured capability. Similarly, the entry into service of such a platform could cause a large increase in measured capability.

Even if one had an agreed list of indicators of capability, there is an issue as to how they should be combined. Additive measures are unlikely to be appropriate. There is a large academic literature on “technologies of conflict” and “conflict success functions” how inputs, relative force levels, influence outcomes, usually expressed as probability of winning.¹⁵ These overlap with standard military operational analysis and investment appraisal procedures.¹⁶ They involve aggregators, which combine the individual elements into an overall measure of force effectiveness. It is clear that you cannot add aircraft and pilots, the capability is the number of aircraft with pilots. The Defence Secretary, Ben Wallace, commented in November 2022 that “I have more F35s than I have pilots at the moment. Part of that is a Catch-22. You do not have enough instructors to train the pilots, so you do not have enough pilots to grow the instructors.”¹⁷

Defence planning emphasises that readiness and capability is not just about availability of platforms and troops but also the systems that surround them, summarised in what in the UK are called the Defence Lines of Development: labelled TEPIDOIL: Training; Equipment; Personnel; Information; Concepts & Doctrine; Organisation; Infrastructure; and Logistics. Lack of logistic support ships has constrained the operations of the two UK aircraft carriers. It is not clear how one would allow for that in an additive capability measure. The academic literature has also investigated issues like capital-labour and quantity-quality substitution in the provision of defence. There can also be substitution with non-defence activities. Examples are the move to using the police rather than the army to maintain order in Northern Ireland and the move to having helicopter search and rescue provided by a private company rather than the military.

¹⁵ See for instance Michelle Garfinkel and Stergios Skaperdas “Economics of Conflict: An Overview” Chapter 22 in Todd Sandler and Keith Hartley, *Handbook of Defense Economics: Defense in a Globalised World*, Elsevier 2007.

¹⁶ See for instance David Kirkpatrick “Choose Your Weapon: combined operational effectiveness and investment appraisal (COEIA) and its role in UK defence procurement, RUSI Whitehall Paper Series 1996, .

¹⁷ House of Lords, International Relations and Defence Committee, oral evidence: Defence concepts and capabilities: from aspiration to reality. 1 November 2022. Q176.

The MOD has regular capability reviews but the list of capabilities changes over time as threats and technologies evolve and the details are often classified. Capabilities are complicated as the MOD says “we recognise that Military Capability is about more than just equipment, it is the combination and integration of equipment alongside our infrastructure, personnel, information, data, concepts, and training.”¹⁸ This makes them difficult to measure. Activities also change over time. Support for Ukraine, in particular delivery of weapons, became an important activity after the Russian invasion, but may not be a permanent one.

In models of conflict success, the productivity of your forces is a complex function of any allied support and the quantity and quality of enemy opposition forces. There have been a number of attempts to construct capability measures that were weighted sums of force elements. This has proved problematic as a measure of the ability to meet military objectives both because of the difficulty of choosing functional forms and weights and because the militarily weak often win wars by adopting asymmetric tactics that prevent the stronger power using their superior military forces effectively.

6. Changing Objectives

The achievement of objectives may provide some indication of outputs and productivity, but again one runs into the issue of sustaining comparable measures over time, since the MOD definitions of the relevant activities, capabilities and objectives change. The Scoping Paper discussed the shift in objectives from those of the Cold War to Counter Terrorism. The objectives listed in Annexe E of the Scoping Paper were: 1. Achieving success in the tasks we undertake. The tasks listed were: operations, other military tasks, contributing to the community and helping to build a safer world. 2. Being able to respond to task that might arise which mainly related to equipment and capabilities. 3. Building for the future. Which included research and the equipment program. In the annual performance report for 2007-8, page 13, there are six targets under these three objectives. One target is judged met, four partly met and one not met (Be ready to respond to the tasks that may arise).¹⁹ Similar “traffic light” signals are used for lower level targets. It is now rare to find such explicit

¹⁸ The Defence Capability Framework, July 2022. This provides a longer term demand signal to industry.

¹⁹ Annual Report and Accounts 2007-8 Volume 1, Annual Performance Report. HC850-1.

judgments of success against pre-specified targets in MOD publications. Instead, there is an emphasis on key achievements.

Following the 2010 Defence Review, three National Security Objectives evolved: to Protect the UK, to Project influence and to Promote Prosperity. While protection and projection are natural objectives of MOD, there was some scepticism as to whether a defence ministry had the skills or tools to promote prosperity. The 2022-3 the MOD Annual Report and Accounts lists key achievements under five priority outcomes: 1. Protect the UK and its Overseas Territories 2. Enhance global security through persistent engagement and response to crises 3. Understand and counter, state and non-state threats 4. Contribute to NATO collective deterrence and Defence 5. Modernise and integrate Defence capabilities by taking a whole force approach to our people and increasing the use of technology and innovation. Some of the reported key achievements are sustainable measures, though difficult to value, such as 40 years of peace in the Falklands. Some vary depending on need. These include Military Aid to Civilian Authorities, explosive ordnance disposal and ceremonial duties such as royal funerals²⁰ or coronations. Having the death of a monarch generate a rise in defence output might be thought undesirable. Creating and supporting jobs is given as a key achievement under priority outcome 5, but this is not an output in the usual national accounting sense.

The changing nature of the army objectives was nicely described by Chief of the General Staff, General Sir Patrick Sanders in oral evidence to the HCDC. He said “the nature of the operation or the task you are going to undertake is always sui generis; it is always unique. I have served in, if you like, four Armies: one that was optimised for the cold war; one that was optimised for interventions, starting with humanitarian interventions in Sierra Leone and the Balkans, and then expanded into interventions in the Middle East; one that reset itself for counter-insurgency or stabilisation operations in Iraq and Afghanistan and, tactically, became extremely proficient and well prepared for it; and now an Army that is optimising for war and for warfighting.”²¹ Measuring the capabilities to meet such changing objectives is challenging, let alone measuring the extent to which the objectives were met.

As in other areas of public services it can be difficult to separate quality change and productivity improvements. In education and health, online lessons and online GP

²⁰ 6,000 members of the armed forces were deployed on ceremonial duties associated with the funeral of the Queen. Annual Report and Accounts 2022-3, p22.

²¹ HC Defence Committee, Oral evidence, 7 November 2023. Q211.

appointments could be regarded as either productivity improvements or reductions in quality relative to in-person equivalents. While it may be more difficult to judge quality in the case of fighter pilots than teachers or GPs, similar issues arise. Would there be a reduction in activity if RAF flying time could be reduced because better simulators are available, or would there be an increase in productivity if the RAF followed the army and allowed (lower paid) NCOs to be pilots.

7. Other approaches

A major role of defence is to prevent attack- and if defence does deter attacks on the UK, then depending on the counterfactual chosen, the value could be very large. The value of the assets not destroyed and lives not lost by preventing a nuclear attack on the UK would be vast. The fact that this type of calculation depends on a speculative counterfactual probably makes this approach unsuitable for national accounts purposes. There is a large literature on the costs of conflict, measured relative to a peaceful counterfactual. It is far more difficult to specify the benefits of peace relative to a counterfactual conflict. This is because there are so many alternative forms the conflict could take, the different types of war the UK might have fought had it not been for the provision of defence. One unpublished study estimated what the cost of a representative conflict would be to the United Kingdom and treated avoiding this outcome as the benefit of defence.

A RAND Europe Study for Dstl²² explored the potential of five methodologies to measure the contribution of defence to UK prosperity in terms of its primary benefits (i.e. protection from external threats). These were:

- a defence analogue to Quality Adjusted Life Years (QALYs) used in health, Protection Adjusted Life Years (PALYs). However, the QALY effect of a drug is easier to calculate than the PALY effect of a tank;
- Value of a Statistical Life (VSL) measures, which again require counterfactuals;
- insurance policy techniques, while treating the defence budget as an insurance premium has attractions, it does not provide an output measure;
- logic models and proxy indicators; and
- discrete choice experiments to judge the willingness of the public to pay for defence. This is discussed further below.

²² *Understanding the Ministry of Defence's Wider Contribution to UK Prosperity: Measuring Defence Output*

These all require dedicated research, and in many cases counterfactuals, so constructing a time series would face difficulties.

A survey based discrete choice experiment was carried out in Germany.²³ The results indicated substantial heterogeneity of preference across respondents and different views about the willingness to pay for different policy options such as an increase in troop numbers, establishment of a European army, an improved air defence system and a reintroduction of compulsory military service. Just as the wellbeing of the population may be measured by a survey, the value of defence could be measured by survey, though it is not clear that either could contribute to the measurement of GDP, because of their subjective nature.

8. Input measures

Output is usually seen as being the product of inputs of factors of production such as labour, capital and intermediate inputs. It may seem that inputs, like numbers employed in different roles and pieces of equipment like tanks, are easier to measure and there is scope to improve the estimates. But there remain formidable difficulties in the case of defence. These are both technical, which arise in the calculation of the MOD balance sheet, like determining rates of depreciation of single use military equipment, where there is no second-hand market, and conceptual: the contribution of defence inputs to the ultimate output security. Human capital, a function of training and experience, is central to military productivity, and measurement while difficult, is perhaps possible. The difficulties include the fact that a military employment contract is quite different from a civilian one.

A large part of expenditure is on the purchase of equipment. This can be regarded as investment in capital stock and the relevant input is the services provided by that capital stock, investment plus depreciated capital stock of the previous period. There appear to be significant differences between the treatment of capital stock in the MOD balance sheet and in the national accounts.

Many issues arise in the measurement of capital stock. There can be a long lag between the expenditure on constructing the equipment and that equipment contributing to capability.

²³ The Value of National Defense: Assessing Public Preferences for Defense Policy Options
Salmai Qari, Tobias Börger, Tim Lohse, Jürgen Meyerhoff, CESifo Working Paper No. 10872. December 2023.

Development costs and other intangible assets are amortised on a straight-line basis over the planned life of the asset.²⁴ Much military equipment is very old and actual life can be greater than planned life for various reasons, such as problems in procuring replacement. On the other hand, losses due to accident or combat, can be large and some decision would be needed to be made about their treatment in national accounts or productivity measures. For instance, the cost of the loss on 17 November 2021 of F-35B ZM152, from the Aircraft Carrier Queen Elizabeth was included in the accounts at £84.398mn.²⁵ Setting asset values for the large land holdings of the MOD is difficult. Measuring capital is hard, measuring the services provided by the asset is even harder, raising all the issues associated with capabilities. It may also be difficult to ensure that services provided by capital owned outright is comparable with services, such as air to air refuelling, provided under a Private Finance Initiative. MOD contracts and relationships are complicated. The 2023 MOD accounts were the first for 13 years to be approved without qualification by the Comptroller and Auditor General. Previously the accounts had been qualified on account of a potentially material omission of assets and associated lease liabilities.²⁶

9. Disaggregated data

To construct good input measures requires more granular data. In a number of my publications I have attempted to conduct econometric analysis of components of the defence budget. This requires consistent historical time series. I have found it difficult to construct these because of changes in definition or presentation of the components. However, what is available may be adequate to improve measurement of inputs.

The MOD reports total expenditure and component shares of expenditure to NATO on a calendar rather than fiscal year basis. The components are under the headings: personnel, equipment, infrastructure, and other. NATO definitions do not match UK definitions so the defence expenditure series do not match. The UK has changed what it reports under NATO definitions and NATO definitions have changed, for instance excluding paramilitary like Gendarmes in 2008. Currently MOD report data back to 2014, but earlier data are available.

²⁴ Annual Report and Accounts 2022-23, p149.

²⁵ Annual Report and Accounts 2022-23, p120. See also Defence Safety Authority, "Service inquiry, Loss of F-35B Lightning ZM152 (BK-18) of 617 RAF, embarked on HMS Queen Elizabeth, 17 November 2021." Crown Copyright 2023.

²⁶ Annual Report and Accounts 2022-23, p126.

In the spreadsheet on Departmental Resources 2023, “Worksheet 3: Defence Expenditure by Commodity Block” there is a fairly detailed breakdown for 2011/12 to 2022/23. Although the data are broadly comparable there are some issues. Other Capital spending increased from £3870m to £8958m from 2021/22 to 2022/23 as a result of increased capital investment following outcomes of the Integrated Review and a £2.3 billion increase associated with transition to lease accounting. The latter was an accounting adjustment and not cash related. Between 20015/16 and 16/17 a large part of R&D was moved from the Cash Resource DEL to the Capital DEL. There are also large variations in Annually Managed Expenditure AME because of changes in Provisions, which can go from large positive numbers to large negative numbers.

Constructing series that are consistent with later definitions is more difficult before 2011/12. Until 2013 there was an annual publication UK Defence Statistics. Although there is some overlap, the categories used in that are somewhat different from those used subsequently. In addition, between 2008/9 and 2009/10 there was a break in the series associated with the introduction of International Financial Reporting Standards, IFRS. Some components series can be run further back. Worksheets 4 and 5 extend equipment and R&D expenditure back to 2003/4 and Worksheet 7, Costs of Operations, to 2004/5. From 2000/01 there was a transition in Government accounting from cash-based accounting to the accruals-based Resource Accounting and Budgeting (RAB) which was implemented in two stages. This caused a major break in many series. To provide a series for total defence spending which is more comparable to the pre-RAB figure Worksheet 1 of Departmental Resources uses Cash/Net Cash requirement.

As noted above the MoD calculation of price indices for index linked contracts may contribute to the measurement of inputs, but faces the difficulties that the calculations have quite different purposes. In recent years the MOD JEDHub Annual Economic Report has provided more data on the contribution of the defence sector to the UK economy.

10. Concluding comments

This paper has reviewed the academic approaches to defence output, and the practical issues with measurement, including the availability of robust military price data, quantifying threat, the impact of alliances, and comparability over time. It has tried to bring out differences in concepts between defence and national accounts, including the various types of inflation and productivity relevant to defence. Specifying the benefits of defence tends to involve

speculative counterfactuals of what might happen without the defence expenditure. There are also problems with input measurement, including the treatment of capital stock depreciation and setting asset values for large land holdings.

The 2008 Scoping Paper contained a serious investigation of a wide range of suggested ways to improve the measurement of output, inputs, and productivity for defence. It raised a range of objections to all the suggested methods. As a result, the input equals output method was retained. It may remain the case that there is no practical way to overcome these objections and provide a measure that meets the Atkinson principles for the measurement of government output and productivity. These include that corroborative evidence should be sought on government productivity, as part of a process of “triangulation” and explicit reference should be made to the margins of error surrounding national accounts estimates.

There is a problem of comparability because the nature of the activities, capabilities and objectives of defence change over time for good reasons as threats, technology and strategy change. There is a danger that the performance measures chosen for output may provide misleading signals or perverse incentives. Suppose certain activities or capabilities are chosen as indicators of output and the threat, technology or strategy changes, as it regularly does. When these activities or capabilities are discontinued to reorient to the new context, measured output will have fallen while the military are fully occupied doing different, perhaps more relevant, things.

Finally, the performance measures that are most important for defence would include elements like: success in operations; maintaining readiness; and effective procurement of equipment that was not late, over budget and that failed to meet requirements. These are difficult to convert into indicators that would match national income accounting criteria.

Appendix.

Atkinson Principles for the Measurement of Government Output and Productivity

A: Government non-market output should, as far as possible, be measured following a procedure parallel to that adopted for market output.

B: Output should in principle be measured in a way that is adjusted for quality.

C: Account should be taken of the complementarity between public and private output.

D: Formal criteria should be established for extending direct output measures to further services.

E: Measures should cover the whole of the UK.

F: Measures of inputs should be as comprehensive as possible, and should include capital services.

G: Criteria should be established for pay and price deflators applied to input spending series.

H: Independent corroborative evidence should be sought on government productivity, as part of a process of “triangulation”.

I: Explicit reference should be made to the margins of error surrounding national accounts estimates.