Appendix 8E. International comparisons of military expenditures: issues and challenges of using purchasing power parities

MICHAEL WARD

I. Introduction

Analysts encounter significant problems when trying to compare one country's expenditure on a group of products, expressed in its own currency, with another country's outlays on the same products valued in a different currency. In the case of military outlays, the problem of comparisons is compounded by the difficulty of identifying the total scope and coverage of official expenditure on 'defence' or 'the military' and obtaining meaningful values for the various items that this comprises.

This section continues with a discussion of why international comparisons are useful and necessary, how they have been undertaken and the specialized nature of military expenditure.[1] Section II looks at how economically relevant international comparisons can be made and compares and assesses the use of exchange rates and purchasing power parities (PPPs) against a historical and theoretical background. In section III the conceptual and empirical issues that need to be resolved to estimate meaningful PPPs are outlined. Sections IV and V look specifically at the problem of comparing military expenditures and relative military burdens in this context and considers the applicability of different comparison methods. Some of the complex practical problems of identifying national military outlays and their composition, and how their absolute size and share of gross domestic product (GDP)—and hence relative absorption of each country's resources—should be calculated, are discussed. Section VI presents the conclusions.

There are a number of reasons why governments and supranational agencies want to make international comparisons. For cross-country poverty analysis and income-related studies, where relative economic well-being, policy priorities and appropriate resource allocations are considered, the reasons are mostly self-evident. In the more political arena of national security and military spending, where economic issues have to be weighed against security concerns, the matter is less clear. The nature of the comparison depends on whether the main interest is in an individual country's capacity to produce military goods, in its total military strength or in how much its government spends on the military each year. Defence analysts want to understand whether comparative assessments of so-called military potential are aligned with a

[1] See also Kravis, I. B. and Lipsey, R. E., *Toward an Explanation of National Price Levels*, Princeton Studies in International Finance no. 52 (Princeton University, International Finance Section: Princeton, N.J., Nov. 1983); Kravis, I. B., Heston, A. W. and Summers, R., *International Comparisons of Real Product and Purchasing Power* (Johns Hopkins University Press: Baltimore, Md., 1978); Kravis, I. B., Heston, A. and Summers, R., *World Product and Income: International Comparisons of Real Gross Domestic Products* (Johns Hopkins University Press: Baltimore, Md., 1982); and Kravis, I. B., Kenessey, Z. and Heston, A. W., *A System of International Comparisons of Gross Product and Purchasing Power* (Johns Hopkins University Press: Baltimore, Md., 1975).

country's capacity to wage war. Others are concerned with how military spending impacts on the economy and the burden it places more broadly on living standards.

Historically, different approaches have been taken to measure comparative military strength. The most traditional has been straight number counts: the total personnel under arms, the number of tanks and guns of a certain type, the number of aircraft in the air force, and so on. Defence analysts have used these to undertake assessments of the relative merits of the armed forces and their equipment. A second approach is to take the value of military expenditure in local currencies and to convert these into US dollar terms using the published exchange rate (defined as the average principal exchange rate relating to the period in question). Exchange rates are subject, however, to the various demand and supply forces, including speculative capital flows, that affect the value of currencies. A third way is to apply purchasing power parities. These effectively transform values in local currencies onto a common price basis using a standard international unit of account. PPP conversions generally work well with most items of expenditure comprising the GDP. International comparisons of, for example, food expenditure by households (often taken as a proxy for actual consumption) can be handled in an undistorted fashion. However, the identification of relevant PPPs and their use to convert more complex categories such as military expenditure into a common international price basis pose both practical and conceptual difficulties.

The specialized nature of military expenditure

Military spending is often unidentified or deliberately obfuscated by governments.[2] Even the overall budget allocated to defence may be a closely guarded state secret. Some outlays on the armed forces may be posted 'off-budget' and some may be secreted away in spending on civil militia, part-time territorial forces, coastguard or customs services and counter-terrorism activities. Some spending in areas such as education, vocational training and medical research may also hide spending on the military. The procedures and practices to account for military expenses vary from country to country. The activities of non-government organizations that sponsor armed forces and guerrilla groups can lead to underestimation of the total extent of a country's military spending.

The focus of this appendix is on comparisons of military expenditure. It describes a more appropriate basis for analysing national expenditures across countries than that afforded by conventional exchange rate comparisons. It advocates the use of purchasing power parities to compare real outlays and argues, implicitly, for the more systematic definition and classification of military spending to ensure that the scope and methodology of comparison are consistent. The objective of the conversion into comparative dollar values is to enable comparison of money spent. Such data can be used to measure the priority that a government attaches to the military or to assess the domestic resources that are devoted to supporting the military. However, since they are input measures, these data carry little information about the output, such as military strength or the capacity to wage war, and cannot be used to assess these outputs, not least because of the wide variation between countries in the input–output relationship for military spending and the intangible nature of critical inputs such as training.

[2] On the transparency and availability of military expenditure data see chapters 6 and 7 in this volume.

Military expenditure is unique among the expenditures reported in GDP or gross national product (GNP) since it does not directly and immediately contribute to any equivalent increase in the economic welfare of households. The reported value of any military outlay is determined by prices that generally reflect the cost incurred directly by government or reimbursed to suppliers to provide what the country's leaders deem to be in the interests of national security and necessary to protect its citizens and their collective assets. Unlike most other items of expenditure, the prices underlying these outlays are not assessed by the market. Costs tend to be determined by recognized institutional processes (such as contract tendering and equipment procurement rules) and historical precedent (the military list). Customarily, an agreed rate of return to the supplier will be added.

II. Features of international comparisons

Two basic problems have to be overcome before reliable estimates of national military expenditures can be produced in internationally comparable terms. The first—a methodological problem that applies to all international comparisons of expenditures—concerns the criteria for choosing an appropriate statistical procedure for converting any reported values expressed in national currencies into some uniform international standard unit of account. The second problem concerns the difficulties surrounding the identification of the scope of military expenditure. This is a question not only of quantities but also of how to determine the prices embodied in officially reported military values, especially given the absence of market prices and equivalent market comparators for many specific and often unique military goods and services.

Methods of international comparison involve: (*a*) comparing actual expenditures and (*b*) comparing the respective shares of such expenditures to some national benchmark, usually GDP or GNP.[3] The former is a pure measure of equivalent money values. The latter is an apparently value-neutral indicator of the national resources that have been devoted to providing a defined outlay. While this may be interesting for internal analysis, it is invalid for international comparisons of relative resource use.

The use of exchange rates and their limitations

Historically, comparisons between countries of goods and services valued in monetary terms have been performed using quoted exchange rates. Outlays, recorded in national currencies, are converted into a commonly recognized international denominator. The US dollar has usually been used as this benchmark.

Exchange rates—specifically, a country's principal official exchange rate—have been adopted because of their convenience. Timely data are always available and the process is clear. The daily behaviour of a country's 'principal exchange rate' is on record in most countries over an extended time period. Foreign exchange data are nominally transparent, effectively independent because they are largely the outcome of financial market forces and, in a sense, neutral. For the purposes of international

[3] GNP—now more correctly referred to as gross national income (GNI)—differs from GDP by adding incomes earned abroad by residents and deducting incomes earned by non-residents in the domestic economy.

comparisons of real economic phenomena, however, exchange rates are not appropriate because they are volatile.

History, particularly in the second half of the 20th century, has shown that exchange rates are unreliable conversion factors for standardizing currency-based variables to a consistent numeraire. Following World War II exchange rates and absolute and relative prices were often in disarray because of resource shortages and different policy stances. New ideas about the respective roles of the public and private sectors in economic affairs had a significant impact on prices and exchange rates. Major currency devaluations (against the US dollar) and other currency depreciations took place among many of the former combatants. However, international trade surged and domestic economies boomed in the period of post-war reconstruction and renewal. Despite a worldwide 'dollar shortage', post-war optimism ran reasonably high in European countries. Most of these countries depended heavily on the machinery produced by a US economy that was relatively less disrupted by the war and each sought expansion through trade- and export-led growth. Exchange rates were not exclusively determined, at least in the short term, by economic events. They became a key policy variable in balancing desires to protect the domestic economy against the need to secure an international competitive advantage. The International Monetary Fund (IMF) tried to establish a regime of fixed exchange rates but throughout the 1950s and 1960s there were regularly devaluations of leading currencies, including the pound sterling, the Italian lira and the French franc. By the early 1970s it was clear that the system of fixed exchange rates had been seriously undermined and could not be continued. In 1974, following the first oil crisis, the IMF formally abandoned the system.

Nevertheless, some countries still pegged their currencies to the dollar, franc or sterling. Sometimes this was for reasons of historical ties and institutional relationships, particularly in the banking network. Some single-currency pegs were later abandoned as the major currencies themselves moved sharply in value, especially when the US dollar began to appreciate strongly in the early 1980s. Instead, a number of countries set the value of their currencies according to a basket of the currencies of their main trading partners weighted to reflect the respective value of their trade. This was a strategy pursued by many developing countries then facing rampant inflation. Capital flight, an outcome of political uncertainty and weaknesses in foreign exchange markets, became more common. Runs on currencies, reflecting speculative and precautionary motives, emerged as a serious problem. In the 1980s the situation gave rise to significant depreciations against the dollar of the currencies of many developing countries. Whenever this happened, dollar-denominated GNP per capita measures published by, for example, the United Nations (UN), the World Bank and the IMF created a 'new' perception about the relative level of development and ranking of countries.

The relevance of exchange rate conversions is now argued mainly on the ground that they depict long-term economic trends. However, economists have customarily attributed more importance to economic 'fundamentals' and international factors in the determination of both the level and movement of a country's exchange rate. Exchange rates do not reflect, consistently and coherently, the structural price differentials that lie at the core of every internationally based value comparison. Even for the 'baseline' country in the denominator, reported exchange rates may fluctuate, as evidenced by the significant fluctuations in the value of the US dollar against most other major currencies from the early 1980s to the present. These fluctuations

represent a major drawback to the usefulness of exchange rates in converting local currencies into a uniform benchmark. The World Bank has devised, for its own operational guidelines, a variant of the standard exchange rate approach to arrive at its published GNP per capita 'Atlas' measures denominated in US dollars.[4] These numbers are widely used but take no account of national price levels, serving only to dampen down observed volatility in exchange rates over the three-year period up to the year in question.

The nature of foreign exchange regimes has changed considerably over the past three decades. In 1975, 87 per cent of all developing countries operated some type of pegged exchange rate. By 1996 this had fallen to well below 50 per cent.[5] Between 1975 and 1996, the exchange rate of various national currencies fell by at least 25 per cent within a year in 116 separate cases. Nearly half of these were countries operating flexible exchange rate regimes.[6] Several countries give only formal recognition to the official principal exchange rate. The IMF regularly notes in its monthly *International Financial Statistics* that some countries operate 'secondary' and 'tertiary' rates that are applicable to certain types of foreign transaction, such as government imports or tourism exchange. Since the behaviour of exchange rates is only loosely related to theory and internal economic conditions, there is little rationale for using them as a basis for making real economic comparisons between countries or over time. For the most part, exchange rate comparisons are appropriate only for international portfolio assessments, bilateral transfers, current financial transactions and international debt settlement. The exchange rate thus determines the real domestic opportunity costs that have to be faced by a given country when meeting any external financial obligation.

Purchasing power parities

Purchasing power parities are the 'shadow' conversion rates that indicate how much, hypothetically, it would cost in a particular country to acquire exactly the same goods and services as $1 buys in the United States (or in whichever country or group of countries is chosen as the baseline reference). Unlike exchange rates, comparisons based on purchasing power parity have the virtue of being base-country invariant. They are also transitive; that is, a country will stand in the same relationship to another country whether a comparison is drawn directly between them or indirectly via a third country. The core principle underlying PPPs is simple: national average expenditures are divided by national average prices to obtain estimates of quantities directly and the quantities in each country are then revalued at uniform average international prices. The practical application of the methodology, however, poses significant problems. An enormous volume of micro price and quantity data pertaining to all categories of private and public goods and services must be collected to perform the requisite calculations and an appropriate aggregation formula has to be chosen.[7]

[4] World Bank, *World Bank Atlas*, 36th edn (World Bank: Washington, DC, 2004).

[5] Caramazza, F. and Aziz, J., *Fixed or Flexible?: Getting the Exchange Rate Right in the 1990s*, Economic Issues no. 13 (International Monetary Fund: Washington, DC, Apr. 1998), URL <http://www.imf.org/external/pubs/ft/issues13/>, p. 2.

[6] Caramazza and Aziz (note 5), p. 5.

[7] Ward, M., *Purchasing Power Parities and Real Expenditures in the OECD* (Organisation for Economic Co-operation and Development: Paris, 1985).

Purchasing power parities and exchange rates

The calculation of PPPs opened up an entirely new way of looking at comparative economic size and national well-being. When introduced, it challenged long-held perceptions about country rankings in economic development—based on exchange rates converted to the US dollar. Partly for this reason, the approach of using PPPs still encounters considerable opposition, even in institutions with an international mandate.

International comparability remains a key objective of the UN statistical system. Clear concepts, conventions and standards of data compilation such as the System of National Accounts (SNA) and classification of goods and services have been laid down.[8] PPPs form part of these standards and, compared with exchange rates, they reflect how prices impact on reported values through differences in price levels within and between countries.

To understand which countries, overall, may be deemed 'cheap' and which are 'expensive', it is only necessary to divide the respective PPP for GDP by the corresponding exchange rate. For example, if the PPP between the United Kingdom and the USA is £1 = $2.15 and the corresponding official exchange rate is £1 = $1.85, then the price level (i.e., 2.15/1.85 = 1.16) indicates that, in general, prices in the UK for equivalent goods are 16 per cent higher than in the USA.

When calculated at a level below that of GDP, PPP ratios can also be considered as measures of the relative price levels of differently defined groups of commodities, such as food or military goods and services. PPPs for different groups of commodities vary and are often quite different from the overall PPP for GDP, another important distinction from the generality of the single exchange rate. The further a specific PPP ratio for a defined spending category is from 1, the more the price level will deviate from the international average for GDP or the relevant subset of GDP expenditures.

PPPs relate to differences in price levels at a particular time (a given state) while a price index measures changes in prices over time (flows).[9] PPPs only provide a measure of comparative real values applicable to a certain time period, such as a year. As such, they represent the spatial, cross-country dimension of price differences at a sector and expenditure category level. In the case of PPPs the economic 'distance' between any two observations, such as between the USA and Kenya, is not fixed, whereas in a conventional price index changes are measured sequentially with reference to fixed reporting dates and ordered chronological intervals. Because disparate and disjoint magnitudes are associated with economic size in cross-section income comparisons since different baskets of goods and services are involved, this factor can give rise to potentially greater problems of distortion in PPP price level estimates than in conventional time series price index measurements. The size of the economic gap helps explain why the GDP of low-income countries rises substantially when PPPs are used rather than exchange rates.

[8] United Nations, Department of Economic and Social Development, *Handbook of the International Comparison Programme* (United Nations: New York, N.Y., 1992), URL <http://siteresources.worldbank.org/ICPINT/Resources/icphbeng.pdf>. The SNA is an international statistical standard for the compilation of national accounts and the measurement of the market economy. The current SNA 1993 will be replaced by the revised SNA 2006. For more information see the website of the United Nations Statistical Division, URL <http://millenniumindicators.un.org/unsd/nationalaccount/method.htm>.

[9] Ahmad, S., 'Harmonization of CPI and PPP: problems and prospects', Eurostat, *Improving the Quality of Price Indices: International Seminar, Florence, December 18–20, 1995* (Office for Official Publications of the European Communities: Luxembourg, 1996), pp. 449–54.

Purchasing power parities and national incomes

Although based on use rather than output, PPP-converted GDP estimates provide an aggregate measure of current production and of the real size of various economies. Evidence from past enquiries shows that, based on international price levels, the real GDPs of poorer countries are much higher than those officially reported in exchange rate terms. Sometimes their economies turn out to be three or four times larger.[10] Because these countries generally have very low price levels in respect of most major categories of expenditure, when an 'average' international price is applied to revalue their goods and services, GDP values are significantly uplifted. The implication is that countries like China and India, with low price levels, in reality manage much larger economies than would be apparent from a current official exchange rate comparison. It also means that important sectors and components of spending in these countries—such as government and, specifically, the military—may be considerably bigger in real PPP terms than is indicated by exchange rates. Significantly, disparities across countries in GDP and constituent spending categories also narrow with the use of international prices.

The difference consequently impacts on measures of 'real income per head'. Even allowing for size of population and differing age and sex components, the inhabitants of some countries which were previously understood to be among the very poorest (including China) are now considered on a PPP basis to be not only much better off but also significantly wealthier than other countries in real terms.[11] These findings are consistent with what is known about relative differences in growth rates over extended time periods and with what has been observed in terms of achievements in social progress independently measured by value-free social indicators.

III. The use of purchasing power parities

The calculation of purchasing power parities

Purchasing power parities are calculated from the combination of binary price ratios between countries of closely similar goods or services. Because weights do not exist for attaching importance to any item, expenditures on similar items are aggregated to form a 'basic heading'. This comprises a common group of like types of goods and services that form a well-defined basket of expenditures making up GDP. The combination of price ratios is then related to the expenditures on these basic headings. The price ratios reflect the corresponding average exchange ratios between countries for that collection of items. As a group, they can be weighted by their respective values in GDP.

Successively, calculations can be made of (*a*) initial item price ratios, (*b*) ratios for generic groups of goods and services, (*c*) ratios for basic expenditure headings and then (*d*) ratios for aggregate consumption categories—such as 'the total final consumption of households', 'the collective consumption of government' and 'gross fixed capital formation'. Progressively, according to all the separate categories of

[10] E.g., the World Bank measures the GNI of China to have been $1416.8 billion in 2003 in exchange rate terms (using its Atlas method) and $6410 billion in PPP terms. World Bank, *World Development Indicators 2005* (World Bank: Washington, DC, 2005), URL <http://devdata.worldbank.org/wdi2005/>, tables 1.1 and 5.7.

[11] World Bank (note 10).

GDP, a structure of cross-country price relationships representing the real rate of price exchange between all types of goods and services of a specific description can be built up. Each spending category is weighted according to its respective importance in a country's GDP, as measured by the respective expenditure outlays. Ultimately, the overall PPP for the whole economy, namely GDP, can be generated in this way. The choice of the formula depends on the primary objectives of the comparison.[12]

The above refers to comparisons in binary form while in practice the calculations are based on a 'multilateralization' of all possible direct and indirect binary ratios between all countries engaged in the comparison for every item in question.

Comparability and 'representativity'

The above methodology deviates from the original PPP notion of pricing an exactly identical basket of items in each country. The relevant 'basket' relates to each respective country's GDP and to the revealed utility preferences and unique structure of spending of final users. This is because comparing a basket containing the same items and pricing it in both Sweden and the UK poses problems. Such a basket would not be characteristic of normal expenditures, as reflected in the actual outlays of either country's GDP. Clearly, going beyond high-income Europe it would be difficult and unrepresentative to price the same bundle of consumption items in a country like China.

This is not just a question of coverage and the necessity for comprehensive pricing. At the core is the more basic issue of the economic relationship between prices and quantities. Prices and the corresponding quantities purchased are normally inversely related, whether it is a question of movements over time or differences in levels. Where a particular item is popular and people thus buy greater quantities of it, the price will tend to be relatively cheaper than other close substitutes. Comparable packets of frozen cod fingers, for example, may be found in both China and Sweden, but in China they will be relatively more expensive because they are neither popular nor bought in large quantities and they are not commonly found in many shops.

This problem also extends to the actual unit size of the quantities bought. It is clear that prices must be related to identically comparable quantities but this cannot be estimated in simple average unit value terms. The price per kilogram of the same quality of rice is quite different depending on whether it is customarily sold in 10-, 25- or 40-kg bags, as is common in Asia, or in 1- or 2-kg packets, as is generally the case in Europe and the USA. Since it is necessary when calculating PPPs to identify items

[12] If, as in the case of military spending, it is important to show respective shares of GDP, then the Geary–Khamis method should be applied at the aggregation level above the basic headings. If it is only necessary to compare internationally what is spent on a specific defence 'item', such as military personnel, then what is known as an EKS aggregation formula can be used. The EKS formula is a multilateral method developed by Ödön Éltető, Pál Köves and Bohdan Szulc that computes the *n*th root of the product of all possible Fisher price indexes between *n* countries. It has been used at the detailed heading level to obtain heading parities and also at the GDP level. Refinements to the EKS method, beginning with the decision to give greater weight to direct binary comparisons and to balanced comparisons of characteristic products in each pair of countries, have been progressively introduced by Eurostat and the Organisation for Economic Co-operation and Development. Éltető, Ö. and Köves, P., 'Egy nemzetközi összehasonlításoknál fellépő indexszámítási problémáról' [On an index number computation problem in international comparison], *Statisztikai Szemle*, vol. 42, no. 5 (1964), pp. 507–18; and Szulc, B. J., 'Price indices below the basic aggregation level', eds R. Turvey et al., *Consumer Price Indices: An ILO Manual* (International Labour Office: Geneva, 1989).

that are representative of people's expenditure choices, there is clearly a conflict between inter-country comparability and what is referred to as national 'characteristicity' and 'representativity' of the items that the statisticians decide to price.

It is essential that exactly the same quantities and qualities are compared when determining each country's respective prices and that derived unit values for specified quantities are not affected by bulk purchases and special packaging. It is also important that there are no obvious or intrinsic quality differences between the various items in the comparison. The ultimate objective is to define the average annual unit price (or value) underlying the different bundles of goods and services expenditures that make up GDP.

IV. International comparisons of military expenditure

This section considers how PPPs can be used for the cross-country comparison of military expenditures. It does not, and cannot, consider comparisons of total military strength or the capacity to wage war. The most that a consideration of military expenditure can do is compare the priorities that governments attach to the military or the domestic resources that are absorbed by the militaries of different countries.

Analysis of military spending

Common to all countries and corresponding to the standard distinction observed in national accounts for other official expenditures, military outlays comprise two main categories: (*a*) direct expenditure on the armed forces and their support personnel, plus expenditure on related consumable items including repairs and maintenance; and (*b*) spending on equipment and buildings. As indicated in these descriptions, each of these categories can be broken down into two component sub-categories. By convention, it is standard national accounting practice to regard all military construction and hardware spending, irrespective of its nature and durability, as current expenditure rather than 'investment' (or, more accurately, gross fixed capital formation). Outlays on buildings that can be used for civil purposes, like military hospitals and schools for the children of military personnel, may be treated differently in the national accounts, although the relevant expenditure will still appear under 'defence' in the annual budget accounts. Strictly speaking, because war is primarily concerned with conquering an opponent and destroying its capacity to wage war, all military equipment—such as guns, tanks, aircraft, ships and so on—should be considered 'expendable'. In a time of war, construction works like blockhouses, gun emplacements, military airfields and dockyards are of strategic significance and subject to destruction. Few military structures or equipment contribute economically to any direct increase in a country's productive potential. Although it could be argued that outlays on military equipment and structures help protect a country, and thus its productive potential, no obvious alternative civil use exists for these items.

Military expenditure of a current nature refers, on the one hand, to payments to armed personnel and support staff and, on the other, to the routine costs of maintaining the military. The first sub-category refers to personal emoluments such as salaries, actual and imputed pensions, benefits, and all other income in kind paid to the regular armed forces and their support staff. Conscripted manpower (usually paid much less) should also be included. Subsidized private accommodation and family

housing and any other transfers of a directly personal nature made to the armed forces should also be added. The second sub-category covers the recurring expenses of housing, feeding and clothing the armed forces on the bases where they are billeted, plus military administration and communications, training and supplies. It includes all forms of regular repair and maintenance and the supply of utility services like electricity and water. Together these two types of outlay constitute the bulk of what the SNA 1993 refers to as 'current government expenditure on defence'.

The second main category of military spending covers what would be regarded elsewhere as capital outlays. In the current SNA these items are 'consumption goods', but there is strong pressure to move them into the capital formation category in the proposed SNA 2006. The category comprises: (*a*) equipment, mostly of a purely military nature, but also computers and vehicles; and (*b*) buildings such as strategic nerve centres, bases and parade grounds, silos, airfields, and storage facilities. The distinction between equipment and buildings is important for calculating relevant PPPs because the respective methodologies are different and because of the need to establish appropriate reference comparisons with other countries. When reference to the military is made under current government expenditure, the latter has to be further divided between the 'collective expenditure of government' and the 'individual household expenditure of government'. Because the role of government varies significantly between countries, particularly in the official provision of social services, consistency dictates that for international comparisons statisticians should distinguish between 'who uses' and 'who pays for' government services. Thus, all government spending on behalf of households gets added to total household consumption while government spending on the military refers only to its public collective responsibility.

In total and in terms of its component structure, military spending is affected by the respective size and strategic location of different countries as well as their allegiance to regional pacts and defence treaties. Such alliances may involve the provision or receipt of military equipment, the engagement of advisers and the use of foreign bases. Traditionally, for example, a portion of the forces of members of the North Atlantic Treaty Organization (NATO) have been stationed abroad. Peacekeeping operations pose a further complication. Inevitably, too, some (imported) equipment used by the home forces will be provided under token cost arrangements or a soft lend-lease procurement agreement that influences the cost of providing defence services.

History and how countries see their military role is also a factor. Countries such as China, which is bordered by 14 other countries, maintain large land forces because they are wary of the potential threats posed by neighbouring countries. The perspective is both domestic and regional. Others, such as the USA, see their role more globally. The USA has only two neighbours but has a long experience of overseas operations as in World War II, the cold war in Europe, and hostilities in Korea, Viet Nam and the Middle East. It has thus tended to focus on maintaining a long-range strategic capability and global mobility that depends heavily on military hardware and maintaining the capacity to strike across vast distances.

Under the present methodology of the World Bank's International Comparison Program (ICP),[13] the government sector, specifically that relating to 'collective' outlays, is taken as a whole, with the consequence that PPPs for military outlays are

[13] The ICP produces internationally comparable price levels, expenditure values and PPP estimates; for more information see the ICP website at URL <http://www.worldbank.org/data/icp/>.

indistinguishable from current government expenditures in general. This creates a potential for distortion, especially if the defence component is large and represents an important part of total government spending, as in the case of both China and the USA. The structure of expenditure on salaries, on the one hand, and goods and services, on the other, is potentially quite different in the case of military outlays compared with the rest of government spending, and even more so if all military spending on equipment and construction works is treated as 'current consumption'. Complications also arise if the average payments made to members of the armed forces, plus the structure of those payments across personnel and ranks, differ significantly from the salary scales of the civilian branches of government.

Quality and efficiency differences

'Quality' differences, broadly defined, are one of the biggest problems encountered when trying to compare military equipment or the skill levels of enlisted personnel in different countries. Given the desire of military leaders to keep one step ahead of potential enemies, quality improvements in military equipment and trained personnel occur continuously. These improvements will be more rapid in rich and technologically advanced countries. The relative professionalism, training and skills of the military forces, however, may be difficult to quantify. When overlaid by such real but unquantifiable factors as nationalism, patriotism and dedication, the question becomes even more problematic. Military effectiveness and readiness cannot be easily proxied by comparisons of real salaries for people of similar age, grade and 'experience', as is the procedure for other government employees. In a different but similar context, the ratio of the price or cost of the standard US rifle, the M-16, cannot be directly compared with the Chinese equivalent, the locally produced Soviet-designed AK-47. Indeed, some military experts argue that, although the AK-47 is simple, it can prove more reliable and robust than its nominally superior but more expensive US counterpart.

Which purchasing power parity rates to use?

There are basically three possible approaches to using PPPs in comparing military outlays between countries. The choice of approach will depend primarily on the availability of data but also on the purpose of the comparison: whether it be the measurement of the resources consumed by the military, opportunity cost of military expenditure or military purchasing power.

1. *Adopt the PPP at the overall GDP level*. This is the intuitive counterpart to applying the official exchange rate covering all goods and services.
2. *Apply a specific PPP for government expenditure.* This may prove more relevant on the grounds that salaries in both government spending as a whole and in the military make up the bulk of the current outlays (two-thirds or more in some cases[14]).
3. *Derive a specific PPP for military spending.* In practice, this can be done by calculating separate PPPs for all the major (and very disparate) elements of military

[14] See, e.g., Omitoogun, W., *Military Expenditure Data in Africa: A Survey of Cameroon, Ethiopia, Ghana, Kenya, Nigeria and Uganda*, SIPRI Research Report no. 17 (Oxford University Press: Oxford, 2003), p. 90.

expenditure, including the use of exchange rates for the purchase of imported military hardware.

The third of these methods is the best for comparing specific expenditures, especially if the expenditures can be decomposed into their major economic categories: personal emoluments, current goods and services (military supplies, transport and communications), purchases of capital-type goods (such as ships and aircraft, both domestic and imported) and defence buildings and construction. Obtaining such detail may not be possible—at least in the same statistical and analytical framework for all countries engaged in the comparison. It would still be necessary to use the GDP- and government-level PPPs to derive share estimates of the real absorption of resources for military purposes.

In the absence of a specific military PPP, the use of a government-level PPP as a proxy for real military outlays is likely to be less satisfactory because the pattern of government collective expenditures in total, except in the case of labour shares, may not match well that of the military. This will be especially the case where annual military spending falls heavily on equipment; to a large extent, this will depend on the strategic perspective of each country. This second method can be useful if the intention is to measure the opportunity cost of military expenditure. It measures what military outlays could have bought if they had been diverted to another area, such as education.

Least satisfactory would be the use of a GDP-level PPP, unless there is some underlying belief that military expenditure represents a national protection 'insurance' cost for preserving national well-being. Nevertheless, in almost all cases this method is still preferable to the use of exchange rates since the difference in real values could be as high as a factor of 4 or even 5.[15] A large difference would undoubtedly occur in a comparison between US and Chinese military spending. The use of the GDP-level PPP would indicate, in real terms, aggregate military outlays in China that are at least three times higher than when expressed in exchange rate terms. This higher figure does not mean that China's military output is that much higher; rather, it means that the real cost to China is higher. As regards the economic burden, the share of national resources devoted to military outlays might not be so different from that calculated using exchange rates.[16]

The way forward with international comparisons seems to be to use relevant PPPs for comparisons of well-defined expenditures and, in the case of military spending, to explore methods for obtaining more detailed information on military outlays in domestic prices. The problem is complicated by the increasing complexity of military procedures, the growing use of weapon systems, and the continually improving quality of military equipment and trained personnel. In some respects, however, this question may be little different from the general one posed by high-technology industry or specialized medical procedures, where similar challenges are faced in comparing their products.

[15] World Bank (note 10).

[16] Kravis, I. B., 'An approximation of the relative real per capita GDP of the Peoples' Republic of China', *Journal of Comparative Economics*, vol. 5, no. 1 (Mar. 1981), pp. 60–78.

Expenditure or output comparisons?

Purchasing power parity calculations are based on the expenditure components of the GDP aggregate. This is primarily because the concern is with the 'equivalent purchasing power' of a country's currency; that is, with how much local 'income' is used in acquiring the goods and services in question. The expenditure method is adopted because it is decomposable, practical and statistically more transparent. Clearly, knowing the share of GDP and national resources that is being allocated to the annual production of military items is also an important concern. The measurement of this phenomenon is conceptually clear in that it requires the calculation of real net output from an estimate of gross output valued in relevant international prices and the deduction of the corresponding intermediate inputs used up by each sector (transacted at a different set of international prices). However, the degree of statistical detail required to perform the calculations is monumental.[17]

The limitations of purchasing power parities for international comparisons of military expenditure

The limitations of PPPs for international comparisons are more practical than conceptual.[18] The problems of using PPPs are both general and specifically related to comparisons of military expenditure as reported in government accounts. They are general because it is not always easy to choose items to be priced for any set of outlays and, particularly for a given basic heading, to price these items at the same types of outlets. The problem is essentially to decide how best to target the process of accurately estimating the average annual unit value underlying the defined group of national expenditures. This means identifying the sources from which people buy most of their goods and services. In the case of military goods, government-owned or -controlled factories set their own prices based on costs while private firms operate within guidelines set by the government. The process of linking relevant prices to such values is far from transparent.

Countries need to resolve a number of sampling and selection issues and agree a common approach. This may be difficult to achieve in the case of the supply and delivery of military goods to foreign governments where transactions are confidential. Furthermore, within any GDP comparison in PPP terms, exchange rates are still used to convert the values of a country's imports and export of goods and services.

Specific problems of defining PPPs for the military concern the difficulty of identifying the items to be priced and then finding the appropriate price for them. There is probably less difficulty in determining the salary of an infantryman with a given number of years of service. A greater problem is identifying the salary of a conscripted soldier and who pays for his services. The real difficulty in pricing, however, relates to the costs of different types of military equipment and construction work and how to define these outlays comparably with other countries. The problem is more

[17] See also the work of the Growth and Development Centre of the University of Groningen under the direction of Professor Angus Maddison. E.g., Ark, B. van and Maddison, A., 'The international comparison of real product and productivity', eds A. Maddison, D. S. Prasada Rao and W. F. Shepherd, *The Asian Economies in the Twentieth Century* (Edward Elgar: Cheltenham, 2002).

[18] For a fuller review of the issues see Castles, I., 'The OECD–EUROSTAT PPP Program: review of practice and procedures', Organisation for Economic Co-operation and Development, Paris, 1997; and Ryten, J., 'Evaluation of the International Comparison Programme', United Nations, Economic and Social Council, Statistical Commission, 16 Nov. 1998, URL <http://unstats.un.org/unsd/methods/icp/>.

complicated the more sophisticated and 'customized' the equipment and facilities. In principle, it might be argued that a MiG-29 or F-16 combat aircraft used in different air forces is the same, but, like many types of modern military hardware, they form part of an overall weapon-delivery system. Even where the equipment is purchased from another country, the true acquisition cost is often concealed. Where identified, the domestic cost of imported military equipment is, by convention, determined using a country's principal exchange rates. However, this is one area where a special exchange rate may be applied to makes the goods in question cheaper for the purchasing country.

While military experts continually weigh up the relative merits of various weapons—the longbow against the crossbow, the Spitfire against the Messerschmitt 109, the MiG-15 against the Sabre F-86—they are not directly comparable Determining a PPP price ratio for any item chosen as a representative product relies on it being possible to price exactly the same item in at least one other country. This is not possible for many items of military hardware because, for strategic reasons, military forces prefer to use their own specific equipment.

There is a greater possibility, both in principle and in practice, that construction work and buildings can be compared. This is because the comparisons can be based on the cost components of common inputs and uniform 'building blocks'. Information about salaries and material costs for standard elements of work such as foundations (footings), skeleton steel frames and retaining walls then forms the basis for such comparisons. In the private sector, where there are many and varied buildings, a detailed bill of quantities can be specified for a standard prototypical building that each country can be asked to price in order to determine, hypothetically, the cost of construction in its own currency.

In current global PPP exercises, military goods and services, military equipment, and defence structures and buildings are not identified separately. In practice, the salaries of military personnel are considered along with the salaries paid to other government employees. The daily running costs of the military and its supplies are also treated like any other government current outlay and are converted as if they represent a mini cross section of all other goods and services used in the economy; that is, they are converted into international values using the PPP at the GDP or government level. The values of imported military goods are converted at the annual exchange rate and this shows up in the country's rest-of-the-world balance. The biggest outstanding problem, however, relates to comparing the costs of domestically produced structures and equipment of a capital nature, such as ships and aircraft. If these are considered in the accounts as government consumption items, then, implicitly, they will be converted into international values using the GDP-level PPP. This is not an exact conversion and it is clearly subject to significant error, although it may not be exposed to any easily discernible bias. This is an area that is currently being researched by the ICP and national accounts experts and recommendations to change existing procedures are expected in the revised SNA 2006.

V. International comparisons of military burden

The choice of currency conversion method influences cross-country comparisons of the economic burden represented by spending on the military. At the level of GDP, PPP converted values show higher levels of national expenditure for lower-income

countries than would be indicated by an exchange rate conversion. These differences are proportionately larger the lower the absolute economic size and average income (GNP per capita) of a country. The same effect will be seen if PPP rates are used for lower component levels of GDP and its sub-aggregates. This means that, if an 'additive' aggregation procedure such as the Geary–Khamis formula—where PPP-adjusted components are summed consistently to give a PPP-adjusted GDP—is used to compile PPPs, the resulting pattern of real outlays in international price terms will diverge from the structure of an economy as understood in its own domestic price terms. These differences are sometimes of considerable strategic importance, as in the case of investment and public administration and, perhaps particularly, military expenditure.

When reviewing spending shares, the use of official exchange rates to similarly convert both GDP and a specific expenditure component will conceal the relative amount of domestic resources absorbed in generating that outlay. Such 'money unit' ratios may appear 'unit free' but they are not 'value free' in an international context. Even within the same country over time, such comparisons may not be relevant to any analysis of real economic change since absolute price levels and the price movements underlying the various expenditure categories may change in different ways. In some cases, price levels may differ also between final users and socio-economic classes.[19]

Comparisons of expenditure sub-components

In calculating PPPs for a country, GDP is divided into approximately 150 well-defined 'basic headings'.[20] In the richer and more complex advanced economies, even more commodity groups may be identified. Below the basic heading level, reliable information about detailed expenditure weights is generally not available. For presentational purposes and to obtain greater reliability, detailed basic heading data are usually condensed into about 50 higher-level expenditure 'groups'.[21] The consolidation of basic headings into larger categories helps to reduce the effects of volatility and errors in individual measurements.

Analysts can determine how much in real terms is allocated to each expenditure sub-component. Separate PPPs related to the underlying international price level of any identified item or group of items, such as the salaries of government personnel or the cost of private motor vehicles, can be calculated. The derivation of PPP-based component shares of GDP provides a realistic insight into the true opportunity costs of that category of spending. Thus, if the ratio of military expenditure to GDP in PPP terms is greater than the same ratio valued in national prices, then it is clear that the country is devoting more of its real resources to that activity than to other types of

[19] Biru, Y., 'The purchasing power of the poor in Zambia', Paper presented at the World Bank Seminar on Prices and Purchasing Power Parities, Washington, DC, Jan. 1999; and Decoster, R., 'Proposal for comparative poverty assessment using purchasing power parities for low income households', Paper prepared for the Development Economics and Data Group, World Bank, Jan. 1999.

[20] E.g., the ICP uses 155 headings. International Comparison Program, 'GDP and the main expenditure aggregates', *ICP 2003–2006 Handbook* (World Bank: Washington, DC, 2004), URL <http://www.worldbank.org/data/icp/>.

[21] E.g., the ICP has 61 groups above the level of its basic headings. International Comparison Program (note 20).

Table 8E.1. Comparison of military burden calculated using different currency conversion methods for select countries, 2000

Figures give military expenditure as a percentage of gross domestic product (GDP) in terms of market exchange rate (MER) and purchasing power parity (PPP).

Country	MER terms[a]	PPP terms[b]	Country	MER terms[a]	PPP terms[b]	Country	MER terms[a]	PPP terms[b]
Africa			China	2.04	3.76	Greece	4.87	4.22
Algeria	3.45	5.77	India	2.35	6.58	Italy	2.09	1.75
Argentina	1.32	1.67	Japan	0.96	0.94	Russia	3.74	3.56
Burundi	5.96	10.69	Mauritius	0.21	0.19	Turkey	5.02	6.12
Ethiopia	9.57	17.81	Pakistan	4.49	8.44	UK	2.48	2.40
Zimbabwe	4.68	4.22	Malaysia	1.70	4.33	*Middle East*		
Americas			Sri Lanka	4.52	13.29	Iran	5.36	6.36
Brazil	1.69	1.84	*Europe*			Israel	8.39	7.48
Canada	1.16	0.84	Armenia	3.56	7.92	Jordan	8.87	18.35
USA	3.07	2.39	Austria	0.83	0.69	Lebanon	5.37	11.38
Asia and Oceania			Bulgaria	2.53	4.51	Syria	5.53	12.84
Australia	1.85	1.64	France	2.58	2.34	Yemen	4.98	16.13

[a] These figures are calculated using the MER for both military expenditure and GDP data, i.e., as if calculated in local currency.

[b] These figures are calculated using government-level PPP rates for the conversion of military expenditure data and GDP-level PPP rates for the conversion of GDP data.

Sources: **PPP rates and exchange rates**: Heston, A., Summers, R. and Aten, B., Penn World Table, Version 6.1, Center for International Comparisons, University of Pennsylvania, Philadelphia, Penn., Oct. 2002, URL <http://pwt.econ.upenn.edu/php_site/pwt_index.php>; **GDP data**: International Monetary Fund, International Financial Statistics Database, URL <http://ifs.apdi.net/imf/>; **Military expenditure data**: Appendix 8A.

spending (see table 8E.1, which presents the military burden of select countries calculated using exchange rates and using purchasing power parities).[22]

Before PPP information became available, when researchers wanted to make international comparisons of detailed expenditure shares they used the price ratios of their own country or measures implicitly based on exchange rates. The same conversion factor used for the country as a whole was thus applied to the lower-level categories of expenditure. However, for categories below the level of GDP, expenditures and prices in each country tend to be determined by very different institutional and market conditions. The use of only one exchange rate in this context to determine the relative importance of these different types of outlays is thus inappropriate.

Government spending

The most important area where the calculation of PPPs has thrown up significant differences in comparative values is government current expenditure and the func-

[22] Summers, R. and Heston, A., 'The Penn World Table (Mark 5): an expanded set of international comparisons 1950–1988', *Quarterly Journal of Economics*, vol. 106, no. 2 (May 1991), pp. 327–68. The Penn World Tables are updated regularly by the Center for International Comparisons at the University of Pennsylvania; URL <http://pwt.econ.upenn.edu/>.

tional allocation of outlays to different departments. This has affected perceptions about the extent of public administration (the bureaucracy), defence (the military) and social services (society building). In these three core areas, direct comparisons using national income shares in national prices are meaningful only in an internal budgeting sense. The degree of involvement of various governments in the provision of, for example, health and education compared with the military varies significantly between countries and so government accounts need to be adjusted from a public 'spending' to a private 'use' basis to provide a more standardized coverage. This can then permit more robust measures of (*a*) the share of government 'proper' in GDP and (*b*) the share of major components of government, like military spending, in total government outlays. This is especially important when considering comparisons between China and the USA because the roles of government and their respective level of involvement in the national economy are clearly very different.

Not surprisingly, it has been found that government's share of GDP tends to be higher in poor countries when PPPs are used because of the dominant role of services and the relatively lower salaries paid in the public sector. In evaluating comparative military outlays, the level of domestic costs significantly affects estimates of the real amount of defence capacity produced each year. It is important to emphasize that PPP calculations are made only with respect to the 'running costs' of the military and not to any comparative evaluation of the total 'stock' of military goods and services; that is, the existing capacity of the military. One important element of the 'stock', however, is the current number of people in the armed forces and this component has to be paid for each year.

In general, the use of PPPs at two different benchmark periods raises other questions of interpretation because the purchasing capacity over the same unit of goods and services changes with changing relative prices during the intervening years and pure time–space consistency in price measurement is difficult to achieve. Such comparisons are also affected by changes in the number and composition of countries involved in the benchmark studies.

VI. Conclusions

This appendix outlines some of the issues involved in trying to compare, in real terms, what is spent each year by countries and, specifically, their governments, on military goods and services. It explains why it is important to use PPP conversion factors in preference to exchange rates to provide real international comparisons when assessing the resource use and the economic burden of military spending. The discussion suggests that, although PPPs can be used for most international comparisons of national expenditures, given the present limited scope and coverage of national data on military outlays and the problems of adjusting these for different quality factors, there are some serious weaknesses in existing PPPs as a basis for comparing military spending. The issue is further complicated by certain lacunae and ambiguities in the conceptual structure of the System of National Accounts relating to the definition of military expenditure. These are being looked at in the revision of the SNA due to be completed in 2006. The drawbacks to PPPs apply particularly to any comparisons involving China or, for that matter, India and any other country not directly involved in the last, 1999/2000 phase of international comparisons where data for missing countries had to be estimated econometrically. This situation is

likely to improve in the foreseeable future when the results of the current round of the International Comparison Program are published in 2006. These will be based on the careful collection and selection of detailed micro data for a wide range of countries, including China and India. Since the ICP's inception in 1975, its country coverage and processes of data compilation have improved considerably.[23]

In the meantime, what should be done to compare annual military outlays? In the specific case of a China–USA comparison where China's exchange rate, although nominally free to float, remains effectively linked to the US dollar within a narrow band of fluctuation, the use of GDP- and government-level PPPs is still preferable to the application of an economically unrepresentative fixed dollar–yuan exchange rate. The fundamental difference is so large that it swamps all other considerations. Better still would be to calculate unique PPPs for each separate component of military spending and apply these to all the relevant expenditure categories. An assessment of the relative resources devoted to the military by the two countries might be better served by a direct binary comparison rather than through a full multilateral PPP exercise. In general, whichever PPP estimates are used, they will probably be less exposed to evident bias than they are subject to significant margins of error. Certainly, comparable estimates of military spending will be less reliable than other similarly deflated components of national expenditure; but this is due as much to the inherent difficulties of identifying military outlays as it is a problem of applying PPPs. Obtaining better estimates of military outlays is the clear priority.

Although PPP conversion is—at least in principle—the preferred method of conversion of military expenditure data to dollar values for purposes of comparisons of resource consumption, it is largely irrelevant for comparisons of military capacity or effectiveness. Assessments of the latter type require more qualitative and detailed analysis, where military expenditure may constitute only one of many variables.

[23] For research in earlier periods see Gilbert, M. and Kravis, I. B., *An International Comparison of National Products and the Purchasing Power of Currencies: A Study of the United States, the United Kingdom, France, Germany and Italy* (Organisation for European Economic Co-operation: Paris, 1954); Gilbert, M. et al., *Comparative National Products and Price Levels: A Study of Western Europe and the United States* (Organisation for European Economic Co-operation: Paris, 1958); and Kenessey, Z., 'International Comparison Program in the 1980's and 1990's', eds D. S. Prasada Rao and J. Salazar-Carrillo, *International Comparisons of Prices, Output and Productivity* (Elsevier Science Publishers: Amsterdam, 1996), pp. 3–29.