



# Lessons from around the world:

## Benchmarking performance in defense

**A first-of-its-kind benchmarking effort compares the productivity and performance of defense ministries across the globe, helping them pinpoint areas of inefficiency and identify the highest-potential opportunities.**

**Scott Gebicke and  
Samuel Magid**

With wars under way in several parts of the globe and many countries' defense budgets suffering drastic cuts, defense ministries are under pressure to do more without spending more. And most defense ministries recognize that they have ample room to improve both the efficiency and effectiveness of their operations. Yet a typical defense ministry has little perspective on what constitutes best practice in defense operations, where its biggest opportunities for saving money or boosting productivity lie, or how it stacks up against its counterparts in other countries in the core areas of defense.

Some would argue that comparing the performance of one defense department to another's is neither achievable nor instructive,

given that countries are in very different political situations and have different priorities and military strategies. Granted, many variables affect the performance of a country's armed forces, and it would be virtually impossible to account for all the complexities and dynamics that come into play. Furthermore, defense ministries make deliberate trade-offs—for example, choosing to pay more for domestically manufactured equipment. That said, defense departments everywhere engage in the same types of operational activities. Our firm belief is that certain aspects of operational performance are indeed comparable across ministries of defense, and that ministries can learn from one another when it comes to delivering more defense output for the same or less input.



In 2008 and 2009, we undertook a first-of-its kind benchmarking effort—one that compares the performance and productivity of defense ministries worldwide. We gathered and analyzed data from 33 countries that account for more than 90 percent of global defense spending, developing a benchmark that we believe yields valuable insights into where and how ministries of defense can become more effective while reducing or maintaining costs. In the simplest terms, the exercise involved analyzing a discrete set of quantitative inputs—namely, publicly available data on the quantity and type of military equipment, number and general classification of personnel, and annual defense budgets disaggregated into key spending categories—and converting them into a set of ratios that measure outputs in three core budget areas of defense: personnel, equipment procurement, and maintenance. Assembling inputs presented a significant research challenge due to wide variability in the quality and quantity of available

data, but defining the inputs was reasonably straightforward; defining and measuring outputs, on the other hand, was a much more complex undertaking (see sidebar, “Our methodology for calculating output,” p. 5).

Our benchmarking results show wide variability across countries in each ratio (Exhibit 1). Once a country has selected a peer group against which to compare itself, it can use these benchmarks to help pinpoint areas of inefficiency and zero in on the highest-potential opportunities.

For the purposes of this benchmarking exercise, we used five straightforward country categories based on types of military strategies: global-force projection (countries with worldwide striking capability), small-force projection (NATO members or countries with a fairly significant presence in international missions), relevant national security threat (countries under attack or threat), emerging regional powers, and non-aligned or neutral countries.

## Exhibit 1 Stacking up

Benchmarking showed wide variations in performance.

Budget area (average % of defense budget)	Key ratios	Range	Average
<b>1 Personnel</b> (45%)	<ul style="list-style-type: none"> <li>• “Tooth to tail” (combat personnel as % of total personnel)</li> <li>• Number of deployed as % of total active troops</li> <li>• Personnel costs per active and other personnel</li> <li>• Personnel costs over military equipment output<sup>1</sup></li> </ul>	<ul style="list-style-type: none"> <li>16–54%</li> <li>1–18%</li> <li>\$800–\$146,000</li> <li>\$2,000–\$218,000</li> </ul>	<ul style="list-style-type: none"> <li>26%</li> <li>5.3%</li> <li>\$44,800</li> <li>\$72,000</li> </ul>
<b>2 Equipment procurement</b> (18%)	<ul style="list-style-type: none"> <li>• Military equipment output<sup>1</sup> over procurement and R&amp;D spending (index)</li> <li>• Procurement spending over active troops</li> </ul>	<ul style="list-style-type: none"> <li>17–330</li> <li>\$1,000–\$536,000</li> </ul>	<ul style="list-style-type: none"> <li>100</li> <li>\$60,000</li> </ul>
<b>3 Maintenance</b> (8%)	<ul style="list-style-type: none"> <li>• Cost of maintenance per unit of military equipment output<sup>1</sup></li> <li>• Cost of maintenance over cost of equipment procurement</li> </ul>	<ul style="list-style-type: none"> <li>\$2,000–\$104,000</li> <li>8.2–446%</li> </ul>	<ul style="list-style-type: none"> <li>\$13,000</li> <li>13%</li> </ul>

<sup>1</sup>One unit of military equipment output is approximately equivalent to one combat-ready unit (eg, a manned and maintained combat vehicle). For more, read “Our methodology for calculating output,” p. 5.

This simplified peer-group categorization was adequate for our initial purposes, but to generate the most useful insights from the benchmarks, a defense ministry must thoughtfully and carefully select a peer group based on its military strategy.

One particularly interesting finding was the variability among countries in the level of joint spending, which ranges from almost 70 percent to 3 percent (Exhibit 2). Not surprisingly, we found that countries that share more functions across the armed services tend to be more efficient. Some countries have recently moved toward increasing their level of joint spending, whether by requiring closer collaboration and coordination among service-specific functions or establishing joint functions. (The article “Big savings from little things: Non-equipment procurement,” p. 34, describes how some countries have centralized procurement of products and services in certain

non-equipment categories; “Supply chain transformation under fire,” p. 50, touches on the United Kingdom’s move from a service specific supply chain to a joint supply chain.)

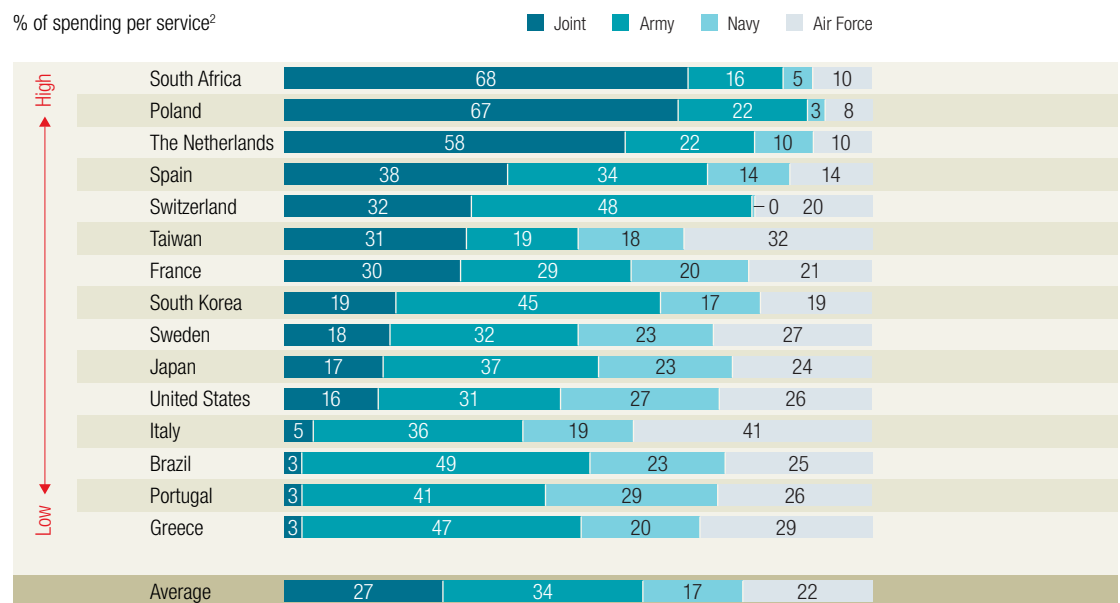
In this article, we highlight some of our findings in each of the three budget areas we benchmarked and offer perspectives on how countries might improve—or have already improved—performance in each area.

**Personnel: Tooth-to-tail and deployment ratios**

From most commanders’ perspectives, the true test of military strength lies in the front line—the “tooth,” in defense industry parlance. The “tail” refers to personnel who perform noncombat functions such as procurement, deep maintenance, accounting, facilities management, or back-office IT. Our benchmarking results show

Exhibit 2  
**Level of joint spending<sup>1</sup>**

Countries that share more functions across the armed forces tend to derive greater efficiencies.



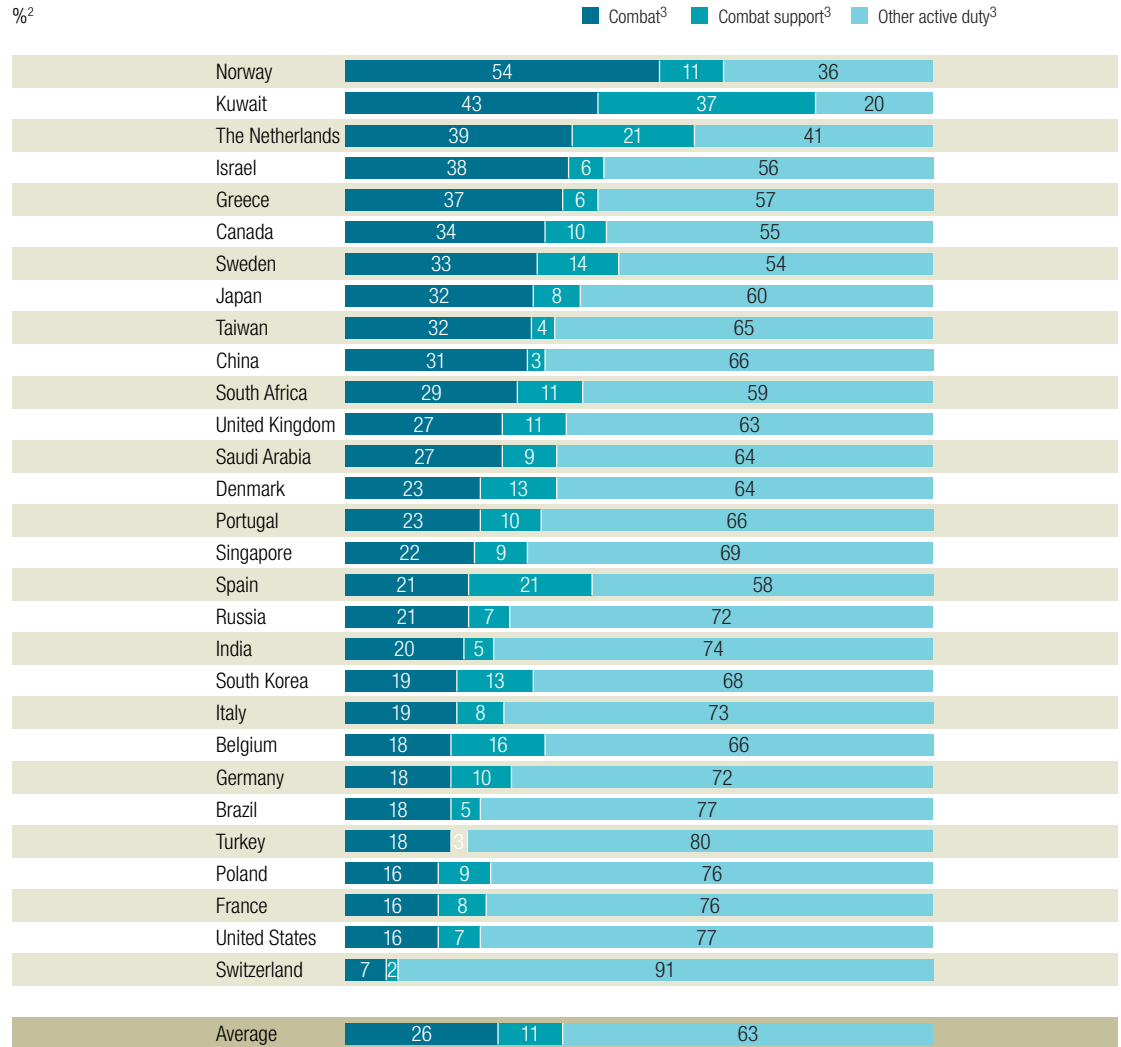
<sup>1</sup>We have removed the benchmark data for Australia as errors in the source data and the methodology as it applied to Australia have been identified.

<sup>2</sup>Figures may not sum to 100%, because of rounding.

Exhibit 3

**‘Tooth to tail’ ratio<sup>1</sup>**

Administrative costs can be reduced without sacrificing fighting power.



<sup>1</sup>We have removed the benchmark data for Australia as errors in the source data and the methodology as it applied to Australia have been identified.

<sup>2</sup>Figures may not sum to 100%, because of rounding.

<sup>3</sup>Combat troops: armor, infantry, reconnaissance, and combat aviation. Combat support: artillery, engineers, and signals. Other active duty: general and administrative functions including HR, IT, procurement, accounting, etc. Reserve personnel not included.

Source: *The Military Balance 2008*, The International Institute of Strategic Studies; McKinsey analysis

stark differences in tooth-to-tail ratios, indicating opportunities to reduce administrative costs in several countries without diminishing fighting power (Exhibit 3).

Some countries are proactively trying to improve their tooth-to-tail ratio. France, for example, is aiming for a dramatic reduction of administrative

personnel through investment in IT systems and outsourcing of certain noncombat operations to the private sector (see “Without taboos: France’s new defense policy,” p. 64).

The defense ministry of a Northern European nation, under pressure to increase military output in the period after the Cold War, set a goal a few

## Our methodology for calculating output

Comparing the performance of one country's armed forces with another's involves both art and science, in part because data on budgets, equipment, and personnel are not always available, reliable, or reported in a comparable way. To develop our benchmarks, our research departments in various countries scoured public data sources and made a number of assumptions to normalize the data. A key part of our analysis was the creation of a new metric for measuring the performance of military equipment. We call our metric "military equipment output," and we used it to calculate some of the key ratios as shown in Exhibit 1 of the article. Military equipment output is a function of four factors: volume, mix of equipment, age of equipment, and overall equipment quality.

**Volume.** To calculate military equipment output, we first gathered data on several countries' active equipment inventory—specifically, how many serviceable units of each type of equipment a country has in each of its armed services (for example, the number of submarines in the navy, the number of main battle tanks in the army). This exercise proved challenging because countries report inventories in many different ways—for example, some include only active equipment while others include equipment for reserves or mothballed equipment.

**Mix.** Then, using the average equipment mix of the United Kingdom and France as our ideal target mix (because both countries have a good balance of army, navy, and air force equipment in all major categories and are sizeable enough but not so large as to skew the data), we assigned a relative value to each type of equipment per armed service—determining, for example, that in the navy an aircraft carrier is

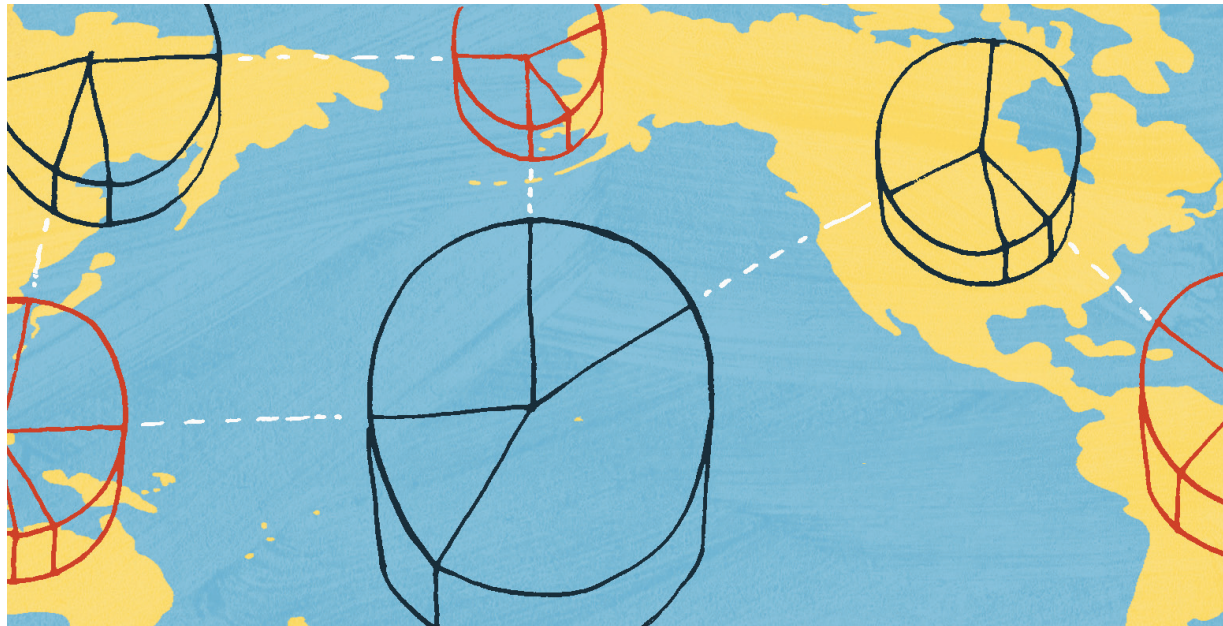
the equivalent of 3.5 submarines or 8 surface combat ships. This allowed us to compare armed services regardless of the composition of their equipment portfolio. We excluded nuclear equipment from the benchmark because it skewed results significantly.

**Age.** Recognizing that there are variations even within the same type of equipment—the F-35 aircraft has significant advantages over older fighter jets like the MiG-19, for example—we also adjusted for age. We determined that a fifth-generation fighter like the F-22 or the F-35, for instance, is equivalent to 3.6 second-generation fighters.

**Quality.** We then took into account a military equipment quality (MEQ) score for each of the armed services in each country, based on rigorous analysis conducted by third-party consultancy Technology Futures. (For more on MEQ, read "From R&D investment to fighting power, 25 years later," p. 70.)

By calculating military equipment output for each of the armed services—the army, the navy, and the air force—we were able to make comparisons across countries. Our benchmark shows, for example, that the US and Russian armies have almost equivalent output levels largely due to the size of the Russian tank fleet, but that the US Navy and Air Force are far superior to their Russian counterparts—a case of American technology trumping the sheer volume of Russia's older platforms and aircraft. The navies of the United Kingdom and France are on par with South Korea's and Japan's, and Israel's air force has about twice the output levels of the air forces of France, Germany, and Brazil.





years ago to increase its tooth-to-tail ratio from 40:60 to 60:40 over three years. It achieved this goal by centralizing formerly duplicative support functions including HR, IT, finance, media and communications, health services, and facilities management. By mapping the functions' activities and resources—what exactly each function did, who did it, and how many people did it in each regiment—and by comparing itself with other public and private-sector organizations, the defense ministry realized that centralization would yield savings of approximately 30 percent per function.

A number of countries have found that one of the hardest parts in a centralization effort is designing the precise division of responsibilities and the interfaces between the centralized service and the various military services. Political and cultural sensitivities come into play as heads of regiments lose responsibility for certain positions and facilities. The need for coordination increases exponentially, particularly because of frequent rotations among military personnel. Individuals

accustomed to tools and processes of their own choosing have to be convinced—and then trained—to use standardized tools and processes.

To ensure the success of a centralization effort, a defense organization must address mind-sets and behaviors. The European defense ministry mentioned earlier held seminars for the top 100 leaders to get their buy-in and to make sure they learned and embraced the new ways of working. To foster collaboration, the ministry also established formal mechanisms; for example, a joint management team, consisting of leaders of each military branch as well as of the centralized functions, participated in an annual prioritization process, ensuring that the most important needs of each branch were well understood and that the centralized service could meet those needs.

Like corporations, defense ministries should seek productivity improvements in administrative functions; in these nonmilitary tasks, productivity growth can and should offset wage growth.

## Exhibit 4

**Deployed forces**

SAMPLE FROM BENCHMARK

Combat forces are under strain in some countries.

	Total active (number of people)	Total deployable <sup>1</sup> (number of people)	Deployed (number of people)	Deployed over total active (%)	Deployed over deployable (%)	Cost per troop deployed (\$ thousands)
United States	1,352,494	N/A	250,000	18.5	N/A	N/A
United Kingdom	185,950	74,750	34,000	18.3	45.5	N/A
The Netherlands	44,636	17,724	3,896	8.7	22.0	68
Finland	10,100	6,000	840	8.3	14.0	216
Sweden	11,574	3,122	950	8.2	30.4	611
France	262,592	42,500	17,485	6.7	41.1	35
Italy	191,152	54,800	11,170	5.8	20.4	N/A
Spain	77,800	39,617	3,344	4.3	8.4	195
Germany	221,185	37,275	8,946	4.0	24.0	172
Greece	135,500	22,182	1,290	1.0	5.8	83

<sup>1</sup>Troops trained and ready to deploy.

Increased productivity in back-office functions can then lead to more favorable deployment rates, as uniformed personnel can be reassigned from support roles to combat roles. A country needs to have many more deployable service members than it might expect to deploy at any one time to account for periods of training and recuperation. In certain countries, combat forces are stretched thin, with deployment rates exceeding 40 percent of potential (Exhibit 4).

These countries have the choice of either reducing deployments—which will essentially mean a loss of fighting power—or shifting a significant number of personnel from administrative roles to combat roles. The latter is clearly the better option.

### Equipment procurement

Countries deliver substantially different levels of military output for the money they spend on equipment procurement. There is a rough correlation between procurement cost per unit of

output and average equipment quality, which naturally raises the possibility of spending large sums in the pursuit of extraordinarily powerful weapons. US Defense Secretary Robert Gates raised this very issue last year when he announced his intent to “pursue greater quantities of systems that represent the ‘75 percent solution’ instead of smaller quantities of ‘99 percent’ or exquisite systems.” The United States is currently at the extreme end of the cost/quality spectrum, delivering very high-quality equipment but at very high cost (Exhibit 5). Once again, variations between countries in the same peer group can be substantial.

In general, countries that make it a point to support their domestic defense industries have higher procurement costs than those that rely on imports. Since this represents a narrowing of the market being considered for purchases, this is not a surprising result. Meanwhile, countries that procure older equipment from the global market tend to have very capable fleets for less money. Of

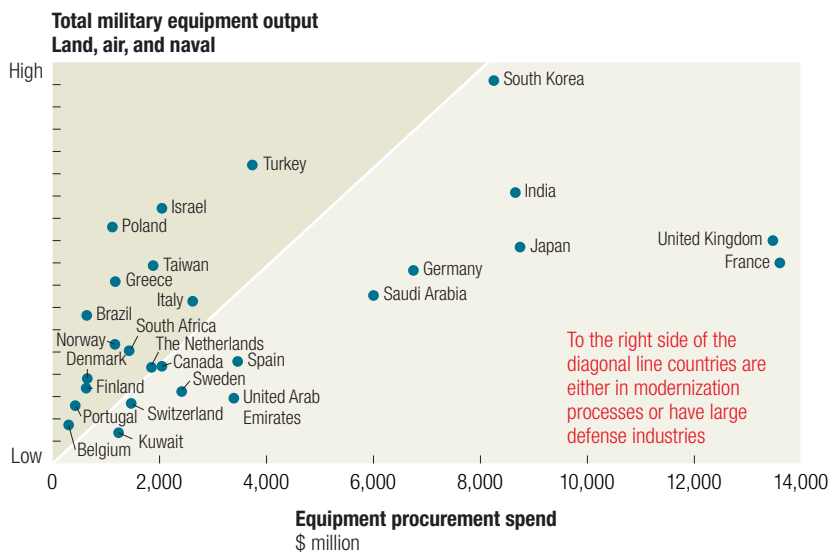
course, one could argue that a strong domestic defense industry is strategically critical to national defense; among other benefits, it gives a country complete control over supply, keeps it from being dependent on foreign providers, and guarantees sovereign protection in critical areas (secure satellite systems, for example). But because maintaining and supporting a domestic defense industry is an expensive proposition and limits financial freedom in other areas, it is critical that countries make sure they develop a strong rationale for their procurement choices by way of a well-defined defense industrial strategy.

Countries with sizable defense industries but declining defense budgets—examples include

Germany, South Africa, Sweden, and the United Kingdom—must evaluate each of their defense subsectors, such as secure communications, missiles, and unmanned aerial vehicles, on two criteria: strategic criticality and commercial viability. Strategic criticality is a qualitative evaluation of the subsector’s importance to military success, of whether there are other countries exporting the product, and of sovereign importance—that is, whether a bespoke product ought to be manufactured domestically for security reasons (as might for instance be the case with encryption software). Commercial viability is a quantitative assessment based on revenue, margins, and cost base, as well as local and global competitiveness. Subsectors that score high on

Exhibit 5  
**Output vs spend**

Governments that support their domestic defense industries tend to spend more for less output.<sup>1</sup>



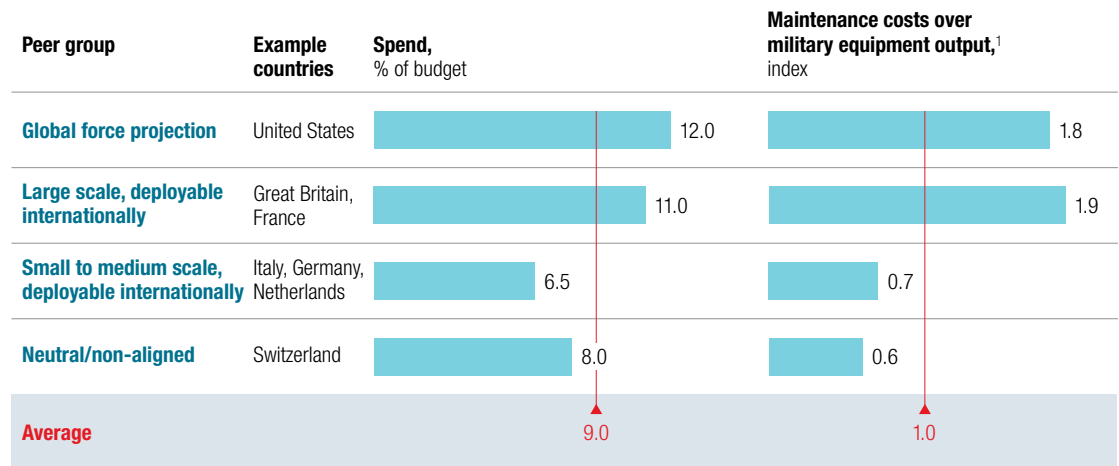
<sup>1</sup>We have removed the benchmark data for Australia, as errors in the source data and the methodology as it applied to Australia have been identified.  
Note: United States, Russia, and China have been excluded due to scale.



Exhibit 6

## Maintenance costs

Maintenance costs vary substantially, both within and among peer groups.



<sup>1</sup>One unit of military equipment output is approximately equivalent to one combat-ready unit (eg, a manned and maintained combat vehicle). For more, read “Our methodology for calculating output,” p. 5.

both criteria ought to be prioritized through R&D funding and export support; subsectors that rate high on only one criterion should receive limited government support; subsectors with low criticality and viability should be considered for divestiture.

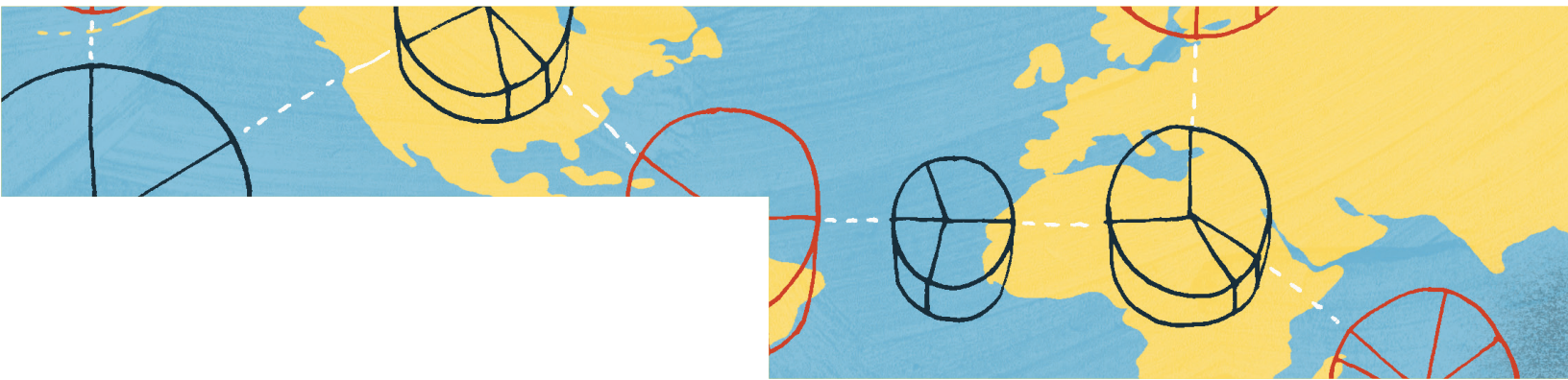
Countries that are increasing their defense spending and looking to grow a nascent domestic industry—India and South Korea, for example—should undertake a similar evaluation, but should of course attempt to assess future commercial viability as opposed to current revenue, margins, and costs. Such an evaluation should be based on comparative advantage and the ability to leverage key capabilities such as engineering talent.

This type of evaluation requires both commercial and analytical skills as well as military strategic-evaluation skills. Defense ministries should create cross-functional teams so that sound commercial and economic analysis can inform equipment-procurement decisions.

### Maintenance

Maintenance costs vary substantially, both within and among peer groups (Exhibit 6). Although the most capable forces naturally have some of the highest levels of expenditure, more detailed investigation reveals a number of drivers that help explain the wide variations that we have observed.

Actual maintenance expenditure is driven by at least four factors. The first driver is vintage. As any owner of a vintage car will immediately understand, forces that continue to operate older equipment often incur much higher levels of maintenance expenditure. There is therefore an implicit trade-off between funds to purchase new equipment and the funds to operate and maintain older equipment. Deferring replacement purchases often saves much less money than people expect, because they fail to account for the high and increasing costs of maintaining the older equipment left in place. The second driver of maintenance expenditure is variety. Forces that operate a wide range of different platforms incur



greater total maintenance expenditure than those with a narrower range, with each platform often requiring specialized technicians, dedicated equipment, and its own spare parts and associated supply chain. Operating several small fleets of different types of helicopter offers a good example. The third driver of maintenance expenditure is readiness. Some countries choose not to maintain their equipment at high levels of readiness, which saves money at the expense of fighting power. You can leave your car in the garage and not service it, but if you need to go somewhere, you can't always count on it to work. Of course, operations in difficult geographies can also substantially increase wear and tear, and hence readiness-related maintenance costs, particularly for land equipment. Repairs for battle damage can be costly as well. The final driver of maintenance expenditure is process efficiency—how efficient someone is at maintaining a given piece of equipment of a given vintage at the required level of readiness. We have looked at all of these factors to try to understand the wide range in expenditure we have observed.

This investigation uncovered massive opportunities for improvement in some countries.

Indeed, our experience working with a number of defense organizations generally indicates a 40 percent to 60 percent potential for increasing the quality and productivity of the maintenance, repair, and overhaul (MRO) function, without increasing costs. We have found that the best-performing military MRO organizations eliminate unnecessary variety, make smart use of outsourcing, excel at contracting, and constantly optimize their maintenance processes.



One of the cornerstones of any benchmarking exercise is the selection of a peer group. Once a defense ministry has chosen its peer group, it can identify the areas in which it most needs to improve and implement best practices to elevate its performance in those areas. The benchmarking results can give valuable directional insight into where the ministry can save money, as well as where it can achieve maximum effectiveness without increasing costs—both critical goals in today's changed world. ○