Economics of Military Outsourcing

Peter MacDonald

Submitted for the degree of Doctor of Philosophy

The University of York

Department of Economics

August 2010

Abstract

This thesis is concerned with outsourcing the production of defence and military activities to private providers. Theoretical approaches to the issue are reviewed and found to make insufficiently specific predictions to be useful in assessing the scope for outsourcing. Instead four empirical analyses are undertaken using publicly available data.

The first estimates the elasticity of substitution between military labour and capital. It builds on Ridge and Smith [169], using data on military wages to allow a more flexible specification. It finds the elasticity likely to be some small positive number, much lower than previous estimates suggest. The second extends the framework to incorporate civilian labour as a third factor of production. It then estimates the elasticity of substitution between military and civilian labour. This elasticity is estimated to be close to zero, suggesting that there is little scope for military outsourcing beyond that which exists without reducing capabilities in some circumstances.

The next approach assesses how far increases in the use of contracting-out and competitive procurement have yielded reductions in UK aggregate military expenditure. The best model suggests that increased outsourcing is associated with lower aggregate military spending in the long-run, but a number of issues arise which raise concerns about the validity of demand for military expenditure models.

The final approach uses an original dataset on control of Ministry of Defence Private Finance Initiative projects to identify the key firms involved in the UK market for military outsourcing contracts, the structure of the industry and its impact on the scope of efficient outsourcing. Whilst there is currently scope for effective competition for contracts in most market segments, this competition is vulnerable to industry consolidation. The key firms involved do not financially outperform their peers in other markets.

Contents

List of Tables

List of Figures

Acknowledgements

Above all I should like to thank Fliss for, amongst other things, her patience and forbearance whilst I have spent far too much time, which should have been spent with her, working on this thesis.

I cannot overstate my debt of gratitude to my supervisor, Professor Keith Hartley. Not only has he been an unfailingly positive source of advice, ideas and encouragement, but he has taken time out of his well deserved retirement from the Department in order to continue my supervision. I am also grateful to Professor Peter Simmons for his advice and direction. This thesis is much the better for their guidance, and I alone am responsible for any remaining errors, omissions or oversights.

Thanks are also due to those who persuaded me to complete my thesis when I wavered.

Financial support from the Economic and Social Research Council in the form of a Ph.D. Studentship is gratefully acknowledged.

Declaration

A large part of Chapter 3 was previously published as MacDonald [126]. The chapter included is an updated version of this paper. Chapter 5 has its roots in parts of Hartley and MacDonald [90]. Although none of the chapter is taken directly from the paper, its foundations are. With these exceptions the work contained in this thesis is original, and was conducted wholly and exclusively by the author.

Chapter 1

Introduction

This thesis is concerned with how activities necessary for the provision of national defence are produced. Defence is a public good. Since it is both non-excludable and non-rivalrous, if provision is left to private markets it will be under produced. In order to avoid this under provision defence is publicly funded. It need not be, however, publicly produced. For each activity necessary for national defence government faces a make-or-buy decision: it may produce these activities publicly; it may pay private firms to produce them on its behalf; or it may use some combination of the two. The central question for this thesis is: what is the most efficient mix of public and private provision of defence?

1.1 Motivation

There are two reasons why this issue is worth studying. Firstly, there is a long term trend towards increasing use of private providers for defence which, in the UK at least, shows little sign of slowing. Secondly, the current difficulties facing the UK's public finances are likely to lead to pressure to make increased use of the private sector in the provision of defence. Both are primarily driven by a desire to produce defence more efficiently, and to reduce the cost of incumbent in-house arrangements.[1] Which activities are likely to be produced most efficiently by private producers, how far efficient outsourcing may be extended, and where the limits (if any) to efficient outsourcing lie are questions suitable for economic analysis.

Over the past 30 years the mix between public and private provision of defence activities in the UK has been transformed. By the end of 1970s the vast majority of defence activities were conducted by government. The Armed Forces produced all military activities in-house. Many of the major defence firms were publicly owned following the Aircraft and Shipbuilding Industries Act 1977. Most civilian activities conducted on behalf of the Ministry of Defence (MOD) were produced by directly employed civilians.

Since then, the defence firms have been privatised, as have parts of the MOD: a siz-

[1] There are other reasons; notably a need to conduct tasks which the Armed Forces cannot perform and, especially in the United States, a desire to circumvent force numbers caps in operational theatres [209, footnote 5].

able proportion of the Defence Evaluation and Research Agency (DERA) was privatised as QinetiQ, and both the Royal Dockyards and the Atomic Weapons Establishment (AWE) are privately operated under Government Owned Contractor Operated (GOCO) arrangements. Many civilian activities have been outsourced and are now conducted by private firms. Even activities which were once considered the sole preserve of the Armed Forces are now produced under contract by private producers. For example, the provision and operation of the RAF's Air-to-Air refueling capability, the training of military pilots, and the provision of military communications satellites are all privately provided. Broadly similar changes have also been seen internationally, especially in the United States.

The use of private firms in defence is not new. Historically, many logistic and supply functions have been undertaken by private enterprise [12, p1]. What has changed is the number and variety of tasks now being contracted out, including many in operational theatres.

Neither the UK nor the US can now deploy their Armed Forces without extensive use of contractors. A measure of the increasing reliance on contractors since the end of the Cold War is provided by the figures for the US use of contractors. At the height of the Gulf War the ratio of contractors to uniformed military personnel was 0.017:1 [34, p138-139]. As of September 2009 the ratio was 0.87:1 in Iraq and 1.63:1 in Afghanistan [176]. In the more benign environment of the Balkans in 2002 the ratio was two contractors for each member of the Armed Forces [209, footnote 5].

In the UK the trend towards increased contracting and use of private provision looks set to continue.[2] In fact the pressures on the public finances in the UK may well cause the trend to increase. The MOD is likely to face real budget cuts of 10–15 percent over the period 2010–2016 [36], and there are no obvious sources of easy savings. Over the past eight years the defence budget has been squeezed by both ongoing operations and budgetary settlements increasing only in line with consumer prices. Over the past 20 years real unit costs of new equipment have grown by at least 2% annually, even excluding development costs having to be spread over falling numbers of units [37]. There appears to be a substantial body of opinion within the Ministry that the solution is increased use of the private sector. The Gray Report [74] on defence procurement recommended outsourcing the procurement and logistics activities of the MOD's Defence Equipment and Support Secretariat. The Defence Support Review by Gerry Grimstone, whilst not yet released, is reported to recommend the expansion of outsourcing and the use of private provision of defence activities [121].

In light of the situation the MOD finds itself in, it is imperative that it operates as efficiently as possible. Clearly, part of this will be selecting the optimal mix between private

[2]Conversely, there have been initiatives in the US to re-examine the cost efficiency of outsourced functions. Congress passed the National Defense Authorization Act for financial year 2008. This requires the Department of Defense (DoD) to consider bringing back in-house (in-source) certain activities which had previously outsourced. The Secretary of Defence stated that the DoD will reduce the number of service support contractors from 39% of the Pentagon's workforce to the pre-2001 level of 26%, and replace them with government employees [70]. The US Army claims to have 'saved $41 million by in-sourcing more than 900 core governmental functions to Army civilians' [71, p13], and expect that 'In-sourcing will continue to result in significant savings to the Army' [71, Addendum A].

and public provision. The identification of those activities most suited to being privately provided, and those which offer the greatest potential savings from outsourcing, is an economic issue. Even if the optimal mix of public and private provision cannot be defined precisely, in order to enable better informed decisions it may be useful to make explicit the costs and benefits of outsourcing various activities.

In principle, the optimal mix of public and private provision in defence is the same issue which must be addressed in the provision of any government activity—it requires one to define the optimal scope of government. The case of defence warrants separate analysis because there are a number of characteristics which are more prevalent in defence activities than in the rest of government.

It is not possible to characterise some single defence activity. A tremendous range of activities are necessary for the provision of national defence. These range from relatively commonplace activities also provided within the civilian economy, such as cleaning of barracks and manufacturing uniforms, through to those unique to defence, including anything conducted on the battlefield. As activities become more idiosyncratically military in nature, so certain characteristics become more important in the make-or-buy decision: the interconnected nature of defence activities; the necessity of specific investments, in both human and physical assets; the prevalence of long term, bilateral monopolistic relationships with opportunism; and uncertainty. None of these is unique to defence. It is the degree to which they are present, and their interaction, which distinguishes the make-or-buy decisions of military activities from those in other public services such as health, education, prisons, etc.

1.2 Objectives

The issue of the most efficient mix of public and private provision of defence is too large a topic to be adequately answered in a thesis. In fact there are very real difficulties with answering the question at all. Theoretical approaches to the issue offer at best incomplete answers (see Section 2.2). Models addressing the issue very quickly become intractable, and where they remain tractable are unable to offer conclusions precise enough to be valuable. Although the theoretical approaches are considered in this thesis, it used empirical analysis to address aspects of the issue. The key questions this thesis seeks to address are:

Question 1 *Which activities does economic theory suggest are best outsourced, and which best provided in-house?*

The alternative is to use empirical evidence to refine the predictions of theory. Because there has been such an increase in the use of private provision over the past 30 years, the questions of when and where does efficiency improve, or are cost savings achieved, are natural candidates for empirical investigation. Unfortunately it is very difficult to address the relevant issues directly because of the paucity of publicly available data, especially at the level of individual outsourcing contracts. Even data aggregated to the level of the MOD is limited, with much of it being 'hidden in a statistical sense within old categories and budget codes' [108].

Although the optimal mix of public and private provision is difficult to address in empirical work directly there are two ways in which the issue may be tackled indirectly, using publicly available data. The first is through exploiting a fundamental difference between public and private provision: the Armed Forces and the MOD, when providing defence activities typically use a mixture of military and civilian labour in combination with capital. Private providers cannot use military labour (as they are unable to sign military labour contracts) and so must substitute either capital or civilian labour for military labour.

An idea of the likely scope for outsourcing in defence may be obtained by answering the following questions:

Question 2 *How easily can capital be substituted for labour in the provision of defence?*

Question 3 *How far can civilian labour be substituted for military labour in the provision of defence?*

The second feasible way to indirectly address the optimal mix is to examine:

Question 4 *Has increased use of private providers led to lower overall defence expenditure?*

The final question arises because, in areas other than defence, one of the key drivers of savings from the use of the private sector has been the introduction of competitive incentives. Given this, in assessing the scope for outsourced provision it is necessary to establish:

Question 5 *How competitive is the market for outsourcing contracts?*

Defence industries, especially those for military equipment, tend to be dominated by a few large firms. Many military activities have very few, if any, potential providers and it may be difficult to arrange effective competition for some military activities. If competition is ineffective or absent then the scope for efficient outsourcing in defence will be lower than in other areas of public services because one of the key drivers of efficiency is missing.

1.3 Methodology

The general approach of this thesis is to address the questions identified above empirically. No attempt is made to model the central issue of which is the most efficient mix of public and private provision of defence theoretically. This is not because theoretical treatments are considered inferior; such treatments could undoubtedly yield valuable insights into problem. There is, though, already a wide range of models which may be applied to the problem (reviewed in Section 2.2) but their predictions tend to be too vague to be readily translated into useful policy advice. Empirical analysis of the issue is preferred because it offers the prospect of allowing more specific conclusions to be drawn about the questions of interest, albeit at the potential cost of loss of generality.

The key difficulties in undertaking empirical work on this issue result from the paucity of available data. In common with lots of work in defence, modelling must be undertaken using very small datasets. There are also frequently differences between the variable required by theory and those available for estimation. Given these difficulties it is essential to ensure that the best possible data is utilised. This thesis uses only publicly available data. In order to use the best available data the approach to data collection taken here is, in descending order of preference, to: Firstly, search as wide a variety as possible of data sources to find a variable which precisely matches that required by theory. Secondly, if the required measure is unavailable to construct it from published sources. Examples of constructed series used in this thesis and the military wage data used in Chapters 3 and 4, the measure of the scale of outsourcing used in Chapter 5, and the dataset of information on PFI project ownership which underpins much of the analysis in Chapter 6. Thirdly, if a variable which precisely matches that required by theory can neither be found or constructed, the closest available measure is used.

Some may consider the available data inadequate to answer the questions of interest. Rather than focusing on the data difficulties this thesis seeks to undertake the best possible analysis, given the available data. Occasionally this approach requires uncomfortable compromises to be made but it does at least offer the prospect of reaching some conclusions, however qualified.

1.4 Structure

The remainder of the thesis is structured as follows: Chapter 2 reviews the literature on military outsourcing, both theoretical and empirical. It derives the research questions outlined above, and seeks to answer Question 1.

Chapter 3 addresses Question 2, building on a study by Ridge and Smith [169] to estimate the elasticity of substitution between military personnel and other military inputs. Data on military wages are used which allows more flexible specification of both the demand and supply equations.

Chapter 4 seeks to answer Question 3. It extends the framework of Ridge and Smith [169] to include three factors of production, including both military and civilian labour. The degree to which civilian labour can be substituted for military is calculated by estimating the elasticity of substitution between civilian and military labour in the military production function. This is used to infer the scope for the use of private (contracted out) civilian labour.

Chapter 5 examines whether higher levels of military outsourcing are associated with lower aggregate defence spending in the UK—addressing Question 4. Three models of the demand for military expenditure including a proxy for the level of outsourcing are estimated using UK time-series data over the period 1979–2007. The results of each are presented and discussed.

Chapter 6 surveys the UK industry for outsourced defence activities, focusing on Question 5. It uses MOD Private Finance Initiative projects as a sample of UK military outsourcing contracts, identifying the key firms involved and the degree of competition. This, augmented by information from other sources, is used identify the major firms in each segment of the outsourcing market. The degree of market concentration in each segment is then examined, and the financial performance of the key firms compared with their peers who are not involved in defence outsourcing.

Finally some conclusions are offered, and directions for future work suggested.

Chapter 2

Literature Review on Defence Outsourcing

2.1 Introduction

The production of final defence output, however defined, requires the production of a wide variety of goods and services. The economic activities required to produce the necessary goods and services range from those seen throughout public services, activities such as providing administrative support, the provision of accommodation, of catering and cleaning services to barracks, right through to activities which are unique to defence. These include many military activities, such as combat operations provided by the Armed Forces, the provision of supplies to operational theatres, and the construction and operation of complex weapon systems.

As a pure public good, defence is collectively provided. Since markets would under produce it must be funded by government, but it need not be produced by government.[1] For each of the activities necessary for the provision of defence government faces a make-or-buy decision. It might either produce that activity in-house, or pay another organisation to produce on its behalf. Since the recent tradition has been to produce almost all defence activities in-house, a decision to use private firms to provide an activity is seen as outsourcing.

The definition of outsourcing used here is broader than has been used by others, especially in the empirical literature. Elsewhere [50, for example] a distinction has been drawn between privatisation and contracting-out. Privatisation tends to involve both the transfer of physical assets from public to private ownership and direct charging of users. Contracting out concerns activities which do not require significant capital assets and are provided directly to, and funded by, the procuring authority. Because almost all activities are cen-

[1] There are relatively minor exceptions to the centralised funding, e.g. in the UK the Met Office provides meterological services and weather forcasting to a large number of public and private clients but remains an executive agency and trading fund of the MOD.

trally funded, such a clear distinction cannot be made in defence. There are many activities which would fit into the category of contracting-out, but for their need for significant assets.

Here, the terms 'outsourcing', 'contracting-out', and 'contracting' are used interchangeably to refer to the involvement of the private sector in providing formerly publicly produced defence activities. They are taken to encompass the entire range of mechanisms through which the use of private provision has been achieved, from market testing, through contracting for availability and capabilities, to Private Finance Initiative (PFI) and Public-Private Partnership (PPP) arrangements and other recent 'partnering' arrangements.

There has been relatively little economic analysis of the issue of military outsourcing specifically, but there is a large literature which might be applied to the issue. Four theoretical approaches which each shed some light on the issue are considered in the next section. The empirical evidence is then reviewed in Section 2.3, dealing first with the evidence from the defence sector and then lessons which might be drawn from the outsourcing of other public services. Finally some conclusions are offered.

2.2 Theoretical approaches

The fundamental issue when considering military outsourcing is which aspects of defence should government provide in-house, and which should be bought in the market. Government faces a set of make-or-buy decisions, and its problem is closely related to the boundaries of the firm problem first raised by Coase [39]. Three of the four approaches discussed below have their roots in the boundaries of the firm problem, but the defence problem is distinguished by differences in the organisational structures and incentives between public and private providers [13, 46, 47, 227], and by the institutional and political factors inherent in the provision of defence [171, 170] including the reliance on public funding.[2]

The Public Choice perspective develops the view of in-house providers as public monopolies. It focuses on the role of competitive incentives in models of bureaucracy, and their ability to restrain rent seeking behaviour of public servants which, it is maintained, makes public monopolies inefficient.

The approaches of transaction cost economics (TCE), property rights theory (PRT), and relational contracting are complementary. They focus on the inefficiencies which arise in both public and private provision of an activity, and how the allocation of ownership can minimise these inefficiencies in a world of incomplete contracts: TCE focuses on uncertainty in ongoing relationships and the need for adaptation [225]; PRT ignores the need for adaptation and focuses exclusively on the hold-up problem and how ownership affects incentives for *ex ante* investment [72, 225]; relational contracting, exemplified by Baker *et al.* [10], returns to the issue of adaptation but, rather than considering formal contracts, focuses on how self-enforcing contracts and reputation affect inefficiency and the role of ownership.

[2] For a summary of the theoretical approaches to, and a survey of the empirical evidence on, make-or-buy decisions within private firms see Lafontaine and Slade [119].

2.2.1 Public choice

Public choice arguments, building on the seminal analyses of bureaucracy of Tullock [205], Downs [54], and Niskanen [150], offer the classical rationale for the competitive tendering of public services. It is distinguished from the other approaches considered by its focus on competitive incentives and bureaucrats' motivation to maximise budgets. For a more extensive discussion see Boyne [29, ch4] from whom this section borrows heavily.

A defence ministry fits Niskanen's [150] definition of a pure bureau, since it involves a monopolist completely controlling the market for the provision of defence who relies entirely on political sponsors for funds rather than individual customers. Delivery by a bureaucracy results in public services being overproduced and produced in a technically inefficient manner.

This inefficiency is caused by the nature of the principal-agent relationship between politicians and public officials (bureaucrats). Public officials are motivated by a number of selfish objectives such as power, prestige, reputation and material rewards. Since all of these interests are served by higher levels of public expenditure, bureaucrats may be thought of as budget maximisers [150]. Bureaucrats can pursue their self-interest because only they, and not their political sponsors, can observe an agency's true need for funds to perform the tasks required of it. This information asymmetry, and their position as a monopoly provider, makes it difficult for politicians to monitor the behaviour of bureaucrats and allows agents to extract rent from their political sponsors [151].

The public choice solution to this inefficiency is the introduction of competition. In the absence of competition, in-house providers act as monopolists and politicians have no yardstick against which to assess efficiency [32, 215]. A competitive bidding process reveals information about alternative unit costs to politicians, changing the balance of power between the politicians and bureaucrats [28]. Creating the appropriate structure of competition partially solves the monitoring problem, allowing the behaviour of officials to be redirected toward the public interest—the pursuit of their own interests can be constrained by institutional arrangements [29].

It is a core public choice proposition that competitive pressures lead to superior performance in the public sector, especially in pure bureaux [28, 29]. This proposition generates testable hypotheses: there should be a significant negative relationship between measures of competition and expenditure, and a positive relationship between competition and efficiency. There is considerable empirical support to these hypotheses, especially the former. This is discussed in Section 2.3.2.

In the public choice framework any benefits of introducing a programme of outsourcing competitions result exclusively from competition itself, rather than inherent differences between public and private provision. However, the rationale for separate models of firms and bureaux is that the two organisational structures work differently. It is not unreasonable to suppose, given they work in different ways, that even in the absence of competitive incentives the same activity undertaken by firm or by bureau will produced differently, and will lead to differences in performance. The public choice literature rather plays down these inherent differences between public and private provision, but they are the focus of the other approaches considered. There are also differences between bureaux facing different institutional environments [227]. These differences may be important because they affect the degree to which effective pressure may be brought to bear on public servants through the introduction of competitive incentives.

That competition, or at least contestability, may be introduced is taken as given. However, there are some defence activities where only limited competition may be constructed. For example, there are no firms who currently have the capability to launch an armoured brigade into battle. Activities such as this are insulated from competition, as are some others (such as the intelligence services) whose work is of such a sensitive nature that the threat of competition is not credible.[3]

The approach also pays little attention to some of the potential difficulties associated with competition such as transactions costs, the erosion of trust between agents and principals, and rent seeking [29]. These shortcomings and how they are determined by the characteristics of the activity in question are addressed below.

2.2.2 Transaction costs

The TCE approach focuses on adaptation over the course of a contract, and its development has been driven largely by the work of Williamson [221, 222, 223, for example]. Each activity is a potential transaction which might be either be performed in-house or contracted for in the market. If the latter course is chosen then a transaction occurs. For every potential transaction comparison is made of the costs and benefits of contracting (transacting), and where the benefits exceed the costs then outsourcing is to be preferred.

For each potential transaction private market provision (here contracting-out) is implicitly assumed to yield some benefit over in-house provision and it is the costs which determine the efficient method of provision. TCE imagines a world with small numbers contracting under imperfect and asymmetric information. In a state of imperfect information contracts cannot be optimal in the full information sense, and bounded rationality and opportunism may lead to parties seeking to exploit an information advantage. The threat of opportunism is exacerbated when there is also asset specificity, leading to the hold-up problem.[4] Transaction costs depend crucially upon, and are increasing in, the frequency of transaction, and the degrees of contractual uncertainty and asset specificity [223].

TCE suggests that activities which are conducted frequently, and exhibit high degrees of

[3]Though there is some scope for competition between bureaux, especially in the example of the intelligence services, but any efficiency gains are likely to be offset by the resulting duplication of capabilities.

[4]If there are small numbers of potential contracting partners, once a contract is signed there is little possibility of finding someone else to fulfil the contract. This being so one of the parties can force renegotiation after the contract is signed, and can impose terms to which the other party has little choice but to aquiesce. Hold-up becomes more of a problem as asset specificity increases, as this reduces the chances of finding someone else to fulfil the contract. The hold-up problem is anticipated by both parties, and reduces the incentives for ex ante investment in specific assets.

uncertainty and asset specificity have high transaction costs which are likely to exceed the benefits of transacting, and are therefore more likely to be best produced in-house. Many defence activities exhibit high degrees of contractual uncertainty and asset specificity, but are transacted infrequently. The importance of uncertainty and the necessity of specific investments mean that once contracts are signed there is a 'fundamental transformation' [226] in the relationship between the procuring agency and the provider:

> 'initial bidding merely sets the contracting process in motion.... Once substantial investments in transaction-specific assets are put in place ... what was a large numbers bidding condition at the outset is effectively transformed into one of bilateral supply thereafter.' [223, p61]

This fundamental transformation means that the small numbers contracting world with imperfect information of TCE describes well the environment in which defence equipment procurement occurs, and so is likely to apply to the market for outsourcing any defence activity which requires substantial equipment (such as the Future Strategic Tanker Aircraft project to supply and operate the RAF's air to air refuelling tanker fleet).

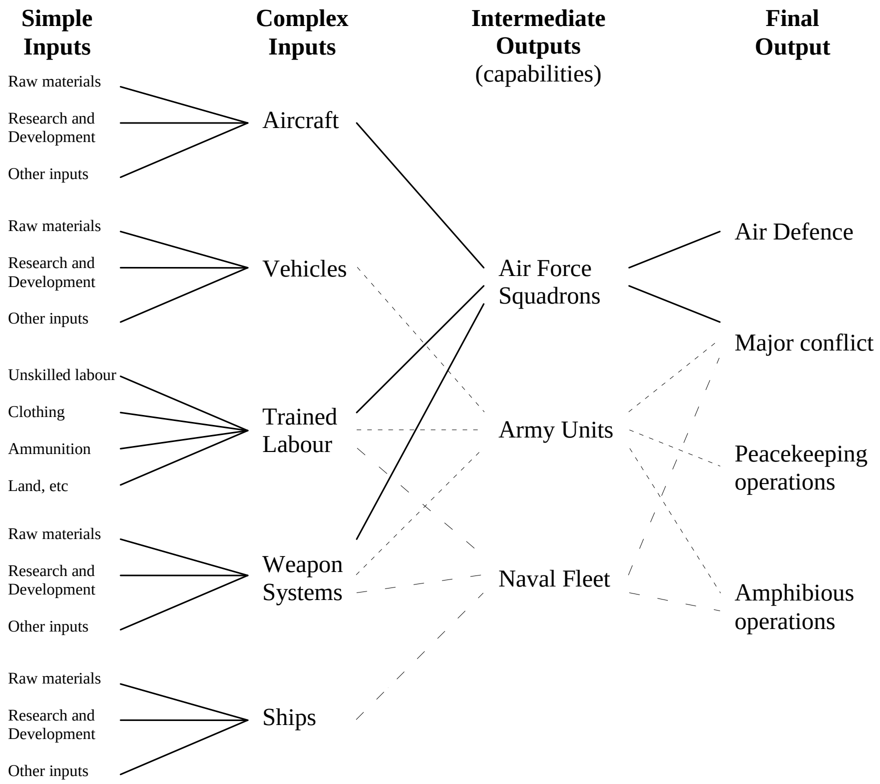
The TCE approach has been applied to the use of private military companies by government as an alternative to standing sovereign forces [61, 62]. Fredland and Kendry [62] distinguish defence inputs and outputs. Inputs such as 'weapons systems, vehicles, ammunition, clothing, etc.' [62, p150] are easier to specify. As a result they incur low transactions costs which are likely to be outweighed by the advantages of contracting-out, and so are best outsourced. Conversely, outputs are best provided publicly as the transactions costs involved in designing and enforcing a contact for defence outputs are high: Investment in new weapons systems is vast, lumpy and full of uncertainty, and private firms are unlikely to be able to procure strategic and major weapons systems. Since most outputs depend upon the use of such equipment the transactions costs involved in contracting would likely be large. The difficulty of writing a complete contract between contractor and government could give the firm unacceptable bargaining power during a crisis (hold-up).[5]

Fredland and Kendry's [62] input-output distinction arises primarily from the assumption that inputs are easier to specify than outputs. Even if true, ease of specification does not necessarily justify such a distinction. Rather, it gives an ordering by which contracting for inputs incurs lower transactions costs than contracting for outputs, but without quantifying the potential benefits of each course, no judgement can be made as to whether the benefits exceed costs for the alternatives.

The process by which defence output is produced is shown in Figure 2.1. Simple inputs are combined into complex inputs such as trained labour and military equipment. These complex inputs are then combined to produce intermediate outputs—the military units which are maintained as standing forces. These standing forces provide a set of capabilities

[5] Although Fredland and Kendry's [62] main conclusion is the input-output distinction, it is qualified by suggesting that there may be some narrowly defined military objectives for which outsourcing is viable, e.g. hiring a PMC to protect embassies (a possibility also suggested by Edmonds [58]).

Figure 2.1: The production of defence output



which may or may not be called upon in any given period. The nature of the final output is generally unknown when decisions are made as to the required capabilities (and so the structure of the standing force), but when requirements become apparent the maintained capabilities are combined as the situation requires. An input-output distinction is over simplified. Some intermediate outputs, such as units providing logistic support, may well be sufficiently easily specified that the benefits of outsourcing exceed the transactions costs, and so are better provided privately. On the other hand, the production of some complex inputs may be better retained in-house. The training of personnel may be one such input. Basic training of individuals might be outsourced, as occurs for basic flying training of pilots in the UK, but unit level training may better retained in-house because it is specific to operational requirements which are unknown at the time of contracting, and because of the importance of different components of the standing force training together.

Fredland [61] extends the TCE based analysis of military outsourcing by incorporating the notion of sovereign transactions [224]. Sovereign transactions are those transactions for which public authority is necessary, loyalty to the state is fundamental, and which may have implications for the security of the state. For such transactions, in addition to asset specificity, contractual uncertainty, and frequency, it is necessary to consider probity (the loyalty and rectitude with which the transaction is delivered), since a breach of probity can put the state at risk. The low-powered objectives characteristic of public bureaucracy may be desirable for sovereign transactions as they deter employees from being non-compliant, or providing misleading information and so reduce the chance of a breach of probity [224]. Applying this to military outsourcing Fredland [61] suggests there may be gains from outsourcing non-combat roles, but that combat roles are best provided by government. Any contract involving combat will be incomplete resulting in renegotiation, hold-up, and opportunistic behaviour that may threaten probity. Efforts to mitigate these hazards will result in high transaction costs. For non-combat roles, such as training, uncertainties are lower, assets are generally less specific and timing is less crucial. This reduces the costs of hold-up and the risks to probity, resulting in lower transaction costs, and so greater scope for outsourcing non-combat roles. The difficulty with this distinction is that the boundary between the two is blurred. As has been seen in Iraq and Afghanistan, in some types of operations any role in-theatre is potentially a combat role. There is also the reliance of combat roles on non-combat roles. The interdependent nature of military activities also leads to scope for opportunism in combat roles being transmitted to non-combat roles such as logistic support and medical facilities.

In applying TCE to military outsourcing the conclusions of both Fredland and Kendry [62] and Fredland [61] revolve around the transactions costs arising from the difficulty in specifying in contracts various military activities, with those which are 'too' difficult to specify being provided in-house. In terms of the activities in Figure 2.1 TCE offers the concrete conclusion that final outputs should be produced in-house. Although the deployment of troops on combat operations is unlikely to be frequently conducted, the levels of uncertainty and asset specificity are high and probity is at stake. In this case the 'choice of organizing mode turns entirely on which mode has superior adaptive properties. [V]ertical integration will ordinarily appear in such circumstances' [223, p78]. This conclusion also holds for all final outputs. TCE also offers explicit conclusions for the production of simple inputs. Since they are both inputs and non-combat activities they should be outsourced: they are procured frequently, probity is a lesser concern and the hazards of uncertainty and asset specificity are less pronounced, so the transactions costs of purchasing on the market will be reasonable.

Although clearly suggesting the existence of limits to the efficient scope for military outsourcing, TCE does not offer clear direction as to precisely where amongst the complex inputs and intermediate outputs of Figure 2.1 they might lie. Whilst it introduces criteria by which the likely size of transactions costs may be judged (uncertainty, frequency, asset specificity and probity) these are too vague to offer precise predictions. This reflects the difficulties of formalising TCE (though some progress is being made [9, 122, 200, 201, for example]) which limits how precise policy recommendations can be and makes empirical testing difficult.

One key limitation of the TCE approach is that it is a maintained assumption that market (here outsourced) provision is more efficient. The source of these high-powered incentives, which presumably result from differences between in-house and market provision, is not explicitly considered. There are a number of reasons incentives may differ: hard budget constraints for private firms [85]; the discipline imposed by capital markets [86], the profit motive or rivalry; or private firms being able to use resources (especially labour) more flexibly than an in-house team [215]. However, the size of the benefits of outsourced provision are likely to depend on the precise nature of the activity being undertaken. Without an explicit theory of how benefits arise and vary with the characteristics of the activity being undertaken, any comparison of the costs and benefits of contracting will be limited.

2.2.3 Property rights approach

The property rights approach differs from TCE in that it focuses on inefficiencies generated solely by incentives for ex ante investment, as opposed to adaptation throughout the duration of the contract [72, 225]. *Ex post* Coasian bargaining is used to assume away problems resulting from on-going relationships which were the focus of TCE. Building on the work of Grossman and Hart [76] and Hart and Moore [82] on the boundaries of the firm, PRT has been widely applied to the differences between government and private provision of public goods and privatisation [25, 60, 81, 84, 177, 178].

It is differences in the ownership of assets which distinguish public and private provision. Asset ownership confers residual rights of control, so influencing incentives to invest. The residual control rights might take the form of access to private information [177, 178, 181]; the right to intervene in production [118, 174]; or the right to control the firm's assets in the event of uncontracted contingencies [25, 40, 84].

The make or buy decision

Applied to the make-or-buy decision under discussion here, the model of Hart *et al*. [84] considers provision of some good or service (defence activity) for which contracts are incomplete because quality is unverifiable (quality is really shorthand for the noncontractible aspects of service provision). Production of the good requires the use of an asset or facility (here perhaps a large capital asset such as an aircraft carrier). Ownership of the asset allows the power of veto over changes made to the asset. The asset is managed by an owner-manager under private provision, or by a government employee under public provision. Whether owner manager or public servant, the manager may exert two types of effort: cost reducing effort, which reduces costs but also reduces quality; and quality improvement effort which improves quality but increases cost. These efforts are observable but not verifiable, but require the consent of the asset's owner. The allocation of ownership determines the incentives for manager effort, and so drives differences between public and private provision (ownership). Hart *et al*.'s [84] model suggests that, relative to public ownership, private ownership always provides greater incentives to cut costs and at least as great incentives to improve quality. However, the incentives for cost cutting may be excessive under private ownership as the manager ignores the negative effects of cost reduction on quality.[6]

For outsourcing the implications are that private ownership (and so provision) is superior when cost cutting has little impact on quality, or when both the opportunities for cost reduction are small and government employees have relatively weak incentives. Conversely, public provision is superior when both the adverse effect of cost reduction on quality is large, and either quality improvement is unimportant or government employees do not have weaker incentives in quality improvement. Applied to weapons procurement, Hart *et al.* [84] conclude that because quality innovation is tremendously important, and incentives of private firms are probably stronger than public employees, weapons production should be outsourced. Fredland and Kendry [62] use the model to reinforce their input-output distinction discussed in Section 2.2.2, but similar criticisms apply again.

The conclusions to be drawn from the Hart *et al.* [84] model depend crucially upon the degree to which quality degradation resulting from cost reducing effort (quality shading) can be effectively controlled by well specified (contractual) performance targets. If quality shading is completely controllable through contractual mechanisms such as performance targets, then the issue should not be examined in an incomplete contracts framework. In defence activities, the inherent uncertainties mean that quality shading is very hard to control. Indeed, much of the weapons procurement literature has arisen precisely because of the difficulty of specifying quality contractually [95, 101, 117, 120, 123]. A related issue, the inadequacies of contract management skills in defence ministries, is discussed in Section 2.3.1.

Dalen and Moen [40] extend the Hart *et al.* [84] model to allow the use of cost-sharing contracts. The private sector's tendency to excessively reduce costs may be controlled through a cost-sharing contract. However, the use of cost-sharing contracts gives contractors incentives to artificially inflate costs (through, for example, cross-subsidisation or managerial perks). If the scope for cost inflation in an activity is high enough then this can more than offset the benefits of controlling quality shading. Public provision is to be preferred if both quality shading is large and important, and the scope for private firms to inflate costs is large. Public provision is preferred in these circumstances because cost inflation can be controlled directly by the government when provision is public.

A serious limitation of the PRT based models of outsourcing is that they rely upon an assumption of owner-managed firms under private provision. This is difficult to justify in the case of defence outsourcing. If ownership and management are separated, then although private cost reducing incentives still do not account for quality shading, the incentives to innovate are no longer clearly greater than those of a publicly employed manager [84, footnote 12].[7] The conclusions of the approach hinge on the incentives of private managers

[6]This conclusion is strengthened by allowing contracting with (non-profit) organisations which have an interest in the success of the project irrespective of their involvement. Their interest in the project relaxes the cost/quality trade-off and weakly strengthens the case against government provision [25], but it is not clear than such organisations exist for defence activities.

[7]One might view the model as contrasting the cases of public provision with agency problem and private provision with no agency problem. The results are driven by a public manager's need to negotiate in order to

being higher powered than those of public managers. More complex ownership structures, and the possibility of regulation (as would be likely for private provision of many defence activities), introduce the possibility of multi-principal inefficiencies [116, 118]. These complicate the conclusions on the scope of outsourcing with the optimal ownership structure being a trade-off between multi-principal inefficiencies and underinvestment.

As an approach to the military outsourcing issue property rights are appealing as it deals explicitly with the differences between public and private provision, something missing from the public choice perspective, and to a lesser degree from TCE. The approach does, though, emphasise the role of *ex ante* investments, ignoring the importance of adaptation through the life of the contract highlighted by TCE. It also lacks rigorous foundations, and although the mechanisms proposed to complete incomplete contracts have not been seen in the real world, their existence is troubling for the approach and suggests it may be missing something important [83].[8]

Although a great deal more formal than TCE, the formality doesn't really help to derive concrete conclusions. Nor does it aid empirical assessment of the model, largely due to the reliance on non-contractible investments which are, by definition, hard to quantify [83]. The approach (or at least as it is presented here) does not provide an explanation of precisely why public ownership could not simply replicate the incentives of the private providers when possible agency relationships are allowed within private providers. There are many reasons why such replication might not be possible; incentives driven by ownership is but one driver of differences between public and private provision. As such, the property rights approach cannot provide a complete description of the factors relevant to the make-or-buy decision.

Another limitation is that the approach views asset ownership as the definition of provision of an activity, but the situation is not this simple. There are examples in defence of private provision of contractors operating government owned assets (e.g. ammunition factories in the US or the Atomic Weapons Establishment in the UK).

Bundling and Public Private Partnerships

So far the government's problem has been viewed as a single make-or-buy decision for each activity necessary for the provision of defence. Actually, many defence activities require the construction of capital assets which are then used to provide a defence service. In these cases the government faces two related make-or-buy decisions: one for the construction of the asset, and a second for the operation of the asset to provide the service required. The

realise gains from effort, whereas a manager-owner does not need to negotiate. Viewed this way the model does not seek to explain differences between public and private provision, but simply imposes them through the agency problem.

[8] Essentially critics contend that there is a tension between assuming an inability to specify contractual outcomes in all possible states of the world, but yet players being able to anticipate their payoffs conditioned on the realisation of the state of the world [203]. There are sophisticated (Maskin) mechanisms such as message games which can be used to obtain the same payoffs as if contingencies were describable. See the special issue of the Review of Economic Studies (1999), 66 (1), or Schmitz [179] for a survey. Never the less, it remains true that those contracts seen in defence contracting are incomplete.

stages are related because decisions taken at the construction stage (especially the quality of the asset) influence the quality of the service which may be provided at the operating stage.

These related make-or-buy decisions are the subject of models of the relative merits of Public Private Partnerships (including the PFI mechanism) and traditional procurement. One of the key features of PPPs is the bundling of the construction and operation elements of public service provision [24, 106]. This feature has been used to distinguish between traditional procurement (which involves purchasing the building and operation stages under separate contracts), and PPPs (which combine both elements in a single contract).

The PRT has been applied to PPPs by Hart [81]. In his model the builder of the asset may invest in the asset in two ways: productive investments (i) which both increase the social benefits and reduce the costs of the operating phase; unproductive investments (u) which reduce both costs and social benefits of operating. Investments are unverifiable, so cannot be contracted upon. Traditional procurement (with separate contracts) leads to suboptimal investment in both i and u, because the builder obtains no benefit from either. Under bundled (PPP) provision the firm internalises the cost saving effects of both investments, resulting in a closer to optimal level of i but overinvestment in unproductive u. The implication is that it is ease of specification which determines whether PPPs or traditional procurement are preferable. If the quality of the service can be more easily specified than the quality of the asset then bundled provision's overinvestment in u may be controlled through appropriate contract specification, and is to be preferred. However, if quality of asset can be most easily specified then unbundled provision should be preferred because the underinvestment in i may be specified in a contract.

Other models of the problem are not strictly property rights based, but seem most appropriately included here. Iossa and Martimort [106] propose a model in which, unlike Hart [81], the unproductive investment, u, has no effect on the social benefit of the project. Productive investments increase the social benefits of the projects but may either increase or decrease operating costs (termed negative and positive externalities respectively). In the model when government is able sign cost-based contracts with firms the preferred form of provision depends upon the sign of the externality. When the externality is negative, bundling increases agency costs and traditional provision may be preferable. When the externality is positive there are some small externalities for which bundled provision is preferred, but more general conclusions cannot be reached.

Bennett and Iossa [19] consider a model in which the builder's investments are non-contractible *ex ante* but verifiable *ex post*. They consider slightly different investments which can be made at both the building and operating stages. Investment at the building stage increases the social benefit of the project and creates either positive or negative externalities at the operating stage. Investment at the operating stage increases social benefits, and reduces operating costs. The key differences between this and Hart's [81] model are that (as with Iossa and Martimort [106]) investments at the operating stage do not reduce but increase social benefit, and that because investments are verifiable *ex post* renegotiation allows contractors to be rewarded for investments which improve social benefit.

Bundling allows the builder to internalise the social benefits and cost reductions of investments. Bundled provision is preferred if cost externalities are positive (so if investments made at the building stage reduce costs of operation), if investments made a operating stage have only a small effect on social benefit, and if the effect of investments on residual value of the asset is large. With a weak externality public provision is to be preferred. There are a large range of cases (combinations of externality and incentives to invest in quality) for which the optimal form cannot be determined.

The most interesting cases for defence are likely to have large cost externalities, and if it is military equipment then the effects of innovation are likely to be great (because these are investments made for operating the equipment) and because when the contract expires equipment is likely to be obsolete and so the residual value likely to be small.

Others have taken a complete contracting approach to modelling the issue. If quality is contractible then for some small positive externality bundling (PPPs) dominates, and as in Hart's [81] PPPs produce more of both cost reducing and quality improving effort. Unbundling is to be preferred when the externality is negative [130]. If however quality is unverifiable then ownership matters. If the externality is positive and the unitary fee is small enough then builder ownership and bundling dominates as the externality is internalised. When the externality is negative unbundling reduces agency costs. Public ownership and unbundling are more likely to be optimal as the unitary fee and uncertainty about the quality of the assets increase.

Rather than a problem of incomplete contracts, the problem might be viewed as one of asymmetric information about the investments being made (moral hazard problem at building stage) and the true operating cost (adverse selection problem at operating stage). Bentz *et al*. [20] consider such a model with complete contracts. Bundling becomes a trade-off between the advantage of forcing a bundled provider to internalise the cost implications of their investments at the building stage, and the disadvantage of precluding the use of separate revelation mechanisms (through separate contracts for each stage) for the building and operation stages. They conclude that bundling is preferable when the costs of improving quality at the building stage are low and initial set-up costs are low. Conversely, when these costs are high unbundled provision is preferred because of the ability to tailor the building and operating contracts to solve the two asymmetric information problems.

Bundling is more beneficial when a better quality of asset can significantly reduce cost at the operational stage (including maintenance cost), and this has been one of the key arguments for the use of PPPs [24, 89]. Another result that consistently comes out of the models is that the case for public provision and ownership is stronger, and likely dominant, in cases where investments made at the construction stage are socially worthwhile, but which increase costs at the operating stage.

The models of bundling do not, however, allow straightforward, practical conclusions to be drawn as to when PPPs and bundling are to be preferred in defence activities. This is partly because of the complexity of the interactions which occur in the real world. Whilst it is unreasonable to expect any model to capture all aspects of an issue, the models above focus on slightly different aspects of the bundling issue, but reach different conclusions, and even propose different criteria for assessing the relative benefits of bundled and unbundled provision. The interactions between activities are complex and cannot be fully captured in models, and the result is that precise or specific practical conclusions cannot be drawn.[9]

The models have two key limitations when applied to the defence problem. Firstly, all of the models represent traditional procurement as placing separate contracts for the building and operating phases with private firms. This does not accurately reflect traditional provision of defence. Instead government bought the asset from private firms, and then performed the operational phase in-house. The failure to model in-house provision (where incentives are not determined in contracts), and the models' heavy reliance on differences in contractual incentives between bundled and unbundled provision, suggest that the conclusions of these models (such as they are) shed only very limited light on the make of buy decisions faced by a defence ministry.

Secondly, bundling reduces flexibility. Contract flexibility is key in areas (such as defence) where preferences and technology change quickly [106]. Unbundled provision has the advantage that the signing of the operating contract may be delayed until costs are realised. Under in-house provision changes may be made at any stage of the operation phase. Although the use of PPPs and bundling improves incentives for cost reducing effort, it also has a cost in terms of reduced flexibility.

Overall, PPPs are more beneficial when a better quality of asset can significantly reduce cost at the operational stage (including maintenance costs), when asset quality has a great impact upon the quality of service, and when demand for the asset is stable and easy to forecast [106]. When considering military equipment the first two conditions are important. Demand for the asset is less stable, but where demand stability can be achieved by specifying reasonable availability requirements then military equipment may benefit from bundled PPP type provision. However, the lack of flexibility inherent in tightly specifying availability and quality (where possible) requirements suggests that for assets intended for use on the battlefield unbundled provision is to be preferred. The alternative of government buying the asset and then tendering for an availability based logistic contract separately has certain advantages in terms of flexibility, especially if the logistic contract is let once the equipment is mature and requirements are well understood.

[9] The complexities introduced by the inter-related nature of the building and operating phases of public service provision suggest that similar difficulties arise from the interdepence of defence activities. This interdependence means that considering separate make-or-buy decisions for each activity in isolation, as each of the approaches considered here does, ignores the interdependence of the activities which may be an important determinant of where the true limits to the use of the private sector outsourcing lie.

2.2.4 Hybrid forms, partnerships and relational contracting

Hybrid forms

Until now we have considered a straight choice between in-house provision and outsourcing. Between these extremes lies a spectrum of possible organisational forms, such as joint ventures and partnerships, combining the public and private sectors. In fact, Williamson [223, p83] suggests that most transactions take a form between the extremes of market and hierarchy.

For many defence activities, where long term relationships are the norm; where uncertainty requires adaptation throughout the contract; and in which investment games are repeated, it is no longer clear cut that ownership should lie with one party or the other [68, 79]. In contrast to the conclusions of of PRT it may be that neither outsourcing nor in-house provision are optimal, but that some joint ownership or partnership arrangement is preferable. In some circumstances joint ownership is to be preferred: if different bargaining solutions are applied [42]; if it is important to ensure that parties invest inside the relationship rather than outside [168]; or if bilateral veto power is useful [172] as may be the case in research joint ventures. It may even be that in long term relationships the importance of reputation (through the value of an ongoing relationship) reduces the problem of hold-up, and that reputation and ownership become substitutes for providing incentives for investment [68].

Relational contracting

The conclusions of PRT and (implicitly) the public choice approach rely on the use of formal contracts, but there are limits to the effectiveness of formal contracts. Their use is efficient if the seller is able to control quality at low cost and if the probability of a breach is low [129], but the enforcement of formal contracts relies upon the use, or threat, of court action to claim damages for poor quality or non-delivery. Many activities which are undertaken on, or close to, the battlefield are unsuited to procurement through formal contracts because neither the costs of controlling quality nor the probability of default are low.

The extreme risks which may be faced in executing a contract for outsourced battlefield activities, the importance of timely delivery, and the interdependent nature of defence activities mean that for many activities the consequential damages which are incurred by government through non-execution of a contract are out of all proportion to the value of the contract. This problem is exacerbated by limited liability and the corporate structures (such as Special Purpose Vehicles/Entities) seen in the delivery of public service contracts, which limit the effects on the parent company of default on any single contract.

These problems have been overcome in the Armed Forces by using military labour contracts, which are almost contracts of slavery. However, private firms cannot engage labour on such terms, and contracts of slavery cannot be used to contract with organisations.

If the choice is between in-house provision (able to use the unique contracts) or outsourcing to a private firm using a formal contract, then in-house provision has a clear advantage in activities which are subjected to the chaos and uncertainty of the battlefield. This advantage results from the limitations of formal contracting—the mechanisms of control over in-house providers are different to those over contracted provision, and sometimes these differences are important.

In situations which involve long term repeated relationships in which reputation and trust are important, the possibility of relational contracting permits an alternative to formal contracting.

Relational contracting requires only that outcomes are observable to both parties (as distinct from observable and verifiable required for formal contracting). Relational contracting reflects the informal mechanisms used in the real world to deal with the incompleteness of formal contracts, and emphasises the dynamic nature of relationships in defence outsourcing—something not well captured by PRT. Contracts are self-enforcing through the value of future relationships. Future relationships may be sufficiently valuable to both parties to ensure that neither wishes to renege, but flexible enough to allow adaptation [10]. This adaptation is crucial in the procurement of military equipment which frequently cannot be tightly specified in advance and so specifications continue to evolve after the initial contract is signed [145].

The use of relational contracting for the provision of public services is considered in Parker and Hartley [157]. Adopting a TCE perspective, they argue that the use of reputation and trust in relationships can reduce the contractual hazard of uncertainty. This reduces the need for formal, detailed contracting so reducing the transaction costs of long term contractual transactions. The implication is that the use of relational contracting extends the range of functions that can be efficiently outsourced, though it does not remove the need for formal contracting entirely as parties have differing objectives. Partnerships are seen as a mechanism for reducing contracting costs, though if formally contracted and not self-enforcing it is hard to see how the problems of hold up and opportunism would be avoided.[10] There is also a risk that partnerships may, in ensuring bilateral monopoly relations endure, restrict the incumbent's future competitive incentives sufficiently to outweigh any initial advantages [156]. The Defence Industrial Strategy makes it clear that in the UK procurement of a wide range of defence activities will in future be procured by 'partnerships' with industry rather than competitively procured in order to retain key industrial capabilities onshore [136]. This represents a shift from the focus on competitive procurement toward noncompetitive partnership based procurement, which means that advantages of outsourcing due to competition (such as those suggested by the Public Choice approach, as well as the implicit competitive auctions of the PRT) can no longer be obtained in many

[10]Molas Gallart and Tang [138] provide some empirical evidence which questions the viability of partnership arrangements between the public and private sectors. Examining conflict that has arisen over intellectual property management during the partial privatisation of the Defence Evaluation and Research Agency they conclude even where there is a small and close knit community and a government policy emphasising public-private partnerships, private industry and governments emerge as clearly separated and, at times, conflicting entities.

areas of defence procurement.

Other more formal approaches to relational contracting have been developed, based on self-enforcing contracts. Baker *et al.* [10] propose a model in which firms have a choice whether to contract on a spot market or with relational contracts, and may adopt either form of contracting within firms and between firms. In such a model the allocation of ownership affects the set of feasible (self-enforcing) relational contracts, through its impact on incentives to renege on relational contracts. In some situations incentives to renege are lower in integrated relationships, in other situations higher. In particular, the military labour contract eliminates the ability of military personnel withdrawing labour if faced with an exceptionally dangerous task: such contracts cannot be signed by private contractors and so face the risk of desertion by their employees. In such circumstances a contractor may have no choice but to default on a contract. Such a model offers a new perspective on the issue of provision in that in different circumstances vertical integration may sometimes facilitate better relational contracts.

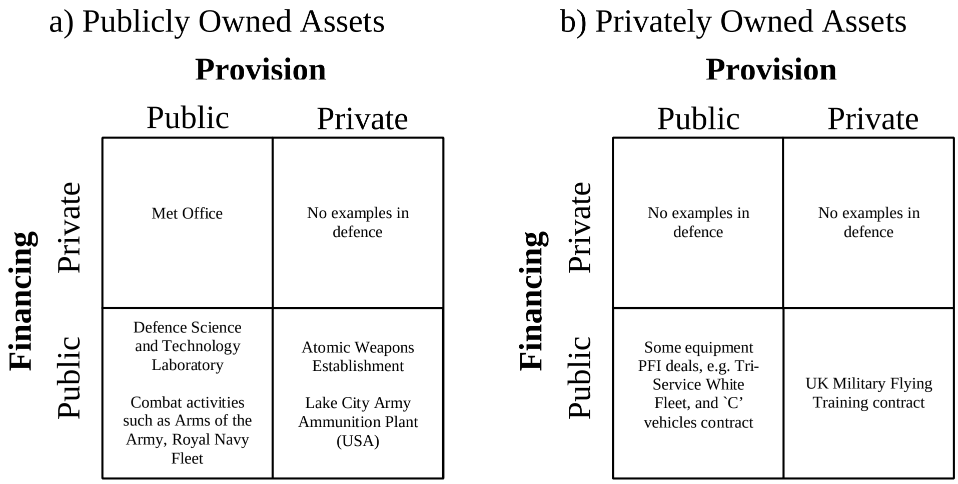
Baker *et al.* [11] consider a related model, but focus entirely on the uncertainty and adaptation aspects of a relationship with no specific investments. Ownership and decision rights are separated to allow hybrid forms and the *ex post* Coasian renegotiation assumption of PRT is relaxed. The resulting optimal governance structure is that which minimises the maximum aggregate temptation to renege, because the party in control simply take self interested decisions. This criterion of the maximum temptation to renege is interesting for military outsourcing, since some defence functions have a high temptation to renege. This is particularly relevant for those services on the battlefield which occasionally face the prospect of catastrophic losses to equipment or personnel, such as contractor support for equipment which might occasionally be required to operate in dangerous circumstances, where the effectiveness of formal enforcement is in doubt. The introduction of relational contracts seems to suggest a new criterion by which the limits to effective outsourcing might be determined, with those that might have high temptation to renege to be provided in-house.

2.2.5 Discussion

Economic theory does not offer clear guidelines as to which defence activities might be most successfully performed by the private sector and which should be publicly produced. By focusing on different aspects of the military outsourcing make-or-buy problem, each of the approaches considered above offers something on the issue of where the limits to defence outsourcing lie, but none appears to offer a complete analysis of the problem.

Figure 2.2 suggests a taxonomy of the ways in which defence activities may be provided. For many activities there are three aspects of the activity which may be either publicly or privately performed: financing, provision or operation of a service, and ownership of specific assets required for service provision. Each of the theoretical approaches considered above focuses on one of these aspects. Being a public good, there is little scope for pri-

Figure 2.2: A taxonomy of the role of private provision of defence



vate financing of defence activities. It is not generally feasible for the ultimate users to be expected to pay directly for military activities, (though the Met Office provides an exception because it produces something which is both militarily and commercially valuable), so public financing is the norm.

TCE is not much concerned with where the ownership of specific assets lies, their existence is sufficient to ensure the fundamental transformation. TCE focuses on the choice between public and private provision or operation of a service. It focuses on adaptation in small numbers contracting and the contractual hazards of frequency, uncertainty, and asset specificity. These suggest a broad distinction between inputs and outputs, though perhaps not as clear cut as suggested. Adding the additional hazard of probity suggests a distinction between combat and non-combat distinction.

Conversely, both PRT and the relational contracting approach take private provision as a given, and focus on public and private ownership of specific assets. The focus of the relational contracting literature is the weakness of formal contract enforcement and the role of reputation. Its conclusions are closely related to those of TCE. Reputation and trust play a role in reducing the contractual hazard of uncertainty, so expanding the role of contracting in defence beyond that suggested by TCE. Self-enforcing contracts also suggest that those activities with a high temptation to renege on the contract should be kept in-house. This is a very similar idea to the hazard of probity, and broadly supports a distinction between combat and non-combat roles.

Since each approach focuses on a different aspect of the problem, their conclusions are complementary, with each suggesting criteria against which the potential for gains from outsourcing might be judged. The focus of PRT on incentives for *ex ante* investment leads to reasonably specific recommendations, which suggest that a much smaller set of activities are retained in-house than under TCE. This is however likely to be too small a set, given the importance of adaptation and renegotiation which are not really focused on in the PRT models.

Overall, the approaches are reasonably consistent in that each broadly concludes that some defence activities are best outsourced, and some are best provided in-house, but the only concrete conclusions to be drawn from the literature are about the activities at the extremes. There is little useful guidance as to where the limits lie between these extremes. Perhaps the most useful contribution of economic theory is highlighting some of the factors relevant to the government's make-or-buy decision.

2.3 Empirical literature

The basic rationale for the involvement of the private sector in the provision of defence, by whatever mechanism, is that contractors are able to produce the same specification of service at lower cost than the existing in-house team. The empirical literature is not, for the most part, concerned with precisely why such savings might arise and has not focused on testing the theoretical considerations discussed in the last section. It has generated some evidence with which to test the public choice notions of competition, and TCE has generated some empirical work, but PRT and relational contracting are difficult to test empirically as they rely on investments and efforts that are difficult to quantify.

The broad issue of outsourcing public services has generated a vast empirical literature. This is partly because economic theory offers few conclusions besides identifying a number of relevant factors; partly because it has focussed on whether cost savings have been achieved in those activities which have been outsourced and these are, in principle, measurable and this is a question that lends itself to empirical analysis; and partly because the political issue of the involvement of private firms in the delivery of public services has generated a great deal of ideological heat [28, p701]. However, it is frequently the case that the information required to undertake a rigorous independent assessment of the costs and benefits of outsourced contracts is simply not available [216].[11] Claimed savings are frequently projected [85] or based on comparison against the hypothetical (e.g. the Public Sector Comparators used for PPPs in the UK).

This section first reviews in detail the limited literature on outsourcing specifically defence activities. It then reviews the broader empirical literature on outsourcing public services, and its implications for defence.

2.3.1 Defence specific empirical literature

In spite of the large literature on the general issue of outsourcing public services, there is little specific to defence. There are a number of reasons for this: i) Until relatively recently

[11] Such information is also sometimes unavailable to even the contracting authority [210, 143, 141]. For example, comparison of the FSTA PFI contract and current arrangements is impossible because the MOD did not record costs of in-house provision in a way that enabled comparison with the proposed contract [145].

defence ministries have generally restricted their outsourcing to activities similar, if not identical, to those being outsourced elsewhere in the public sector, and so conclusions from studies in other areas applied straightforwardly. ii) Defence specific, and especially military, activities tend to be conducted only once, or at most a few times, in each country (generally either collectively by the defence ministry or by each of the Armed Forces). The type of data this generates—a small number of observations of each of a large number of different activities—does not lend itself easily to econometric analysis, and leads to difficulties in producing an adequate benchmark for the counterfactual form of provision. iii) Perhaps the greatest difficulty in empirical work on this issue is to ensure that when comparing forms of providers one is comparing like-with-like (this issue, discussed at greater length in Section 2.3.2, is complicated by the uncertainties inherent in many defence activities and the importance of quality).

The latter two reasons mean that much of the evidence specific to defence is based on case studies as this is the most straightforward way of avoiding the need to compare activities with very different characteristics. This does however, limit the ability to apply their conclusions more generally for fear of the conclusions being case specific. The papers considered below are detailed in Table 2.1.

The first defence activities to be outsourced in the late 1970s and 1980s were essentially civilian activities, through schemes such as the MOD's programme of market testing to a lesser degree DoD's A-76 competitions for commercial activities were essentially civilian activities. There is now good evidence that these programmes resulted in sizable cost savings. Perhaps the most compelling evidence of cost savings comes from other areas of the public sector, and this evidence is considered in the next subsection. Studies of early defence outsourcing initiatives also report evidence of savings 'of between 26% and one third for defence-related services in Canada, United Kingdom, and Australia' [8, 105]. The Australian DoD's Commercial Support Program yielded estimated savings of 36% [48], although it represented only a fraction of the defence services contracted. Based on only 6 case US studies, Gates and Robbert [69] report large savings on personnel costs (between 34% and 59%) arising from competition for defence services. In a much broader analysis of over 2000 US DoD competitions for commercial defence services between 1978 and 1994 Snyder *et al*. [189, 190] found savings of around 30–35%. The larger part of these savings (76%) result from contracting with private contractors, rather than the incentive effects of competition on the in-house team (24%). Savings of a similar magnitude in the same type of DoD competitions were found by Gansler [65] who concludes that *expected* savings averaged around 31% between 1975 and 1993, rising to 42% between 1994 and 2001, though whilst in agreement with other studies these figures appear to be based on a small number of case studies.

Whilst unlikely to overturn the evidence of large savings, these estimated savings from these programmes may be overstated for two reasons: Firstly (and this criticism may be leveled at almost any study of this problem but is no less valid for it), it is not clear how well they ensured that the specifications of the work conducted before and after competition are

Table 2.1: Summary of the defence specific empirical literature

Study	Savings reported	Country	Sample	Comments
Domberger *et al*. [48]	38%	Australia	Analysis of the Commercial Services Program	Savings resulted from reduced personnel and increased use of civilian labour. Changes reduced surge capacity.
Industry Commission [109, Appendix F]	38%	Australia	Case study of F-111 aircraft maintenance	Savings largely from reduced labour requirements. Unable to disentangle the effects of reduced surge capacity and efficiency.
Gates and Roberts [69]	30–60%[a]	USA	6 case studies of A-76 competitions	Savings are real and enduring where contractors used, but unverifiable for in-house teams.
Snyder *et al*. [189, 190]	30–35%		Over 2000 A-76 competitions 1978–1994	Competition 24%, use of private rather than public providers 76%
Domberger *et al*. [51]	24%	New Zealand	Case study of vehicle maintenance and warehousing	Contract offered savings of 35%, but reduced in first two years by renegotiation of original contract.
Kirk and DePalma [113]	Unknown	USA	13 case studies of PBL contracts	Insufficient information to judge performance of contracts.
Gansler [65]	33%	USA	5 case studies of competitive sourcing	Improved performance from outsourcing. Concludes from other studies that average savings were 31% during 1975–1993 and 42% between 1994–2001
Gansler and Lucyshyn [66]	44%	USA	Over 1100 A-76 competitions 1994–2004	Much of savings derived from reducing workforce: average reduction of 38%. Only 5% of jobs competed resulted in involuntary separation.
Gansler and Lucyshyn [67]	58%[a]	USA	Case study of support services to Air Force Base	Public Team won bid. Labour costs reduced 58%.
Kulmala *et al*. [115]	13%	Finland	Case study of transport equipment management	Costs rise initially then fall below in-house level.
National Audit Office [141]	Significant	UK	Case study of support for fast jets.	Contracting for availability has resulted in improved availability and significant savings.
National Audit Office [143]	Unknown	UK	8 case studies of defence PFI projects	Generally on time and budget, but concerns over management of changes and data availability.

Notes: [a] savings are to personnel costs only the same. Secondly, a large part of the savings arise from using fewer personnel to provide the service in question [66, 69], but it is not possible to know what happened to the DoD labour freed up by the process, and so whether personnel savings are translated into savings for the overall defence budget.[12]

More recently the use of private contractors in military activities has expanded far beyond the scope of the earliest programmes. Contractor provided services are now essential to the functioning and deployment of the Armed Forces of both the UK and the US. There is far less evidence on savings resulting from contractors undertaking these, sometimes uniquely military, tasks, and their nature is such that the evidence on the initial programmes cannot be readily applied.

Three studies have performed case studies of maintenance facilities and programs (Kulmala *et al.* [115] in Finland; Domberger *et al.* [51] in New Zealand; and Industry Commission [109, Appendix F] in Australia) . These find savings from the use of private providers of 13%, 24% and 38%, respectively, each concluding that quality of service is either maintained or improved. These savings result largely from more flexible use of resources, especially labour. A key source of personnel savings was from using civilian labour instead of military: civilian labour is not only cheaper, but is not required to participate in exercises and training (this can consume up to 40% of military labour's time [51]).

Although all three maintenance studies are clear about cost savings achieved, they raise three issues which may have wider relevance: Firstly, all indicate significant transactions costs associated with ongoing costs of monitoring and administration (as well as bidding and transfer costs in Industry Commission [109, Appendix F]). Although these costs are incurred under both in-house and outsourced provision, being very difficult to measure they are accounted for in very few empirical studies. Secondly, they provide evidence of renegotiation of the original contract reducing the level of anticipated savings. Renegotiation is unavoidable in long term contracts (and changes also occur under in-house provision), even when contracting for the support requirements of well understood systems, but it does impose further administration costs [115] and risks eroding the possible savings from contracting-out. Domberger *et al.* [51] found that, consistent with the theoretical predictions of hold-up, renegotiations in the first two years of a six year contract reduced savings from 37% under the original contract to 24%. Finally, it is hard to distinguish whether savings are being achieved through the elimination of organisational slack, or reduced surge capacity. Prior to contracting the Australian F-111 maintenance program worked with excess capacity, but subsequently have worked with limited excess capacity, and sudden increases in workloads may be difficult to cope with [109, Appendix F]. If some of the savings identified empirically are due to reduced surge capability, then comparison of in-house and contracted provision is no longer comparing like-with-like. The nature of many defence

[12] Gansler and Lucyshyn [66] find that although DoD competitive sourcing competitions result in an average 44% reduction in labour requirements, only only around 5% of postitions are involuntarily separated. The remainder are given voluntary redundency or redeployed. The 58% personnel savings reported in Gansler and Lucyshyn [67] resulted in no redundencies, but rather the reassignment of 848 military personnel.

services is such that separating peacetime and surge requirements is difficult [33]. Whereas the military tends to maintains surge capacity permanently, contractors rely on contract labour to provide surge capacity and potential capability is hard to assess until tested [33].

Evidence of the relative cost performance of more recent forms of contracted provision such as contracting for availability (Performance Based Logistics (PBL) contracts in the US) and the PPP/PFI arrangements for equipment services is sparse. The National Audit Office (NAO) [143] review ten case studies of MOD PFI projects to deliver defence services and find that generally the services delivered are satisfactory, and delivered on time and on budget, but no comparison is made against other forms of provision, or even against the Public Sector Comparator. When reporting on the moves towards Contracting for Availability in support for fast jets they [141] report that even the MOD lacks the demand and usage data to make a comparison, or even to negotiate affordable contract prices. A review of 13 US PBL contracts [113] finds that the amounts specified in contracts underestimate the actual expenditures, but also that there is that there is insufficient data to determine whether these contracts are actually fulfilling their objectives.

Kirk and DePalma [113] also conclude that the monitoring of the performance of these contracts is inadequate. This is a recurring theme in the literature, and there are serious doubts about the ability of defence ministries to manage relationships with contractors. If cost savings are to be achieved robust contracts must be designed, implemented, and effectively monitored by the public sector [32, 49].[13] Many of the problems arising in efforts at outsourcing arise because government failed to properly oversee the contracted out function is performed effectively and efficiently [65]. In the US the GAO has frequently raised concerns over the ability of the DoD to oversee and manage service contractors recently stating that 'the lack of well-defined requirements, difficulties employing sound business practices, and workforce and training issues hinder efforts to effectively manage and oversee contracts and contractors. These factors ultimately contribute to higher costs, schedule delays, unmet goals, and negative operational impacts' [214] (see also, amongst others, [208, 211, 212, 213]). Similar concerns exist in the UK: the NAO [141] recommended the MOD 'assess whether it has sufficient commercial, cost modelling and project management skills to develop the commercially viable support solutions and negotiate contracts, given the increasing complexity and likely volume of industrial logistics support', and suggested that management of PFI projects could be improved and was creating risks to value for money [143, p8].[14] As well as official reports there is also anecdotal evidence of poor contractor management of contractors (see for example Simons [182] on US difficulties in Iraq, Mobley [137] on the Balkans, and Guistozzi [77] on Afghanistan).

Although it is easy to cite anecdotal evidence of performance failures of contractors,

[13]This is as true for partnership arrangements and other novel contracting arrangements as it is for traditional contracting, since what is formalised in contract which serves as the basis for negotiation or legal dispute should those arrangements or partnerships break down.

[14]The concern over project management of outsourcing contracts reflects the concerns raised for decades over management of defence procurement, especially of major weapon systems. See, for example, [74, 103, 145, 139], or almost any *Major Project Report* such as [140].

such problems also arise under in-house provision. Recent deployments have highlighted failures of in-house (military) provision of logistic support in both the UK [144] and the US [158]. Problems will arise under any form of provision, that is the nature of the military environment. However, much of the rationale for the expansion of private provision of defence services is expected cost savings. The concerns over ineffective monitoring and management cast doubt on how likely expected savings will actually be realised.[15] Unlike the theoretical considerations reviewed in Section 2.2, the empirical evidence suggests that choice facing government is between two imperfect solutions to the make-or-buy decision: The choice is not between ideal competitive regime and imperfect in-house arrangements. Nor is it between ideal in-house provision and imperfect competition, but rather the real choice is between two imperfect forms of provision.

Overall, there is good evidence that cost savings have been seen from the earliest initiatives for contracting-out defence services, though the estimated savings may be overstated, but there is paucity of evidence on the relative cost performance of more recent initiatives. The theoretical approaches suggest that relatively simple activities will yield savings, and this is confirmed by the empirical evidence, but is it not currently possible to assess whether savings are also being achieved as the activities outsourced become more uniquely military in nature.

There is no clear evidence on whether more recent efforts have yielded cost savings or not. Whilst the activities now being outsourced, because of their complexity, offer greater scope for innovation than activities such as cleaning and catering, this scope for innovation also represents scope for quality shading, and so more recent activities entail greater transaction costs. One might expect the MOD to first outsource those activities which it felt offered the greatest scope for savings—cherry picking [215]—and that subsequent outsourcing will yield progressively lower savings until when no savings are achieved the limits to efficient outsourcing have been reached. The empirical evidence is not sufficient to assess the validity of this argument. It is not possible to assess whether recent initiatives have yielded lower savings than the initial efforts, or even whether they yield savings at all. Although one might suspect there are economic limits to the scope of outsourcing in defence, it is not possible to determine where they lie.

In 1993 Uttley [215] raised three questions which remain pertinent and largely unanswered: (*i*) What are the costs of administering and monitoring the contractorisation/market testing processes (the analagous costs for in-house provision are also not well understood)? (*ii*) Does contractorisation lead to a decline in quality? (*iii*) In comparing public and private bids for activities, is like being compared with like?

There is scope for more work on the first of these questions, with TCE suggesting the transaction costs are the determinant of the limits of efficient outsourcing, but management costs are also incurred under in-house provision. Whilst there is some evidence of the costs

[15]Because management issues are endemic they should be accounted for in the comparison of options for provision, perhaps using a methodology similar to the Optimism Bias adjustment used in the assessment of large capital projects [97].

of monitoring and overseeing private contracts, there is no work on the costs of in-house teams. This is an important gap, as it is the relative size of the monitoring and bidding costs which is most relevant—i.e. high transaction costs associated with outsourced contracts are only relevant if the in-house provision of the same activity is relatively cheap, and vice versa.

The second question remains important, but it is not likely to prove possible to address completely in defence activities. Quality is multi-dimensional, and if extensive changes are made to an activity when switching to private provision, some aspects of quality may increase, some may decrease, but even if all aspects of quality are measurable one still requires an weighting system to reduce quality to a single comparable measure.

The third question is relevant to defence, especially the issue of whether when changes are made as provision switches to private provision any cost savings arise from true efficiencies, or rather from reduced surge capacity. This issue is likely to be especially relevant in logistic and maintenance activities which now see widespread involvement of the private sector. Although general conclusions on this issue will be hard to draw, if sufficient information is available detailed before and after case studies may go some way to providing an answer.

A fourth pertinent question, arising from TCE and the maintenance case studies is how far renegotiation of operational contracts erodes savings expected when the contract was signed? This question could be answered using detailed case studies (as was done by Domberger *et al*. [51]) but these must be conducted on completed, rather than ongoing, contracts. There must, however, be a distinction drawn between changes instigated by the contractor (hold-up) and those instigated by the government changing specifications. When renegotiation is caused by government, the changes in costs which result must be compared against those costs that would have been incurred under an in-house team.

The four questions above focus on detailed assessment of individual contracts, which requires extensive use of data and information which is difficult to obtain. There is also scope for empirical investigation of the broader outcomes of the rising use of private provision of defence output. This might be done using readily available aggregate data.

2.3.2 Empirical evidence from areas other than defence

Whilst there has been relatively little work in defence, there is a voluminous empirical literature on the use of private firms to provide public services in other areas of the public sector. By way of illustration, Hodge [100] finds 129 empirical studies by 2000. A comprehensive review of this literature is beyond the scope of this chapter. Instead, this section relies largely on surveys and reviews. Most of the evidence relates to pure contracting-out of largely labour based activities which do not require large capital assets for their provision. Given this most of the evidence considered may be applied most readily to the earlier efforts at contracting in defence.

In any empirical comparison of providers of public services the two key issues of cost and quality are inextricably linked: whether cost changes are the result of improved efficiency cannot be assessed without controlling for differences in quality. On costs there is a consensus that contracting for public services has resulted in cost savings. There is, though, considerable variation in estimated savings [28, 50, 94, 109], and there remains some debate as to their magnitude [110]. In a meta-analysis of 135 estimates of savings Hodge [100] found an average 11% saving from the use of contractors, but estimates range from savings of 80% to cost increases of 290%. Of the reviews considered, only Boyne [29] finds the evidence of cost savings unconvincing because he considers the evidence fatally undermined by the issues discussed below.

The evidence on whether contracting increases, decreases or maintains the quality of service provision is less clear cut, and examples of each have been seen [109]. Although lower costs are not, as a general rule, accompanied by poorer quality service [100], there is evidence that in some circumstances quality may suffer [110]. Whether lower quality is important depends on the preferences of society, but the literature has not clearly identified the circumstances in which quality deterioration might be expected to happen. Domberger and Jensen [50, p75] suggest that it 'could be a problem of contract design and therefore preventable', but this rather ignores the nature of quality which is that is it difficult to measure and therefore difficult to specify in contract, though similar difficulties arise with in-house provision.[16]

Although the broad consensus is that cost savings can frequently be achieved without reductions in service quality, there remain a number of unresolved issues raised by the empirical literature which have some bearing on the potential for extending the use of private provision of defence activities.

Difficulty of ensuring comparing like-with-like

A lack of data on the non-contractible aspects of the service provided, the difficulties of dealing with joint product outputs, poor measurement and collection of information prior to tendering competitions, and changes made to service specification when tendering all hamper empirical work's ability to ensure that when comparing public and private provision, one is comparing like-with-like [50, 110].

The primary difficulty is determining whether observed cost changes result from true efficiencies, or from quality shading. This affects all empirical analyses to some degree. Measures used to control for quality, only partially capture it (e.g. the frequency and location of collection in studies of refuse collection [52, 198]). Partial measures of quality can lead providers to focus on measurable aspects of quality at the expense of unmeasured aspects—as seen following the introduction of competition in the NHS where waiting times have reduced at the expense of unmeasured and unobserved quality [167]. Others have compared standards before competition with the service specified in contract rather than

[16] Even if quality reductions are, in principle, preventable the evidence of poor contract management in defence ministries discussed in the Section 2.3.1 above suggests that this cannot be relied upon when comparing forms of provision.

the service actually provided [29]. Whilst this is unsatisfactory, the measurement of actual quality is subjective, and is likely to rely upon the (possibly partisan) views of service managers. Even if quality differences could be accurately captured, any changes in cost and quality must be compared according to society's preferences if a judgement as to which form of provision is to be reached.

One area in which public and private provision do differ is the mechanisms through which contractors and in-house teams are controlled. Contractual mechanisms may reduce control over quality of service because of the difficulties of specification. Real time control can be crucial in defence as services move towards the battlefield. Contractors are outside of the conventional military chain of command, and the weaker mechanisms to ensure timely delivery of a contracted service may affect operational effectiveness [33]. Similar issues have been seen in hospital cleaning where 'contracting-out of cleaning appears to have made it more difficult for ward managers and matrons to control' [104, para 16] because contracted cleaning is controlled by central contract rather than being directed by those at ward level.

Comparisons must use appropriate measures of cost

Some of the costs arising from contracted service provision, notably the transactions costs of tendering, monitoring and oversight are difficult to measure and have not been well captured in existing empirical work [28, 29, 100, 109]. As mentioned in the discussion of the defence maintenance case studies above, accurate comparison requires all relevant costs are included. Estimates of these transactions costs under contracted provision vary from 2–3% up to 20% [100], and most falling at the lower end of this range. Omitting costs of this order is unlikely to overturn the broad conclusion of savings from contracting but will cause savings to be overstated [109]. Of course, management costs also occur under in-house provision, but there is no evidence of their likely size. Transactions costs (particularly the costs of monitoring) will increase as the importance of noncontractible elements of the service under consideration rises. In defence, this is likely to occur as one moves close to the uncertainty and chaos of combat, and as more resources must be devoted to overcoming control issues eventually transaction costs are likely to outweigh potential savings. In these circumstances in-house, military providers' use of military employment costs allows them the economise on the transactions costs of ensuring control.

For PPP arrangements the transactions costs arising from the 'need to find the right private sector partner, and to negotiate, monitor and renegotiate a long-term contract giving him the right incentives to strive for service quality while containing costs' [217, p115], mean that PPPs are dearer to set up and follow up than in-house provision. In the UK it is estimated that the transactions costs of PPP arrangements for the procurement phase alone (that is excluding any monitoring and renegotiation costs) incurred by government are between 1% and 7% of the capital value of the project, but including the costs incurred by bidders (both winners and losers) gives a total figure of 'well over 10%' [55, p13]. For PFI projects in the UK the cost of advisors alone averages 2.6% of a project's capital value

[142], reflecting the difficulties of negotiating long term service agrements. Similar costs are likely to be incurred in negotiating outsourced services, especially as those activities being outsourced become more complex.

A second issue with expenditure data is that many studies compare actual costs of service provision before competition with the estimated costs afterwards [29], as opposed to realised costs. With more recent contracts being up to 30 years this is unavoidable, it is not clear that estimated costs will actually be realised since renegotiation (be they due to hold-up or changing requirements) must be expected. This is a particular problem for PFI/PPP arrangements and, to a lesser degree, Performance Based Logistics contracts because of the long contracts required to justify specific investments.

Savings are not uniformly distributed across activities

Although average savings are around 8% to 14% over all services [100] and figures of between 10% and 20% have been widely cited, there are large differences between estimated savings across activities. These differences are seen in Table 2.2. The most significant savings are concentrated in a few key areas—notably cleaning, refuse collection and maintenance—the areas which have seen the largest number of studies, so biasing (upward) the apparent average saving [29]. That the most frequently studied areas are those offering the greatest scope for savings is no accident (and not only because researchers seek significant results). Researchers face similar problems as contracting bodies in having to control for quality. Those areas which are easiest to control for quality have seen the greatest amount of work, but are also the areas in which quality is easiest to specify in contract, and so offer the least scope for quality shading. That savings are greatest where there is least scope for quality shading is consistent with the broad propositions of the theoretical literature.

Table 2.2 clearly shows that there is no generally applicable rule of thumb as to the level of savings which might be achieved, rather each activity (or even each contract) contains unique characteristics which determine the probability of achieving savings [109, p127]. However, the literature has remarkably little to offer as to which factors or characteristics drive the differences between activities. It appears that those activities seeing the largest savings are those which are relatively simple to specify, so offering the least scope for hold-up and renegotiation, and for which there are potential suppliers.

Boyne [28] suggests the causes of variation are: the market power of producers; whether revenue is raised directly from the public or centrally through subscription; the strength of competition between political parties; and the intensity of competition between public authorities. That military activities are funded centrally, produced only once in each country, and in markets where firms have significant market power suggests limited scope for the public choice type competitive forces producing savings in military activities. Civilian activities which were the focus of the earliest outsourcing efforts offer the greatest scope for competitive pressures, and this may have driven the savings seen.

Table 2.2: International evidence of savings from contracting by category of service provided

	Number of % estimates available	Average % cost saving for category	Significance level of effect
Cleaning	7	-30.2	<.0001
Corporate Services	1	+24.1	ns
Engineering Works	2	-24.2	ns
Fire	1	-17.6	ns
Health	1	+3.5	ns
Maintenance	3	-30.5	<.0001
Multiple	1	-19.9	<.0001
Other	3	-5.9	ns
Parks & Recreation	2	-7.5	ns
Police & Security	2	-8.5	ns
Refuse Collection	13	-19.3	<.0001
Training	2	+0.9	ns
Transport	2	+16.1	ns

Source: Hodge [100] Tables 7.6 and 7.7.

Water treatment is an area not considered by Hodge [100], but which shares a number of characteristics of some defence activities: it requires continual investment in specific assets, features long term contracts and limited competition (competition is for the market monopoly and there are very high rates of contract renewal because of the informational advantage the specific assets confer on the incumbent). Even though there is little scope for quality shading because of detailed statutory water quality requirements, Bel and Warner [17] find the overwhelming majority of studies find no cost difference between public and private providers (3 of 16 studies considered found private savings) in their review of the international evidence. Water treatment does not have the same degree of uncertainty as some defence activities, but the evidence in Bel and Warner [17] suggests that there is more to achieving cost savings through the use of private producers than controling quality-shading.

Do savings result in long term reductions in expenditure?

The transactions cost literature suggests that, when committed to long term relationships involving specific assets, opportunistic behaviour by firms may result in hold-up. The result is that expected savings are gradually eroded through successive renegotiation. A similar situation may arise if competitive bidding results in the winning firm submiting too low a bid (the so called winner's curse) and then requiring an increased settlement in order to avoid bankruptcy (something the procuring authority has little choice but to acquiesce to if it no longer has the capacity to provide an in-house service). If expected, rather than actual, expenditures are used to estimate cost savings, any erosion of savings over the course of the contract caused by renegotiation and hold-up will not be detected.

The empirical evidence supports the existence of short-run savings, at least in some areas, but there is less evidence on whether such savings are permanent. Jensen and Stonecash [110] suggest that the savings for outsourcing may be transitory, but this is not something which has been examined in detail. Direct evidence on the hold-up hypothesis is thin, but the case studies of defence maintenance contracts noted above are consistent with the existence of hold-up. Some evidence of reduced savings over time is to be found in the refuse collection literature: Szymanksi and Wilkins [199] and Szymanski [198] find the trend for initial savings from competitive tendering reversed around four years after contract award coinciding with the first renegotiation of the contracts. This is consistent with initial underbidding, but the effect reduces rather that reverses initial savings. Bel and Costas [16] find that more recent competitions contracts yield greater savings than those signed before 1990, but there is no way to determine whether this is the result of savings being eroded over time (as they suggest), or due to improved contracting and tendering processes.

There has been little work to determine whether the savings seen in expenditure on individual activities are being passed back to the taxpayer through lower overall expenditure, or are retained by public authorities [28, 29]. The focus of the literature has been on individual contracts rather than the effect of the contracting process on overall public expenditure. There is some limited evidence that savings do not result in lower expenditures, but rather are absorbed within the organisation [1]. Some evidence of this was found in the defence literature in Section 2.3.1, with a lack of clarity over whether labour reductions are reassigned elsewhere, or result in lower overall employment. As Hodge [100] notes there is clearly scope for more work on this issue.

What are the sources of savings?

The empirical evidence strongly supports the key public choice proposition of cost reductions resulting from the introduction of competition into the provision of public services. In achieving cost savings 'the role of competition cannot be overstated' [100]. Domberger and Jensen go so far as to state that 'there should be no doubt that as long as *ex-ante* competition (or the threat of *ex-ante* competition) is maintained, efficient outcomes should be obtainable' [50, p75]. When the effects of competition are distinguished from those due exclusively to the form of provision used, much of the observed cost reduction is attributable to the competitive effects of the contracting process rather than the organisational form [18, 52, 198, 199]. Much of the evidence for this comes from studies of refuse collection. The evidence for greater savings from the use of private contractors rather than in-house providers is weaker: some find significantly higher (10% higher) savings from contracts awarded to private firms [18, 198], but others find no significant difference [52, 44, 199].

A limitation of empirical work on the effects of competition has been that competition has been considered a binary variable—contracts are either competitively tendered and so competitive, or not tendered and so uncompetitive. Instead, one might expect that the potential for savings is larger where there is more intense competition for contracts. Dijkgraaf and Gradus [45] included measures of market concentration as a proxy for the degree of local competition for contracts concluding that private collection is between 10% and 20% cheaper than municipal collection on average, but this difference depends crucially on the level of local competition and in the presence of local monopoly private collection may increase costs. This role for the intensity of competition has implications for some defence activities, where the market is oligopolistic at best.

If the savings seen in empirical work are truly due to competitive incentives, then defence activities for which effective competition cannot be arranged offer less scope for savings from outsourcing than have been seen from the earliest programs. Competition is difficult to arrange in some activities because of the small number of potential providers, and the large capital investments required. A further difficulty in PPP type arrangements is that the high tendering and bidding costs arising from the complexity of these contracts limits the power of *ex ante* competition to provide incentives for efficient production [217].

For example, the recent Future Strategic Tanker Aircraft (FSTA) PFI contract attracted only two bidding consortia one of which withdrew before the contract was let [145]. This resulted in four years of noncompetitive negotiation with the remaining consortium because the MOD was committed to a privately provided PFI solution, but without being able to engineer the effective competition necessary to obtain the benefits suggested by the Public Choice literature, and implicitly assumed by PRT and to a lesser degree TCE. As the defence activities being outsourced become both more idiosyncratic and requiring large capital investments in specific assets, so the difficulties of arranging effective competition increase. The Defence Industrial Strategy retaining certain key industrial capabilities in the UK, effectively creating national champions, exacerbates the difficulties of arranging competition by committing the MOD to buying from these national champions. In these areas Williamson's fundamental transformation occurs before the contract is signed.

Boyne [28] distinguishes contestability (decision to seek external tenders) from competition (receipt of external bids). This is perhaps a useful distinction for defence, as contestability may be possible to arrange in areas where effective *ex ante* competition could not occur. When dealing with small numbers of potential providers, contracting for an activity removes the possibility of in-house teams competing for future competitions as capabilities lost, reducing effective future competition, but may also remove the threat of contestability.

Estimates are subject to a selectivity bias

There are two sources of selectivity bias, which are not generally accounted for in empirical work. The first results from the widespread use of survey data in the literature. Those authorities experiencing savings may be more likely to respond to surveys. The second arises because the choice of producer is not random, but rather related to the expected costs of each type of provider [64, 100, 152]. Only those bids where a private contractor is cheaper are accepted and so, all other things being equal, one would expect to find empirical savings from private provision [100].[17] The solution to this problem is to explicitly include the choice of provider in the econometric model but, with the notable exceptions of Ohlsson [152] and McDavid [125], this has not been done.

The selectivity bias is most problematic where competition is voluntary (as is the case when considering extensions to the use of the private sector to provide defence services), but in refuse collection studies undertaken after the introduction of Compulsory Competitive Tendering find results in line with the rest of the literature [18, 198].

Public and private provision operate under different technologies

The final issue raised relates to the methodology most widely used in the very influential refuse collection literature based on a two factor Cobb-Douglas production function, $Q = AL^{\alpha}K^{\beta}$. Following Stevens [195], and assuming cost minimisation estimating a log-linear cost function of the form:

$$\ln C = c_0 + c_1 \ln w + c_2 \ln Q + u.$$

c_0 is a vector of factors exogenous to producers but affecting their productivity and commonly includes the location and frequency of pickups, population density, meteorological variables, and a dummy variable indicating private provision which is of primary interest here. This dummy variable is assumed to capture all differences between public and private provision, and is interpreted as percentage savings attributable the use of private providers. However, public and private producers work in different ways (e.g. contractors have been observed to use more side and front loading vehicles with smaller crews [21]). If they operate different technologies, each form of provision will have its own cost functions, and the dummy variables approach is misspecified [152].

This matters if the lessons from refuse collection are to be applied elsewhere, as differences in the cost functions of public producers and private firms (if any) may then be used identify the characteristics determining the probability of achieving savings from contracting.

Few papers have allowed for different cost functions. Stevens [195] herself allowed for the possibility, but the restrictions arising from the use of a single cost function were not rejected. Others have found evidence of differences. Ohlsson [152] finds differences between the cost functions of public and private providers in Sweden. He concludes that public production is 6% cheaper but, crucially, applying the Stevens [195] dummy variable methodology would lead to the erroneous conclusion that public production is 13% more expensive. Similar differences in cost functions have been found in Dutch data by Dijkgraaf

[17] Hodge [100, p107] cites a Government Accountability Office [207] report which found that private sector bidders that won contracts had bids 39% lower than their public competitors, whereas had bids on average 33% higher in those cases where the public competitor won. Were only those contracts won by the private sector considered a saving of 39% would be found, but due to the selectivity bias rather than lower productivity of public production.

and Gradus [44].[18]

These differences are unlikely to overturn the evidence of savings in refuse collection, but they suggest that there is scope to gain a better understanding of precisely how the savings identified arise.[19] A better understanding of how savings are driven by the differing characteristics of individual contracts within the same activities, would enable better forecasts in other areas of where the use of the private sector is likely to successful and what that limitations to the use of the private provision (if any) are. Because they have been so influential, it may be worthwhile to reanalyse the large UK datasets on refuse collection [18, 52, 198, 199], to investigate whether they support differences between the cost functions, and what the differences might be.

2.3.3 Discussion of empirical literature

Although there is strong evidence the savings have been achieved, the vast majority of the evidence available relates to the earliest efforts at contracting in defence. Although a broad consensus seems to have been reached, especially in areas other than defence, many pertinent questions remain unanswered. There is rather less evidence relating to the more recent initiatives, especially those which have resulted in the use of contractors to provide services closer to the battlefield and those using novel contracting arrangements such as contracting for availability or capability. This is largely because of a lack of the necessary data and the inherent difficulties of undertaking empirical work to assess more recent initiatives. Assessments are hampered by the long contracts involved and any comprehensive assessment (especially if hold-up and renegotiation are to be detected) cannot be made until a contract has run its course. Ongoing contracts are also routinely affected by commercial confidentiality objections to releasing information,

The ways in which the use of the private sector to provide public services have been extended—incorporating areas involving specific assets and serious scope for shading on non-contractible aspects of contracts—make rigorous assessment difficult. There are very real concerns about how well the relative efficiencies of public and private provision can be addressed in these areas, because of the difficulty of accurately controlling for quality shading. Even conclusions of the studies relating to the earliest contracting efforts must be somewhat qualified by their, at best, partial controlling for non-contractible elements of the service provided. As Boyne [28, p710] puts it, 'measures of expenditure and quality are poor operationalizations of the theoretical constructs', but there is no obvious solution to

[18] Piers et al. [166] also allow for different technologies, estimating separate production functions for public and private providers. They find public labour more efficient, but that the efficiency of capital differs: at lower outputs private capital is more efficient, but at higher outputs public capital dominates.

[19] Ohlsson [152] finds that distance from the disposal site increases private, but not public, costs, and public costs increase with the number of pick-up points per tonne, whereas private costs do not. Dijkgraaf and Gradus [44] found that public costs increase in the number of inhabitants per pick-up point (proxy for amount of rubbish collected at each point) whereas this does not increase private costs. Together they suggest that small collection rounds, a long way from the disposal point may be better suited to municipal provision (consistent with Ganley and Grahl's [64] assertion that the technologies for rural and urban collections are different).

this problem. At the moment, it must be accepted that all studies tackling this issue will be affected to some degree.

The discussion in Section 2.3.1 raises questions about the ability of defence ministries to negotiate, monitor and enforce effective contracts. These concerns reinforce the limitations of formal contracting raised by the relational contracting literature. These concerns also increase the risk of hold-up and renegotiation of contracts once signed. There is some evidence that this is occuring in defence maintenance contracts [51, 109, 115].

There is good support for the public choice proposition that the introduction of competitive tendering reduces the costs of in-house providers. However, competition has tended to be treated as something which is present or not. There is little empirical evidence on how important the intensity of competition is, despite it being long recognised as one of the factors influencing the likelihood of savings from contracting being realised [50]. That evidence which exists indicates that competition cannot be taken for granted, and the degree to which effective competition is present in markets affects the level of savings achieved from the contracting-out process [44]. There is scope for more empirical work on the importance of the intensity of competition, preferably in an area other than refuse collection to allow some understanding of the interaction between the characteristics of the activity being contracted for and the effects of competition.

The evidence considered in Section 2.3.2 demonstrates that there are significant differences in the savings in various areas of public services. This is consistent with both the notion that the ease with which a service can be specified of specifying details of service in contract is potentially important and with the differences in savings being driven by differences in the characteristics of the activity undertaken. There is little empirical evidence on the precise nature of the differences between public and private provision, but there is scope for more work in this line to be undertaken. Ohlsson [152] raises an important methodological issue with the influential refuse collection literature. Estimating separate cost functions for public and private producers may shed some light on the differences in how production is undertaken, and the strengths and weaknesses of each form of provision. This would most usefully be done in an area other than refuse collection, but it would even be useful to re-examine the extensive datasets for refuse collection in the UK in order to discover whether pooling of cost functions is supported or not.

There is also some scope for progress on the issue of whether the early promise shown by the extension of contracting has lead to long term reductions in expenditure. Some activities have been privately provided for sufficiently long that the time is ripe for an assessment of whether the predictions of hold-up and renegotiation eroding initial savings from contracting raised by TCE hold true. It is not currently possible to determine whether this is a significant issue in real contracts, but such evidence as exists suggests that in some contracts savings have been eroded by renegotiation.

Progress might also be made specifically in defence on the issue of whether the savings seen in individual contracts result in reduced overall expenditures. The public choice hypothesis that greater competition is associated with lower expenditure might be tested using long-run data. This issue might be addressed using the types of aggregate data readily available in defence, though two issues will need to be overcome: firstly, the perennial issue of quantifying defence output must be addressed since changes due to fluctuations in output must be controlled for; secondly, effective measures of competition must be obtained. Chapter 5 makes some progress in this direction.

2.4 Conclusion

The economic literature on military outsourcing, whether theoretical or empirical, does not identify clear limits to outsourcing, nor which characteristics or features determine those limits. The theoretical literature suggests that limits exist, with relatively simple activities better outsourced, but is not able to identify precisely where the limit lies. Probably due to its origins in the boundaries of the firm, the focus of the literature has been on the make-or-buy decision. The decision over which capabilities to maintain in a standing defence force has not examined. With the rapidly developing markets for private military services and the growing number of private military companies, the possibilities for purchasing (some) military activities as and when required is (for some activities) becoming a viable alternative to training and equipping standing forces. The implications of this have not been explored. Nor is the market for military activities well understood. There has been little exploration of the supply side of the market for outsourced defence activities. Many empirical questions remain unanswered, and although data are not easily available, tackling these questions may be more feasible now that some activities have been outsourced for over twenty years.

The key questions which remain unanswered are:

1. Where are the limits to the effectiveness of formal contracting? Does the importance in many military activities of flexibility and of control by commanders on the ground limit the use of private providers? This question is related to the difference between civilian and military labour contracts, and the extent to which capital and civilian labour (both of which are available to private providers) can be used to perform roles previously undertaken by military labour.

2. Are savings at contract level translated into reduced aggregate expenditure?

3. Do observed savings result from efficiencies, quality shading, or reduced surge capacity? If efficiencies result from competition, what savings may be expected in activities where effective competition cannot be arranged? How competitive are markets for military outsourcing?

4. What are costs of negotiating and monitoring contracts for private provision, and how do these compare with the costs of managing in-house provision? On a related note, has opportunistic behaviour been seen in existing contracts, and have expected savings been eroded? How do the costs of altering contracted out activities compare with the costs of changing in-house provision?

5. Does contracting-out lead to quality shading exist and can it be reduced by the bundling of assets and services?

6. What precisely are the differences in the way public and private providers operate—are their technologies different, and if so how are differences related to the characteristics of the activity being undertaken? Can such differences determine which activites are best provided by each form of provision?

Many of the issues raised will require extensive datasets which are unlikely to be obtained in defence, and any work in defence will have to take a different approach to that followed in other areas of public services. There is, though, some scope for investigating the degree to which hold-up and renegotiation affect the efficiency of outsourcing by looking at whether cost savings are maintained once a procuring agency loses its in-house capability and expertise. The issue of whether the cost savings seen in earlier contracts are also present in the more complicated activities, such as training, which are increasingly being outsourced is an important issue, but cannot be addressed with the currently available data.

The next two chapters (3 and 4) address Question 1. First, in Chapter 3 the ease of substitution between capital and military labour is examined. It estimates the elasticity of substitution between military capital and labour within the military production function. Next, Chapter 4 focuses on substitution between civilian and military labour.

Question 2 is then addressed in Chapter 5 which analyses the relationship between the level of outsourcing and aggregate military expenditure in the UK. Question 3 is addressed in Chapter 6 which conducts a survey of the industry for outsourced contracts, and examines how competitive it is.

Chapter 3

Substitution Between Capital and Labour in Defence

3.1 Introduction

Since private companies cannot hire labour using military labour contracts, any defence activity which is outsourced must rely upon either capital or civilian labour to perform the tasks which are undertaken by military labour under in-house provision. The outsourcing of any activity which, when provided in-house, relied upon military labour for its provision necessarily entails substitution of either capital or civilian labour for military labour. This chapter deals with substitution between military labour and capital (substitution between civilian and military labour is the focus of Chapter 4), and investigates the extent to which military capital may be substituted for military manpower in the U.K. Armed Forces. It does so by estimating the elasticity of substitution between military capital and labour. The starting point is Ridge and Smith [169]. They used U.K. time-series data over the period 1953–87 to estimate the aggregate elasticity of substitution between equipment and manpower. In order to account for the possible simultaneous equations bias inherent in single equation estimates it was estimated using the reduced form of a supply and demand framework. Although their results are qualified by data problems, particularly the lack of data on military wages and the price of military physical capital, they estimated that the elasticity of substitution was around unity. Smith *et al.* [187] also estimate the elasticity using data for a cross section of countries in 1976, and found it to be around 0.6 and significant, but this was estimated using only a demand function, so may be tainted by simultaneous equations bias.

Here, by obtaining data on military wages, the structural equations are estimated directly. The results suggest that the elasticity of substitution is somewhat lower than the value of unity suggested by Ridge and Smith [169], and likely to be closer to zero than one. The framework adopted also enables estimates to be obtained of the elasticity of the supply of military manpower with respect to unemployment, which is in the region of 0.05 to 0.1, and to military and civilian wages.

The next section introduces the model; Section 3.3 outlines some problems with the approach; and Section 3.4 describes the data. Some estimation issues are discussed in Section 3.5; the results are presented in Section 3.6; the robustness of the results is then checked by re-estimating the model over a longer sample; and finally some conclusions are offered.

3.2 Model

The approach taken here builds on Ridge and Smith [169]. Motivated by forecast recruiting difficulties in the 1990s, they obtained an estimate of the elasticity of substitution between military labour and capital without requiring data on either the price or stock of capital, which are both unavailable. Production of military output is represented by an aggregate production function describing the technology by which output is produced from factors of production. It is assumed to be a two factor, constant returns to scale, constant elasticity of substitution production function:

$$M = A\left[\alpha L_M^{\rho} + (1-\alpha) K^{\rho}\right]^{\frac{1}{\rho}}. \tag{3.1}$$

M is defence output (represented by military expenditure), L_M is service personnel, K is capital (taken to be all other factors including civilian labour), and their respective prices are P_M, W_M and R. The elasticity of substitution between military labour and capital is $\sigma_{KL} = \frac{1}{1-\rho}$.

If the military are cost-minimisers, if output (M) is given exogenously by government and if input prices are exogenous then the first order condition for cost-minimisation from L_M is:

$$W_M = \mu\alpha L_M^{\rho-1} M^{1-\rho}, \tag{3.2}$$

where μ is the Lagrange Multiplier. Applying the Envelope Theorem, the optimal value of μ may be taken to be marginal cost. If the price of defence output (P_M) is also taken as being equal to marginal cost, then (3.2) yields the factor demand equation for military labour:

$$\ln\left(\frac{M}{L_M}\right) = \ln A + \sigma_{KL}\ln\left(\frac{W_M}{P_M}\right) \tag{3.3}$$

The price of military output is never realised, but (3.3) may be rewritten as:

$$\ln\left(\frac{M}{L_M}\right) = \ln A + \sigma_{KL}\ln\left(\frac{W_M}{P_Y}\right) + \sigma_{KL}\ln\left(\frac{P_Y}{P_M}\right)$$

and if the ratio between the price of military output (P_M) and general prices in the economy (P_Y) is constant then the logged factor demand for military labour is:

$$l_m = \alpha_0 + m - \sigma_{KL} w_m \tag{3.4}$$

in which lower cases represent logged values, and w_m is logged real military wages.

The requirement that price equals marginal cost is not trivial. Were defence privately produced, then with constant returns to scale it is equivalent to assuming zero economic profits and competitive markets (Berndt [23, p. 454]). When produced by government, output is never sold and P_M is never determined, but the price equals marginal cost condition may be satisfied if the level of defence output is determined so as to equate the benefit to society of the marginal unit of output and its cost.

The labour demand (3.4) allows estimation of the elasticity of substitution between military labour and military capital, σ_{KL}, without requiring data on the price or stock of capital. Clearly, the elasticities between labour and the various components of K will differ, but since K does not appear directly in (3.4) the choice of K has no operational role and so taking K to be all other factors should have only second order effects (Smith *et al*. [187]). Smith *et al*. [187] estimated a labour demand similar to (3.4) using data for a cross-section of countries in 1976, obtaining an estimated σ of 0.6 which was significantly different to zero.

Military wages and the number of service personnel will be jointly determined by supply and demand in the market for military labour. In order to avoid simultaneous equations bias in estimates of 3.4 we, in common with Ridge and Smith [169], adopt a military labour supply function of the form [7]:

$$l_m = \alpha_1 + \eta n_m + \beta(w_m - w_a) + \theta z. \tag{3.5}$$

β is the wage elasticity between civilian (w_a) and military wages, n_m is the population of relevant age groups, and z is other factors affecting supply. In principle w_m differs between (3.4) and (3.5) with the latter comprising only payments to service personnel, whereas the former also includes the costs of recruitment, retention, retirement and death benefits, training, housing, etc. [173]. Overall employment costs may be as much as double basic wage costs [219], but the difference is ignored here.

At this point we depart from the analysis of Ridge and Smith [169]. They use the demand (3.4) and supply (3.5) to obtain a reduced form equation for military labour. This allows estimation of the elasticity of interest without requiring data on military wages. Here we follow a different approach—by constructing a data series for the level of military wages we are able instead to estimate the demand and supply functions directly.

Implicit in the supply function (3.5) is an assumption of instantaneous adjustment of the level of military labour. Since it takes some time to join the forces and, once engaged, service personnel are committed to relatively long term contracts, this assumption is avoided by incorporating a partial adjustment process of the form:

$$l_{mt} = (1 - \delta)l^*_{mt} + \delta l_{mt-1} + \lambda(l_{mt-1} - l_{mt-2}). \tag{3.6}$$

Here l^*_{mt} represents the (logged) desired level of service personnel, δ represents the normal carryover from the last period's supply, and λ captures the effect of recent additions not being able to leave.

Taking l^*_{mt} in (3.6) as being given by (3.5) gives the supply function:[1]

$$l_{mt} = \alpha'_1 + (1-\delta)\eta n_m + (1-\delta)\beta\left[w_m - w_a\right] + (1-\delta)\theta z + (\delta+\lambda)l_{mt-1} - \lambda l_{mt-2} \qquad (3.7)$$

in which both δ, $\lambda \in$[0, 1], with δ expected to lie at the upper end of the permissible range, and be larger than λ. It is this supply function and the demand function given by (3.4) which are used as the structural equations in later sections, with variables' expected signs as they appear in (3.4) and (3.7). There are no identification issues with the model outlined here.

To obtain the initial econometric model an additive stochastic term is appended to equations (3.4) and (3.7), and a trend included in each. The trend term in the supply function is intended to proxy for systematic changes in tastes with respect to military service [7]. That in demand might be thought to account for the possibility of a trending component in the efficiency parameter (A).[2]

3.3 Problems with approach

There are a number of problems with the approach outlined above, primarily arising from the adoption of a production function based approach and what form the production function should take or from the assumptions used in deriving the structural equations.

The adoption of any production function is problematic because of the difficulties associated with measuring military output. Since none of the more appropriate measures proposed were available, military expenditure is used to represent output.[3] A CES production function is also used but this does not allow changes in the elasticity of substitution, for example because of changes (or perceived changes) in the strategic environment. It is also possible that the existence of increasing or decreasing returns to scale invalidates CES, though some justification is found in the results of Smith *et al.* [187] who, using a cross-section of countries, could not reject the hypothesis of constant returns to scale.

We have also assumed a production function of only two factors, with K representing

[1] This supply function differs slightly from that in Ridge and Smith [169]. They first obtained the reduced form supply equation and then allowed for some partial adjustment in the reduced form equation.

[2] It should be noted that due to the structure of the constant in (3.4), a significant trend may also encapsulate trends in δ or in η. If a significant trend is picking up a trend that is truly in η, then our CES production function is invalid. Alternatively if the trend is truly contained in δ then we face the prospect of changing parameter values in the supply function. Unfortunately we have no way of determining from which component of the constant term any significant trend derives. Here we proceed on the basis that any significant trend component results form the efficiency parameter, γ, but must be aware of the possibility that it might also be derived from either η or δ, invalidating the model, or the estimation technique respectively.

[3] The true output of defence might be considered peace; security; (the present value of lifetime incomes of) lives saved [173]; levels of 'readiness' of forces [218]; or capabilities [88, 135]. Howver, many aspect of defence output are unmeasurable, and there is no series which captures its multifarious nature [4]. Being a weighted index of inputs, expenditure may not be a good proxy for output, but it is the best available and has been used before [169, 187].

Table 3.1: Definition of variables used

Variable	Label	Definition
R	LM	Logged number of UK military personnel
M	ME	Logged real defence expenditure
w_m	WM	Logged real military wages
w_a	WA	Logged civilian real wage for male manual workers
n_m	POPM	Logged UK population of 15–19 year olds
z	UN	Unemployment rate for males of working age
	GDUM	Gulf War dummy takes value of 1 in 1991
	FDUM	Falklands Conflict dummy takes value of 1 in 1982
	PCW	Post-Cold War dummy takes value of 1 for all years after 1989
Trend	T	Takes value 1 in first year, 2 in second, etc.

Any variable suffixed with PCW is that variable in only the post Cold-War period. It is created by multiplying the variable by PCW, e.g. WMPCW = WM*PCW, and so only takes a non-zero value in the years after 1989.

all inputs other than R in a single index rather than true capital. This of course makes any results model dependent, but incorporating more factors necessitates abandoning CES, and the adoption of a more general functional form. This would require unobtainable time-series data on the stock and price of capital.

The second major strand of criticism surrounds the assumptions required to derive the demand for military manpower. In particular it relies heavily on the questionable assumption of cost-minimisation in the military. Owen [153] provides some evidence of inefficiency through over-manning in UK forces, and the lack of cost-minimisation incentives may make this assumption inappropriate. Departures from cost-minimising behaviour will also reduce incentives for factor substitution [173] and cause lags in response and slow adjustment [187]. The assumption that military expenditure is exogenously set by the government might also be queried and some form of haggling suspected. If so, then military expenditure is endogenous and the results below are biased.

Although the possible problems of the approach are numerous, none can be addressed given the current level of available data. Given this, the results below may be best viewed as a useful if not infallible 'best guess'.

3.4 Data

The data on the variables used for estimation are annual observations over the fiscal years 1970 to 2000. The variables' construction and sources are detailed in Section A.1 of Appendix A. Their plots and summary statistics are provided in Section A.2. The variables used are defined in Table 3.1 (z encompasses all of the last four series).

The population variable demanded by the theoretical model is that of the appropriate age groups: that is of both recruits (generally 17–21) and re-engagements (aged up to 45 or 50). The population of 15–19 year olds is used since no more suitable variable (such as population of 15 to 24 year olds) in available. Unemployment is included in z because if the real wage in the civilian labour market does not instantly adjust downwards in response to excess supply, actual unemployment is likely to result, forcing job hunters into the recruitment office [7]. The Falklands and Gulf War dummies are included because it is received wisdom that recruits are more forthcoming during (and immediately after) high-intensity conflicts. One would expect that the compatibility or otherwise, of military life with family life or the number and type of deployments would be important for retention, but such a series is unobtainable.[4]

3.5 Estimation issues

Due to the trending nature of the variables there is potential for multicollinearity to complicate inference. Although not an infallible measure of the effects of multicollinearity the R_j^2 (the R^2 from the artificial regression of variable j on a constant and all other variables) for both demand and supply are presented. Those for demand are $R_{ME}^2 = 0.00, R_{WM}^2 = 0.00$, and those for supply are: $R_{POPM}^2 = 0.78, R_{WM}^2 = 0.87, R_{WA}^2 = 0.88, R_{UN}^2 = 0.34, R_{LM_{t-1}}^2 = 0.98, R_{LM_{t-2}}^2 = 0.98$. These suggest that inference from the demand equation (i.e. the estimate of σ_{KL}) should be unaffected, but that from supply should be treated with a degree more caution than it might otherwise have been. The results of estimation having dropped an observation, not reported here, also support this.

Prior to estimation Hausman's specification tests are used to assess whether simultaneity is an issue in the model [23, 124]. For the demand equation we obtain the predicted values of WM_t ($\widehat{WM}_t$) from the reduced form equations of the system, and then estimate by OLS:

$$LM_t = c + c_1 t + c_2 ME_t + c_3 WM_t + c_4 \widehat{WM_t} + u_t$$

Testing the hypothesis that $c_4 = 0$ is a test of the null that WM_t and u_t are uncorrelated, which is equivalent to a null of exogenous WM_t. This test gives an asymptotically valid t-statistic of 3.09, so we may reject the hypothesis of exogeneity. The analogous test for exogeneity of LM_t in the inverse demand equation gives a statistic of 1.93 and although not significant (compared to $N(0,1)$) at the 5 percent level, should be accepted as significant since the test is essentially a pre-test. The corresponding statistic for the exogeneity of WM_t in the supply equation is 6.02, so we again conclude that WM_t is endogenous.

Having established that the sample suggests that LM and WM are endogenous, OLS estimation of equations (3.4) and (3.7) would be inconsistent. Therefore, estimation is conducted using Theil's Two Stage Least Squares (2SLS), which gives consistent estimates and

[4] Hosek and Totten [102] find that in the US, reenlistment is related to the number and type of deployments. Non-hostile deployment has a positive effect on reenlistment and hostile deployments either have little effect, or a positive effect for those in the Army. Jaffry *et al.* [107] find that in the Royal Navy one of two deployments at sea increases the probability of exit, but following a heavy deployment schedule the probability of exit is reduced.

is quite robust to multicollinearity and specification errors.

In principle, system estimators might obtain more efficient estimates, but given the highly aggregated, simplistic nature of the model here (and in particular the *ad hoc* nature of the supply function), such efficiency gains may be more than offset by inconsistency of the estimates of all equations being caused by misspecification in only one equation [92].

As with any estimator, the use of 2SLS is not without its problems. Since the estimators are only consistent, any inference can only be asymptotically valid. However, 29 observations is too small a sample to have a high degree of confidence in the validity of asymptotic results. Although its small sample properties are in general unknown, Monte Carlo studies have indicated that the small sample properties of 2SLS are superior on most criteria to other estimators [112]. However, Phillips [165] concluded that the approach to the asymptotic normal distribution is slow in sample size for 2SLS and Bound *et al.* [26] suggested that with typical sample sizes, increasing the number of over-identifying restrictions can cause severe biases in 2SLS estimates.

3.6 Results

The second and third columns of Table 3.2 present the results of estimating equations (3.4) and (3.7) by 2SLS, but these fail to account for the possibility of a structural break at the end of the Cold War (especially in the demand for service personnel).[5] Allowing for a structural break in 1990 involves adding two new endogenous variables into the system (LMPCW and WMPCW) and many new predetermined variables. The results for the general equations are presented in the final two columns of Table 3.2.[6] Wooldridge's [229] asymptotically valid LM tests are used to test the null hypotheses that the coefficients of all dummies in each equation are equal to zero. The tests are complicated by uncertainty as to whether there is a break in both equations and its implications for the instruments available for estimation. Given this, the null hypothesis that all dummies are insignificant in the demand equation must be tested assuming both that there is, and that there is not, a structural break in the supply equation. If there is a break in supply, the test gives an LM statistic of $nR_u^2 = 29.00$ asymptotically distributed as $\chi^2(4)$ under the null of no break. If there is no such break in supply then $nR_u^2 = 13.01$. Since both tests reject the null we may conclude that the data reject there being no break in demand. Testing for a break in supply on this basis gives $nR_u^2 = 24.66$, asymptotically distributed as $\chi^2(8)$ under the null, which is again rejected

[5]The use of 1990 is a little arbitrary since, although the Berlin Wall fell in 1989, adjustment to the post-Cold War environment took time. In the UK the Armed Forces were still undergoing major changes as late as 1995 [133, 134].

[6]Although structural equations determining LMPCW and WMPCW are unknown, if one is prepared to assume that any predetermined variables on the right hand side of their structural equations are already contained in either the demand or supply equations estimated in Table 3.2, then the identification of equations (3.4) and (3.7) is unaffected and 2SLS remains viable.

Table 3.2: Estimation of structural equations allowing for structural break

29 observations (dependent variables LM) 1970–2000				
	Simple model		Structural break allowed	
	Demand	Supply	Demand	Supply
CONSTANT	-1.8084 (2.4665)	-4.45909 (3.7350)	18.6301 (0.8253)***	18.8779 (2.9020)***
PCW	–	–	-21.9461 (3.7579)***	-57.2759 (10.4809)***
T	-0.0294 (0.0049)***	-0.0078 (0.0059)	-0.000015 ((0.00003)	-0.0126 ((0.0026)***
TPCW	–	–	-0.0198 (0.0131)	-0.0043 (0.0267)
ME	0.5716 (0.0713)***	–	-0.3384 (0.0605)***	–
LMEPCW	–	–	1.15841 (0.2361)***	–
WM	0.8781 (0.3961)**	0.8226 (0.5826)	0.1059 (0.1437)	0.2258 (0.1001)**
WMPCW	–	–	-0.0233 (0.6569)	0.9618 (0.8709)
POPM	–	0.2233 (0.1254)*	-	-0.4827 (0.0934)***
POPMPCW	–	–	-	1.3872 (1.0149)
WA	–	-0.2924 (0.3410)	-	-0.0058 (0.0967)
WAPCW	–	–	-	3.1352 (1.4941)**
UN	–	-0.0008 (0.0044)	-	0.0121 (0.0019)***
UNPCW	–	–	-	-0.0239 (0.0356)
LM[-1]	–	1.3419 (0.2788)***	-	0.6664 (0.1507)***
LM[-1]PCW	–	–	-	0.1405 (0.2131)
LM[-2]	–	-0.4443 (0.2965)	-	-0.6466 (0.1517)***
LM[-2]PCW	–	–	-	1.2432 (0.3444)***
FDUM	–	-0.0339 (0.0359)	-	-0.0040 (0.0110)
GDUM	–	-0.01547 (0.0353)	-	-0.0575 ((0.0573)
RSS	0.05429	0.01624	0.0149488	0.0005536
AIC	-3.167	-3.960	-4.283	-6.787
		Diagnostics		
Model test[a]	$\chi^2(2)$=24.7351***	$\chi^2(8)$=28.0468***	$\chi^2(4)$=14.7417***	$\chi^2(4)$=26.2512***
RESET test	$\chi^2(2)$=19.7461***	$\chi^2(2)$=22.0922***	$\chi^2(2) = 4.1850$	$\chi^2(2)$=16.5498***
OverID restrictions[b]	$\chi^2(6)$=16.1849**	-	$\chi^2(10)$=22.81***	-
SC test[c]	t-stat of 0.951	t-stat of 0.582	t-stat of 1.083	t-stat of 1.881*
Het tests[d]	$\chi^2(4) = 4.1789$	$\chi^2(4)$=11.6348**	$\chi^2(4) = 4.6546$	$\chi^2(4) = 3.0908$
	$\chi^2(2) = 3.9788$	$\chi^2(2) = 1.0323$	$\chi^2(2) = 2.6217$	$\chi^2(2) = 0.0169$

Notes: Standard errors in parentheses. *, ** and *** indicate significance at the 10%, 5% and 1% levels, respectively. [a]Asymptotically valid of restrictions that all regressors except constant and trend are irrelevant [228]. [b]Baasman-Sargan tests of overidentifying restrictions. [c]Appropriate asymptotically valid Breusch-Godfrey LM tests [230], testing for AR(1) errors. [d] Top statistic is from Koenker's test with PCW, T, ME and WM used as regressors, second statistic is Wooldridge's [230] test.

at the 1 percent level, indicating the presence of a structural break in both the supply and demand equations.

For the models allowing for a structural break in Table 3.2, both estimated equations include a number of individually insignificant variables, and as such the estimates may not be efficient. We note, however, that the demand results give an estimated σ of around 0.1 in each period though neither is significant, nor do tests support the over-identifying restrictions. To obtain efficient estimates the simplest, data consistent model is obtained by excluding insignificant variables. This is done by applying to the supply and demand equations the most restrictive set of restrictions not rejected by the data at the 5 percent level, that is, by finding the most parsimonious model consistent with the sample data. The results for these preferred specifications are presented in Tables 3.3 and 3.4.

When testing, care must once again be taken of restrictions on one equation affecting the instruments available for, and so the estimates of, another equation. We first test the supply restrictions assuming both the unrestricted and the restricted (Table 3.3 (A)) forms of demand giving LM statistics of 0.88 and 5.49, respectively, both asymptotically distributed as $\chi^2(6)$) so neither null is rejected. We may now assume a restricted supply when testing the restrictions on demand restrictions, giving an LM statistic of 7.31 asymptotically distributed as $\chi^2(4)$ under the null. We cannot, therefore, reject the restricted forms presented in Tables 3.3 and 3.4 and it is on these that the following analysis is based.

Table 3.3 presents the results of estimating the demand function both excluding (A) and including WM (C). Although demand of form (A) is the simplest data consistent demand function, Breusch-Godfrey type tests indicate (AR(1)) serially correlated errors. Such serial correlation invalidates the LM tests used to derive the specification, as does heteroskedasticity, for which there is weak evidence. Therefore, we proceed on the basis that the demand function estimated as (C), is the best model we have, despite WM being individually insignificant.

Further discussion of the demand function below concentrates on the 2SLS estimates (C), with the OLS estimates (D) presented for comparison because 2SLS need not be preferable to OLS. In small samples such as we have here 2SLS often has large finite sample bias [23], and when specification errors are a more serious problem than simultaneity, OLS is preferable to 2SLS [165]. Further, if WM is truly irrelevant as suggested in the results (C), there is no interaction between the two equations of the system so OLS is both consistent and efficient. The similarity of the estimates in C and D may be thought to lend a greater degree of confidence to the estimates. It may, however, be that both biases are large, but in the same direction. As all the estimates are so similar, that which follows is little affected whether WM is included or not, or whether OLS or 2SLS estimates are used.

For the Cold War period the results in Table 3.3 suggest the following model for the demand function (standard errors below in parentheses):

Table 3.3: Original demand and supply estimates

Dependent variable LM, 29 observations 1970–2000	A	B	C	D
Constant	12.7806 (0.0129)***	12.7806 (0.0139)***	11.1866 (1.0518)***	11.7890 (0.5484)***
PCW	−21.2853 (1.4812)***	−21.2849 (1.5953)***	−21.2674 (1.4241)***	−21.2739 (1.5274)***
T	−0.0076 (0.0012)***	−0.0076 (0.0013)***	−0.0113 (0.0027)***	−0.0099 (0.0018)***
LMEPCW	1.1194 (0.0782)***	1.1194 (0.0843)***	1.1178 (0.0752)***	1.1184 (0.0807)***
WM	−	-	0.3589 (0.2368)	0.2233 (0.1235)
RSS	0.0174998	0.202998	0.0155282	0.0178646
AIC	-4.299	-4.151	-4.350	-4.210
Diagnostics				
Model test	$\chi^2(2) =$ 25.4021***	$F(1,4) =$ 176.51***	$\chi^2(3) =$ 26.2801***	$F(3,24) =$ 72.93***
RESET test	$\chi^2(2) = 0.5003$	$\chi^2(2) = 0.5003$	$\chi^2(2) = 1.9569$	$\chi^2(2) = 0.7739$
Overidentifying restrictions	$\chi^2(8) =$ 24.9397***	-	$\chi^2(7) =$ 24.6862***	-
SC test	$t = 2.212$**	BG: $t = 2.004$* DW: 1.1443**	$t = 0.312$	BG: $t = 0.831$ DW: 1.5722
Het tests[a]	$\chi^2(4) = 8.922$* $\chi^2(2) = 5.6716$*	$\chi^2(4) = 8.9215$* $\chi^2(2) = 5.6714$*	$\chi^2(4) = 5.1216$ $\chi^2(2) = 2.1239$	$\chi^2(4) = 7.0608$ $\chi^2(2) = 2.7285$

Model A is the simplest data consistent model by 2SLS. Model B is the simplest data consistent model by OLS. Model C is Model A with WM added. Model D is Model B with WM added.

Notes: See also notes to Table 3.2.

$$E\left(r_t\left|_{t\leq 1989}\right.\right) = \underset{(0.0129)}{12.7806} - \underset{(0.0012)}{0.0076}t + \underset{(0.2369)}{0.3589}w_{mt} \tag{3.8}$$

The coefficient on w_m is opposite to that expected, but is not significant at any conventional level. The significance of the trend term indicates that there has been a trend in the constant term, assumed to be in the efficiency parameter, γ. Real military expenditure is not included and this is unexpected. A possible explanation may be that due to prior commitments (e.g. to allies), the UK had limited scope for reducing service personnel in response to tighter budgets.
For the post-Cold War period we have an estimated demand function of the form:[7]

$$E\left(r_t\left|_{t>1989}\right.\right) = \underset{(2.0815)}{-10.0808} - \underset{(0.0027)}{0.01133}t + \underset{(0.0782)}{1.11945}m_t + \underset{(0.2369)}{0.3589}w_{mt} \tag{3.9}$$

We now have a significant coefficient on real military expenditure and cannot reject the hypothesis that this coefficient is equal to one (asymptotic t-statistic of 1.5268), thus the sample data provides no evidence against CRS in the military production function in the post-Cold War period. The coefficient on the trend term has also increased in recent years.

The data does not reject stability of $\hat{\sigma}_{KL}$ through the Cold War and post-Cold War periods. In both (3.8) and (3.9) WM is insignificant. There are however difficulties with testing the hypothesis that $\hat{\sigma}_{KL} = 0$ in this framework. The CES production function is degenerate at the limit as $\rho \rightarrow -\infty$. The limiting case as $\sigma_{KL} \rightarrow 0$ is a Leontief production function with no substitution possibilities, but it is not strictly possible to test in this framework for $\sigma_{KL} = 0$. Instead the exclusion of w_{mt} in the specification search should properly be interpreted as suggesting that σ_{KL} is at the lower end of the permissible scale. We cannot for instance, reject the hypothesis that $\hat{\sigma}_{KL} = 0.05$ at the 5 percent level. There is also the possibility of measurement error, which would bias the coefficient on WM toward zero in both periods.

OLS estimates are also presented for the supply function in Table 3.4 since it is known that the supply function is misspecified, but whether this misspecification is more serious than the simultaneous equations bias is unknown. Overall, both regressions seem reasonable, though unsurprisingly both RESET statistics are significant at the 10 percent level, giving weak evidence of misspecification. That the over-identifying restrictions are rejected here, as well as in Table 3.3, may suggest that some of the excluded instruments belong in the estimated equation.

The two sets of estimates are again very similar and the analysis below again concentrates on the 2SLS estimates. For both periods the signs on the population variable are not those expected. This may indicate that recruitment effort over-adjusts for expected changes

[7]The estimates of parameters for the post-Cold War period are simply the sum of the estimated Cold War period coefficient, β_2, and the post-Cold War coefficient, β_2. It follows that: $Var\left(\beta_1+\beta_2\right) = Var\left(\beta_1\right) + Var\left(\beta_2\right) + 2Cov\left(\beta_1,\beta_2\right)$. Here we assume that $Cov\left(\beta_1,\beta_2\right) = 0$, such that $S.E.\left(\beta_1+\beta_2\right) = \sqrt{\left\{SE\left(\beta_1\right)\right\}^2 + \left\{SE\left(\beta_2\right)\right\}^2}$. Again this is presented below in parenthesis.

Table 3.4: Preferred specification of Supply

29 observations (dependent variable LM)		
	2SLS	OLS
Constant	17.9971 $(3.2967)^{***}$	18.0032 $(4.2973)^{***}$
PCW	-49.2235 $(9.8114)^{***}$	-50.3074 $(12.9576)^{***}$
T	0.0088 $(0.0015)^{***}$	-0.0088 $(0.0019)^{***}$
TPCW	-0.0491 $(0.0085)^{***}$	-0.0491 $(0.0110)^{***}$
WMPCW	2.6449 $(0.4074)^{***}$	-0.4846 $(0.1325)^{***}$
POPM	-0.4847 $(0.1017)^{***}$	2.6662 $(0.5333)^{***}$
WAPCW	4.6360 $(1.1633)^{***}$	4.7669 $(1.5367)^{***}$
UN	0.0115 $(0.0022)^{***}$	0.0115 $(0.0028)^{***}$
UNPCW	-0.0744 $(0.0112)^{***}$	-0.0754 $(0.0147)^{***}$
LM[-1]	0.6540 $(0.1209)^{***}$	0.6495 $(0.1582)^{***}$
LM[-2]	-0.4867 $(0.1457)^{***}$	-0.4827 $(0.1902)^{***}$
LM[-2]PCW	1.2775 $(0.3159)^{***}$	1.3084 $(0.4162)^{***}$
RSS	0.0017188	0.0029307
AIC	-5.534	-5.534
	Diagnostics	
Model Test	$\chi^2(1) = 26.3583^{***}$	F(8,17]=20.4972***
RESET test	$\chi^2(2) = 5.8633^{*}$	$\chi^2(2) = 5.8531^{*}$
SC test	$t = 0.206$	BG: $t = 0.266$
		DW: 1.9654
Het tests	$\chi^2(4) = 9.1585$	$\chi^2(4) = 9.034$
	$\chi^2(2) = 1.1006$	$\chi^2(2) = 1.0462$

Notes: See notes to Table 3.2.

in eligible population, or perhaps that POPM is not a good proxy for the relevant population.

For the Cold War period the results in Table 3.4 give an estimated supply function of:

$$E\left(r_t\left|_{t\leq 1989}\right.\right) = \underset{(3.2968)}{17.9971} - \underset{(0.0015)}{0.0088}t - \underset{(0.1017)}{0.4847}n_{mt} + \underset{(0.0022)}{0.0115}un_t + \underset{(0.1209)}{0.6540}l_{mt-1} - \underset{(0.1457)}{0.4867}l_{mt-2}$$

which yields estimates of $\hat{\lambda} = \underset{(0.1457)}{0.4867}$ and $\hat{\delta} = \underset{(0.1894)}{0.1674}$, lower than expected *a priori*, with $\hat{\delta}$ not significant at conventional levels. The trend term is significant suggesting a downward trend in the propensity to serve in the armed forces. If this is due to the difference between the conditions in military and civilian employment, it suggests that the X-factor was systematically failing to compensate sufficiently for the hardships of military rather than civilian employment. The estimate of $\hat{\beta}$ is equal to zero by virtue of having been excluded in the model selection process. The same applies to the coefficients on w_{mt} and w_a, and their equality is not inconsistent with the restrictions required in Section 3.2. That the estimated β is zero may indicate a low but positive true value of beta, such as 0.1, and is not inconsistent with Ridge and Smith's finding β to be around 0.2. The estimated elasticity of unemployment suggests that during the Cold War as unemployment increased by one percent, so *ceteris paribus* supply will increase by around 1 percent.

For the post-Cold War period the estimated supply function is:

$$E\left(r_t\left|_{t>1989}\right.\right) = -\underset{(10.3505)}{31.2264} - \underset{(0.0086)}{0.0579}t - \underset{(0.1017)}{0.4847}n_{mt} + \underset{(0.4074)}{2.6449}w_{mt} + \underset{(1.1633)}{4.6360}w_a - \underset{(0.0114)}{0.0629}un_t + \underset{(0.1209)}{0.6540}l_{mt-1} + \underset{(0.3478)}{0.7908}l_{mt-2}$$

A test of whether the coefficients on post-Cold War real civilian and military wages are exactly opposite yields a statistic of F(1,16)=26.88, rejecting the restrictions suggested by the model at the 1 percent level. The coefficient on the trend term indicates that the propensity to serve in the forces has been falling quicker in recent years (potential recruits may have been put off by the large number of high profile redundancies during the early post-Cold War period), so too the elasticity of unemployment is reduced but is still significant.

The coefficient on real military wages suggests that a 10 percent increase in military wages will increase supply of military manpower by around 26 percent. Counter-intuitively the estimated elasticity with respect to real civilian wages suggests that a 10 percent increase in civilian wages will increase supply by around 46 percent. No explanation of this is offered. We also have an unexpected sign on l_{mt-1} which yields estimates of $\hat{\lambda} = \underset{(0.3478)}{-0.7908}$ and $\hat{\delta} = \underset{(0.3682)}{1.4448}$. $\hat{\lambda}$ is significant and negative, whereas we expected a value in the range

[0, 1]. Although $\hat{\delta}$ is also outside the expected range we cannot reject the hypothesis that $\delta = 1$ (t-statistic of 1.2080) or for that matter $\delta = 0.9$. One other noteworthy feature of Table 3.4 is the change in the coefficients of the lagged dependent variables, which go from $0.6540l_{mt-1} - 0.4867l_{mt-2}$ during the Cold War to $0.6540l_{mt-1} + 0.7908l_{mt-2}$ in the post-Cold War period. Since the sum of the post-Cold War coefficients are greater than unity, the process is explosive, though since the post-Cold War sample is so short it is difficult to know how much weight to put on this.

Although overall the results are not as clear cut as might be hoped, estimating the structural equations does suggest that the elasticity of substitution between capital and labour is close to zero, and not significant. The results are not, however, inconsistent with a low positive value. From the structural supply equation during the Cold War period we find the wage elasticity close to zero, an elasticity of unemployment of around 0.01, and that the propensity to serve in the armed forces is falling throughout the sample period.

3.7 Robustness of results

The results of the previous section were previously published in MacDonald [126]. Because they are not as clear cut as might be hoped, and there is some weak evidence of misspecification in supply especially, this section checks their robustness. It does so by re-estimating the models over a longer sample 1970–2006 in order to check that the preferred models of demand and supply obtained above are not sample specific.

The first step is to check whether military wages are endogenous in the longer sample. The Hausman tests of the null of exogeneity for demand and supply yield statistics of $\chi^2(7) = 0.43$ and $\chi^2(16) = 4.03$, respectively. Neither is significant at any conventional level, and so over the longer sample we are not able to reject the null of exogeneity of military wages. Given this, for the restimation only the results of OLS estimation are presented. Whilst the small sample makes one wary of relying on asymptotic results, both the tests and the fact that the OLS and 2SLS estimates were very similar in the shorter sample above provide some reassurance that little is lost through using OLS results alone.

The estimated unrestricted and preferred models of both demand and supply are presented in Table 3.5. The unrestricted demand model is much the same as previously, with the exception that the sign on ME is reversed and now positive. The coefficient on ME for the post-Cold War period is $0.779\,2$ (SE=0.1786) is lower, but not significantly different to the previous value at the 5% level (though it is at the 10% level). The diagnostic tests presented in Table 3.5 raise concerns about the specification of the unrestricted demand model. Heteroskedasticity, serial correlation and RESET tests all giving significant statistics. For this reason robust standard errors are reported.

The preferred demand model is also broadly similar to that in Table 3.3. Testing the required restrictions yields an insignificant statistic of F(3, 27) = 0.73.[8] The addition of a

[8] When tested, the restrictions necessary for the preferred demand model of Section 3.6 are rejected—F(4, 27)

Table 3.5: Re-estimated structural equations

35 observations (dependent variables LM)				
	Demand		Supply	
	Unrestricted	Preferred	Unrestricted	Preferred
CONSTANT	12.3634 (1.6799)***	14.7965 (0.0202)***	10.7914 (2.7385)**	3.1809 (1.0581)
PCW	−5.2137 (2.4228)***	−8.0187 (0.5449)***	8.9288 (14.323)	–
T	−0.0102 (0.0022)***	−0.0078 (0.0012)***	−.01011 (0.0041)**	−0.0118 (0.0025)***
TPCW	−0.01359 (0.0035)***	−0.0166 (0.0015)***	0.0076 (0.0163)	–
ME	0.0946 (0.1209)***	–	–	–
MEPCW	0.6846 (0.1315)***	0.7726 (0.053)***	–	–
WM	0.1521 (0.1327)	–	0.0792 (0.0633)	–
WMPCW	−0.1911 (0.2152)	–	−1.7607 (1.6452)	0.0002 (0.0010)
POPM	–	–	−0.7074 (0.4628)	0.1779 (0.1714)
POPMPCW	–	–	0.2186 (0.8493)	–
WA	–	–	0.3641 (0.1265)**	0.2802 (0.0759)***
WAPCW	–	–	0.6394 (1.2588)	–
UN	–	–	0.0801 (0.0292)**	0.0356 (0.0125)**
UNPCW	–	–	0.0408 (0.1640)	–
LM[-1]	–	–	1.3090 (0.1922)***	1.1888 (0.1024)***
LM[-2]	–	–	−0.6993 (0.2547)**	−0.6154 (0.0907)***
LM[-1]PCW	–	–	−0.0078 (0.0031)**	−0.0043 (0.0005)***
LM[-2]PCW	–	–	−0.00004 (.0025)	–
FDUM	–	–	−0.0532 (0.0098)***	−0.0410 (0.0079)***
GDUM	–	–	−0.0424 (0.0189)**	–
AFGDUM	–	–	0.0828 (0.0819)	0.0202 (0.0101)*
IRAQ	–	–	0.0350 (0.0138)***	–
RSS	0.0167725	0.0187165	0.0026203	0.0041148
AIC	-152.1921	-154.3537	-193.1681	-195.3724
		Diagnostics		
Model test	F(7,27)= 722.83***	F(4,30)= 630.51***	F(19,15)= 478.93***	F(10,24)= 926.26***
RESET test	F(3,24)=4.80***	F(3,27)=3.26***	F(3,12)=2.26	F(3,21)=1.92
SC test	F(1,26)=5.385***	F(1,29)=10.843	F(1,14)=24.445***	F(1,23)=3.757*
Het tests[a]	$\chi^2(4)$=23.14***	$\chi^2(4)$=16.77***	$\chi^2(4) = 1.07$	$\chi^2(4) = 6.87$
	$\chi^2(2)$=11.29***	$\chi^2(2)$=11.39***	$\chi^2(2) = 0.06$	$\chi^2(2) = 2.64$

Notes: See notes to Table 3.2.

break in the trend is new, and in the model for the longer sample the trend post-Cold War has a coefficient of -0.244 (SE of 0.0019) lower (more negative) than that previously reported, but still negative and significant.

The heteroskedasticity tests suggest the presence of unspecified heteroskedasticity so robust standard errors are reported. The RESET statistics on both the restricted and unrestricted estimates are significant, suggesting that the demand relationship is misspecified and is not properly captured by the models presented in Table 3.5. In spite of these tests it still appears that military wages are not relevant to the demand for military labour. From this we conclude that the estimated elasticity of substitution found over the longer dataset is again zero, and the earlier finding that $\widehat{\sigma} = 0$ appears to be robust.

The unrestricted supply in Table 3.5 is more changed from the previous estimates in Table 3.2 than was demand. The serial correlation test statistic suggests autocorrelated errors, so once again robust standard errors are reported, but the discussion is focussed on the preferred model.

The restrictions necessary for the preferred model yield an insignificant statistic of F(10, 15) = 2.00, so cannot be rejected.[9] The estimated supply model in the final column of Table 3.5 is similar in many respects to supply estimates of Section 3.6. Since there remains some weak evidence of serial correlation, the standard errors reported are again robust. The trend remains negative and significant, but is now weaker than previously. The key difference is that military wages are no longer significant. Whilst this accounts for the results of the Hausman tests of exogeneity, this is unexpected. The coefficient on civilian wages is still, counter-intuitively, positive and significant, albeit lower than previously, suggesting that a 10% increase in civilian wages leads to a 2.8% increase in the supply of military labour. Whilst lower than the previous 46%, it remains a curious result and perhaps raises the suspicion that all is not well with the estimated model. The other obvious difference is that the Falklands dummy is now significant, but the negative coefficient suggests that the conflict reduced rather than increased, as expected *a priori*, the supply of military labour. In contrast, between them Iraq and Afghanistan appear to have a weakly positive effect on supply. This difference is consistent with the supply effect of conflict being specific to each conflict.

Overall, from re-estimating the models, the supply function of Table 3.4 appears to be less well specified than does demand, since there were a larger number of changes to the specification and coefficients when re-estimated. The estimated coefficients of the demand function appear to be reasonably robust, but the RESET statistics from re-estimating over the longer sample suggest that the form used here does not properly capture the demand for military labour. Nevertheless the conclusion that military wages are not relevant in the demand for military labour stands, and so too does the conclusion that the elasticity of substitution between military labour and capital in the military production function is zero.

= 50.71 which is significant at the 1% level.

[9]The restrictions necessary to obtain the original preferred model of Table 3.4 are rejected at the 1% level—F(8, 15) = 9.11.

3.8 Conclusion

The estimation results presented above are far from perfect, and conclusions to be drawn must be qualified by the number of potential problems with both the approach and the estimated models. In particular the *ad hoc* supply function appears to be less well specified than demand, and doubts as to the accuracy of the military wage/cost data also raises questions.

Of the assumptions used in obtaining the military demand, two raise concerns. The first is the assumption of cost-minimisation. The second is the inclusion of only two factors in the production function. Aside from obtaining better quality, more disaggregated data which would clearly allow more accurate estimates to be obtained, approaching the issue from a disequilibrium perspective using switching regressions (possibly as an alternative to assuming lags) may also yield results. More accurate estimates may be obtained by allowing for more factors of production.

Nevertheless, we have found no evidence to support Ridge and Smith's [169] proposition that the elasticity of substitution between military capital and labour is around unity. Instead, we have found some evidence of a value closer to zero, a finding which does not appear to be specific to the dataset used and is also found when the models are re-estimated over a longer sample period. These results suggest that the scope for substituting capital for military labour in the military production function is limited. Given this, if activities are to be outsourced, private producers must rely on using civilian labour to perform the roles undertaken by military labour when produced in-house. The scope for substitution between military and civilian labour is the focus of the next chapter. It extends the framework of Section 3.2. Civilian labour is added as a third factor of production, which allows the substitution possibilities between the two types of labour to be examined.

Chapter 4

Substitution Between Military and Civilian Labour

4.1 Introduction

Military employment contracts are distinctive. They specify little besides the possibilities for leaving the Armed Forces. Although voluntarily entered into, in some respects they are contracts of slavery. For the duration of the contract Servicemen have no say over when and where they serve, or what they are asked to do. Such contracts have evolved because they are advantageous in some circumstances, but only government may employ military labour. Private organisations cannot sign (or rather cannot enforce) military empoyment contracts.

The ability to hire military labour is one respect in which public and private provision of defence activities differ. When provided in-house defence activities are typically provided by a mix of military and civilian labour, but private providers are able to use only civilian labour. Given this, outsourcing military activities involves a substitution of factor inputs, with civilian labour replacing military labour.

This chapter extends the framework used in Chapter 3 to include a third factor, civilian labour. This allows the elasticity of substitution between civilian and military labour to be estimated within the MOD. An estimate of this elasticity may shed some light on the scope for outsourcing, or contracting-out, functions previously conducted by military personnel: a high elasticity suggests that the scope for efficient outsourcing is large; a low estimate suggests that, without adversely affecting output or capabilities, the scope of outsourcing is limited. An accurate estimate of the elasticity of substitution between military and civilian labour may also allow forecasts of the Armed Forces' labour costs to be improved.

Two previous papers have considered substitution between labour types in military production functions.[1] Smoker [188] examined the cost implications of the mix of military and

[1] More generally, the issue of substitution between labour types tackled here is related to the literatures on labour aggregation functions [27, 53, for example] and substitution between more and less educated labour [38, 111, for example]).

civilian labour in the context of the Department of Defense's program of civilianization during the mid-1970s. Although concerned largely with in-house labour rather than outsourced labour he finds that replacing a military post with a civilian job frequently leads to cost savings, but in 40% of cases actually increases the costs of performing that role. Albrecht [2] considered the efficient allocation of experienced and inexperienced military manpower in seventeen United States Air Force occupational specialities. Based on a production function of similar form but at a more disaggregated level, difficulties arose primarily because of an inability to accurately assess the relative productivity and substitutability of various categories of military labour. This chapter is also similar to the approaches of Albrecht [2] and Ciccone and Peri [38], though applied in a different context.

The chapter proceeds as follows: first civilian labour is incorporated; next, the approach to estimation is discussed, the dataset described, and results presented; Section 4.6 discusses the results and some difficulties of the approach adopted; and finally some conclusions are offered.

4.2 Incorporating civilian labour

There are a number of ways in which civilian labour may be incorporated into the military production function (3.1). The simplest would be to adopt a straightforward three factor CES function, but this would restrict the substitution possibilities between any pair of factors to be equal—imposing equal elasticities of substitution between military and civilian labour, military labour and capital, and civilian labour and capital. The most flexible approach would be to adopt a flexible cost function which avoids imposing arbitrary restrictions on the production technology. In principle, this would allow a system of cost share equations to be estimated, and the elasticities in each period to be estimated. Such an approach is not possible due to the lack of data on the price of capital.

Instead, civilian labour (L_C, with price W_C) is introduced by adopting a production function in which labour types and capital enter in a (weakly) separable way. The potential difficulty with this approach is the decision as to how to group factors [59, 80]. The choice is arbitrary but in this case both military and civilian labour are equally able to use capital, and so the following form follows naturally from the problem:

$$M = f\left[K, g\left(L_M, L_C\right)\right]. \tag{4.1}$$

Civilian and military labour are first combined in the labour aggregation function, g, and the resulting aggregate labour input then combined with capital (redefined to be all factors other than military and civilian labour) to produce output. The advantage of adopting a separable production function is that the elasticity of interest can be obtained from the relative labour demand without requiring data on capital. By restricting attention to relative labour demands, the issue reduces to that of specifying the labour aggregation function, g.[2]

[2] A similar approach is used by Ciccone and Peri [38] to justify focusing on a labour only aggregate pro-

Of course, the adoption of a separable production is not without cost. It imposes three restrictions which, because of the absence of data on capital, are not testable: the ease of substitution between civilian labour and capital, and military labour and capital must be equal; they must be unaffected by the amount of capital; and differences in degree of complementarity between capital and labour types are obscured [27]. If actual technology is non-separable, conclusions based on separable production technology will be invalid.

The labour aggregation function, g, is taken to be a CES function such that:

$$M = f\left[K, B\left[\eta L_M^{\theta} + (1-\eta) L_C^{\theta}\right]^{\frac{1}{\theta}}\right] \tag{4.2}$$

with the elasticity of substitution between military and civilian labour, $\sigma_L = \frac{1}{1-\theta}$. A CES function is chosen because, although it may be desirable to allow the elasticity of substitution between labour types to change over the sample period, the small sample size (34 observations) makes this unfeasible. Its selection is consistent with Bowles [27], Dougherty [53], and Ciccone and Peri [38].

If the military are cost-minimisers, the level of output is exogenously given by government, and that input prices are exogenously given the relative demand for military labour may be written as:

$$\frac{L_M}{L_C} = \left(\frac{\eta}{(1-\eta)}\right)^{-\sigma_L} \left(\frac{w_M}{w_C}\right)^{-\sigma_L} \tag{4.3}$$

and the logged relative demand function:

$$(l_m - l_c) = a - \sigma_L (w_m - w_c). \tag{4.4}$$

This logged relative factor demand is the basic estimating equation used to estimate σ_L.[3] It is preferred to estimating the (non-linear) demands for military and civilian labour because it does not require specification of the production function, f, and because it avoids the need to make a competitive markets assumption, that $\mu = P_M$, used in the derivation of the demand for military labour (3.4). In principle it also avoids the need to include measures of defence output, M.

4.3 Approach to estimation

The logged relative demand for military labour (4.4), with an additive stochastic error term appended, is used to estimate σ_L. Such an error specification might arise if the optimal

duction function when estimating the long-run substitutability between more and less educated labour in the US.

[3]Equation (4.4) may be obtained in a number of ways. The method presented is the least restrictive but precisely the same relative demand for military labour may be obtained by using a simple three factor CES production function, by using a Sato type two-level CES production function, or by assuming two separate types of capital one used exclusively by civilian labour and the other used by military labour.

labour demands are chosen with error [23]. Single equation estimators are used because, although desirable [22, 80], the absence of data on capital precludes joint estimation of the three relative demands.

There are two major difficulties associated with estimating σ_L using (4.4): the possible endogeneity of the wage ratio, and missing dynamics arising from the inability of the Armed Forces (because of the time taken to train recruits and the lengthy employment contracts) to act in accordance with the instantaneous adjustment assumption implicit in (4.4).

The endogeneity of wages, or the wage ratio, is the same problem faced by Chapter 3. The difference here is that the system determining wages is not completely known, especially the determination of civilian wages. So long as valid instruments can be found, Instrumental Variables offers a feasible means to obtain consistent (if not unbiased) estimates of (4.4).

Two sets of instruments are used. Since the data are time-series, the natural candidates for instruments for the wage ratio are lagged values of the wage ratio [75]. so the first set of instruments used is two lagged values of the wage ratio. The second set of instruments used are based on the supply sides of the civilian and military labour markets. The supply of military labour is taken to be of the form (3.5), and the exogenous elements n_m, w_a and un are adopted as instruments. The supply function for civilian labour to the MOD is not specified, but is assumed to involve the population relevant for the supply of civilian staff to the MOD (n_c). This is also included as an instrument.

Demand for military labour does not adjust quickly—it takes time to train recruits, there is a need to maintain a mix of experience within the rank structure, and employment contracts are relatively long. It is reasonable to expect there to be some dynamic structure to the demand for military labour, which will be transmitted to the relative demand but is missing from (4.4). This raises doubts as to whether lagged values of the wage ratio will be valid instruments. If there are dynamics missing, then they will likely manifest themselves in significant serial correlation test statistics.

In order to deal with the issue of missing dynamics, the model is initially estimated by OLS. OLS is used primarily to infer the dynamic structure of relative demand, but although not consistent it may also provide useful estimates. It has been suggested that because of its favourable variance properties, when estimating factor demands with small samples it may best to ignore the simultaneous equations bias and rely on OLS estimates [22].[4]

It is also not a foregone conclusion that the wage ratio is endogenous. In the market for military labour when tested over the longer sample in Chapter 3 wages did not appear to be endogenous, perhaps because wages are set independently by the AFPRB. In the market for civilian labour, the MOD is but a small employer in the market and might be thought of as facing a perfectly elastic supply of civilian labour. If this is the case then the wages of MOD civilians are exogenously determined in the overall labour market. If both civilian

[4]The logic is that, due to OLS's favourable variance properties, in small samples the bias from using an inconsistent OLS estimator are smaller than the finite sample biases of a consistent IV estimator. This is especially true if weak instruments are used.

and military wages are exogenous then OLS will be a consistent and unbiased estimator.

4.4 Data

The data used for estimation are time-series for the UK over the period 1970 to 2003. Neither military wage data nor MOD expenditure data are available prior to 1970, and changes to the way civilian staff are counted within the MOD and to how civilian wages are calculated mean that the sample cannot be extended beyond 2003. The construction and sources of the series used is described in Appendix A. Table 4.1 briefly describes the series and labels used in estimation and how they correspond the variables in Section 4.2. The series are plotted, and their summary statistics provided in Section A.3 of Appendix A.

The variable L_C in the relative demand for military labour (4.4) is total civilian labour inputs into production. This includes both directly employed civilian labour and civilian labour inputs from contracted-out defence activities. Data on the number of staff involved in the production of contracted out defence activities does not exist, so the number of directly employed civilians is used here as a proxy for total civilian labour inputs. This difference may be important, especially towards the end of the sample. Military expenditure is used to represent military output, but the caveats in Footnote 3 in Chapter 3 still apply.

Table 4.1: Definition of variables used

Variable	Label[a]	Description
$l_m - l_c$	LMLC	Logged ratio of military to civilian personnel
$w_m - w_c$	WMWC	Logged ratio of real military wages to real wages of MOD civilian staff
m	ME	Logged real defence expenditure
		Supply side instruments
w_a	WA	Logged real civilian wage for males (all industries)
n_m	POPM	Population of males aged 15–29
n_c	POPC	Population of working age
un	UN	Unemployment rate for males of working age
-	CAS	Deaths as proportion of the number of military personnel

4.5 Results

4.5.1 OLS estimation

The results of estimating the logged relative demand for military labour (4.4) by OLS are presented in Table 4.2.[5] In order to allow a general to simple specification search to be undertaken the first column presents a model of the relative labour demand augmented with lags of both the labour and wage ratios. These lags are included to account for the difficulties faced by the MOD in adjusting to their desired level of military labour. Two additional variables are also included: CAS is the proportion of personnel who die each

[5] All estimates were obtained in Stata 9.2 using the ivreg2 routine [15].

Table 4.2: OLS estimation of logged relative demand

Dependent variable LMLC, OLS estimates

	General model	Preferred model	Simple model	Simple model with trend
Constant	-7.750 (1.918)***	-5.632 (1.597)***	0.418 (0.030)***	0.009 (0.045)
Trend	0.011 (0.003)***	0.007 (0.002)***	–	0.022 (0.002)***
LMLC(-1)	0.761 (0.185)***	0.794 (0.073)***	–	–
LMLC(-2)	-0.057 (0.168)	–	–	–
WMWC	-0.109 (0.139)	–	-1.629 (0.331)***	0.164 (0.249)
WMWC(-1)	0.061 (0.143)	–	–	–
WMWC(-2)	0.225 (0.130)*	–	–	–
CAS	-34.143 (41.216)	–	–	–
ME	0.328 (0.081)***	0.238 (0.067)***	–	–
Observations	32	33	34	34
$\overline{R}^2$	0.978	0.977	0.41	0.85
		Diagnostics		
Model Test	F(8,24)=177.1***	F(3,30)=460.6***	F(1,32) =24.23***	F(2,32)=94.3***
RESET test	F(2,21)=0.625	F(2,27)=0.397	F(2,30)=1.770	F(2,29)=8.724***
SC test[a]	F(1,21)=2.813	F(1,28)=0.010	F(1,30)=71***	F(1,30)=102.5***
Het tests[b]	F(8,23)=0.184	F(3,29)=0.176	F(1,32)=0.246	F(2,31)=0.098
	F(8,23)=0.172	F(8,24)=0.444	F(2,31)=0.198	F(5,28)=1.98

Notes: * significant at 10% level, *** at 5%, and *** at 1%. [a]Breusch-Godfrey LM tests for AR(1) errors, replacing missing residuals with zeros [41]. [b]Top row are Breusch-Pagan-Godfrey tests using regressors, bottom row are White's tests.

year, and is used instead of dummy variables to account for the number and intensity of deployments undertaken. Logged military expenditure (ME) is included as a measure of output. Its coefficient has no specific interpretation, but significance would raise questions as to the validity of use of a separable production function.

Whilst the diagnostic tests do not raise concerns about the specification of the general model, there are a number of insignificant regressors. In order to present the most efficient estimates possible the second column of Table 4.2 present the simplest model consistent with the sample data. A test of the necessary restrictions gives a statistic of 1.0081 (distributed as F(5,23) under the null of no restrictions), and so are not rejected at any conventional level, but any further exclusions are rejected.

Since the sample straddles the end of the Cold War, the possibility of a structural break at the end of the Cold War must be considered, especially since there is evidence in Chapter 3 of a break occurring between 1989 and 1990. However, for neither the general nor preferred models do Chow tests of parameter stability find evidence of a structural break at the end of the Cold War.

Before discussing the preferred model it is worth mentioning the estimates of the simple models in the final columns of Table 4.2. In the estimated simple model the coefficient on WMWC is of the expected sign and suggests an elasticity of around 1.6 but the serial correlation test statistic is significant. Whilst this may result from truly serially correlated errors it is more likely to be caused by misspecified dynamics, and provides some support for the inclusion of lagged terms in the general model. Once a trend is included in the simple model WMWC becomes insignificant suggesting that the significance of WMWC in the simple model was picking up its trending nature. The significant serial correlation test statistic remains.

That the serial correlation test statistic for the preferred model is not significant suggests that the slow adjustment of the level of military labour appears to be fully captured by the first lagged dependent variable, since the second lag is not retained. The other diagnostic tests do not suggest problems with the specification of the preferred model, but only the lagged relative employment term and military expenditures are significant. Although the fact that none of the wage ratio terms are retained appears to suggest that σ_L is equal, or close, to zero, that wages have no effect is unexpected and perhaps a little incredible. The reservations about the validity of testing for $\sigma_L = 0$ in this framework discussed in Section 3.6 also apply here, so one must conclude that the exclusion of the wage ratio terms suggests that σ_L is small and close to zero but positive.

That the estimated coefficient on ME is positive and significant is also a little troubling given the adoption of a separable CES production function, since it suggests that as output increases so the relative demand for military labour rises. If true this suggests that not only is the assumption of separability inappropriate, but that the technology of defence is non-homothetic and the functional form assumed here is inappropriate invalidating any of the models estimated. Even if ME is not a good proxy for output, the significance of its coefficient suggests that the relative demand for military labour increases during times of increased defence spending which is still troubling.

Overall, the OLS results appear to be well specified, and were it not for the possibility of the wage ratio being endogenous these results would stand. However, there are grounds to suspect that the OLS estimator is inconsistent so the next two subsections estimate the expanded model using Instrumental Variables. The lag structure, albeit arrived at through (the possibly inconsistent) OLS estimator, is retained.

4.5.2 IV estimation using lagged values as instruments

The results of estimating the expanded model of Table 4.2 using lagged values of the wage ratio as instruments are presented in the left-hand column of Table 4.3.[6]

Although lagged values are the natural candidates for instruments, they need not be good instruments. If they are not strongly correlated with the wage ratio—weak instruments—

[6]Two lags are used as instruments for WMWC. The lagged values of WMWC included as regressors are taken as being predetermined, and since two lags are included as regressors it is the third and fourth lags that are used as instruments.

Table 4.3: Instrumental Variables estimates

Dependent variable LMLC, Instrumental Variables estimates

	IV (lags)	IV (supply)
Constant	-9.456	-7.882
	$(2.034)^{***}$	$(1.668)^{***}$
Trend	0.012	0.011
	$(0.0034)^{***}$	$(0.002)^{***}$
LMLC(-1)	0.744	0.762
	$(0.216)^{***}$	$(0.160)^{***}$
LMLC(-2)	-0.145	-0.075
	(0.200)	(0.146)
WMWC	-0.572	-0.250
	$(0.295)^{*}$	(0.185)
WMWC(-1)	0.178	0.112
	(0.186)	(0.134)
WMWC(-2)	0.146	0.241
	(0.165)	$(0.114)^{**}$
CAS	-47.856	-32.382
	(48.842)	(35.761)
ME	0.403	0.334
	$(0.103)^{***}$	$(0.070)^{***}$
Observations	30	32
RSS	0.0286106001	0.0250135831
Diagnostics		
SC test[a]	$\chi^2(1)$ =1.044	$\chi^2(1)$ =1.249
Het tests[b]	$\chi^2(9)$ =1.232	$\chi^2(11)$ =11.381
	$\chi^2(2)$ =2.122	$\chi^2(2)$ =3.294
B-S test[c]	$\chi^2(1)$ =0.149	$\chi^2(3)$ =3.425

Notes: * significant at 10% level, *** at 5%, and *** at 1%. WMWC instrumented using excluded instruments WMWC(-3) and WMWC(-4) for lags; POPM, POPC, WAM, UN for supply side instruments.
[a] Sargan's [175] test for AR(1) errors recommended by [164] [b]Top row are Pagan and Hall's [155] statistics, bottom row Pesaran and Taylor's [164] HET_1 test. [c]Basmann–Sargan test of overidentifying restrictions.

any IV estimates relying upon them may be seriously biased even in large samples [26, 78, 193], and may lead to misleading estimates of significance [132]. A suggested rule of thumb as to whether weak instruments are a concern is whether F-statistic from an F-test of the joint significance of instruments in a first stage regression is less than 10 [193, 196]. By this measure the lagged values are weak instruments (F-statistic is 5.45). The model estimated here seems to be identified, but not weakly identified raising doubts as to the accuracy of the estimated coefficients.[7]

Although the instruments are weak, testing for the endogeneity of the wage ratio suggests that the wage ratio is endogenous. A Durbin-Wu-Hausman test gives a statistic which is significant at the 5% level suggesting that the OLS estimates are inconsistent.

[7]Identification can be tested using Anderson's test, and Cragg-Donald tests. Both reject underidentification at the 1% level (test statistics of 10.59 and 16.36 respectively, both distrubuted as $\chi^2(2)$ under the null). However, the Cragg-Donald Wald F-statistic is 5.455. When compared with Stock and Yogo's [197] critical values this statistic indicates a true rejection rate of more that 25%, suggesting that the weakness of lagged values as instruments is likely to be causing significant biases in the estimated coefficients.

Looking at the estimates, the IV estimates using lagged values as instruments are similar to those obtained by OLS estimation. The lack of evidence of serially correlated errors provides reassurance that the lagged values are valid instruments. The apparent absence of conditional heteroskedasticity justifies the use of IV over a more general estimator such as GMM.[8]

The only notable difference between the two sets of estimates is that using lagged values as instruments it is the coefficient on WMWC, rather than the coefficient on the second lag of WMWC, which is significant at the 10% level. It is again of the expected sign, less than one and reasonably close to zero. ME again appears to be significant.

The coefficient of primary interest here is that of WMWC. Since weak instruments may cause asymptotic t-tests to perform poorly [147] a conditional likelihood ratio test is used to jointly test the significance of the coefficient on WMWC and the overidentifying restrictions. This test has correct size with weak instruments, is robust to misspecification of the first stage regression, and dominates the commonly used Anderson-Rubin test [131]. It returns a p-value of 0.0246, providing evidence against the null that the coefficient on WMWC is equal to zero.

Lagged values appear to be weak instruments for the wage ratio. Although Hausman tests suggest that these IV estimates are preferable to the OLS results of Table 4.2 the weakness of lagged values as instruments the estimates are likely to be biased, and the standard errors misleading. None the less, there is some evidence to suggest that (contrary to the OLS results above) the coefficient on WMWC is non-zero, implying a small (but different to zero) elasticity of substitution between civilian and military labour.

Overall the results are very similar to the OLS estimates with a number of insignificant regressors. Testing for the most parsimonious model not rejected by the data results in the same preferred model as in Table 4.2. An IV Wald test [229, p. 98] of the required restrictions returns a small sample F-statistic of 1.13 (distributed as F(5,21) under the null), and fails to reject the restrictions at any conventional level. Any further restrictions are rejected. Again neither the wage ratio, nor its lags, appear to be significant and again the significant coefficient on ME raises questions as to the validity of the separable production function. This preferred model contradicts the result of the conditional likelihood ratio test. Overall, perhaps these results do not demonstrate whether the elasticity of interest is exactly zero or close to zero, the evidence here does suggest that if not zero, the elasticity is not large, certainly below one.

4.5.3 IV estimation using supply side instruments

The results of using instead the supply side instruments (listed in Table 4.1) are presented in the final column of Table 4.2. These are also weak instruments (F-statistic from the first stage regression is 3.88), weaker instruments even than the lagged values used in the pre-

[8]Under conditional homoskedasticity both IV and GMM are consistent and asymptotically efficient, but GMM may have poor small sample properties [14, 93].

vious subsection. Perhaps because of this, a null that the wage ratio is exogenous cannot be rejected (Durbin-Wu-Hausman statistic of 1.808 distributed as $\chi^2(1)$ under the null of exogeneity.

When supply side instruments are used there is no reason to prefer IV estimates to the OLS estimates in Table 4.2. They are nevertheless discussed briefly. As when lagged values are used as instruments the model appears to be identified (Anderson canonical correlations test statistic of 13.97 and Cragg-Donald statistic of 24.81, both distributed as $\chi^2(4)$ under the null of non-identification), but not weakly identified (the Stock-Yogo [197] critical values suggest high absolute and relative biases). Diagnostic tests do not suggest the presence of serially correlated or heteroskedastic errors.

The estimated coefficients are strikingly similar to both the OLS and lagged value instrumental variable results. The only notable difference with the results of using lagged values as instruments is that, as with the OLS estimates, it is the second lag of the wage ratio that is significant at the 10% level. As with the previous subsection the simplest data consistent model is the same as the preferred OLS model of Table 4.2, but a conditional likelihood ratio test cannot reject a null that the wage ratio is irrelevant to the relative demand for military labour.

4.6 Discussion of results and limitations of approach

Three sets of estimates were obtained in the previous section. Each is flawed. Both of the IV estimates suffer from weak instruments which are likely to lead to biased estimated coefficients and invalid inference. If the wage ratio is endogenous, and the IV estimates using the strongest instruments suggest it is, then the OLS estimates presented in Table 4.2 are inconsistent. Although flawed the similarity of the estimates lends them more credibility than would be the case were they contradictory. For all, model selection suggests that the wage ratio is not relevant to the relative demand for military labour. In spite of this the conclusion that σ_L is zero is not supported by the conditional likelihood ratio tests in either of IV estimates.

The conclusion that σ_L is positive but close to zero is consistent with the results above, albeit based on a sample of just over 30 observations. Such an elasticity suggests that there is little scope for contracting-out to increase the efficiency of defence provision. Such a low elasticity is surprising. It seems reasonable to expect at least some substitution possibilities between military and civilian labour. Even if they do not produce exactly the same output, one might expect that there is some scope for different labour types to undertake slightly different activities which contribute to overall output in different ways, but which leave total capabilities unchanged [2]. One explanation might be that although the true elasticity allows for some substitution, the MOD fails to take advantage of this—it does not cost-minimise.

Alternatively, the elasticity may truly be close to zero. If this is the case, then an extensive program of outsourcing may lead to reduced capabilities in the outsourced activities. If the true elasticity is zero then outsourcing leads to civilian labour being used in the circumstances for which military labour contracts were developed, and so leads to some reduction in capabilities. This does not mean that the substitution resulting from outsourcing does not lead to reduced costs, but simply that some of these savings may be achieved at the cost of reduced capabilities which must be recognised.

There are a number of limitations to the estimates obtained in Section 4.5, and some of these raise doubts about whether the estimated elasticity of σ_L close to zero is accurate. These might be usefully divided into into factors related to the data used, and factors related to the model adopted in Section 4.2.

Of the reasons related to data, perhaps the most important is the difference between the measure of civilian labour demanded by L_C in equation 4.4 (total civilian labour used in the production of defence) and the measure used in estimation (civilian labour directly employed by the MOD). The rise in outsourced provision over the sample period means that the estimate produced above is an underestimate of the true elasticity, but the magnitude of the difference between the estimated and true values of σ_L is unknown. What is certain is that some cases of sucessful substitution are not captured (e.g. the successful use of civilian labour in the provision of support for fast jets [141]). There is scope for more case studies of those activities where substitution between military and civilian labour has occured.

Another issue related to the data used are that directly the number of directly employed MOD civilians fell sharply during the 1980s due to the programme of privatisations, such as that of the Royal Dockyards. This fall may account for the trending nature of L_M/L_C, but is unrelated to relative wages. The final data related factor arises from the way in which military wages are determined. They are set by the AFPRB at the level of comparable civilian occupations plus some X-factor. This means that there is not a large amount of variability in the wage ratio, and it may be that this causes the insignificance of the parameter on the wage ratio and its lags.

The issues related to the derivation of the relative demand (4.4) and stem from the assumptions made, especially that of cost-minimisation. The military are commonly perceived as being inefficient, at least anecdotally (for instance in equipment procurement). Indeed, the most common justification for contracting-out is that the private sector can produce efficiently, whereas the public sector cannot. If the assumption of cost-minimisation cannot be maintained then the σ_L in the relative factor demand, and so the estimated value of zero, cannot be interpreted as the elasticity of substitution between military and civilian labour.

The significance of the military expenditure coefficient in all of the estimated models also raises doubts about the validity of the relative demand estimated. Their significance suggests that the true technology may be non-homothetic. If this is the case, the separable CES production function adopted here is inappropriate and the estimated parameter on the wage ratio in the relative demand does not accurately reflect the elasticity of substitution between civilian and military labour.

4.7 Conclusion

This chapter has calculated three estimates of the elasticity of substitution between civilian and military labour. All three are similar and suggest that the elasticity is close to or equal to zero, but there are a number of issues which raise doubts as to the accuracy of these estimates: they are based on a small sample; both sets of instruments used to obtain the IV estimates are weak instruments; the data used for estimation may not be close enough to the variables demanded by the model to allow accurate estimates to be obtained; and the approach adopted here relies on a number assumptions which may not be valid.

An elasticity of zero or close to zero suggests that the military production technology has well defined and distinct roles for civilian and military labour, and with relative wages irrelevant to the determination of the relative levels of each type of labour. A similar situation was observed in the US DoD during the 1970s [188]. The results obained are consistent with the MOD employing a core of civilian labour and allows military labour to fluctuate in response to variations in defence output. The positive and significant coefficient on military expenditure may be explained if military labour increases with output (or increases with output at a faster rate than does civilian labour). This is curious since civilian labour is easier to adjust, requiring as it does shorter employment contracts and less training.

Perhaps the most fruitful extension of the work in this chapter would be to obtain data on the price of capital. This would allow a more flexible approach to be adopted, avoiding the imposition of separability, and allow the form of the military production function to be inferred from the sample data rather than imposed as it is here. Unfortunately such data are not available, so the next chapter takes a different approach to identifying the scope for military outsourcing. It examines the long-run relationship between aggregate defence expenditure and the level of outsourcing.

Chapter 5

Outsourcing in Models of UK Demand for Military Expenditure

5.1 Introduction

Given the difficulties with improving the substitution based approach of the two preceding chapters, this chapter takes a different approach. It examines whether higher levels of military outsourcing are associated with lower aggregate defence spending in the UK. Whilst there is reasonably good evidence of savings at the level of individual contracts, certainly for the earlier outsourcing initiatives in defence (see Section 2.3.1), there remains some debate over the source of these savings. For more recent programmes, empirical investigation of the cost-effectiveness of outsourcing of military activities is hampered by a lack of publicly available data, making statistical analysis impossible.

Even if there were data available, the difficulties of ensuring one is comparing like-with-like is especially acute in defence (see also Section 2.3.2). The most compelling approach applied to other areas of public services—using large cross-section and panel datasets in which there are a sufficiently large number of examples of public and of outsourced private provision to allow both comparison of the two forms of provision and service quality to be accounted for—is not possible for most defence activities. Each activity necessary for the provision of national defence tends to be produced or procured only by the defence ministry, so in each country there is only one example of each activity being produced (cf. the 423 local authorities in England and Wales each providing a refuse collection, providing 423 observations of refuse collection in each year with which to construct a useful dataset). The construction of an international cross-sectional dataset is prevented by the large variations in the way defence is structured and provided between countries, and because it would contain too few examples of private provision of an activity.

Since the available data are aggregated, reported either for a service or tri-service, this chapter takes an approach closer to that of Albin [1]. Rather than considering whether public or private provision is more efficient in specific contracts or activities, Albin [1] used a cross-section of Australian councils to examine whether those with higher levels of outsourcing also have lower rates (expenditures per head). The approach taken here is to use time-series data for the UK to examine whether increased use of outsourcing is associated with lower aggregate expenditure on defence. This is done by estimating a demand for military expenditure model which includes a measure of the level of outsourcing. Although the use of time-series for a single country avoids the need to account for international differences, this problem is replaced by the need to account for threats and the level of output changing over time.

5.2 Demand for military expenditure model

Our starting point, in common with many other demand for military expenditure analyses, is Smith's [184, 185] welfare maximising approach.[1] Welfare is given by:

$$W = w\left(S, C\right). \tag{5.1}$$

S is security and C non-defence expenditure. Welfare is maximised subject to a budget constraint:

$$Y = P_M M + P_C C. \tag{5.2}$$

Y is aggregate real income and P_M and P_C are the prices of military and civilian output, respectively. A second constraint is a production constraint in the form of a security function:

$$S = s\left(M, E, O, Z\right) \tag{5.3}$$

Security is a function of military expenditure, M, the strategic environment, E. O allows differences in the way defence is produced to be reflected. Here O is taken here to be the level of outsourcing. It belongs in the security function because if outsourced provision of certain activities results in cost savings, then as the level of outsourced production rises, so more security may be produced for a given expenditure. Z encompasses any other factors affecting security. Maximising (5.1) subject to (5.2) and (5.3) gives a general demand for military expenditure:

$$M = f\left(Y, P_M, P_C, E, O, Z\right) \tag{5.4}$$

Three specific forms of the general demand (5.4) have been widely used in the literature, each modelling a different measure of military expenditure, namely, the level of expenditure, the logged level and the share of GDP devoted to defence.[2] They are considered below in turn.

[1]Examples of others taking this approach are the models for various countries in Hartley and Sandler [91], Solomon [192] for Canada, Nikolaidou [149] for the EU15, Dunne and Perlo-Freeman [57] for a selection of developing countries, and Gadea et al. [63] for 15 NATO countries.

[2]Logged shares has also been used by Dunne and Perlo-Freeman because it 'worked far better than a linear specification' [57, Footnote 13], but it is not clear on what basis it was decided to model logged shares.

5.2.1 Levels

The simplest form of (5.4) is a simple linear function (for example [63]). It is the least widely used of the three, but has the virtue of modelling the variable one is actually concerned with (the level of military expenditure) directly [173].

If the welfare function (5.1) is of Stone-Geary form [127, 186]:

$$W = a_0 \ln C + (1 - a_0) \ln S,$$

and the security function (5.3) linear:

$$S = a_1 + a_2 M + a_3 E + a_4 O + a_5 Z,$$

then welfare maximisation then yields a linear demand for military expenditure:

$$M = \alpha_0 + \alpha_1 (Y/r) + \alpha_2 E + \alpha_3 O + \alpha_4 Z, \tag{5.5}$$

where $\alpha_0 = -a_0 a_1/a_2$, $\alpha_1 = (1 - a_0)$, $\alpha_2 = -a_0 a_3/a_2$, $\alpha_3 = -a_0 a_4/a_2$ and $\alpha_4 = -a_0 a_5/a_2$.

Since data on prices of military output are not available for the UK over the sample period, it is not possible to estimate (5.5). This is commonly overcome by assuming a constant ratio between civil and military prices, $P_C/P_M = r$, to obtain the linear demand:

$$M = \alpha_0 + \alpha_1' (Y/P_C) + \alpha_2 E + \alpha_3 O + \alpha_4 Z, \tag{5.6}$$

where $\alpha_1' = (1 - a_0)$ and all other coefficients unchanged. The assumption of a constant price ratio may not be an innocuous assumption: in Canada, where data is avaialble, the ratio does vary and is a determinant of Canada defence expenditures [192].

Care must be taken over the interpretation of the coefficients obtained from estimating (5.6), since the parameters of the underlying welfare and security functions cannot be recovered. Never-the-less if, when estimated, O is found to have a negative effect on military expenditure, this suggests that greater outsourcing increases security through improving the efficiency with which defence is produced.

5.2.2 Logged levels

The log linear form of (5.4) frequently produces a better fit than the linear model [149, 192]. It also has the virtue of offering a convenient interpretation of the estimated coefficients as elasticities. It may be derived by assuming a two factor CES welfare function [184]:

$$W = b_0 \left(b_1 C^{-b_2} + (1 - b_1) S^{-b_2} \right)^{-1/b_2}$$

and a Cobb-Douglas security function:

$$S = b_3 M^{b_4} E^{b_5} O^{b_6}.$$

The first order conditions for maximising the welfare function (5.1) subject to the budget constraint (5.2) and security function (5.3) may be written:

$$\frac{\partial W}{\partial S}\frac{\partial S}{\partial M} = \frac{P_M}{P_C}\frac{\partial W}{\partial C}.$$

Inserting the partial derivatives for the particular forms used here and substituting for S using the security function yields:

$$\frac{(1-b_1)}{b_1} b_4 C^{b_2+1} = b_3^{b_2} M^{b_2 b_4+1} E^{b_2 b_5} O^{b_2 b_6} \frac{P_M}{P_C}.$$

Rearranging and taking logs gives a log-linear demand for military expenditures:

$$\ln M = \beta_0 + \beta_1 \ln C + \beta_2 \ln E + \beta_3 \ln O + \beta_4 \ln (P_M/P_C)$$

where $\beta_0 = [\ln ((1-b_1)/b_1) + \ln b_4 - b_2 \ln b_3] / (1+b_2 b_4)$, $\beta_1 = \frac{1+b_2}{(1+b_2 b_4)}$, $\beta_2 = \frac{-b_2 b_5}{(1+b_2 b_4)}$, $\beta_3 = \frac{-b_2 b_6}{(1+b_2 b_4)}$, and $\beta_4 = \frac{1}{(1+b_2 b_4)}$. Assuming a constant price ratio $P_C/P_M = r$ and appending a vector of other shift variables, **Z**, yields the estimable log-linear function:

$$\ln M = \beta_0' + \beta_1 \ln C + \beta_2 \ln E + \beta_3 \ln O + \boldsymbol{\beta}_5 \mathbf{Z} \tag{5.7}$$

which may be estimated. In equation (5.7) $\beta_0' = \beta_0 + 1/((1+b_2 b_4) r)$ and the other coefficients are unchanged. This model is frequently estimated with national income, Y, replacing C in equation (5.7) [149, 192, for example], an approach also adopted here.

5.2.3 Shares

The final form of (5.4) considered is a model of the share of GDP allocated to defence spending [185, for example]. It may be derived as one of the Deaton-Muellbauer share equations from an AIDS model with only two goods:

$$M/Y = \gamma_0 + \gamma_1 \ln Y + \gamma_2 \ln P_C + \gamma_3 \ln P_M + \boldsymbol{\gamma}_4 \mathbf{Z}. \tag{5.8}$$

Z is a vector of shift variables, one of which is taken to be O, representing the level of outsourcing. Once again, equation (5.8) cannot be estimated because it includes military prices. Since there are only two goods, imposing homogeneity means $\gamma_2 + \gamma_3 = 0$, which combined with the assumption of a constant price ratio yields:

$$M/Y = \gamma_0' + \gamma_1 \ln Y + \boldsymbol{\gamma}_4 \mathbf{Z}. \tag{5.9}$$

with $\gamma_0' = \gamma_0 + \gamma_2 \ln r$. The absence of data on military prices makes this restriction untestable, but estimation of (5.9) is at least feasible. The shares model can be further simplified if one is also willing to impose homotheticity, $\gamma_1 = 0$, and the defence share of GDP share of military may be modelled as a linear function of the shift variables in $\mathbf{Z}$. Since both the strategic environment, E, and the method of production, O, are amongst the shift variables, the following model may be estimated:

$$M/Y = \gamma_0' + \gamma_1' \ln E + \gamma_2' O + \boldsymbol{\gamma}_4 \mathbf{Z}. \tag{5.10}$$

This specification is preferred to (5.9) because it is a more parsimonious model than those for logs and levels above; it does not include national income on the right-hand side. Because of the very small dataset used for estimation this parsimony is valuable.

5.3 Choice of model

There are no strong grounds to choose between the three models of the previous section. One way in which they do differ is in their implications for the shape of the Engel curves for defence spending. If one had a strong view on the shape of the Engel curves, this would provide grounds on which to choose between the three models but it is not clear, *a priori*, whether the slope of the Engel curves should be increasing, decreasing or constant. Given this, there are no theoretical reasons to prefer one form over the others; the data requirements are also the same. Previous authors using these models have tended to choose between them either arbitrarily or to pick the measure of military expenditure which best fits the data. Here, such arbitrary choices are avoided by estimating and presenting the results of all three models.

There are, though, choices to be made in finding real world variables which accurately reflect the rather vague theoretical variables required by each of the demands (5.6), (5.7) and (5.10). The difficulty is not finding candidate variables—see, for example, the variety of models in Hartley and Sandler [91]—rather, because sample size restricts analysis to relatively parsimonious specifications, the difficulty is choosing between variables each with valid claims to inclusion.

The requirements for Y and P_C are straightforward with national income and the RPI used, respectively. Since defence is normal good, the coefficient on Y is expected to be positive.

The strategic environment, E, may be split: the effects of allies on the UK, and the effect of the threat from (potential) enemies. NATO defence spending is used to capture Allies' impact on the UK's security (also used by [63, 90, 149, 192]). Increased spending by allies increases the UK's security for a given level of spending, so has a positive effect in the security function (5.3), but the expected sign on its coefficient in (5.4) will depend on whether the UK is a free-rider within the alliance, contributes proportionately, or acts as a leader spending more than proportionately. Quantifying the threat, or perceived threat, to the UK is more difficult. For the Cold War period USSR military spending is used [90, 184], which is expected to increase UK military spending as it responds to perceived threat. Since the end of the Cold War the threats facing the UK are more diverse and harder to quantify.

In the absence of better measures, dummy variables are used to reflect major deployments during the period: the Gulf War, Afghanistan and Iraq. These are expected to increase defence spending.

The production variable, O, is represented by the proportion of defence production outsourced. If outsourcing leads to cost and efficiency savings then demand for military expenditure will fall *ceteris paribus*, and the coefficient on an outsourcing variable should be negative.

The final variable in the demands (5.6), (5.7) and (5.10) is Z which allows for any other variables, especially domestic political considerations. The variables used here are population and government spending in areas other than defence.[3] Population is included largely because others have used it [56, 149]. Even if defence is truly non-rivalrous, the coefficient on population need not be zero: a positive coefficient may be capturing any scale effects, or the benefits of defence arising from per capita consumption rather than the national income included here [185, p72].[4]

Government spending affects the allocation of non-military expenditures between public and private uses, and so might be thought of as shifting the parameters of the welfare function. If defence benefits from broader government spending, the sign of the estimated coefficient will be positive [128]. Alternatively, if government spending represents the opportunity cost of defence spending, the sign will be negative [149].

5.4 Data

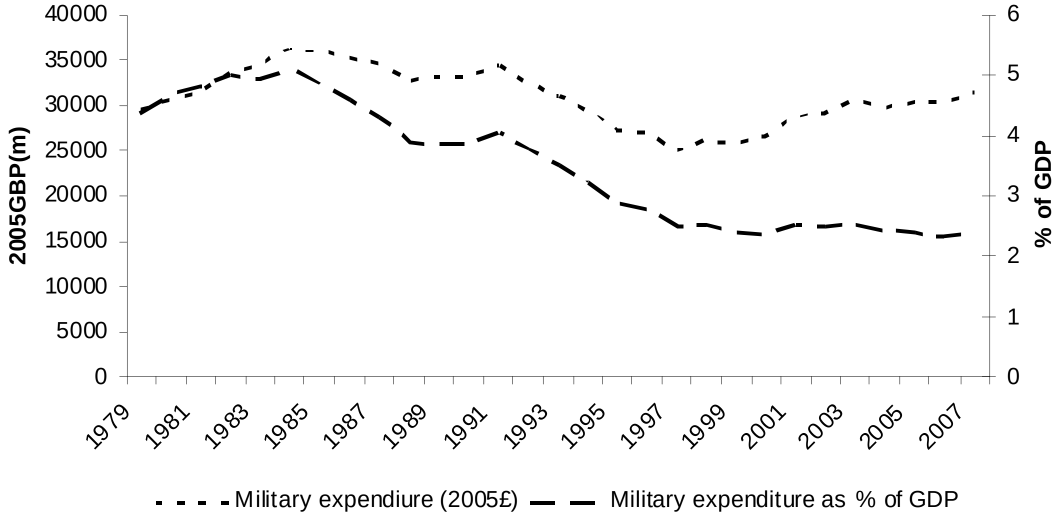
Equations (5.6), (5.7) and (5.10) are estimated using a sample of annual time-series data over the period 1979–2007. Figure 5.1 plots the level of military expenditure (LME) and defence share of GDP (SME). The fluctuations in these series are very similar. They rise until 1983–4, fall during the immediate post-Cold War period, and then level out or gradually rising between 1998 and 2001. The share of defence in GDP, SME, falls steadily until it levels off in the late 1990s. The other variables used are plotted (and summary statistics provided) in Section A.4 of Appendix A.

Table 5.1 details the variables used in estimation and their sources. Most are straightforward and require no explanation beyond that in the previous section. However, there is no publicly available measure of the level of outsourcing production, O. Instead the proportion of MOD contracts let by competitive tender (OUT) is used as a proxy. The rationale is twofold. First outsourcing an activity requires a competitive tendering process and so increases the proportion of contracts let by competitive tender. Secondly, the use of competitive procurement and outsourcing are elements of the same broad policy initiative, which sought to increase the use of competition and the involvement of the private sector in

[3] Another natural candidate would be a dummy variable reflecting the political party in government. Others are favoured here because a political dummy has not been found significant in the UK [90].

[4] Scale effects will arise even if a pure public good since as population increases, cost per capita decreases, and so demand may rise if the good is normal [149].

Figure 5.1: Graph of real miltary expenditures and defence share of GDP



the provision of public services. Alternative proxies for the level of outsourced production, and the effect of their use on the results below are considered in Appendix B. The main conclusions are not sensitive to the use of OUT as a measure of outsourcing.

5.5 Estimation methodology

At first glance the estimation of equations (5.6), (5.7) and (5.10) is a straightforward exercise, but it is complicated by uncertainty about the stationarity, or otherwise, of some of the key series used. Unit root tests on the series used (albeit over a small sample) are presented and discussed in Appendix C. They fail to conclusively determine the stationarity of the series LME and ME, and well as some other regressors. Because of this uncertainty an Autoregressive Distributed Lag (ARDL) approach [162, 163] is used to obtain the estimated long-run demands for military expenditure. The more widely used Engel-Granger two-step method is not feasible given the uncertainty over the order of integration of the regressors.

The key advantage of the ARDL approach is that it may be applied irrespective of whether regressors are I(1), I(0), or some combination of the two.[5] The approach comprises two stages: first, test for the absence of a long-run (possibly cointegrating) relationship between the variables using bounds tests. If a relationship is found, the second stage is to estimate an ARDL model and find the simplest model consistent with the data.

To test for a long-run relationship suppose that military expenditure can be approximated by a VAR(ρ) augmented with a vector of dummy variables. Estimate the following

[5] It has been widely used in the estimation of demand of military expenditure models [56, 63, 149, 192, 231, for example].

Table 5.1: Definition of variables used

Variable	Label	Description
M	LME	The level of UK military spending
	ME	The natural logarithm of the level of military expenditure
	SME	Military spending as a share of GDP
Y	GDP	UK real GDP
O	OUT	Level of outsourcing
E	SPILL	NATO military spending excluding the UK (US spending for shares models)
	THREAT	Military spending of USSR during the Cold War
	CWDUM	Cold War dummy
	GDUM	Gulf War dummy
	AFGDUM	Dummy for Afghanistan deployments
	IRAQ	Iraq dummy
Z	GOVT	Real government non-defence expenditure
	POP	Resident population of the UK
	T	Linear trend
	TBREAK	Linear trend beginning in 2001

Notes: Any variable prefixed *LN* denoted the natural logarithm of than variable. Similarly the prefix *PR* denoted that variable as a proportion of GDP.

by OLS:

$$\triangle M_t = \alpha_0 + \sum_{i=1}^{\rho} \alpha_i \triangle M_{t-1} + \sum_{i=1}^{\rho} \boldsymbol{\beta}_i' \triangle \mathbf{x}_{t-i} + \boldsymbol{\gamma}' \mathbf{d}_t + \delta_0 M_{t-1} + \boldsymbol{\delta}' \mathbf{x}_{t-1} + \varepsilon_t \tag{5.11}$$

where M_t is military expenditure, $\mathbf{x}_t$ is a k vector of the other elements of the underlying VAR(ρ), $\mathbf{d}_t$ is a vector of dummy variables, and $\boldsymbol{\delta}' = (\delta_1, \delta_2, ...\delta_k)$. An F-test of the null hypothesis $\delta_0 = \delta_1 = ... = \delta_k = 0$ is a test of the hypothesis of no long-run relationship. The test is valid irrespective of whether the regressors are I(1) or I(0), but depends crucially on an assumption of serially uncorrelated errors. The distribution of the F-statistic is non-standard and depends on k.[6] If the null of no relationship is rejected, and any relationship not thought spurious, then equation (5.11) is a valid model and model selection techniques may then be applied to obtain a more parsimonious specification. Minimising AIC is used as the model selection criterion here.

The flexibility of the ARDL approach comes at a cost. As with VAR modelling, the number of parameters increases rapidly as the complexity of the model increases. Given the small sample available here, a balance must be struck between allowing lags, ρ, to adequately capture the dynamics of defence expenditure and allowing a richer models with more regressors, k. We are restricted to very simple models with few regressors and less lags than would be ideal.

[6]The distribution, and so the critical values, also depend on the mix of I(0) and I(1) regressors. The value used here is the critical value for the case in which all regressors are I(1). This is the upper bound of the true critical value.

5.6 Results

In the absence of a compelling *a priori* reason to prefer one model over the others, the ARDL methodology is applied to (5.6), (5.7) and (5.10), and the estimated models the levels, logged levels and shares of military expenditure are presented below in turn.

5.6.1 Modelling level of military expenditure

The starting point for modelling the levels of military expenditure is a relatively broad model, in the sense that we have relatively large k but small ρ. In equation (5.11) the elements of $\mathbf{x}$ are GDP, OUT, GOVT, POP, SPILL, and THREAT. The elements of $\mathbf{z}$ are CONST, AFGDUM, CWDUM and IRAQ as well as a trend (T) and a trend from 2001 (TBREAK) in order to allow a trend break in 2001.

The long run effects of GDP on military spending are expected to be positive since defence is considered a normal good. These long run effects are likely to operate with a lag, so the short run effects may be lower. OUT is expected to have a negative long run effect, since increased levels of outsourced provision are expected to reduce overall expenditures. Again this long run effect may take time to become apparent so the short run effects may be weaker. The effects of GOVT on military spending are expected to be positive since the expectation is that military expenditures follow the ebbs and flows of broader public spending. If POP has any effect it is expected to be positive because of the scale effects of defence as a public good. The long run effect of both SPILL and THREAT are expected to be positive. SPILL because of the UK's role as a leader in NATO, THREAT because increased spending by the USSR are expected to be matched, albeit with a lag, by the UK. These lags may mean that THREAT's short run effects are lower than the long run. CWDUM is expected to have a positive long run effect as it should reflect the decreased spending in the immediate post-Cold War period (the peace dividend). Neither AFGDM nor IRAQ are anticipated to have long run effects, but in the short run the increased spending on these operations is expected to be picked up with a positive coefficient on both.

The model is estimated over the period 1981–2007 with $\rho = 1$ (so allowing for a maximum of two lags of each variable in the ARDL model in levels).[7] The statistic for a test of a long-run relationship is 5.8932 against a 1% critical value of 4.978, rejecting the null of no relationship. However, the testing regression shows evidence of serially correlated errors.

If this is reflecting true autocorrelation in the error term, then it would invalidate the bounds testing procedure. It is more likely to result from poorly captured dynamics. Given the long term nature of defence procurement programmes, and the need to build capabilities before they can be utilised, allowing only two lags in the underlying ARDL model is unlikely to fully capture the dynamics.

Clearly, in an ideal world one would allow many more lags in the underlying ARDL

[7] The largest available sample for each model is used in estimation. Although this means the sample size changes between models, the alternative of ignoring some observations in the models with fewer lags was thought to be unjustifiable given the extremely small sample.

Table 5.2: Tests for long-run relationships in models of the levels of military expenditure

N=26 1982–2007	Model LEV-A	Model LEV-B	Model LEV-C	Model LEV-D	Model LEV-E
Elements of $\mathbf{x}$	OUT, GOVT, SPILL, THREAT	GDP, OUT, GOVT, SPILL	GDP, GOVT, SPILL, THREAT	GDP, OUT, GOVT, THREAT	GDP, OUT, SPILL, THREAT
	Tests for long-run relationship				
	F(5, 7)= 3.336	F(5, 7)= 7.562***	F(5, 7)= 1.263	F(5, 7)= 6.911***	F(5, 7)= 6.041***
	Serial Correlation tests[a]				
1 lag	F(1, 6)= 11.131**	F(1, 6)= 5.7545*	F(1, 6)= 7.748**	F(1, 6)= 8.720**	F(1, 6)= 2.446
2 lags	F(2, 5)= 7.488**	F(2, 5)= 8.062**	F(2, 5)= 35.134***	F(2, 5)= 13.099**	F(2, 5)= 1.101=
3 lags	F(3, 4)= 4.052	F(3, 4)= 4.499*	F(3, 4)= 19.657***	F(3, 4)= 8.546**	F(3, 4) 0.890

Notes: *, **, and *** denote significance at the 10%, 5% and 1% levels respectively. For each model $\mathbf{d}$=(CONST, AFGDUM)$'$, and T and TBREAK included. [a]Breusch-Godfrey LM tests on the testing regression.

model in order to better capture the dynamics of defence spending. Allowing for many lags and any conceivable regressor would allow a truly general to specific specification search to be undertaken, but the small sample size and consequent trade-off between k and ρ makes this impossible. If more lags are to be included, it requires the use of simpler models with fewer explanatory variables. Including a third lag in the underlying ARDL model requires two elements of $\mathbf{x}$ to be dropped. POP is dropped because the scale effects of population changes are thought likely to be small; indeed, population was found to be insignificant in Nikolaidou's [149] model for UK military expenditure. There is no clear candidate for the second, though GOVT and THREAT (because it represents threats over less than one third of the sample period) are thought to be the most likely candidates. These are not, however, such compelling candidates as POP, so all of the possible specifications resulting from excluding POP and one other element of $\mathbf{x}$ are considered. The dummies are simplified to $\mathbf{d}$ =(CONST, AFGDUM)$'$, T and TBREAK are retained.

Table 5.2 presents the long-run relationship and serial correlation test statistics for each of these revised models of the level of military expenditure with three lags allowed in the underlying ARDL. The model corresponding to omitting POP and THREAT is LEV-B, LEV-E omits POP and GOVT, respectively.

Both LEV-B and LEV-E yield significant test statistics in the test for a long-run relationship, but LEV-B appears to suffer serially correlated errors. Of all the models considered, only LEV-E appears to have both a significant relationship statistic and no evidence of serially correlated errors.[8]

[8]Examination of the individual t-statistics would be interesting and may yield some insight into the nature of the relationships between variables. However, the F-statistics for testing for a long run relationship have non-standard distributions; their critical values depend on the stationarity or otherwise of each of the variables included. It is likely that the critical values for t-tests on individual variables would also be non-standard, but

The test statistics in Table 5.2 might be interpreted in two ways. Firstly, long-run military expenditures are driven by the interaction of GDP and OUT. Each of the models including both rejects the null of no relationship. Of these, LEV-E shows no evidence of serially correlated errors and is valid. The significant statistics of the serial correlation tests on models LEV-B and LEV-D are picking up the omitted variables THREAT and SPILL, respectively. The estimated error correction form and long-run parameters for LEV-E are presented in the first column of Table 5.4 and the long-run parameters in Table 5.5.

A second, perhaps more compelling, interpretation of the results in Table 5.2 is that the preponderance of significant serial correlation statistics indicates that the dynamics of military spending are still not properly captured by three lags. Unfortunately, the sample is such that allowing a fourth lag in the ARDL necessitates the use of models with only three regressors in $\mathbf{x}$. This is probably too few to accurately model defence spending, but for completeness the relationship and serial correlation statistics are presented in Table 5.3 for models allowing four lags. Still having no grounds to favour the exclusion of one regressor over the others, statistics for each of the models which result from excluding an additional variable from the models in Table 5.2 are presented in Table 5.3.

The tests in Table 5.3 show that whilst all models yield significant relationship test statistics, all but two (including LEV-K which results from excluding POP, GOVT and THREAT) also show evidence of serial correlation, invalidating a bounds testing approach. The error correction representation of models LEV-J and LEV-K are presented alongside LEV-E in Tables 5.4 and 5.5, but are not thought to be reliable models of military spending. Overall, the statistics in Tables 5.2 and 5.3 suggest that given the trade-off between ρ and k, the sample is too small to find an adequate long-run model of military spending: there are either insufficient lags or the model is insufficiently broad. No adequate compromise between the two competing demands is possible given the sample size.

Table 5.4 presents the short-run estimates of the three specifications which in Tables 5.2 and 5.3 exhibit both significant long-run relationship test statistics, and no evidence of serially correlated errors. Both LEV-J and LEV-L are just identified, so diagnostic statistics are unavailable. They should not be relied upon, but incidentally in the corresponding long-run estimates presented in Table 5.5 neither appears to explain the long-run evolution of military expenditure. Model LEV-E still contains far too many parameters relative to the sample size for comfort, but does at least allow the calculation of some diagnostic statistics. None of these cause concern. The short-run model seems reasonable, but none of the long-run parameters are significant. That not a single coefficient in any of the estimated long-run models presented in Table 5.5 is significant suggests that none of models estimated is able to explain even a fraction of the long run level of military expenditures. It suggests that the traditional demand for military expenditure models do not work well in the UK when

as far as the author is aware have not been derived. Since in the absence of reliable critical values it is not clear how to interpret t-statistics, they are not presented.

Table 5.3: Tests for long-run relationships in further revised levels models

N=25 1983–2007	Model LEV-F	Model LEV-G	Model LEV-H	Model LEV-I	Model LEV-J	Model LEV-K	Model LEV-L	Model LEV-M	Model LEV-N	Model LEV-M
Elements of **x**	OUT, GOVT, SPILL	GOVT, SPILL, THREAT	OUT, GOVT, THREAT	OUT, SPILL, THREAT	GDP, GOVT, SPILL	GDP, OUT, SPILL	GDP, OUT, THREAT	GDP, GOVT, THREAT	GDP, SPILL, THREAT	GDP, OUT, THREAT
Tests for long-run relationship										
	F(4, 5)= 9.169***	F(4, 5)= 7.548**	F(4, 5)= 5.882**	F(4, 5)= 10.289***	F(4, 5)= 9.213***	F(4, 5)= 6.723**	F(4, 5)= 9.637***	F(4, 5)= 5.864**	F(4, 5)= 18.063***	F(4, 5)= 11.892***
Serial Correlation tests										
1 Lag	F(1, 4)= 7.316*	F(1, 4)= 1.541	F(1, 4)= 0.456	F(1, 4)= 2.063	F(1, 4)= 0.622	F(1, 4)= 3.571	F(1, 4)= 2.282	F(1, 4)= 0.286	F(1, 4)= 2.607	F(1, 4)= 0.284
2 lags	F(2, 3)= 23.738**	F(2, 3)= 19.933**	F(2, 3)= 12.112**	F(2, 3)= 38.157***	F(2, 3)= 3.327	F(2, 3)= 55.746***	F(2, 3)= 4.671	F(2, 3)= 2.057	F(2, 3)= 9.076*	F(2, 3)= 8.312*
3 lags	F(3, 2)= 31.149**	F(3, 2)= 8.892	F(3, 2)= 9.870*	F(3, 2)= 32.709**	F(3, 2)= 8.684	F(3, 2)= 62.050**	F(3, 2)= 5.228	F(3, 2)= 22.784**	F(3, 2)= 23.652**	F(3, 2)= 5.240

Notes: See Table 5.2

Table 5.4: Short-run estimates of error correction representation for levels

	Model LEV-E		**Model LEV-J**		**Model LEV-L**	
Elements of **x**	GDP, OUT, SPILL, THREAT		GDP, GOVT, SPILL		GDP, OUT, THREAT	
	N=26, 1982-2007		N=25, 1983-2007		N=25, 1983-2007	
	Coefficient	**Standard Error**	**Coefficient**	**Standard Error**	**Coefficient**	**Standard Error**
DRME(-1)	-1.330*	0.653	1.467	1.870	-0.134	0.157
DRME(-2)	-0.635	0.422	1.863	1.070	-0.048	0.226
DRME(-3)	-	-	0.916*	0.410	0.241	0.159
DRGDP	-0.075	0.043	-0.004	0.046	0.018	0.018
DRGDP(-1)	0.0142	0.020	0.068	0.106	-0.002	0.018
DRGDP(-2)	0.030	0.023	0.081	0.058	0.003	0.014
DRGDP(-3)	-	-	0.024	0.047	0.089***	0.018
DOUT	-16554.7***	4776.7	-	-	3307.1	2909.5
DOUT(-1)	27046.7**	10574.6	-	-	11691.4	7198.5
DOUT(-2)	11438.2	6203.5	-	-	-7365.3	7652.3
DOUT(-3)	-	-	-	-	-6931.8	3659.6
DGOVT	-16554.7	4776.7	0.034	0.175	.216**	0.063
DGOVT(-1)	27046.7	10574.6	0.049	0.079	-0.009	0.025
DGOVT(-2)	11438.2	6203.5	-0.004	0.090	-.122***	0.027
DGOVT(-3)	-	-	0.061	0.128	-.130**	0.037
DSPILL	-.094**	0.037	-0.015	0.047	-	-
DSPILL(-1)	.063**	0.022	-0.045	0.037	-	-
DSPILL(-2)	-	-	-0.023	0.022	-	-
DSPILL(-3)	-	-	-0.018	0.023	-	-
DTHREAT	0.043	0.034	-	-	-	-
DTHREAT(-1)	-.140**	0.047	-	-	-	-
DTHREAT(-2)	-.092**	0.033	-	-	-	-
CONST	102490.2**	31539	93561.2	78912.5	83447.6***	19022.4
AFGDUM	-2972.7	1701.7	601.486	4405.1	1194.6	997.435
T	4753.8*	2173.8	2705.5	2813.4	-987.334	991.368
TBREAK	3490.4***	748.662	2992.6	3162.6	954.0991	1048.2
ECM(-1)	0.233	0.667	-2.21	2.120	-1.350	0.268
			Diagnostics			
SC test[a]	F(1, 3)= 0.024288		N/A		N/A	
RESET[b]	F(1, 3)= 0.015553		N/A		N/A	
Het test[c]	F(1, 24)= 0.15722		N/A		N/A	
AIC	-186.8352		-188.284		-158.4085	
RSS	488868.5		809134.1		74137.9	

Notes: Prefix D denotes the first difference of a variable, suffix (- *i*) denotes the *i*th lag. ECM is the error correction term. [a]Godfrey-Breusch-Pagan LM tests on the underlying ARDL model for AR(1) errors. [b]Using the square of the fitted value. [c]Kroenker's test for heteroskedastic errors using the squared fitted values and is asymptotically valid in the presence of non-normal errors.

Table 5.5: Long-run estimates of error correction representation for levels

	Model LEV-E		**Model LEV-J**		**Model LEV-L**	
Elements of $\mathbf{x}$	RGDP, OUT, SPILL, THREAT		RGDP, GOVT, SPILL		RGDP, OUT, THREAT	
	N=26, 1982-2007		N=25, 1983-2007		N=25, 1983-2007	
	Coefficient	**Standard Error**	**Coefficient**	**Standard Error**	**Coefficient**	**Standard Error**
GDP	0.529	1.642	-0.048	0.023	-0.043*	0.013
GOVT	-	-	-0.043	0.106	0.088	0.051
SPILL	0.587	1.536	0.027	0.020	-	-
OUT	269396.2	758353.7	-	-	-11198.6	5202.7
THREAT	-1.096	3.162	-	-	-	-
CONST	-440121.2	1325635	42335.2	13134.7	61801.5***	6160.9
AFGDUM	12765.7	33917	272.164	1787.7	884.696	876.641
T	-20414.1	60928.6	1224.2	1272.9	-731.222	851.308
TBREAK	-14988.7	41660.2	1354.1	2294.8	706.609	664.942

Notes: See Table 5.4

applied to the levels of military spending.

Model LEV-E is the best model specification for the level of military expenditures, but it is a poor model. This may be because the regressors used do not actually drive the levels of military expenditure in the long-run; because we have been unable to capture the long-run relationships (perhaps because of the simplistic representation of the strategic environment adopted); or most likely because the small sample is simply insufficient to identify the true long-run relationship. Irrespective of the reason it is not possible to find a well specified model of the level of military expenditures based on equation (5.5)—even by data mining and trying all possible combinations of regressors. As such it is not possible to determine, by modelling the levels of military expenditure, whether the level of outsourcing affects military spending in the long-run.

5.6.2 Estimation of logs model for military expenditure

The starting point for estimating a model of logged military spending based on (5.7) is analogous to the staring point for the levels model above. The elements of $\mathbf{x}$ in equation (5.11) are LNGDP, LNOUT, LNGOVT, LNPOP, LNSPILL, and LNTHREAT. The elements of $\mathbf{z}$ are CONST, AFGDUM, CWDUM, and IRAQ. T and TBREAK are included, and initially $\rho = 1$. This model yields a statistic of 4.6559 in a test for long-run relationships against a 5% critical value of 4.088, so rejecting the null of no relationship. As in the levels case, there is strong evidence of serial correlation in the errors.

Again taking this as evidence of mis-specified dynamics rather than true serial correlation, three lags are allowed, and the number of regressors reduced. The relationship and serial correlation tests for the models analogous to those in Table 5.2 are presented in Table

Table 5.6: Tests for long-run relationships in models of logged military expenditure

N=26 1982–2007	Model LN-A	Model LN-B	Model LN-C	Model LN-D	Model LN-E
Elements of **x**	LNOUT, LNGOVT, LNSPILL, LNTHREAT	LNGDP, LNOUT, LNGOVT, LNSPILL	LNGDP, LNGOVT, LNSPILL, LNTHREAT	LNGDP, LNOUT, LNGOVT, LNTHREAT	LNGDP, LNOUT, LNSPILL, LNTHREAT
	Tests for long-run relationship				
	F(5, 7)= 8.336***	F(5, 7)= 12.148***	F(5, 7)= 7.321***	F(5, 7)= 17.420***	F(5, 7)= 13.177***
	Serial Correlation tests				
1 lag	F(1, 6)= 2.108	F(1, 6)= 4.650*	F(1, 6)= 8.729**	F(1, 6)= 6.663**	F(1, 6)= 3.220
2 lags	F(2, 5)= 1.118	F(2, 5)= 4.978*	F(2, 5)= 4.027*	F(2, 5)= 3.220	F(2, 5)= 4.084*
3 lags	F(3, 4)= 6.291*	F(3, 4)= 27.738***	F(3, 4)= 2.890	F(3, 4)= 4.176	F(3, 4)= 11.347**

Notes: See Table 5.2.

5.6. Whilst each has significant relationship test statistics, all also show strong evidence of serially correlated errors so a further lag is allowed for, but at the cost of further simplifying the models under consideration. Relationship and serial correlation test statistics for the logged equivalents of the models in Table 5.3, are presented in Table 5.6. These allow four lags in the underlying ARDL.

Once again the test statistics from the tests for long-run relationships are significant, but most of the models show strong evidence of serial correlation invalidating the bounds testing procedure for testing relationships. Only three models show no evidence of serial correlation: models LN-G, LN-H, LN-K, and LN-M. The estimated long- and short-run error correction representation of each of these models are presented in Tables 5.8 and 5.9.

Of the results presented in Tables 5.8 and 5.9, Models LN-G and LN-M are only just identified, and none has enough degrees of freedom to calculate diagnostic statistics. LN-M is without a single individually significant coefficient in either the short- or long-run. It warrants no further discussion. LN-G is little better. The error correction term does at least seem relevant though only LNOUT is individually significant, and then only at 10% level. It is not a satisfactory model of logged defence expenditures, though it is worth noting that the coefficient on LNOUT is negative.

LN-K is the preferred model of logged military expenditures, but it still has too few degrees of freedom to calculate diagnostics. As such the model cannot be considered anything like reliable, but for completeness the long-run coefficients are interpreted: The estimated coefficient on LNOUT is negative and significant, as in the shares model. LNGDP is positive as expected *a priori* and significant at the 10% level. This is at odds with Hartley and MacDonald [90] who estimate a similar model (though without LNOUT) using similar data over a longer period; they find LNGDP to be insignificant. The sign on TBREAK is positive and significant reflecting the increases in expenditure since 2001.

Table 5.7: Tests for long-run relationships in further revised logged models

N=25 1983–2007	Model LN-F	Model LN-G	Model LN-H	Model LN-I	Model LN-J	Model LN-K	Model LN-L	Model LN-M	Model LN-N	Model LN-O
Elements of **x**	LNOUT, LNGOVT, LNSPILL	LNGOVT, LNSPILL, LNTHREAT	LNOUT, LNGOVT, LNTHREAT	LNOUT, LNSPILL, LNTHREAT	LNGDP, LNGOVT, LNSPILL	LNGDP, LNGOVT, LNOUT	LNGDP, LNOUT, LNSPILL	LNGDP, LNGOVT, LNTHREAT	LNGDP, LNSPILL, LNTHREAT	LNGDP, LNOUT, LNTHREAT
Tests for long-run relationship										
F(4, 5)	10.287***	9.889***	7.748***	9.219***	9.904***	71.187***	5.800**	7.271***	10.252***	5.773**
Serial Correlation tests										
1 lag, F(1,4)	1.598	0.871	4.378	2.735	0.684	1.808	1.120	0.096	0.773	0.575
2 lags, F(2, 3)	6.458*	2.789	5.408	26.880**	4.326	10.696**	10.731**	0.201	18.140**	273.820***
3 lags, F(3, 2)	3.599	1.385	2.404	15.473*	4.273	9.081	139.293***	0.819	8.512	137.435***

Notes: See Table 5.2.

Table 5.8: Table of logs short-run estimates

	Model LN-G		**Model LN-H**		**Model LN-K**		**Model LN-M**	
1983-2007	LNGOVT,		LNOUT,		LNGDP,		LNGDP,	
N=25	LNSPILL, LNTHREAT		LNGOVT, LNTHREAT		LNGOVT, LNOUT		LNRGDP, LNTHREAT	
	Coeff.	**Std. Error**	**Coeff.**	**Std. Error**	**Coeff.**	**Std. Error**	**Coeff.**	**Std. Error**
DLNRME(-1)	1.669*	0.843	0.428	0.301	0.871**	0.213	0.498	1.704
DLNRME(-2)	1.774**	0.487	0.819	0.160	1.024***	0.166	2.106	2.079
DLNRME(-3)	-	-	0.493	0.137	0.583**	0.109	0.709	0.607
DLNRGDP	-	-	-	-	0.255	0.413	4.160	7.575
DLNRGDP(-1)	-	-	-	-	-1.797	0.337	-3.434	4.306
DLNRGDP(-2)	-	-	-	-	-0.814	0.278	-0.267	2.846
DLNRGDP(-3)	-	-	-	-	-	-	-5.754	8.455
DLNOUT	-	-	-0.152*	0.066	-0.003	0.017	-	-
DLNOUT(-1)	-	-	0.196**	0.068	0.310***	0.036	-	-
DLNOUT(-2)	-	-	0.170**	0.053	0.178**	0.033	-	-
DLNOUT(-3)	-	-	0.056	0.036	0.050*	0.019	-	-
DLNGOVT	-1.556	1.058	0.575*	0.261	0.369	0.160	-2.740	4.055
DLNGOVT(-1)	1.528*	0.597	-0.805**	0.303	-0.883	0.234	-4.081	5.392
DLNGOVT(-2)	-2.335*	0.936	-0.746	0.446	-1.109	0.185	2.706	5.769
DLNGOVT(-3)	0.221	0.288	-0.4	0.336	-0.725	0.138	-3.203	3.602
DLNSPILL	-0.170	0.719	-	-	-	-	-	-
DLNSPILL(-1)	-0.709	0.491	-	-	-	-	-	-
DLNSPILL(-2)	0.242*	0.117						
DLNSPILL(-3)	-0.851***	0.189						
DLNTHREAT	0.778***	0.203	-0.007	0.005	-	-	-0.047	0.056
DLNTHREAT(-1)	-1.188	1.495	0.000	0.010	-	-	0.012	0.072
DLNTHREAT(-2)	-0.662	1.024	0.002	0.009	-	-	-0.017	0.075
DLNTHREAT(-3)	-0.829	0.587	-0.003	0.005	-	-	0.023	0.022
CONST	-6.335	9.619	9.303	5.670	-7.943	10.493	-116.046	134.366
AFGDUM	0.119	0.046	0.038	0.017	0.021	0.011	0.190	0.349
T	-0.063	0.060	-0.049*	0.024	-0.115	0.040	-0.375	0.430
TBREAK	0.002	0.016	0.044**	0.015	0.064	0.006	0.075	0.104
ECM(-1)	-1.654	0.918	-1.700***	0.285	-2.209***	0.243	-1.430	1.733
SC Test			N/A		F(1, 2)=4.5791		N/A	
RESET			N/A		N/A		N/A	
Het test			N/A		F(1, 23)=0.48000		N/A	
AIC	81.6087		95.742		102.6234		59.9873	
RSS	0.0003679		1.10E-04		6.85E-05		0.0019151	

Notes: See Table 5.4

Table 5.9: Table of logs long-run estimates

	Model LN-G		Model LN-H		Model LN-K		Model LN-M	
1983-2007	LNGOVT,		LNOUT,		LNGDP,		LNGDP,	
N=25	LNSPILL, LNTHREAT		LNGOVT, LNTHREAT		LNGOVT, LNOUT		LNRGDP, LNTHREAT	
	Coeff.	**Std. Error**	**Coeff.**	**Std. Error**	**Coeff.**	**Std. Error**	**Coeff.**	**Std. Error**
LNRGDP	-	-	-	-	0.666*	0.240	6.339	5.460
LNTHREAT	0.672	1.297	-0.003	0.009	-	-	-0.036	0.091
LNOUT	-	-	-0.211*	0.062	-0.134***	0.015	-	-
LNGOVT	0.469	0.343	0.414	0.303	0.457**	0.132	0.817	2.175
LNSPILL	0.686***	0.067	-	-	-	-	-	-
CONST	-3.830	3.939	5.472	3.566	-3.595	4.469	-81.170	57.916
AFGDUM	0.072	0.032	0.022	0.011	0.010	0.005	0.133	0.146
T	-0.038	0.018	-0.029	0.013	-0.052**	0.014	-0.262	30.146
TBREAK	0.001	0.010	0.026	0.011	0.029***	0.002	0.053	0.119

Notes: See Table 5.4

Although LN-K is preferable to the other models of Tables 5.8 and 5.9, it is not compelling. The poor performance of similar models, and the lack of diagnostic tests make it impossible to draw firm conclusions based upon it. Again, the possibility of simply having identified a spurious regression cannot be discounted. Even mining the data fails to find a convincing model of logged military expenditure. Notwithstanding the small sample period used, the results here raise concerns that the traditional demand for military expenditure models do not fit UK data—not only for the level of military spending (presented in subsection 5.6.1), but also for the most commonly used (logged) specification of equation (5.7).

5.6.3 A model of the defence share of GDP

Given the trade-off between ρ and k, modeling the defence share of GDP (SME) has one significant advantage over the others. It obviates the need to include a measure of national income, so saving degrees of freedom.

Once again we begin with an initial model analogous to those of the two preceding subsections. To begin $\mathbf{x}$ in equation (5.11) contains OUT, PRGOVT, POP, SPILL, and THREAT, and $\mathbf{d}$ contains CONST, AFDGUM, CWDUM, IRAQ, T and TBREAK. Allowing two lags in the underlying ARDL ($\rho = 1$) the long-run relationship test yields a statistic of 2.2563 against a critical value of 4.329. The test cannot reject the null hypothesis of no long-run relationship. There is also strong evidence of serial correlation in the errors from the testing regression, calling into question the validity of the testing procedure. Again progress is made by including an extra lag ($\rho = 2$), with POP omitted from $\mathbf{x}$ and the vector of dummies, $\mathbf{z}$, simplified to (CONST AFGDUM)$'$ by dropping CWDUM and IRAQ. The test for a long-run relationship in the revised model and the associated serial correlation test statistics are presented as Model PR-A in Table 5.10.

Table 5.10: Tests for long-run relationships in models of shares of military expenditure

N=26 1982–2007	Model PR-A	Model PR-B	Model PR-C	Model PR-D	Model PR-E
Elements of **x**	OUT, PRGOVT, SPILL, THREAT	OUT, PRGOVT, SPILL	PRGOVT, SPILL, THREAT	OUT, PRGOVT, THREAT	OUT, SPILL, THREAT
			Tests for long-run relationship		
	F(5, 7)= 6.984***	F(4, 5)= 9.7563***	F(4, 5)= 13.1388***	F(4, 5)= 3.9130	F(4, 5)= 5.1787**
			Serial Correlation tests		
1 lag	F(1, 6)= 6.8888**	F(1, 4)= 0.46281	F(1, 4)= 5.6174*	F(1, 4)= 3.0558	F(1, 4)= 2.6890
2 lags	F(2, 5)= 4.2083*	F(2, 3)= 1.5023	F(2, 3)= 2.7058	F(2, 3)= 5.8176*	F(2, 3)= 2.5806
3 lags	F(3, 4)= 3.3489*	F(3, 2)= 4.6981	F(3, 2)= 2.0632	F(3, 2)= 10.8097*	F(3, 2)= 2.1835

Notes: See Table 5.2

For Model PR-A the test for a long-run relationship in Table 5.10 rejects the null of no relationship at all conventional levels of significance, but there is also some evidence of serial correlation in the errors suggesting that the dynamics of military expenditure are still misspecified. Since there is no clear cut candidate variable to be dropped, the models resulting from omitting each remaining element of **x** in turn are considered. The test statistics for long-run relationships in these models are presented in Table 5.10. Given that NATO membership is a cornerstone of UK defence policy, Model PR-D is thought the least likely to explain military expenditures. Each of the remaining models (PR-B, PR-C, and PR-E) are consistent with Nikolaidou [149] whose estimated logged model found only income and spill-ins from NATO allies are significant. Model PR-E is closest to Smith's [184, 185] models of the UK.

As expected model PR-D is unable to reject a null of no long-run relationship and this model is not considered further. The tests for Models PR-B, PR-C and PR-E all reject the null of no long-run relationship and don't appear to be suffering serially correlated errors. For each, the underlying ARDL(4,4,4,4) are estimated, the preferred lag lengths selected by minimising AIC and the estimated error correction forms and long-run parameters presented in Tables 5.11 and 5.12, respectively.

In Table 5.11 the estimated Model PR-E has significant test statistics for serial correlation. If the result of true serial correlation then the bounds testing approach is not valid. If the result of poorly captured dynamics, then the estimated model is not reliable. Even without the serial correlation test statistics, none of the long-run parameters are significant which contradicts the result of the test for a long-run relationship. This would appear to be due to the variables omitted in the model selection process. Asymptotically or in a large sample this would not matter. That such small changes affect the significance of the error correction

Table 5.11: Short-run estimates of error correction representation for shares

	Model Model PR-B		**Model PR-C**		**Model PR-E**	
Elements of **x**	OUT, PRGOVT, SPILL		PRGOVT, SPILL, THREAT		OUT, SPILL, THREAT	
	N=25, 1983-2007		N=25, 1983-2007		N=25, 1983-2007	
	Coefficient	**Standard Error**	**Coefficient**	**Standard Error**	**Coefficient**	**Standard Error**
DPRRME(-1)	1.74*	0.726	-0.752	0.664	-0.195	0.333
DPRRME(-2)	1.331**	0.447	0.411	0.439	1.232***	0.249
DPRRME(-3)	-	-	0.699**	0.207	1.037***	0.328
DOUT(-1)	2.54	1.477	-	-	-	-
DOUT(-2)	1.095	1.000	-	-	-	-
DPRGOVT	0.053	0.031	0.064	0.034	-	-
DPRGOVT(-1)	-0.140*	0.062	-0.040	0.047	-	-
DPRGOVT(-2)	-0.095**	0.040	-0.035	0.051	-	-
DPRGOVT(-3)	-	-	0.033	0.059	-	-
DSPILL	-0.089	0.184	-0.396	0.404	-0.500**	0.198
DSPILL(-1)	-0.316*	0.159	-0.161	0.313	-0.358	0.202
DSPILL(-2)	-0.237	0.134	0.193	0.207	-	-
DSPILL(-3)	-0.107	0.112	-	-	-	-
DTHREAT	-	-	-0.206E-5	0.118E-5	-0.209E-5**	0.857E-6
DTHREAT(-1)	-	-	-0.240E-4	0.129E-4	-0.149E-4***	0.462E-5
DTHREAT(-2)	-	-	-0.133E-4	0.9883E-5	-0.847E-5**	0.285E-5
DTHREAT(-3)	-	-	-0.847E-5	0.548E-5	-0.309E-5	0.224E-5
CONST	-2.645	3.365	-0.858	3.361	5.102***	1.003
AFGDUM	0.385**	0.132	0.115	0.150	0.179*	0.0858
T	-0.006	0.049	-0.006	0.082	-0.110***	0.025
TBREAK	-0.251	0.164	0.013	0.111	0.124*	0.058
ECM(-1)	-2.307	0.869	-0.390	0.675	-0.557	0.350
			Diagnostics			
SC test	F(1, 4)=0.058787		F(1, 2)=0.74017		F(1, 6)=7.1090**	
RESET	F(1, 4)=0.55906		N/A		F(1, 6)=3.1872	
Het test	F(1, 23)=0.45017		F(1, 23)=0.47479		F(1,23)=0.34011	
AIC	33.4841		43.9400		36.1272	
RSS	0.020288		0.0074899		0.019270	

Notes: See Table 5.4

Table 5.12: Long-run estimates of error correction representation for shares

	Model PR-B		**Model PR-C**		**Model PR-E**	
Elements of **x**	OUT, PRGOVT, SPILL N=25, 1983-2007		PRGOVT, SPILL, THREAT N=25, 1983-2007		OUT, SPILL, THREAT N=25, 1983-2007	
	Coefficient	**Standard Error**	**Coefficient**	**Standard Error**	**Coefficient**	**Standard Error**
PRGOVT	0.089***	0.014	0.338	0.637	-	-
SPILL	0.328**	0.099	-2.341	5.789	-0.892	1.048
OUT	-1.776**	0.577	-	-	3.985	2.380
THREAT	-	-	0.641E-4	0.142E-3	0.287E-4	0.214E-4
CONST	-1.147	1.188	-2.198	10.697	9.154	5.541
AFGDUM	0.167**	0.065	0.294	0.341	0.322	0.194
T	-0.003	0.022	-0.014	0.197	-0.198	0.107
TBREAK	-0.109**	0.040	0.034	0.308	0.223	0.221

Notes: See Table 5.4

term, and so the estimated long-run relationship, raises doubts as to the robustness of any model estimated using this sample. It might be suggestive of spurious relationships.

Model PR-C appears to be a very poor model: in spite of a test rejecting the null of no long-run relationship in Table 5.10, it has only one individually significant regressor in the short-run model, no significant long-run parameters and with too few degrees of freedom to calculate diagnostic test statistics. It is clearly not a good model, especially in the long-run. It does not warrant further discussion.

Model PR-B in Tables 5.11 and 5.12 is the best model of shares. The diagnostic tests highlight no issues, and the model appears reasonable. It is by far the best model of all those considered in this chapter. Never-the-less, there are a number of issues which cast doubt on the reliability of the estimated PR-B as a model of demand for military expenditure. It arises from a small sample and leaves too few degrees of freedom to be confident in the estimates. Given the performance of most of the other models, the possibility of a spurious relationship cannot be discounted.[9] Perhaps the most important determinants of military expenditure (the strategic environment) are not properly captured here because these are not easily quantified. The trade-off between ρ and k means attention must be restricted to very simple models of long-run expenditures.

It is the long-run estimates presented in Table 5.12 which provide the estimated coefficients of the demand (5.7). Of primary interest here is the effect of OUT on the defence share of GDP. The estimated coefficient is negative and significant. This suggests that higher levels of outsourcing are associated with lower defence shares of GDP (even if one does not believe that OUT truly reflects the level of outsourcing, the literal interpretation of the results is that increased competitive tendering in defence procurement is associated with lower a

[9] Some reassurance is provided by re-estimating the model over the periods 1984-2007 or 1983-2006. The lag orders selected and the sign and magnitude of the long run parameters are similar in these cases.

defence share of GDP). It appears that increasing the level of outsourced production by 1% results in the long-run in a 1.8% reduction in the defence share of GDP, *ceteris paribus*. Given the sample period, care must be taken with ascribing causality. The rise of outsourced production partially coincides with the post-Cold War drawdown. The negative coefficient on OUT may reflect a contraction of UK defence spending over and above that of the USA (the effect of which is captured by SPILL).

The estimated coefficient on PRGOVT is positive, and suggests that military expenditure follows the path of general government expenditure, but at a slower rate. The estimated coefficient on SPILL also positive. A 1% increase in the US defence share of GDP is reflected in a 0.33% increase in the UK share. This simply reflects similar defence interests and responses to common threats. The sample is dominated by the post-Cold War drawdown and then post-2001 expeditionary deployments.

The negative and significant coefficient on TBREAK (the post-2001 trend) is interesting. AFGDUM suggests a constant upward shift in defence as a proportion of GDP in 2001, but the coefficient of TBREAK suggests the start of a downward trend. This seems curious, given the increased expenditure over the post-2001 period. However, if SPILL is reflecting output through the shared deployments to Iraq and Afghanistan, and these have been increasing then the Armed Forces are becoming more efficient, or this coefficient reflects the recently much reported overstretch (in that the same output is produced with falling expenditure).

5.7 Conclusion

The best model estimated above is a model of defence share of GDP, namely, PR-B. It is the only vaguely satisfactory model. It suggests that outsourcing is relevant to the determination of long-run military expenditure—increased outsourcing is associated with lower aggregate military expenditure. Other factors which are significant are spill-ins from allies (positive), and non-defence government spending as a proportion of GDP (positive). There is also evidence for a break in both trend and constant in 2001.

Although superficially this model appears to be satisfactory, albeit with caveats, confidence in the model is undermined by a number of issues:

1. The approach taken to estimating a model of military expenditure is essentially data mining. Every possible model given the chosen parameters was estimated. Those which best passed various diagnostic tests and give the neatest estimates were presented. In spite of this approach we have found only one superficially satisfactory model. That other very similar models were found to be misspecified or contain no long-run relationship raises the possibility that the preferred model is based on a spurious relationship.

2. Some of the key variables required in original demands (5.6), (5.7) and (5.10) are not easily quantified. Some of the variables required are not well captured in the series

used in empirical implementation. This is especially true of the measures of strategic environment.

3. Only a very small sample was available. The changes in the strategic environment in the sample period (especially the end of the Cold War) mean that a sample of only 25 or 26 observations is probably inadequate to isolate the long-run determinants of military expenditure.

The second and third issues certainly apply to some degree. The first raises more fundamental questions about the use of demand for military expenditure models. That no convincing model of UK military expenditure has been found, in spite of trying all possible combinations of regressors casts some doubt on the validity of previous models in the literature, since authors only present the best fitting model. Here, were only the preferred model presented it may appear to be sound, but it is clearly not.

Obtaining a model of UK demand for military expenditures whose specification one is confident in would require a much larger dataset, even if attention were restricted (as here) to parsimonious long-run specifications. It would also require improved measures of the strategic environment. Neither of these requirements is likely to be met soon, so some circumspection must be exercised when modeling empirically demand for military expenditures using the type of models which have become traditional.

Given the difficulties of improving on these estimates of the effect of outsourcing on aggregate military spending, the next chapter takes a different approach. It focuses on the degree of competition in markets for outsourcing contracts in the UK, and how far the savings atributable to the introduction of competition (highlighted in subsection 2.3.2 of the literature review) might be expected to be realised when outsourcing military activities.

Chapter 6

UK Market for Military Outsourcing Contracts

6.1 Introduction

Thus far in this thesis the structure of the market for military outsourcing and the firms involved in undertaking outsourced defence tasks, have been little mentioned. More generally, although broader defence industries and their constituent firms have been widely studied [30, 31, 87, 114, for example], the market specifically for outsourced activities has attracted little attention. One of the key issues identified in Chapter 2 is the role of competition and competitive incentives in driving the savings attributed to outsourcing where they have been seen. The structure of the market(s) for military outsourcing contracts is a crucial determinant of the efficacy of competitive incentives.

The structure of the industry has a direct impact upon the scope for efficient outsourcing in defence. If the industry is structured in such a way as to provide little effective competition for outsourcing contracts, then savings achieved from outsourcing military activities will be lower than from a fiercely competitive market. As such, from the buyers perspective, private provision will be more efficient than public for fewer activites, and the optimal scope for outsourcing lower than were the market perfectly competitive. Even in a monopoly market private provision may be more efficient (productive efficiency) than public provision, but any efficiency savings arising from switching from in-house to outsourced provision will accrue as rents to firms rather than to the MOD as cost savings.

The UK is, along with the US, one of the countries which makes greatest use of outsourced providers in defence. In the UK around 42% of spending on defence support services is paid to outsourced providers, compared with an average of 19% in other European countries [3]. Despite its importance, it is difficult to estimate the scale of MOD outsourcing. Information on outsourced activities is difficult to acquire. This is partially because of commercial confidentiality, but largely because in the public accounts information is 'hidden in a statistical sense within old categories and budget codes' [108, p. 7]. Although the MOD reports the level of goods and services bought from the private sector, it is not possible to distinguish between spending on activities which used to be undertaken in-house (and so are now outsourced) and those items which have always been bought on the market.

There are two existing estimates of the scale of MOD outsourcing. A review of the public services industry estimates that the MOD spent £10.1bn procuring services, and that this grew at 3.6% annually over period 1996–2008 [108]. AMR International [3] estimate that in 2007 the MOD spent £13.4bn on Defence Support Services, and that the market may be broken into four segments:

Market segment	MOD spend (2007)	% outsourced
Facilities Management	£1,271m	97%
Maintenance, Repair and Overhaul	£3,419m	38%
Training	£239m	10%
Information Technology and Communications	£737m	87%

Source: AMR International [3]

Little is currently known about either the market for military outsourcing, or the industry which provides outsourced activities. This chapter uses firm level data to: identify the key firms involved in MOD outsourced activities; to analyse the structure of the outsourcing industry in the UK; and to examine whether there is effective competition for outsourced defence activities.

Identifying the structure of the industry, and the firms involved is difficult for the same reasons as outlined above. The approach taken here is to focus on one mechanism for outsourcing which is relatively well documented: the Private Finance Initiative (PFI). Using an original dataset of ownership of MOD PFI contracts through time, the key firms involved are identified. This is then augmented with information from existing sources, and the degree of competition within the market, and the relative financial performance of key firms involved, is analysed.

6.2 The PFI market

6.2.1 MOD PFI portfolio

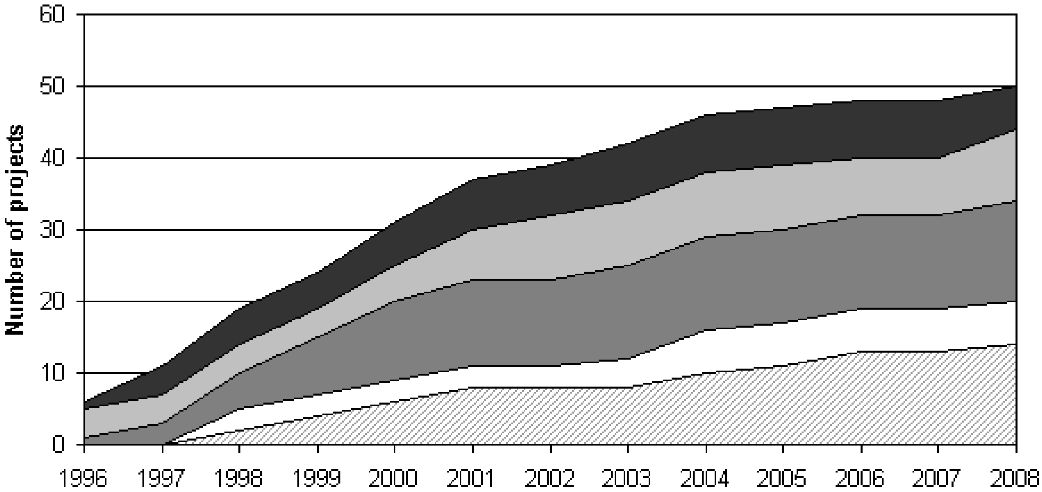
Between 1996 and the end of 2008 the MOD had signed 57 PFI contracts, of which 50 remained active at the end of 2008 with a combined capital value of £9.41bn.[1] Figure 6.1 summarizes the development of the MOD's PFI portfolio since it began in 1996, showing in each year the number of active projects in each of five categories (a), and their combined capital values (b).[2] Between 1996 and 2002 the number of signed projects grew rapidly, but

[1] Active defined as having reached financial close, but not having reached contract end. The capital value of a project is the size of the private partner's capital investment where available, otherwise the value of the public sector comparator is used. It is the only consistently reported measure of the size or value of PFI projects. It is not a measure of cost to government, nor of payments to the private firms. By way of illustration, capital value of all MOD PFI projects signed was £9.543bn against forecast payments of £46.3bn between 1997 and 2041, and the MOD is committed to annual payments averaging £1.569bn between 2009/10 and 2019/20.

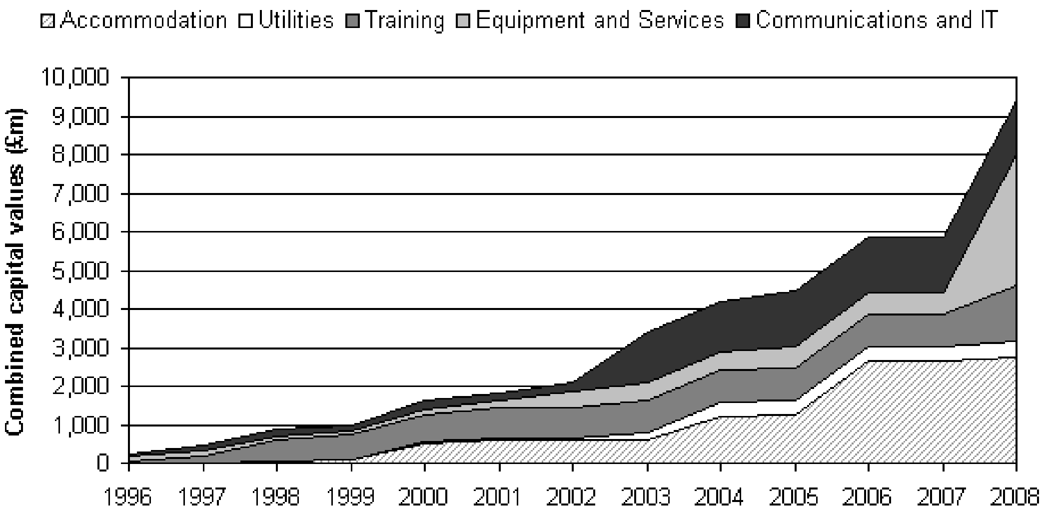
[2] The allocation of projects to categories is a little arbitrary, since some projects contain aspects of more than one category, but does give a reasonable indication of the MOD's portfolio. This arbitrariness is especially true

Figure 6.1: Graphs of MOD's portfolio of PFI projects by both (a) the number of projects and (b) combined capital value

(a)



(b)



Source: Author's calculations using data obtained from numerous sources including HM Treasury PFI signed projects list (September 2009), augmented with information from MOD documentation and annual reports and resource accounts, company websites and reports and accounts, and answers to parliamentary questions.

Table 6.1: MOD PFI projects by category at December 2008

Category	The % of total MOD PFI projects in each category by:		Examples
	Numbers[a]	Capital values[b]	
Accommodation	28%	29%	Refurbishment of MOD main building; Provision of 145 family quarters for RAF Cosford and Shawbury.
Utilities	12%	4%	Provision of water and serwerage facilities for tidworth garrison; Upgrading power station and provision of guaranteed power to RAF Fylingdales.
Equipment and Services	20%	36%	Supply and maintenance of civilian vehicle fleet; Provision and operation of heavy equipment transporters in peacetime and on operations
Training	28%	16%	Facilities for type and continuation training, and mission rehearsal for medium lift helicopters; Provision of Hawk aircrew training simulators
Communications and IT	12%	15%	Provision of email system for the RAF; Provision of Defence Fixed Telecommunications System
Total	100%	100%	

[a]The number of projects in each category as a percentage of the total number of projects. [b]Of the total capital value of projects, the percentage attributable to projects in each category

the lower graph shows them to be relatively low value projects. Since 2002, growth in the number of projects has slowed, but rapid increase in the cumulative capital value reflect the increasing size of signed projects. There are three projects which dwarf the remainder: Skynet 5 signed in 2003 with a capital value of £1.1bn, the Allenby/Connaught accommodation project signed in 2006 with a capital value of £1.3bn, and the FSTA contract signed in 2008 with a capital value of £2.7bn. Between them these projects account for around half of the total capital value of projects signed.

The mix of signed projects at the end of 2008 is shown in Table 6.1. It is clear that the most important sectors are Equipment and Services, and Accommodation, followed by Training and Communications and IT.

6.2.2 Ownership and control of MOD PFI projects

Identifying the firms which control MOD PFI projects is no simple task. There are two reasons for this. Firstly, because of the corporate structures within which the contracts tend to be held. PFI contracts are almost always awarded to Special Purpose Vehicles (SPVs) which are corporate entities created solely for the purpose of holding the contract. Ownership of the SPV is typically split between a consortium of firms which each undertake an aspect of the contract: a firm that constructs whichever asset is required, a firm that maintains and

in distinguishing between Training and Equipment projects since some contracts involve both the provision of equipment and training personnel in the use of that equipment, e.g. the Defence Helicopter Flying School provides both flying training and the helicopters used for that training.

operates the asset to provide the service, and a firm who finances construction of the asset. Each of the firms involved normally control their holding through several layers of holding company, which may themselves be joint ventures. Tracing control of contracts is further complicated by the active secondary market in stakes in PFI projects which means frequent changes in control over time.

The second difficulty arises because the documentation for MOD PFI contracts, whilst good relative to other outsourcing contracts, is patchy, frequently wrong or out of date, and often contradictory. The ownership data listed by HM Treasury [98] and by PartnershipsUK [154] often differs, and occasionally both are different to the information listed in companies' annual returns to Companies House (especially for the proportions of shares held by each owner).

In order to ensure the accuracy of the ownership and control data used, an original dataset was created. For each contract the SPV was identified using various sources (HM Treasury [98]; PartnershipsUK database [154]; company reports and accounts and parliamentary questions), and ownership traced at the end of each calendar year using annual returns from Companies House (accessed through FAME database). In this way each holding was traced back to (wherever possible) its ultimate owner.

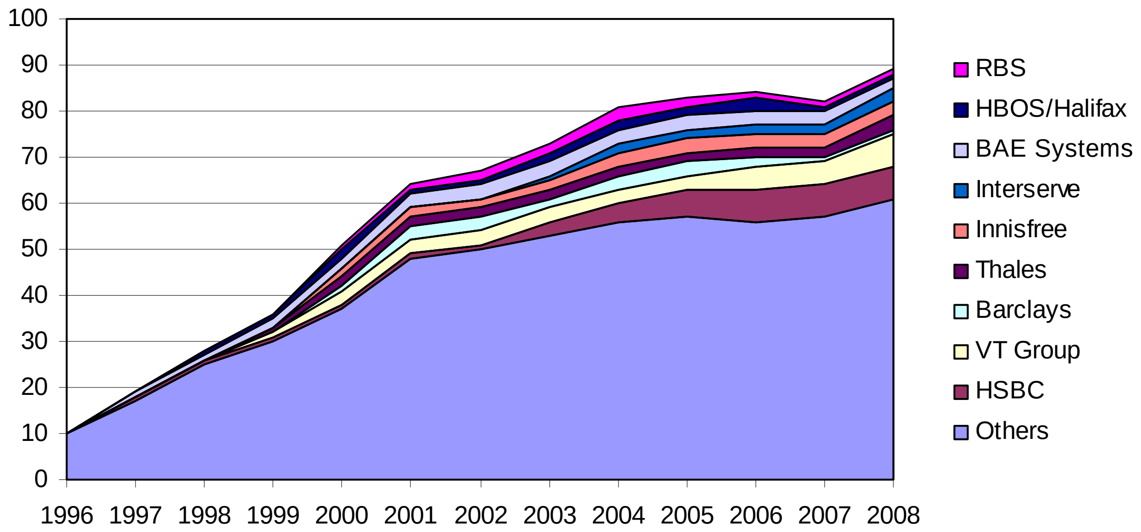
There is no entirely satisfactory measure which fully captures a firm's control of contracts in the PFI market. It is clear from Figure 6.1 that measuring projects by capital values or by the number of projects yield very different pictures of the MOD portfolio. This is also true when measuring the holdings of the firms. One natural measure is to count the number of projects a firm is involved in (called stakes henceforth), but this assigns the same weight to both a small holding in a small project and complete ownership of a large project. The second obvious measure, which accounts for project and stake size, is to split the capital value of a project according to ownership and then sum each firm's share of the capital values (called capital values henceforth). However, this measure is vulnerable to being dominated by the few large projects.

Both measures are shown in Figure 6.2 which shows the evolution of ownership of PFI projects over the whole market. When ownership is measured by stakes (a) it is striking how disbursed ownership appears. Far more so than when ownership is measured by capital values (b). By this measure ownership has become much more concentrated since 2002, but there is still no firm, or group of firms, which dominates the market.

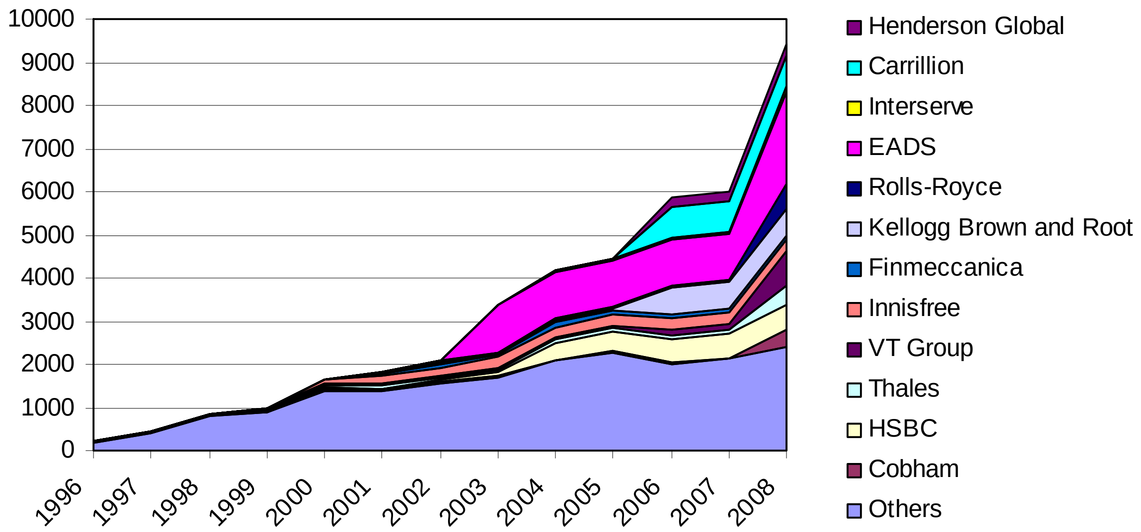
Figure 6.2 highlights the importance of banks and other financial firms in the ownership of PFI projects. Their presence is simply a quirk of the structure of the PFI mechanism (because it involves the project financing which is not actually an outsourced activity). Here we are concerned with the market for the provision of outsourced services and, since financial institutions do not operate in this market, they are not considered further. Table 6.2 identifies the firms (excluding financial institutions) with the greatest involvement in MOD PFI projects, measured by both the number of stakes in projects held and by the shares of project

Figure 6.2: Ownership of MOD PFI projects by (a) stakes held and (b) combined capital values

(a)



(b)



Source: As per Figure 6.1. Notes: In (a) firms which have never had stakes in more than two projects are combined in Others, in (b) firms with total shares of project capital values of less than £100m are combined into Others.

Table 6.2: Key firms controlling MOD PFI contracts at December 2008

Firm	Project stakes	Capital values (£m)
VT Group	7	827
Thales	3	428
Interserve	3	101
EADS	2	2155
Carillion	2	709
KBR	2	598
Rolls-Royce	2	574
Cobham	2	393

Source: As per Figure 6.1. Criteria for inclusion are stakes in more than one project and stakes of capital values exceeding £100m. In July 2010 VT Group were taken over by Babcock.

capital values attributable, at December 2008.[3] These are taken to be the key firms involved in the provision of outsourced services through PFI contracts. Although they only account for 25% of the total number of stakes (36% if stakes held by financiers are excluded), between them they account for 61% of the total capital values (74% if stakes held by financiers are excluded).

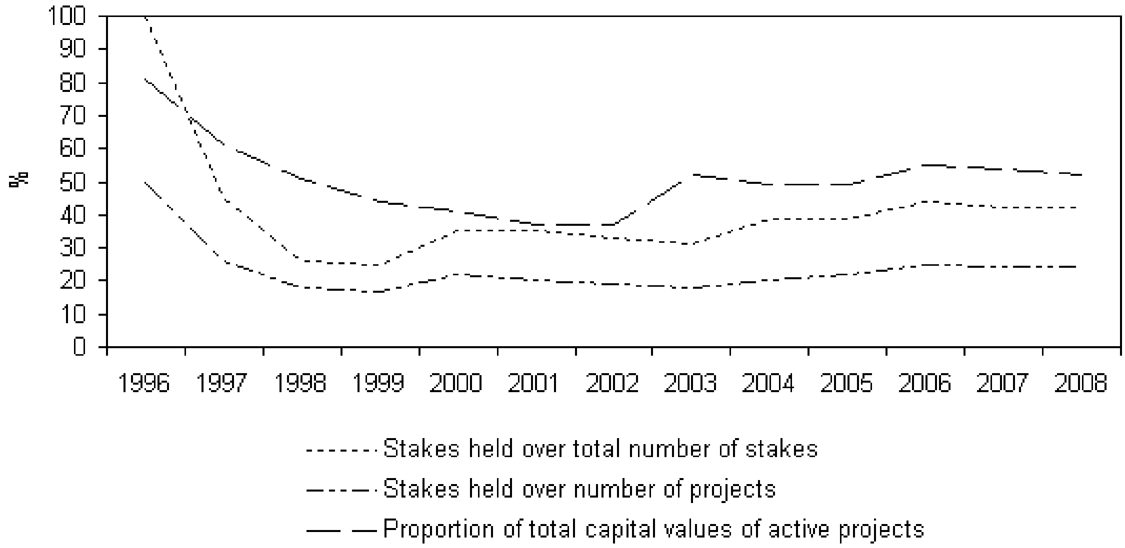
The evolution of market concentration might be seen in Figure 6.2, but is better shown in Figure 6.3. This shows three measures of the influence of the largest five firms: two by the number of stakes they hold (number of stakes held divided by the total number of projects, and divided by total number of stakes) and the proportion of total capital values allocated to the largest firms. Whilst none is a perfect measure of market concentration, all follow a broadly similar pattern. Concentration first falls from an initially high level reflecting the increasing number of projects leading to more diversified ownership, then since 2000–2002 market concentration has been gradually increasing, but ownership remains relatively diverse.

Given the variety of activities provided through PFI contracts, it is hard to maintain that it is a single market in which all contracts are competed for by all firms. The categories of project considered in Section 6.2.1 offer a much closer representation of the industries in which one might expect firms to be competing with each other.

The key firms in each of these market segments are shown in Table 6.3, as well as some measures of the degree of control these firms have in each sub-market. It shows that, at least to some degree, different firms are key in the different segments of the outsourcing market. In Accommodation and Utilities, although the main players are present in the largest projects, they have stakes in only a quarter of the projects. This segment is a little skewed by the inclusion of utilities contracts which are undertaken by utilities companies who have no other involvement in outsourcing. The 25% figure is likely an understatement,

[3]The financial firms excluded which would also have been featured were HSBC, Innisfree and Henderson Global Investors.

Figure 6.3: Five firm concentration ratios for all MOD PFI projects



but is broadly reflective of the situation. In the Training segment the three largest firms seem to have considerable holdings, but do not occupy a dominant position. A similar situation prevails in the Equipment and Services market. When measured by stakes held control in the IT and Communications segment is diverse. EADS's dominance in the capital value figure is because the Skynet 5 project is out of all proportion to the others. Overall, although in each market segment there are large players, none of them hold such market power as to dominate, or to raise concerns that effective competition should not be possible to arrange.

6.3 The broader outsourcing market

PFI projects represent only part of defence outsourcing activity—in 2007/8 expenditure on the operating elements of off-balance sheet projects and on-balance sheets was £1.3bn (£1.5bn in 2008/9), but a fraction of the entire market. Although used for a wide variety of activities, PFI contracts do not capture all outsourced activities. Rather, they are focused on those activities which require large up-front capital investment. Those aspects of the industry providing less capital intensive services (such as facilities management, catering and cleaning) are under represented in, or omitted from, Table 6.3.

In order to capture these elements of the market for outsourced defence activities we rely upon three alternative sources of information: AMR's survey of the Defence Support Services industry [3], the MOD's list of firms receiving payments of more than £500m in financial year 2008/09 [206, Table 1.17a in 2009 edition], and Perlo-Freeman and Sköns [159]. The most useful of these is the evidence provided by AMR International's survey of the defence support services market [3]. They list the largest firms in each of four categories of

Table 6.3: Key firms in PFI market segments at December 2008

Sector	Key Firms	Market share by capital value[a]	% of projects key firms have stakes in
Accommodation and utilities	Carillion KBR Interserve	62%	25%
Training	VT Group (now Babcock) Thales BAE Systems	38%	54%
Equipment and Services	VT Group (now Babcock) EADS Rolls-Royce	63%	40%
IT and Communications	EADS	76%	20%

Source: As per Figure 6.1. Notes: [a]The proportion of the capital values shares held by non-financial firms controlled by the key firms. VT Group were taken over by Babcock in July 2010.

Table 6.4: Key firms in defence support services market in 2007

Market segment	Key contractors	Market share of key contractors	Estimated number of competitors
Facilities Management	Carillion Babcock Sodexo	>75%	>10
Training	VT Group (now Babcock) BAE Systems Lockheed Martin	N/A	>35
Maintenance, Repair and Overhaul	BAE Systems VT Group (now Babcock) Babcock	35%	>15
Information Technology and Communications	EADS EDS	≈40%	>60

Source: AMR International [3]. Notes: VT Group were taken over by Babcock in July 2010.

the defence support services market, and although it is not entirely clear how they arrive at their list, it provides a useful benchmark which includes the soft services absent from the PFI market. Their findings are summarised in Table 6.4.

Comparing the firms listed in Table 6.4 with those from Table 6.3, the category Facilities Management in Table 6.4 compares most closely to Accommodation and Utilities category in Table 6.3. Carillion features in both, and so appears to be a large provider of outsourced activities. Table 6.4 also highlights the roles of Babcock and Sodexo.

Babcock is a large provider to the MOD [206, 2009 edition]—they are one of top 6 providers to the MOD, and around 60% of group revenues come from MOD support (2008/09) through marine and defence divisions [202]. Perlo-Freeman and Sköns [159] cite Babcock as one of the largest UK suppliers of military services. It was not picked up through analysing PFI contracts partly because a sizable part of the business with the MOD is operating some of the Royal Dockyards as a GOCO. A further reason for Babcock's inclusion is their takeover of VT Group in July 2010. This will make the combined company perhaps the largest UK supplier of defence outsourced services.

Sodexo specializes in the provision of soft services (those support services required for the operation of a facility such as cleaning, catering, estate management) and provides services worth between £100m and £250m to the MOD, but Serco—a similar firm also focusing on the provision of soft services—undertakes work for the MOD valued at over £500m [206, 2009 edition]. It is also listed as one of the top four UK firms providing military services by Perlo-Freeman and Sköns [159]. The firms most involved in providing MOD outsourcing of accommodation and facilities management comprise Carillion, Babcock, Serco, and Kellogg Brown & Root. Serco is omitted because defence services account for only 3% of Sodexo's revenues [191] (cf. 25% for Serco).

Consider next the Training category. The training currently provided is focused on training of aircrews, with all basic pilot training privately provided (both fixed and rotary wing), as well as simulators for Hawks, Apache, Tornado GR4 and Sentry E3D. Since only around 10% of training needs are privately provided [3] there is scope for a large increase in this category. Tables 6.3 and 6.4 both feature VT Group and BAE Systems, and there is one entry each for Thales and Lockheed Martin. Thales is one of the largest providers of PFI to the MOD, whereas Lockheed Martin seems to be included in Table 6.4 on account of its commitment (albeit large) to the UK Military Flying Training System contract. For this reason the largest firms in the Training market are taken to be VT Group, BAE Systems and Thales.

The Equipment and Services category in Table 6.3 corresponds most closely to the Maintenance, Repair and Organisation category in Table 6.4. Again the VT Group and BAE Systems feature, but the PFI market suggests that both EADS and Rolls-Royce are heavily involved. These both have a large stake in the FSTA project which was signed after Table 6.4 was compiled. Given this, the largest suppliers of Equipment and Services are taken to be VT Group, BAE Systems, EADS, Rolls-Royce, and Babcock.

The final category to be considered is IT and Communications. EADS features in both Tables, as well as EDS in Table 6.4. EDS is not included in the list of key players because most IT projects have never been conducted in-house, and so should not be considered outsourced.

There is one other firm which is heavily involved in providing outsourced services for the MOD which has not yet been mentioned—QinetiQ. It is one of the top ten suppliers to the MOD [206, 2009 edition], it is listed as the largest supplier of military services and outsourced research by Perlo-Freeman and Sköns [159], and its work clearly meets the definition of outsourcing since its work was in-house before DERA was broken-up and parts of it privatised as QinetiQ.

Overall, the analysis of PFI projects provides a useful starting point in identifying the key firms in the outsourcing market, but some firms and outsourcing activities are under-represented in the PFI market. By correcting for this under representation through using the other available sources of information on the outsourcing market we have identified the key players in the outsourcing market in 2007/08 as being: **VT Group**, **BAE Systems**, **EADS**, **Rolls-Royce**, **Babcock**, **Thales**, **Serco**, **Carillion**, **Kellogg Brown & Root** and **QinetiQ**.

6.4 Market concentration

As well as identifying the key firms in each segment of the outsourcing market, Tables 6.3 and 6.4 also present some measures of market concentration in each of the segments, or rather measures of the degree of dominance of the key firms. Accommodation and Utilities/Facilities Management is the least competitive segment. AMR estimate there are more than 10 potential competitors for contracts, but two of the three measures of concentration for this segment suggest the key firms (listed in Tables 6.3 and 6.4) control more than 60% of the market. In each of the other sectors the key firms appear to control around 40% of the market, perhaps a little more in Training. Even in the Equipment and Services/Maintenance, Repair and Overhaul segment, which might be expected reflect the structure of the wider defence industries in which firms exercise market power, AMR estimates there to be more than 15 potential competitors.

Overall, none of the market segments is totally dominated by the key firms. There is no monopoly provider to any of the sectors. The Accommodation/Facilities Management segment is the most concentrated, but not to such a degree that it raises questions about the potential degree of competition in the market. Although this segment would appear to have few barriers to entry, an oligopolistic market structure is suggested by the fact that the largest three or four firms controling 60% of the market.

In the other segments the key firms also control a significant proportion of the market, but there remains scope for considerable competition for contracts, not least from potential entrants. It is difficult to tell simply from looking at market concentration, but given that there is a single buyer—the MOD—in this market one would expect that the MOD will be able to exercise a degree of market power and organise effective competition. There may, though, be very specialized areas and activities for which this is not possible.

The apparent competitiveness of the market(s) may be altered were there to be consolidation within the industry. Babcock's 2010 take-over of VT Group combined two of the largest defence outsourcing firms, and is a step in this direction. The ability to arrange effective competition may be fragile. If further consolidation is seen, and especially if the scale of consolidation which has occurred in defence industries in the post-Cold War period [87, 114] were to be repeated, the industry could quickly come to resemble the defence equipment market. This market is monopolistic or oligopolistic, and the award of contracts is carefully managed in order to retain key capabilities on-shore. Both of these factors constrain the ability of the MOD to arrange effective competition.

6.5 Financial performance of the key players

The previous section showed there is a significant degree of concentration (at least 40% market control by the top three or four firms) in each of the segments of the defence outsourcing markets. This section examines the financial performance of the key firms involved. Ideally, one would wish to examine whether the provision of outsourced activities is more or less profitable than each firm's business in other (non-defence outsourcing) market segments or business lines. However, firms do not report financial information in sufficient detail to allow this level of analysis. It is not even possible to obtain measures of financial performance for the relevant divisions within firms—there is no way to allocate central costs, employees, and balance sheet items between divisions or market segments.

The alternative, pursued here, is to compare the performance of those firms most heavily involved with industry benchmarks in an attempt to establish whether the firms involved are more profitable than their peers. This is, of course, less satisfactory than comparing results within firms, since for most of the firms involved outsourced services represent only a fraction of their business, but it is the only feasible approach.

Table 6.5 displays some measures of financial performance for the key firms identified above. Value added is a measure of the wealth generated by a company. Both value added and profit are measures of the productivity of a firm, with value added being a broader measure. Margin (operating profit as a percentage of sales) and value added as a percentage of sales are measures of how effectively revenue is converted into profit and wealth generated. Asset turnover and return on capital employed are measures of how efficiently financial capital is employed in the generation of wealth and profit.

These firms are of three basic types: those involved in support services—VT Group, Babcock, and Serco; those whose primary or original business is construction—Carillon and KBR; and the remainder from the defence and aerospace industry.[4] The financial performance of firms in each of these categories is considered separately below. The defence and aerospace firms amongst them might be expected to show higher value added and profit per employee because of the capital intensive nature of their activities.

In order to provide benchmark figures against which to compare firms' performance Table 6.6 presents the relevant sectoral averages for the largest UK and European firms. The figures in Table 6.6 are comparable with the figures for financial year 2007/8 in Table 6.5. If the key firms involved are more profitable than the sectoral benchmarks, this would indicate that competition for outsourcing contracts is not intense enough to erode profits on these contracts to normal levels.

Support Services Each of the firms in this category derive a significant proportion of revenues from the defence sector. By profit per employee and margins VT Group and

[4]Whilst VT Group was once a defence and shipbuilding firm it has pursued a strategy of transforming itself into an outsourced services provider. This culminated in the sale of its holding in BVT Surface Fleet to BAE Systems in October 2009, and changing its listing on the LSE from Aerospace and Defence to Support Services on 21 December 2009 [220].

Table 6.5: Measures of financial performance of key outsourcing firms

	Year to:	Defence revenues	Revenue (£m)	Employees	Operating profit (£m)	Value Added (£m)[a]	Profit per employee	Margin[b]	Value added per employee	Value added as % of sales	Asset turnover[c]	ROCE[d]
VT Group	31-Mar-09	>46%[e]	1,096	12,431	39.0	423.6	£3,137	3.6%	£34,076	38.6	167.5%	6.0%
	31-Mar-08		845	11,829	49.5	403.8	£4,185	5.9%	£34,136	47.8	161.1%	9.4%
Babcock	31-Mar-09	60%	1,901.9	16,389	133.1	786.8	£8,121	7.0%	£48,008	41.4	256.6%	18.0%
	31-Mar-08		1,555.9	15,002	110.2	628.7	£7,346	7.1%	£41,908	40.4	190.3%	13.5%
Serco	31-Dec-08	25%	3,124	42,684	156	1,517	£3,655	5.0%	£35,540	48.6	185%	9.2%
	31-Dec-07		2,811	41,445	133	1,370	£3,219	4.7%	£33,056	48.7	272%	12.9%
Kellogg Root	31-Dec-08	N/A	881	7,565	69	357.7	£9,121	7.8%	£47,284	40.6	168.2%#	18.7%#
& Brown*	31-Dec-07		995	9,661	59	526.4	£6,107	5.9%	£54,487	52.9	143.1%#	9.9%#
Carillion	31-Dec-08		5,205.8	35,639	88.0	1132.5	£2,469.2	1.7%	£31,777	21.8	348%	5.9%
	31-Dec-07		3,951.7	30,746	65.5	857.7	£2,130.4	1.7%	£27,896	21.7	423%	7.0%
Thales[f]	31-Dec-08		9,907	63,248	588.5	4,488.2	£9,305	5.9%	£90,717	45.3	217.3%	12.9%
	31-Dec-07		8,369	61,195	465.6	3,732.6	£7,609	5.6%	£89,615	44.6	182.1%	10.1%
BAE Systems	31-Dec-08		16,671	90,000	1,700	7,073	£18,888	10.2%	£78,588	42.4	112.2%	11.4%
	31-Dec-07		14,309	83,000	1,038	5,572	£12,506	7.3%	£67,132	38.9	133.3%	9.7%
QinetiQ	31-Mar-09	73%[g]	1,617	13,882	132	918	£9,473	8.1%	£66,158	56.8	103%	8.4%
	31-Mar-08		1,366	13,470	76	774	£5,672	5.6%	£57,483	56.7	128%	7.2%
EADS[f]	31-Dec-08		33,843	118,349	2,168.3	10,822.1	£18,322	6.4%	£116,900	32.0	109.7%	7.0%
	31-Dec-07		26,628	116,493	-22.5	7,364.5	-£193	-0.1%	£92,881	27.7	94.5%	-0.1%
Rolls-Royce	31-Dec-08		9,082	22,500	855.0	3,257.0	£38,000	9.4%	£144,756	35.9	140.0%	13.2%
	31-Dec-07		7,435	22,900	514.0	2,737.0	£22,445	6.9%	£119,520	36.8	117.3%	8.1%

Sources: Company reports and accounts. Notes: [a] For consistency with the BIS database data presented in Table 6.6 this is taken to be the sum of Employment costs, operating profit, and depreciation, amortisation and impairment. [b]Operating profit as percentage of turnover. [c] Turnover as a percentage of long term capital employed defined as total assets less current liabilities. [d] Return on capital employed calculated as operating profit divided by long term capital employed. [e] not possible to calculate from accounts,but paid >£500m by MOD [206, 2009 edition]. [f]Reported in Euro, converted to Sterling average of relevant monthly rates from [96]. [g] 23% of business is from providing managed services in EMEA area * Figures taken from BIS database as not listed. # Figures constructed from accounts for Kellogg Brown and Root Limited and subsidiary undertakings which accounted in 2008 for 62% of revenues and 88% of profits. It is the group which includes Aspire Defence Holdings.

Table 6.6: UK and European sectoral average measures of financial performance

Sector	Number of firms	Sales (£m)	Employees	Operating profit (£m)	Value Added (£m)	Profit per employee (£)	Margin	Value added per employee (£)	Value added as % of sale
Support Services									
UK top 800 firms	119	1,005	13,663	88.9	461	6,395	10.9	82,269	57.0
European 750 firms	48	3,547	46,167	294.6	1,407	6,382	12.3	70,885	53.3
Aerospace and Defence									
UK top 800 firms	16	3,897	24,233	255.1	1,484	10,525	8.2	56,982	40.2
European 750 firms	11	9,665	43,100	411.2	3,483	9,542	7.2	82,845	40.9
Construction									
UK top 800 firms	34	1,390	7,244	73.5	386	9,841	6.9	64,653	30.4
European 750 firms	42	8,791	45,375	871.5	2,960	19,206	11.1	72,883	35.3

Source: Constructed from data in BIS Value Added Scoreboard [43]. Notes: Figures refer to financial year 2007–2008. See also notes to Table 6.5.

Serco appear to be broadly similar in terms of profitability. Babcock is significantly more profitable, generating higher profits and value added per employee, and higher margins. It is also the firm most heavily committed to defence outsourcing.

Comparing their performance against the sectoral averages in Table 6.6, all three generate lower margins and value added per employee. Babcock alone appears more profitable (per employee) than the benchmarks. It may be that the differences between firm performance and the benchmarks result from the inadequacy of the support services average as a benchmark. The sector includes a wide range of firms, from employment agencies and law firms through to consulting and accounting firms, very few of which operate in comparable markets to the firms of interest. It is likely that this is the explanation of the difference between the three firms' value added per employee and the benchmark figure.

Overall, care must be taken not to rely too heavily on only two years figures: the firms in the support services sector do not appear to be excessively profitable relative to their peers. In fact, both the VT Group and Serco seem less profitable. Babcock is most profitable, and is also the firm with the highest proportion of revenues coming from defence outsourcing, but it is not clear that these two facts are linked. They may be explained by the different mixes of work undertaken by the firms.

Construction Since there are only two construction firms considered it is difficult to draw firm conclusions. In terms of profit per employee KBR is around the UK average, and Carillion lower. Both are substantially below the European average. Much the same may be said of margins. Both firms generate substantially less revenue per employee and value added than either the UK or European averages.

Carillion, when compared to the benchmarks, and even to KBR, looks to be a far less profitable business. This is not because they are more heavily involved in support services than most. In fact, the support services side of its business is more profitable than the construction side. One might suspect that this reflects the onset of recession, but it is also seen in the 2007/08 figures which pre-date the recession. Both firms seem less profitable than their peers in construction, but this is because the construction sides of their business are less profitable than other construction firms: for both firms the provision of support services—including defence outsourcing—is their most profitable business segment.

The only example found of direct reporting of the returns of a PFI contract is KBR which presents figures for the Colchester Garrison contract. In 2008 KBR's stake in this project produced a margin (on sales) of 26% (2007: 17.2) and a ROCE of 26.5% (2007: 28.1%). Whilst this is clearly more profitable than KBR's overall performance, care should be exercised in relying on figures for only two years of a 35 year contract.

Defence and Aerospace Firms The difficulty with the defence and aerospace firms is that whilst defence accounts for a sizable proportion of all the listed firms' business, much of it is the construction of equipment rather than outsourced activities per se. Although it is not possible to obtain precise figures, with the exception of QinetiQ, outsourced activities are a minority interest.

As might be expected of firms primarily involved in the capital intensive business of construction military equipment and aircraft, value added per employee is much higher than for the firms involved in outsourcing from other sectors. The same is true of profit per employee.

The performance of the firms in the sector is varied. With the exception of Rolls-Royce, the firms involved seem to have lower margins than both the UK and European sectoral averages (except BAE in 2008). Their levels of value added per employee are higher than for UK firms in the sector, but in broadly in line with European average. This raises the question of whether UK defence firms are under rewarded for the value they create.

Of the firms considered Rolls-Royce is the most profitable (by profit per employee and margins), but it would not be true to say that the firms most heavily involved in outsourcing are more profitable than their peers.

Overall, the key firms in the outsourcing market do not seem to be outperforming their peers. The sole exception to this is Babcock, which appears to be outperforming both its competitors within the market, and its peers outside. Babcock is also most heavily involved in providing defence support services, in the sense that it is the firm which is most dependent upon outsourced activities. Although it is tempting to assign some significance to this coincidence, it cannot be justified on the basis of the available evidence.

There is a limit to what can be inferred from firm performance about the structure and degree of competition in the market. The use of financial data at the firm level, whilst unavoidable, means that the relative profitability of undertaking outsourcing activities cannot be established. If disaggregated data were obtained, a great deal could be learned, but this would require the cooperation of the firms involved. Even if such data could be obtained, care must be taken not to rely too much upon a single year's figures. Almost all contracts are performed over more than one year, some for up to 35 years, and the structure of many contracts is such that the profit of the private provider is only realised towards the end of the contract [204, for example].

6.6 Conclusion

This chapter has contributed to the rather scant literature on the military outsourcing market. Using an original dataset on PFI projects, as well as other sources, analysed the structure of the growing military outsourcing industry, in particular identifying the key firms involved in the provision of outsourced activities for the MOD.

These firms control a significant proportion of each segment of the market for outsourced contracts, especially Accommodation and Facilities Management. never-the-less, there is currently sufficient scope for competition to ensure that, in most activities, auctions for outsourcing contracts are competitive. The key firms involved do not appear to be more profitable than their peers, though it has not been possible to compare firms' performance in outsourced contracts against their other activities.

The competition which exists may be fragile, and even relatively minor consolidation within the industry may jeopardize it. The take-over of VT Group by Babcock is a step in this direction, and the active secondary market for project stakes offers an alternative mechanism through which it might be achieved. Over time, the market may well come to resemble the military equipment procurement industry with its monopolistic-oligopolistic structure, strategic behaviour and firms holding market power.

If information from firms on the financial performance of individual contracts could be obtained, a more satisfactory analysis of the relative profitability of outsourcing contracts could be undertaken. The PFI market may also be able to shed some light on the hypothesis that the intensity of competition is reduced by the possibility of contracts being obtained by buying out the winning bidder rather than participating in the tender process, though this would need data on the consortia which bid for contracts be obtained. Were this to be seen, any efficiency savings from the use of private firms would accrue to the winning bidder as rent rather than to the MOD as savings.

Chapter 7

Conclusions and Directions for Further Research

7.1 Main findings

This thesis makes the following contributions:

- It improves upon previous estimates of the elasticity of substitution between military labour and capital in the military production function (Chapter 3). This elasticity is close to zero, much lower than previously thought.

- It also provides an estimate of the elasticity of substitution between military and civilian labour in the military production function (Chapter 4). This elasticity is also low and close to zero. Together, these results suggest that there is little scope for factor substitution in the military production function.

- The long-run effect of the level of outsourcing on aggregate military expenditure is estimated (Chapter 5). The best model suggests that increased levels of outsourced production in defence are associated with lower expenditure, but the difficulties encountered raise questions about the validity of the whole class of demand for military expenditure models.

- Finally, it adds to the somewhat sparse knowledge of the industry providing outsourced military activities (Chapter 6). An original dataset is used to identify the key firms involved in each segment of the UK market, and to calculate measures of market concentration. The financial performance of these key firms is compared with benchmarks from outside the outsourcing industry, and finds they do not generate excessive profits.

The introduction posed a number of questions which this thesis seeks to answer. Question 1 asks which activities does economic theory suggests are best outsourced and which are best performed in-house? Economic theory provides few clear guidelines with which to answer this question. The various approaches considered in Chapter 2 each tackle aspects of the issue, but none offers a complete analysis. The conclusions which emerge are insufficiently specific to be usefully applied to the question.

The empirical literature offers a much clearer, albeit partial, answer to the question. There is clear evidence if an activity:

1. does not require the use of large specific assets;
2. is straightforward to specify in contract;
3. has stable and predictable requirements;
4. has many potential providers; and
5. does not rely crucially upon timely contract enforcement,

then substantial savings (in the order of 10–20%) may be realised from outsourcing that activity. Unfortunately these conditions only apply to a small minority of defence activities, not least because of the importance of specific military equipment in many activities. Many of the more recently outsourced activities do not fulfil the criteria above and, since theoretical analysis cannot offer guidance, empirical analysis of the relative efficiencies of public and private provision of these activities is needed. This requires more detailed data than are currently publicly available.

Sufficient time has now passed to allow an assessment of the performance of completed outsourcing contracts to be undertaken. This would allow identification of the actual efficiency and cost savings achieved in the earliest contracts, removing the need to rely upon projected savings. Comparison of the two would allow the question of whether renegotiation erodes expected savings to be answered, and the magnitude of this effect (if it exists) to be measured.

Question 4 is addressed in Chapter 5. Firm conclusions are hard to draw because of the difficulties in estimating a long-run relationship, but it appears that increased outsourcing is negatively and significantly related to aggregate military expenditures. The difficulties of estimating the long-run relationship mean that this tentative result will be hard to improve. The difficulties are principally data based—most notably constructing accurate measures of the strategic environment—so applying a similar methodology to countries other than the UK (perhaps the US) may allow these to be overcome.

Questions 3 and 2 are concerned with the substitution of other factors of production for military labour, since military labour is unavailable to private firms undertaking outsourced military activities. The possibilities for substituting capital for military labour are examined in Chapter 3. The estimated elasticity is close to zero, suggesting there is little scope for contractors to use capital rather than military labour. Chapter 4 tackles the possibility of using civilian rather than military labour. The estimated elasticity is also around zero. Together these chapters suggest that there is little possibility for factor substitution in the military production function. It is not possible to costlessly replace military labour with contracted civilian labour. The price is likely to be reduced flexibility. There are, though, limits to how far this conclusion can be applied. What is actually estimated is the MOD's past behaviour on substitution. If private firms are able to use resources more flexibly—which is frequently cited as one of their advantages—they will have greater substitution possibilities than the MOD.

The final question posed was how competitive are the markets for outsourcing contracts? This was addressed in Chapter 6. The structure of the market for outsourcing contracts is such that there is the scope to arrange effective competition for contracts, but this conclusion is vulnerable to consolidation within the industry. The key firms in the market control a large proportion of each of the market segments, but in the Accommodation and Utilities segment has an oligopolistic structure. The key firms involved do not earn extra-normal profits when compared to their peers undertaking similar activities in other industries.

7.2 Proposals for further research

Overall, although important, the scope for efficient outsourcing in defence is not well understood. This thesis has made some progress on the issue, but much remains to be done. In particular the limits to the use of outsourcing, which are driven by the limits to the effectiveness of formal and relational contracts, are not well understood. Nor are the differences between public and private provision sufficiently well understood to be able to make informed judgements as to where scope for increasing efficiency through extending private provision of defence activities remains.

The best prospects for identifying the precise differences in the way public and private provision operate lie outside of defence, though any conclusions may apply directly. One avenue through which the differences might be explored is by re-analyzing the large scale datasets used in the analysis of refuse collection whilst allowing for public and private producers to operate under different technologies.

Within defence there is a need for more work in assessing the relative cost performance of more recent outsourcing initiatives. It is unlikely that large datasets (such as those used in the refuse collection literature) will be able to be constructed, but detailed case studies may offer some insights. This is especially true of some of the more recent forms of contract such as contracting for availability and PBLs. These are increasingly being used to outsource defence activities and there is remarkably little information as to their performance.

The arrangements for the maintenance of military aircraft in the US may offer scope for conducting analysis of these contractual arrangements since there is a statutory requirement (Title 10, §2466 of the US Code) for at least 50 percent of maintenance work to be conducted by in-house depots. This has led to a mix of public and private arrangements for maintenance. Whilst it is not the case that precisely the same activities, for the same aircraft, are being undertaken under both in-house and contracted arrangements, it would allow analysis of common factors which drive maintenance costs across aircraft types under each form of provision.

There is also scope for empirical work to improve estimates of the transactions costs of both outsourced provision of activities and the costs of managing in-house arrangements. Accurately quantifying the relative costs of arranging and overseeing various arrangements for provision is the only way to enable comprehensive assessments of the relative cost performance of public and private provision. There is currently little information available on the transaction costs of outsourced provision; there is even less on the management costs of in-house provision—any progress made in quantifying these costs would be valuable.

The estimates of substitution possibilities might be refined by considering separate elasticities for each service. The RAF in particular, because it operates from bases remote from the battlefield, may have greater scope for the use of civilian labour, and greater scope for the use of outsourced provision. Further disaggregation may also yield insights, but the data necessary to conduct such analysis is unlikely to be available.

Given the results in Chapter 5, further work on the demand for military expenditures is required. There is scope for work similar to that of Chapter 5 to be conducted for countries other than the UK, in order to examine whether the questions raised as to the models' validity in the UK also exist in other countries. Re-estimating the models which have been previously estimated on UK military spending using the longest possible dataset (rather than being restricted to samples for which series on outsourcing are available) would clarify whether the concerns raised here are sample specific, or whether the approach is truly flawed.

Perhaps the greatest improvements to the demand for military expenditure models would be to develop some grounds on which to choose between the possible specifications *a priori*. Were one specification clearly preferable, it would remove the element of discretion to choose the form which best fits the data. This would improve on the current situation in which there are many models in the literature which apparently fit well, but which use a wide range of specifications each chosen on the basis of best fit. One direction in which progress in choosing between the various specifications is their differing implications for the shape of the Engel curves for defence spending.

International comparisons of the extent of military outsourcing would improve knowledge of the sector. In the UK there is scope for improving the measurement of the size of the outsourcing market, and its structure. Chapter 6 is based largely on the market for PFI contracts. A broader approach incorporating all of the various mechanisms by which outsourcing is achieved would improve understanding of the market, as would examining a larger number of more tightly defined sectors within the overall market.

Finally, one of the key concerns about the use of contractors in military activities is the reduced control which may be exercised over them by commanders on the ground. Case studies of the use of contracted activities in battlefield environments, and of private military companies, may help to establish whether this concern is justified. An assessment could be made as to whether private producers actually engage in the opportunistic behaviour suggested by TCE, or whether this can be controlled by relational contracting and the value of on-going relationships.

Appendix A

Detailed Description of Data

This appendix comprises four sections. The first details the definitions, sources and construction of the data used in Chapters 3 and 4. The final three provide the summary statistics, correlations and plots of the data used in Chapters 3, 4 and 5 respectively, in order to aid the reader's understanding of the information contained within the data.

A.1 Description of data used in Chapters 3 and 4

This section outlines the definitions, sources and details of construction of the data series used in estimation in Chapters 3 and 4. Where reported for financial years series are converted to calendar years by being taken as the observation for the calendar year in which the financial year begins, e.g. data for financial year 1970/71 are used as if relating to calendar year 1970. In Chapter 3 suffixed PCW is that variable for the post-Cold War period, e.g. $WMPCW = WM * PCW$. Any variable suffixed CW in Chapter 4 is that variable during to Cold War period only, e.g. $WMCW = WM * CWDUM$.

Afghanistan dummy (AFGDUM) Dummy variable reflecting deployments in Afghanistan: takes value of 1 in years 2001–2007, and 0 in other years.

Operational tempo (CAS) The number of deaths of Armed Forces personnel in a period divided by the number of service personnel. It may be interpreted as the (unconditional) probability of a serviceman dying in each period.

Cold War dummy (CWDUM) Dummy variable reflecting the Cold War: takes value of 1 in years 1970–1989, and 0 in other years.

Falklands Conflict dummy (FDUM) Dummy variable reflecting the Falklands Conflict: takes value of 1 in 1982, and 0 in other years.

Real national income (GDP) GDP of the UK (ONS Series BKTL) deflated using GDP deflator (ONS Series YBGB).

Gulf War Dummy (GDUM) Dummy variable reflecting the Gulf War: takes value of 1 in 1991, and 0 in other years.

Real government expenditure (GOVT) Total government expenditure, less defence spending, deflated by GDP deflator to 2005 prices (Source: ONS data and undeflated LME series).

Iraq War dummy (IRAQ) Dummy variable reflecting deployments in Iraq: takes value of 1 in years 2003–2007, and 0 in other years.

Logged number of MOD civilian employees, l_c (LC) Reported figures for MOD civilian staff as at 1st April each year, including both UK based and locally engaged staff. Prior to 2001 the staff in Trading Funds (such as the Met Office) were included in the reported total, so to ensure consistency these staff have been added to the reported total from 2001. Prior to 1995 all part-time personnel were counted as half of full time, after April 1995 they counted as the proportion of full time hours that they work—on average 60%. Data taken from Annual Abstract of Statistics (AAS) [5, years 1980, 1984, 1992, 2000, 2004] and UK Defence Statistics [206, years 2002 and 2006]. Where different issues contain different figures the later (presumably revised) figures are preferred.

Logged number of service personnel, l_m (LM) The reported strengths of the UK regular Armed Forces as at 1st April each year. Excludes personnel locally entered overseas (such as Gurkhas), mobilised reservists and FTRS. Data taken from UK Defence Statistics [206, years 1992, 1998, 2001, 2002, 2006, 2008], Statement on the Defence Estimates Volume II [194, years 1980, 1982, 1987] and AAS [5, 1976 edition].

Levels of real military spending M (LME) The military expenditure figures for periods 1970–2000 are the reported total military expenditure at current/outturn prices, taken from the AAS [5] and Defence Statistics 2006. The introduction of Resource Accounting and Budgeting (RAB) in financial year 2001/02 makes expenditure figures for 2001 to 2003 incomparable. Rather than use the reported Departmental Expenditure Limits, Net Cash Requirements are used. Net Cash Requirement is the actual money that the MOD requests from the Government in order to fund its activities, and is broadly comparable to the old expenditure series. The series is deflated by the RPI in Chapter 3 and by the Final Consumption Expenditure by General Government Deflator (ONS series YBFT) in Chapters 4 and 5.

Logged real military spending, m (ME) Logged level of the military expenditure series, LME.

Primary measure of the level of outsourcing (OUT) The proportion of the number of MOD headquarters contracts in each year offered by competitive tendering (Source: UK Defence Statistics [206]).

Level of outsourcing measured by contract value (OUTV) The proportion of the total value (cf. number used in OUT) of MOD contracts let by competitive tendering (Source: UK Defence Statistics [206]).

Level of outsourcing measured by contracts let by reference to market forces (OUTM) The proportion the total number of MOD contracts let by either competitive tendering or by reference to market forces in each year (Source: UK Defence Statistics [206]).

Post-cold War dummy (PCW) Dummy variable for the post-Cold War period: takes value of 0 in years 1970–1989, and 1 in subsequent years.

Logged total UK population (POP) ONS mid-year estimates of the resident UK population (Source: AAS [5, 2009 edition])

Logged population of potential recruits, n_m (POPM) In Chapter 3 this is the population of 15–19 year old males. In Chapter 4 this is the population of 15–29 year old males. Both are the ONS mid-year estimates of the home (resident in later periods) population (Source: AAS [5], and `www.statistics.gov.uk`). Where figures reported for a year differ between issues of AAS, the latest figures are preferred—presumed revised.

Population of 15 to 65 year old males and 15 to 60 year old women, n_C (POPC) Mid-year estimates of the home (resident in later periods) population of the entire working age population—males aged 15–65 and women aged 15–65. This is used as the population relevant for the supply of civilian labour to the MOD. Sources are as for the population of 15 to 29 year old males above.

Use of PFI projects in defence (RPFI) The real capital value of public private partnership and PFI projects operational in each year. For a definition of capital value see Footnote 1 in Chapter 6.

Military spending as share of GDP (SME) Military spending as proportion of GDP, calculated from undeflated GDP and LME series.

Spill-ins from allies (SPILL) For the levels and logged models in Chapter 5 total NATO (excluding the UK) military spending in constant 2005 US dollars was used. In the shares of models of Chapter 5 US military spending as a proportion of US GDP is used (Source: SIPRI Yearbooks [183]).

Military spending by the USSR during the Cold War (THREAT) In years 1979–1989 this takes the value of USSR military spending in constant 2005 US dollars. For years after 1989 it takes a values of zero (Source: SIPRI Yearbooks [183]).

Trend, t (T) Linear trend over the appropriate sample period.

Unemployment rate, u_t (UN) The UK unemployment rate for males aged 16 and over, ONS series MGSY. Data are averages for calendar years (constructed from averaging monthly three-month rolling averages), based on the Labour Force Survey, and available annually from 1971. The observations for 1968–1970 are taken from the AAS [5, 1974, Table 150].

Logged wages from alternative employment, w_a (WA) In Chapter 3 this is the average annual wage for full-time male manual workers on adult rates in all industries unaffected by absence as at April each year. In Chapter 4 it is the average annual wage for full-time male manual and non-manual workers on adult rates in all industries unaffected by absence as at April each year. It is constructed as an average of the reported manual and non-manual weekly wages reported in the New Earnings Survey (NES). Up until the 1990s the NES only published average wages for various categories of manual and non-manual workers. From these annual manual and non-manual figures a weighted average of the manual and non-manual average wages for a given category of worker was constructed. This appears to match the methodology used by the ONS to produce their average male/average all worker figures. Figures from 2003 are taken directly from the NES. The series are deflated using the RPI.

Logged wages of MOD civilian staff, w_c (WC) The average annual wage for all full-time workers in the closest SIC category to Public Administration. It was constructed as above from the reported manual and non-manual average weekly wages for males and females in the most relevant industry. From 1970 to 1982 SIC 1968 was used and the most appropriate industry was Public administration (SIC 1968 Order XXVII). From 1983 to 1994 SIC 1980 was used and the most appropriate industry was Public Administration, national defence and compulsory social security (SIC 1980 code 91). From 1995 to 2003 SIC 1992 was used and Public Administration and Defence; Compulsory Social Security was the most appropriate industry. Although the changes in SICs introduce discontinuity into the series, the differences in wages and sample size suggest that these discontinuities are not serious. One potential difficulty is that workers are classified by activity rather than sector, and as such many of the MOD civilian staff are not represented in the series, e.g. workers in the Royal Dockyards (an important category of MOD workers early in the sample) are not represented.

Logged real military wages, w_m (WM) This series is the average basic wage paid to all Army ranks introduced on 1st April each year (with the exception of 1971 which were introduced on 1st August). The average wage for 1970–2000 is constructed by weighting the pay at each rank [6, 146] by the rank structure of the Armed Forces in 1990. The wage for each rank is the wage for that band which had most personnel in through 1990s. For some years in 1970s wages for Generals not reported so have assumed that they are paid 5/3rds as much as Brigadiers, but the weight assigned to Generals is only 0.000664 so this makes little difference to the series. From 2001, due to the introduction of a new pay system (Pay 2000) the construction of the weighted average is not possible. Observations for 2001–2006 are constructed by adjusting the April 2000 average wage according to the Military Salaries Index: all ranks [206, years 2004 to 2008]. The series is then deflated using the RPI.

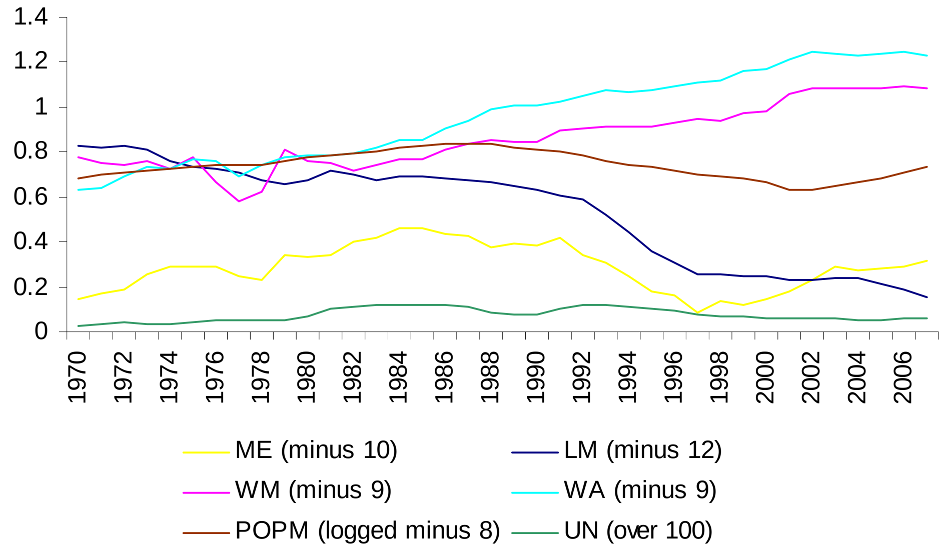
A.2 Summary statistics for the data used in Chapter 3

Table A.1 below details some of the summary statistics of the data used in Chapter 3, and Figure A.1 plots the key series used.

Table A.1: Summary statistics and correlation coefficients of the data used in Chapter 3

	ME	LM	WM	WA	POPM (logged)	UN
Mean	10.286	12.536	9.863	9.960	7.10e+07	7.389
Std. Dev.	0.103	0.227	0.139	0.201	1.49e+07	3.038
Min	10.087	12.157	9.583	9.635	3.56e+07	2.6
Max	10.462	12.829	10.089	10.249	1.03e+08	12.3
Pairwise correlation coefficients						
ME	1					
LM	0.414	1				
WM	-0.252	-0.895	1			
WA	-0.176	-0.931	0.938	1		
POPM	0.076	-0.839	0.739	0.872	1	
UN	0.533	0.022	-0.021	0.157	0.336	1

Figure A.1: Graph of key variables used in Chapter 3



A.3 Summary statistics for the data used in Chapter 4

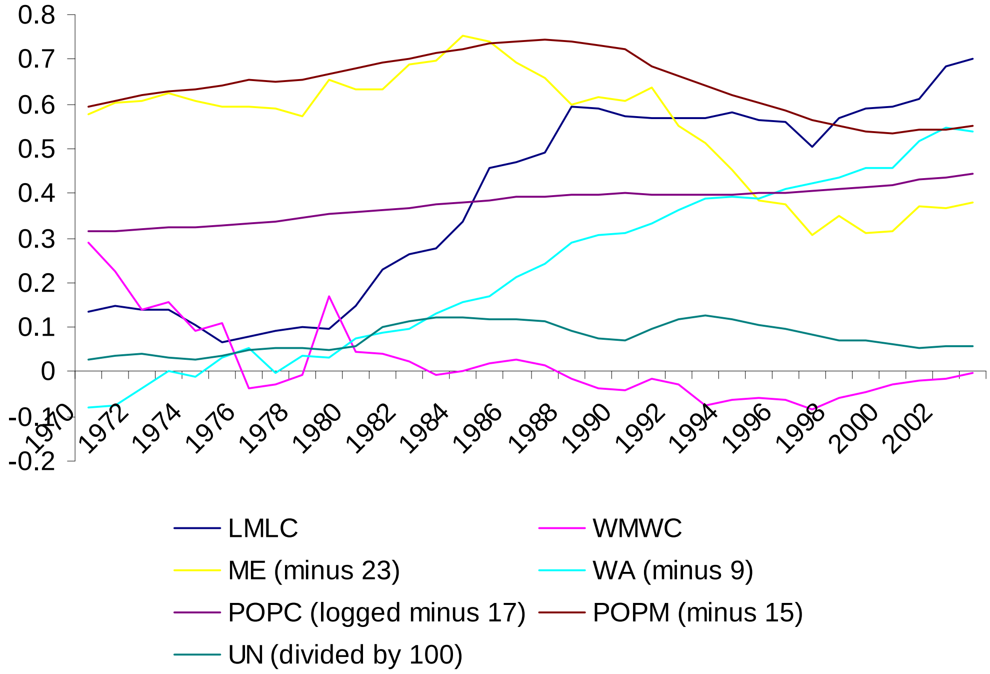
Table A.2 below details some of the summary statistics of the data used in Chapter 4, and Figure A.2 plots the key series used.

Table A.2: Summary statistics and correlation coefficients of the data used in Chapter 4

	LMLC	WMWC	ME	WA	POPC	POPM	UN
Mean	0.388	0.019	23.548	9.226	3.53e+07	6237068	7.676
Std. Dev.	0.219	0.088	0.134	0.196	1283109	419372	3.229
Min	0.066	-0.086	23.305	8.918	3.31e+07	5568600	2.6
Max	0.702	0.290	23.751	9.545	3.76e+07	6885000	12.6
Pairwise correlation coefficients							
LMLC	1						
WMWC	-0.657	1					
ME	-0.547	0.416	1				
WA	0.944	-0.731	-0.721	1			
POPC	0.942	-0.715	-0.569	0.962	1		
POPM	-0.189	0.027	0.869	-0.392	-0.221	1	
UN	0.485	-0.571	0.172	0.373	0.471	0.470	1

Notes: The differences between ME in this table and Table A.1 are due to them being deflated to different base years. Other differences are due to the different sample sizes.

Figure A.2: Graph of key variables used in Chapter 4



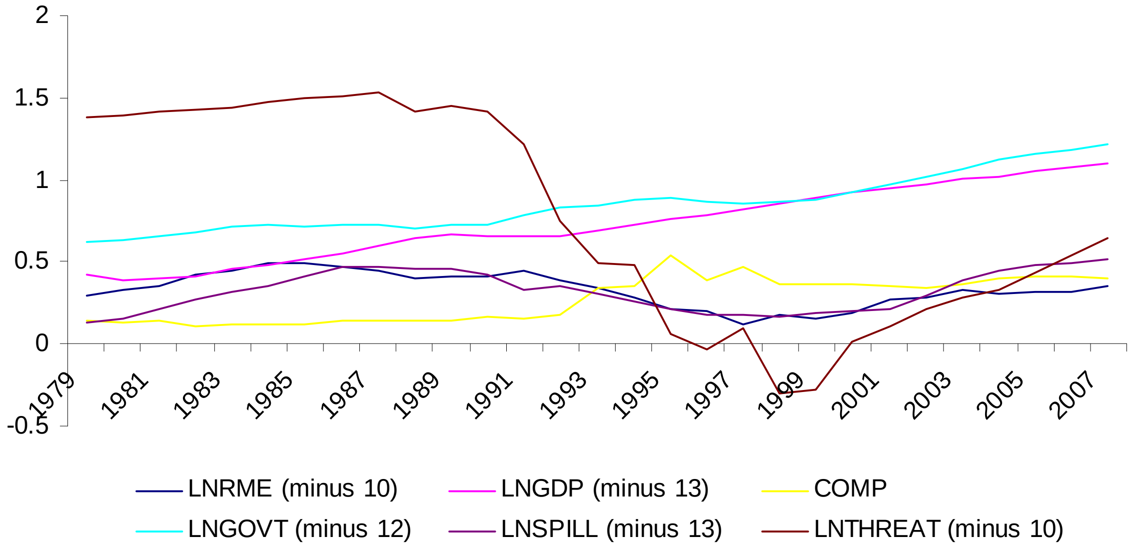
A.4 Summary statistics for the data used in Chapter 5

Table A.3 below details some of the summary statistics of the logged data used in Chapter 5, and Figure A.2 plots the key series used. Although the chapter used the levels, logs and shares of the variables, presenting only the details of the logged data is felt to give the reader sufficient information on the relationships between the variables.

Table A.3: Summary statistics and correlation coefficients of the data used in Chapter 5

	ME	LNGDP	COMP	LNGOVT	LNSPILL	LNTHREAT
Mean	10.334	13.729	0.270	12.852	13.322	10.773
Std. Dev.	0.104	0.224	0.134	0.171	0.123	0.645
Min	10.123	13.389	0.11	12.619	13.130	9.694
Max	10.497	14.102	0.54	13.213	13.515	11.531
Pairwise correlation coefficients						
ME	1					
LNGDP	-0.519	1				
COMP	-0.783	0.820	1			
LNGOVT	-0.373	0.945	0.780	1		
LNSPILL	0.632	0.239	-0.162	0.324	1	
LNTHREAT	0.871	-0.762	-0.902	-0.660	0.376	1

Figure A.3: Graph of key variables used in Chapter 5



Appendix B

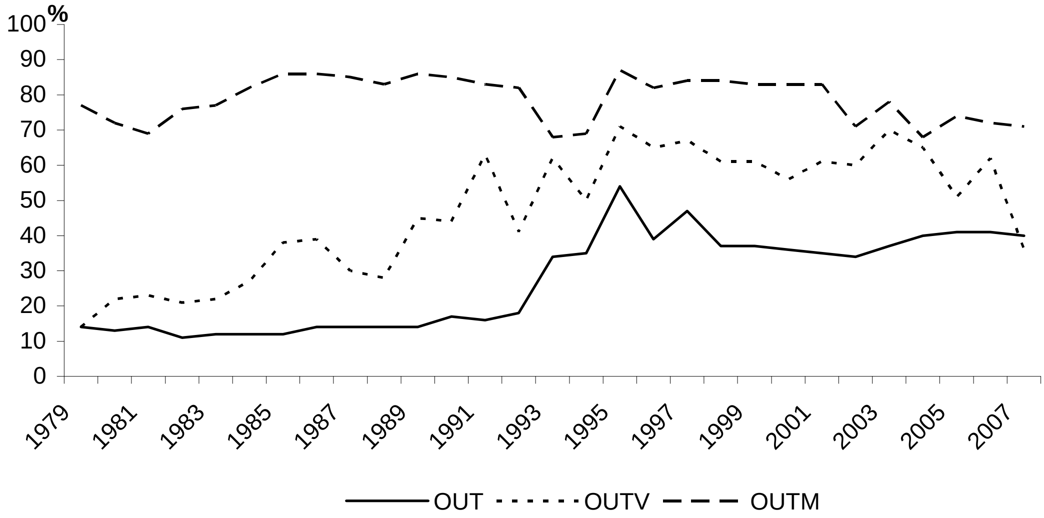
Alternative Proxies for the Level of Outsourcing in Defence

Chapter 5 used the proportion of the total number of MOD contracts let by competitive tendering (OUT) as a measure of the level of outsourced production. This Appendix examines the sensitivity of the results of Chapter 5 to the use of OUT, and considers two alternative measures of. Both are closely related to OUT. The first is the proportion of the total value (cf. number) of MOD contracts let by competitive tendering (OUTV). The second is the proportion the total number of MOD contracts let by either competitive tendering or by reference to market forces (OUTM).

Figure B.1 plots each of the candidate proxies for the level of outsourcing. OUTV is more volitile than OUT because it is suceptable to being distorted by a small number of large contracts in any given year, but it better reflects the number of large contracts allocated by competition. OUTV has fallen since 2003, perhaps reflecting the use of non-competitive procurement of Urgent Operational Requirements for operations in Iraq and Afghanistan. OUTM is a broader measure than OUT. It is able to capture any shift from non-contestable forms of procurement (such as cost-plus contracts) towards more contestable if not completely competitive procurement methods (such as contracts let by reference to market forces). It is relatively stable over the sample suggesting procurement has switched from being contestable to competitive, rather than previously uncontested areas being exposed to competition. OUT is used in Chapter 5 because it is both less influenced by a small number of high value contracts than is OUTV and able to capture any movement away from 'contracts let by reference to market forces' into 'contracts let by competitive tendering' (which is lost in the broader OUTM).

The remainder of this Appendix presents the results of estimating the defence shares model (5.10) using the each of the alternative to OUT in turn. Precisely the same strategy is pursued here as was taken in the original estimation of the shares model in Subsection 5.6.3. In the initial model, $\mathbf{x}$ in equation (5.11) contains RPFI, PRGOVT, POP, SPILL, THREAT and a measure of outsourcing, and $\mathbf{d}$ contains CONST, AFDGUM, CWDUM and IRAQ. T and TBREAK are also included again. One lag is included ($\rho = 1$), and the model estimated over

Figure B.1: Graph of alternative proxies for the level of military outsourcing



the period 1981–2007. If this initial model is misspecified, simpler models allowing more lags are considered.

B.1 Value of competitively let contracts

A test of the existence of a long-run relationship in the initial model yields a statistic of 3.093 against 5% critical value of 4.329, and cannot reject null of no relationship, though there is again strong evidence of serially correlation errors in the testing regression. Given this, the relationship and serial correlation tests on the models analagous to those of Table 5.10 are presented in Table B.1. Model PR-C excludes the proxy for the level of outsourcing and remains unchanged from Table 5.10. Without exception, the models in Table B.1 show evidence of serially correlated errors invalidating the bounds testing approach. As a result

Table B.1: Tests for long-run relationships using OUTV as proxy for outsourcing

	Model PR-A	Model PR-B	Model PR-D	Model PR-E
Elements of **x**	OUT, PRGOVT, SPILL, THREAT	OUT, PRGOVT, SPILL	OUT, PRGOVT, THREAT	OUT, SPILL, THREAT
	N=26, 1982–2007	N=25, 1983–2007	N=25, 1983–2007	N=25, 1983–2007
	Tests for long-run relationship			
	F(5,7)=10.026***	F(4, 5)=12.616***	F(4, 5)=12.2062**	F(4, 5)=3.321
	Serial Correlation tests (LM tests)			
1 lag	F(1, 6)=6.192**	F(1, 4)=4.999*	F(1, 4)=2.635	F(1, 4)=11.856**
2 lags	F(2, 5)=2.758	F(2, 3)=24.444**	F(2, 3)=6.175*	F(2, 3)= 17.156**
3 lags	F(3, 4)=1.716	F(3, 2)=13.255*	F(3, 2)=16.506*	F(3, 2)=7.642

Notes: See Table 5.2

we are unable to find an adequate model of the defence share of GDP using OUTV as a proxy for the level of outsourcing.

Although possibly invalid, the results of re-estimating the preferred model (PR-B) of Chapter 5 using OUTV as the measure of outsourcing are presented in the second column of Tables B.2 and B.3. It is not a good model; it explains little of the long-run variation in military spending. However, although the long-run coefficients are largely insignificant they are broadly similar, in sign and magnitude, to those of the original estimates of Subsection 5.6.3. The coefficient on OUT is smaller but still negative.

B.2 Competitive tendering plus referenced to market forces

When OUTM is used rather than OUT the initial model yields a test statistic for a long-run relationship of 2.5883 against 5% critical value of 4.329. Again a null of no long-run relationship cannot be rejects, but there is evidence of serial correlation.

Table B.4 presents the statistics analogous to those in Table 5.10. Only model PR-B (the preferred model in the main body) exhibits both a significant relationship test statistic and no evidence of serial correlation, appearing to support the selection PR-B as the preferred model in Table 5.10.

The results of re-estimating this model using OUTM rather than OUT are presented in the final columns of Tables B.2 and B.3. It is a poor model, which fails completely to explain long-run variations in military expenditure; not a single regressor is individually significant, even in the short-run. Using OUTM causes the signs of all but one of the long-run coefficients relative to the OUT model of Tables 5.11 and 5.12. Furthermore, it has too few degrees of freedom to calculate diagnostic statistics. That the model PR-B with OUTM is so poor raises further doubts as to the robustness of the model in Chapter 5. It lends support to the suggestion that the preferred model in Chapter 5 results from identifying a spurious relationship.

Table B.2: Short-run estimates of preferred model using alternative measures of OUT

	Model PR-B OUT_j=OUTV		Model PR-B OUT_j=OUTM	
1983–2007 N=25	Coeff.	Std. Error	Coeff.	Std. Error
D SME(-1)	0.760**	0.268	-1.776	1.859
D SME(-2)	1.493***	0.152	-0.291	1.473
D SME(-3)	0.584***	0.168	1.56	0.886
DOUT	0.003*	0.001	-0.055	0.043
DOUT(-1)	0.011***	0.003	0.062	0.041
DOUT(-2)	0.004	0.002	0.037	0.025
DOUT(-3)	-	-	0.029	0.020
DPRGOVT	-0.006	0.018	0.247	0.237
DPRGOVT(-1)	-0.010	0.019	-0.010	0.098
DPRGOVT(-2)	-0.044**	0.013	-0.276	0.192
DPRGOVT(-3)	-	-	-0.584	0.506
DSPILL	-0.378***	0.080	0.320	0.791
DSPILL(-1)	-0.449***	0.103	-0.026	0.471
DSPILL(-2)	-0.048	0.071	-0.734	0.653
DCONST	5.770	2.394	16.500	10.872
DAFGDUM	0.299	0.073	0.146	0.284
DT	-0.121	0.038	-0.118	0.068
DTBREAK	0.008	0.093	0.241	0.299
ECM(-1)	-1.527	0.286	0.755	1.582
	Diagnostics			
SC Tests - 1 lag	F(1, 4)=1.5426		F(1, 2)= 7.9836	
2 lags	F(2, 3)=5.7439*		N/A	
3 Lags	F(3, 2)=4.7321		N/A	
RESET	F(1, 4)=0.46682		N/A	
Normality	CHSQ(2)=3.9954		N/A	
Het test	F(1, 23)=0.050593		F(1, 23)=0.66742	
AIC	51.4209		29.0382	
AdjR^2	0.96404		0.69391	
RSS	0.0048311		0.024673	

Notes: See Table 5.4

Table B.3: Long-run estimates of preferred model using alternative measures of OUT

	Model PR-B		Model PR-B	
	OUT_j=OUTV		OUT_j=OUTM	
1983–2007	Coeff.	Std.	Coeff.	Std.
N=25		Error		Error
	long-run estimates			
PRGOVT	0.015	0.018	-0.101	0.453
SPILL	0.249*	0.107	1.320	1.494
OUT_j	-0.002	0.002	0.251	0.434
CONST	3.779	1.935	-21.85	34.671
AFGDUM	0.196***	0.043	-0.193	0.734
T	-0.079*	0.032	0.156	0.291
TBREAK	0.005	0.062	-0.320	0.375

Notes: See Table 5.4

Table B.4: Tests for long-run relationships using OUTM as proxy for outsourcing

	Model PR-A	Model PR-B	Model PR-D	Model PR-E
Elements of **x**	OUT, PRGOVT, SPILL, THREAT	OUT, PRGOVT, SPILL	OUT, PRGOVT, THREAT	OUT, SPILL, THREAT
	N=26, 1982–2007	N=25, 1983–2007	N=25, 1983–2007	N=25, 1983–2007
	Tests for long-run relationship			
	F(5, 7)=4.3670*	F(4, 5)=8.6490***	F(4, 5)=2.6982	F(4, 5)= 8.6756***
	Serial Correlation tests (LM tests)			
1 lag	F(1, 6)=4.0801*	F(1, 4)=3.7436	F(1, 4)=0.26359	F(1, 4)=3.5231
2 lags	F(2, 5)=4.5270*	F(2, 3)=1.9069	F(2, 3)=4.8847	F(2, 3)=20.1195**
3 lags	F(3, 4)=2.8373	F(3, 2)=2.1461	F(3, 2)=217.0236***	F(3, 2)=11.0604*

Notes: See Table 5.2

Appendix C

Unit Root Tests on Regressors Used in Chapter 5

This Appendix presents the results of unit root tests on the variables used in Chapter 5. It is intended to demonstrate that there is some uncertainty about the stationarity, or otherwise, of the variables used, and so justifying the use bounds testing and ARDL.

Table C.1 presents the results of unit root tests on the three measures of military expenditure and their differences without allowing for the existence of a trend break in 2001. These should be treated with a degree of scepticism since the sample is rather too small to yield reliable results from unit root tests which rely on asymptotic results and have notoriously low power, even for large samples.

For the three measures of military expenditure there is also a concern that they may contain structural breaks. Visual inspection of the series in Figure 5.1 on page 97 suggests that all three may have a break in their trends around 1998–2001. An obvious explanation of such a break, if in 2001, would be the change in (perceived) strategic environment following the terrorist attacks of 9/11.

If no trend break is allowed, then the level of military spending (LME) has no overall trend. The applicable unit root tests in the third column of Table C.1 offer no clear conclusion as to LME's stationarity. Neither the KPSS test nor the others reject their contradictory null hypotheses. Allowing for one break at an unknown point, the Zivot-Andrews test yields no significant t-statistics, though any break towards the end of the sample may not be detected because the test truncates 15% of observations at either end. Forcing any trend break to occur in 2001, Perron's [160, 161] IO model test cannot reject the null of non-stationarity with break.[1] Whether tested with or without a break, it is not clear whether LME is stationary.

Logged military expenditure (ME) appears to have a slight overall trend in Figure 5.1. The unit root tests in the fourth column of Table C.1 are inconclusive (though perhaps sug-

[1] The t-statistic from Perron's [161] innovational outlier Model C, with lag length 4 (selected by the iterated t-test method [148] from maximum of five) is -2.5901 against the 5% critical value of -4.37046 calculated from Carrion i Silvestre *et al.*'s [35] surface response function.

gestive of a fractionally integrated process). A Zivot-Andrews test yields no significant t-statistics, but they are falling (towards negative critical value) as the break point is moved towards 2003. Forcing the break to be in 2001, a Perron test yields t-statistic of -2.12976, so cannot reject the null of non-stationarity with break in 2001.

When no break is allowed, the defence share of GDP (SME) appears to be trending. The test statistics in the fifth column of Table C.1 suggest trend stationarity. The Zivot-Andrews t-statistics are significant in both 1997 and 2001. Again forcing any break to be in 2001, the Perron IO test with one period for adjustment statistic is -1.4141, so again cannot reject null of non-stationarity with break in 2001. Indeed none, of the measures of military expenditure can reject a null of nonstationarity about a trend with a break in 2001.

GDP is trending, so the lower tests in Table C.1 are appropriate. It appears not to be stationary; the KPSS test rejects its null of I(0) about a trend. Although the Dickey-Fuller test rejects its null of nonstationarity, the three more refined tests cannot.

GOVT should certainly include a trend: when plotted it appears explosive. The tests presented in the seventh column of Table C.1 suggest a nonstationary process about the trend; the KPSS test rejects its null of trend stationarity and the other tests cannot reject their null of non-stationarity.

POP is also trending, but in this case the unit root tests offer no clear conclusion; the DFGLS and KPSS reject their contradictory null hypotheses. The remaining tests yield positive statistics, even having allowed a trend, perhaps suggesting an explosive process.

OUT displays a clear positive trend through the sample period. The statistics in the ninth column of Table C.1 suggest a nonstationary process about a trend; the KPSS test rejects its null of trend stationarity, and the remaining tests cannot reject their nulls of non-stationarity.

SPILL follows a broadly similar path to LME, so may either have a slight trend or be level. If a trend is allowed, most of the tests suggest a nonstationary process about a trend, but the ADF test rejects this null. The tests without a trend, but with a constant, offer no clear conclusion either; The KPSS and ADF tests suggest stationarity, but the remaining tests cannot reject their null of nonstationarity. Irrespective of whether a trend is included or not, the unit root tests cannot determine conclusively the stationarity of SPILL.

The final column relates to THREAT, but the tests ignore the huge break in that variable at the end of the Cold War and so are not thought reliable. Unit Root tests accounting for the possibility of a structural break are not conducted because of the uncertainty already established for other series, especially SPILL.

Although the stationarity of some variables can be determined, many cannot, including the three measures of military spending. Even for those for which the tests are conclusive, the size of the sample and the low power of the tests used mean it is sensible to adopt an estimator which is consistent irrespective of whether the variables used are stationary.

Table C.1: Unit root tests on regressors used in Chapter 5

1985–2007	5% critical value	LME	ME	SME	GDP	GOVT	POP	OUT	SPILL	THREAT
					Constant , no trend					
DF Test	-3	(0) -1.710	(0) -1.605	(0) -2.853*	(0) -4.151***	(0) 2.984	(0) 8.444	(0) -1.546	(0) -0.248	(0) -1.436
ADF Test[a]	-3	(7) -2.888*	(7) -2.810*	(5) -2.671*	(7) 4.161	(6) 3.528	(7) 2.726	(4) -1.059	(4) -3.694**	(3) -1.982
PPerron[b]	-3	(2) -1.727	(2) -1.656	(2) -2.533	(2) 0.974	(2) 2.094	(2) 6.063	(2) -1.499	(2) -0.859	(2) -1.412
DFGLS[c]	-	-2.972***	-2.875**	-1.677	-0.635	0.747	2.621	-1.238	-0.968	-1.693
		(7, -2.285)	(7, -2.285)	(7, -2.285)	(8, -2.347)	(8, -2.347)	(4, -2.436)	(2, -2.427)	(1, -2.483)	(3, -2.368)
KPSS[d]	0.463	(2) 0.488**	(2) 0.462*	(2) 0.88***	(2) 0.947***	(2) 0.862***	(2) 0.943***	(2) 0.678**	(2) 0.233	(2) 0.674**
					Constant and trend					
DF Test	-3.6	(0) -0.470	(0) -0.495***	(0) -4.587**	(0) -4.692***	(0) -0.613	(0) 5.534	(0) -2.027	(0) 0.060	(0) -0.037
ADF Test[a]	-3.6	(7) -1.809	(7) -1.784	(3) -3.245*	(7) 1.605	(7) 2.758	(7) 2.108	(2) -2.269	(7) -3.975**	(3) -1.654
PPerron[b]	-3.6	(2) -0.685	(2) -0.732	(2) -4.589***	(2) -1.101	(2) -0.947	(2) 5.386	(2) -2.103	(2) -0.499	(2) -0.414
DFGLS[c]	-	-2.993	-2.902**	-2.857	-0.884	-1.222	-3.133**	-2.432	-2.877	-2.280
		(7, -2.820*)	(7, -2.820)	(3, -3.197)	(8, -2.808)	(1, -3.414)	(4, -3.009)	(2, -3.313)	(3, -3.197)	(3, -3.197)
KPSS[d]	0.146	(2) 0.21**	(2) 0.206**	(1) 0.115	(0) 0.203**	(2) 0.191**	(2) 0.221***	(2) 0.160**	(1) 0 .233***	(2) 0.244***

Notes: Conducted using a consistent sample 1985–2007 (23 observations), with a maximum permissible lag 8 (in accordance with Schwert's [180] criterion: k_{max} =int$\{12 * (23/100)^{0.25}\} = 8$). The exception is OUT which has maximum lag of 5 because of the shorter series available. *, **, and *** denote significance at the 10%, 5% and 1% levels respectively. [a] Lag length selected using Ng and Perron's [148] sequential t-testing method. [b] using Newey-West standard errors and lags calculated by int{4(T/100)2/9}. [c] Critical value presented in parenthesis. The test has undistorted size under the null in the presence of a neglected level or trend break. [d] Using automatic bandwidth selection with the Quadratic Spectral kernel used to weight the empirical autocovariance function [99].

List of Abbreviations

2SLS	Two Stage Least Squares
AAS	Annual Abstract of Statistics
ADF	Augmented Dickey-Fuller
AIC	Akaike Information Criterion
AIDS	Almost Ideal Demand System
AR	Auto Regression
ARDL	Autoregressive Distributed Lag
AFPRB	Armed Forces Pay Review Body
BAE	BAE Systems
CES	Constant elasticity of substitution
CRS	Constant returns to scale
DIS	Defence Industrial Strategy
DoD	Department of Defence (US unless otherwise specified)
FTRS	Full-time Reserve Service
GAO	Government Accountability Office (US)
GDP	Gross Domestic Product
GMM	Generalised Method of Moments
GOCO	Government owned contractor operated
IV	Instrumental Variables
KBR	Kellogg Brown & Root
MOD	Ministry of Defence (UK)
NAO	National Audit Office (UK)
NES	New Earnings Survey
OLS	Ordinary Least Squares
ONS	Office for National Statistics (UK)
PBL	Performance Based Logistic (contract)
PFI	Private Finance Initiative
PPP	Public Private Partnership
PRT	Property Rights Theory
RAB	Resource Accounting and Budgeting
RAF	Royal Air Force
ROCE	Return on Capital Employed

RPI	Retail Prices Index
RSS	Residual Sum of Squares
SC	Serial Correlation
SIPRI	Stockholm International Peace Research Institute
SPV	Special Purpose Vehicle (corporate entity)
TCE	Transactions Cost Economics
UK	United Kingdom
US	United States of America
VAR	Vector Autoregression

References

[1] Albin, S. (1992) Bureau Shaping and Contracting Out: The Case of Australian Local Government. Public Policy Discussion Paper 29. Canberra: ANU.

[2] Albrecht, M. (1979) *Labor substitution in the military environment: Implications for enlisted force management*. Rand Report R-2330-MRAL. Santa Monica: The Rand Corporation.

[3] AMR International (2009) *European Defence Services Report 2008*. London: AMR International.

[4] Anagboso, M., and A. Spence (2008) *Scoping Paper on Possible Improvements to Measurement of Defence in the UK National Accounts*. Newport: Office for National Statistics. Available at `http://www.statistics.gov.uk/articles/nojournal/Defence08v1.pdf`.

[5] *Annual Abstract of Statistics* (various years 1974–2004). London: HMSO/TSO.

[6] Armed Forces Pay Review Body (various years 1972–2006) *Annual Report*. London: HMSO/TSO.

[7] Ash, C., B. Udis and R. McNown (1983) Enlistments in the all-volunteer force: A military personnel supply model and its forecasts. *The American Economic Review* 73(1) 145–155.

[8] Badelow, C. (1998) Discussion in *Commercial Support Program Benchmarking Seminar*, Canberra, 29–30 September. Cited in [100].

[9] Bajari, P., and S. Tadelis (2001) Incentives Versus Transaction Costs: A Theory of Procurement Contracts. *RAND Journal of Economics* 32(3) 287–307.

[10] Baker, G., R. Gibbons and K. Murphy (2002) Relational contracts and the theory of the firm. *Quarterly Journal of Economics* 117 39–83.

[11] Baker, G., R. Gibbons and K. Murphy (2006) *Contracting for Control*. Unpublished manuscript. Los Angeles: University of Southern California.

[12] Bannerman, G. (2008) *Merchants and the Military in Eighteenth-Century Britain*. London: Pickering & Chatto.

[13] Baron, D.P. (1993) Defense Procurement: Politics, Management, and Incentives. In *Incentives in procurement contracting*. Pew Studies in Economics and Security, edited by Leitzel, J. and J. Tirole, pp 7–22. Boulder and Oxford: Westview Press.

[14] Baum, C., M. Shaffer and S. Stillman (2003) *Instrumental variables and GMM: Estimation and testing*. Boston College Working Paper No. 545.

[15] Baum, C., M. Shaffer and S. Stillman (2007) *ivreg2: Stata module for extended instrumental variables/2SLS, GMM and AC/HAC, LIML and k-class regression*. http://ideas.repec.org/c/boc/bocode/s425401.html

[16] Bel, G., and A. Costas (2006) Do public sector reforms get rusty? Local Privatisation in Spain *The Journal of Policy Reform* 9(1) 1–24.

[17] Bel, G., and M. Warner (2008) Does privatization of solid waste and water services reduce costs? A review of empirical studies. *Resources, Conservation and Recycling 52 1337–1348.*

[18] Bello, H., and S. Szymanski (1996) Compulsory Competitive Tendering for Public Services in the UK: The Case of Refuse Collection. *Journal of Business Finance and Accounting* Vol. 23(5 and 6) 881–903.

[19] Bennett, J., and E. Iossa (2006) Building and managing facilities for public services. *Journal of Public Economics* 90 2143–2160.

[20] Bentz, A., P. Grout and M. Halonen (2004) What Should Governments Buy form the Private Sector—Assets or Services? Unpublished manuscript available at `http://www.development.wne.uw.edu.pl/uploads/Courses/institutional_grout.pdf` accessed June 2010.

[21] Berenyi, E. and M. Stevens (1988) Does privatisation work? A study of the delivery of eight local services. *State & Local Government Review* 20 (1) 11–20.

[22] Berndt, E. (1976) Reconciling Alternative Estimates of the Elasticity of Substitution. *The Review of Economics and Statistics* 58(1) 59–68.

[23] Berndt, E. (1991) *The Practice of Econometrics: Classic and Contemporary*. Reading; Addison-Wesley.

[24] de Bettignies, J-E., and T. Ross (2004) The Economics of Public-Private Partnerships. *Canadian Public Policy* 30(2) 135–154.

[25] Besley, T. and M. Ghatak (2001) Government versus Private Ownership of Public Goods. *Quarterly Journal of Economics* 116 1343–1372.

[26] Bound, J., D. Jeager and R. Baker (1995) Problems with IV estimators when the correlation between the instruments and endogenous variables is weak. *Journal of the American Statistical Association* 90 443–450.

[27] Bowles, S. (1970) Aggregation of Labor Inputs in the Economics of Growth and Planning: Experiments with a Two-Level CES Function. *The Journal of Political Economy* 78(1) 68–81.

[28] Boyne, G.A. (1998a) Competitive Tendering in Local Government: A Review of the Theory and Evidence. *Public Administration* 76 695–712.

[29] Boyne, G.A. (1998b) *Public Choice Theory and Local Government: A Comparative Analysis of the UK and the USA*. Basingstoke: Palgrave.

[30] Braddon, D. (2005) What lies beneath? Who owns British defence contractors and does it matter? *Frontiers in Finance and Economics*. 2 (1) 68–91.

[31] Braddon, D. (2008) Hidden depth: tracing corporate ownership and its implications for the UK defense industry. *The Economics of Peace and Security Journal* 3 (1) 63–71.

[32] Braddon, D., P. Dowdall and A. Kendry (1996) Organizational Reform, Market Testing and Defence. In *Privatization: Social science themes and perspectives*, edited by D. Braddon and D. Foster, pp 95–127. Aldershot, U.K.: Dartmouth.

[33] Camm, F. (1996) *Expanding Private Production of Defense Services*. National Defence Research Institute Report. Santa Monica: RAND.

[34] Camm, F., and V. Greenfield (2005) *How Should the Army Use Contractors on the Battlefield? Assessing Comparative Risk in Sourcing Decisions*. Santa Monica: RAND.

[35] Carrion i Silvestre, J., A. Sanso i Rossello and M. Ortuno (1999) Response surfaces estimates for the Dickey-Fuller unit root test with structural breaks. *Economics Letters* 63 279–283.

[36] Chalmers, M. (2009) *Preparing for the Lean Years*. Future Defence Review, Working Paper No. 1. London: Royal United Services Institute.

[37] Chalmers, M. (2010) *Capability Cost Trends: Implications for the Defence Review*. Future Defence Review, Working Paper No. 5. London: Royal United Services Institute.

[38] Ciccone, A. and G. Peri (2005) Long-Run Substitutability Between More and Less Educated Workers: Evidence from U.S. States, 1950–1990. *Review of Economics and Statistics* 87(4) 652–663.

[39] Coase, R. (1937) The nature of the firm. *Economica* 4 386–405.

[40] Dalen, D., and E. Moen (2003) *The proper scope of government when costs are contractible*. CEPR Discussion Paper No. 3992.

[41] Davidson, R. and J. MacKinnon (1993) *Estimation and Inference in Econometrics*. Oxford: Oxford University Press.

[42] De Meza, D. and B. Lockwood (1998) Does asset ownership always motivate managers? Outside options and the property right theory of the firm. *Quarterly Journal of Economics* 113 361–386.

[43] Department for Business Innovation & Skills (2009) The 2009 Value Added Scoreboard: The top 800 UK and 750 European companies by Value Added. London: Department for Business Innovation & Skills.

[44] Dijkgraaf, E., and R.H. Gradus (2003) Cost Savings of Contracting Out Refuse Collection. *Empirica* 30 (2) 149–161.

[45] Dijkgraaf, E., and R.H. Gradus (2007) Collusion in the Dutch Waste Collection Market. *Local Government Studies* 33 (4) 573–588.

[46] Dixit, A.K. (1997) Power of Incentives in Public Versus Private Organizations. *American Economic Review* (Papers and Proceedings) 87 378–382.

[47] Dixit, A.K. (2002) Incentives and Organisations in the Public Sector: An Interpretive Review. *The Journal of Human Resources* 37 696–727.

[48] Domberger, S., C. Hall and M. Jeffries (1995) Competitive Tendering and Contracting in Commonwealth Government Agencies: 1995 Survey Findings. Appendix B in *Competitive tendering and contracting by public sector agencies*. By the Industry Commission. Report No. 48. Melbourne: Australian Government Publishing Service.

[49] Domberger, S., C. Hall and E. Li (1995). The determinants of price and quality in competitively tendered contracts. *Economic Journal* 105 1454–1470.

[50] Domberger, S., and P. Jensen (1997) *Contracting out by the public sector: theory, evidence and prospects*. Oxford Review of Economic Policy 13 (4) 67–78.

[51] Domberger, S., P. Jensen and R. Stonecash (2002) Examining the magnitude and sources of cost savings associated with outsourcing. *Public Performance & Management Review* 26 (2) 148–168.

[52] Domberger, S., S. Meadowcroft and D. Thompson (1986) Competitive tendering and efficiency: the case of refuse collection. *Fiscal Studies* 7 (4) 69–89.

[53] Dougherty, C. (1972) Estimates of Labor Aggregation Functions. *The Journal of Political Economy* 80(6) 1101–1119.

[54] Downs, A. (1967) *Inside Bureaucracy*. Boston, MA: Little Brown.

[55] Dunkin, G. and T. Valila (2005) *Transaction Costs In Public-Private Partnerships: A First Look At The Evidence*. Economic and Financial Report 2005/03. Luxembourg: European Investment Bank.

[56] Dunne, J.P., E. Nikolaidou and N. Mylonidis (2003) The demand for military spending in the peripheral countries of Europe. *Defence and Peace Economics* 14(6) 447–460.

[57] Dunne, J.P. and S. Perlo-Freeman (2003) Demand for military spending in developing countries. *International Review of Applied Economics* 17(1) 23–48.

[58] Edmonds, M. (1999) Defence Privatisation: From State Enterprise to Commercialisation. *Cambridge Review of International Affairs* 13 114–129.

[59] Fallon, P. and P. Layard (1975) Capital-skill complementarity, income distribution, and ouput accounting. *Journal of Political Economy* 83(2) 279–301.

[60] Francesconi, M., and A. Muthoo (2006) Control rights in public-private partnerships. Mimeo, University of Essex.

[61] Fredland, J. (2004) Outsourcing Military Force: A Transactions Cost Perspective on the Role of Military Companies. *Defence and Peace Economics* 15 205–219.

[62] Fredland, J., and A. Kendry (1999) The privatisation of military force: economic virtues, vices and government responsibility. *Cambridge Review of International Affairs* 13, 147–164.

[63] Gadea, D., E. Pardos, C. Perez-Fornies (2004) A long-run analysis of defence spending in the SPILL countries. *Defence and Peace Economics* 15(3) 231–249.

[64] Ganley, J., and J. Grahl (1988) Competition and Efficiency in Refuse Collection: A Critical Comment. *Fiscal Studies* 9 80–89.

[65] Gansler, J. (2006) Moving toward Market-Based Government: The Changing Role of Government as the Provider. In *Competition, Choice and Incentives in Government Programs*, edited by J. Kamensky and A. Morales. Oxford: Rowman & Littlefield.

[66] Gansler, J., and W. Lucyshyn (2006a) Competitive Sourcing: What Happens to Federal Employees? In *Competition, Choice and Incentives in Government Programs*, edited by J. Kamensky and A. Morales. Oxford: Rowman & Littlefield.

[67] Gansler, J., and W. Lucyshyn (2006b) Implementing Alternative Sourcing Strategies: Four Case Studies In *Competition, Choice and Incentives in Government Programs*, edited by J. Kamensky and A. Morales. Oxford: Rowman & Littlefield.

[68] Garvey, G. (1995) Why reputation favors joint ventures over vertical and horizontal integration: a simple model. *Journal of Economic Behaviour & Organisation* 28 387–397.

[69] Gates, S., and A. Robbert (2000) *Personnel Savings in Competitively Sourced DoD Activities: Are They Real? Will They Last?* National Defence Research Institute Report MR-1117-OSD. Santa Monica: RAND.

[70] Gates, R. (2009) DoD News Briefing With Secretary Gates From The Pentagon, 6th April 2009. Transcript available at http://www.defense.gov/transcripts/transcript.aspx?transcriptid=4396 accessed 10th June 2010.

[71] Geren, P., and G.W. Casey (2009) A statement on the posture of the United States Army 2009. Submitted to the Committees and Subcommittees of the United States Senate and House of Representatives, 1st Session, 111th Congress, May 2009.

[72] Gibbons, R. (2005) Four formal(izable) theories of the firm? *Journal of Economic Behaviour & Organisation* 58 200–245.

[73] Godfrey, L.G. (1988) Misspecification tests in econometrics: The Lagrange multiplier principle and other approaches. Econometric Society Monographs No. 16. Cambridge: Cambridge University Press.

[74] Gray, B. (2009) Review of Acquisition for the Secretary of State for Defence. London: Ministry of Defence. Available at http://www.aof.mod.uk/aofcontent/downloads/gray/gray_report.pdf accessed June 2010.

[75] Greene, W. (2003) *Econometric Analysis*, Fifth Edition. Upper Saddle River, NJ: Prentice Hall

[76] Grossman, S., and O. Hart (1986) The Costs and Benefits of Ownership: A Theory of Vertical and Lateral Integration. *Journal of Political Economy* 94 691–719.

[77] Guistozzi, A. (2007) The privatizing of war and security in Afghanistan: future or dead end? *The Economics of Peace and Security Journal* 2 (1) 30–35.

[78] Hahn, J. and J. Hausman (2002) Notes on bias in estimators for simultaneous equations models. *Economics Letters* 75(2) 237–41.

[79] Halonen M. (2002) Reputation and the allocation of ownership. *Economic Journal* 112 539–558.

[80] Hammermash, D. (1993) *Labor Demand*. Princeton, NJ; Princeton University Press.

[81] Hart, O. (2003) Incomplete Contracts and Public Ownership: Remarks, and an Application to Public-Private Partnerships. *Economic Journal* 113 C69–76.

[82] Hart, O., and J. Moore (1990) Property Rights and the Nature of the Firm. *Journal of Political Economy* 98 1119–1158.

[83] Hart, O., and J. Moore (2007) Incomplete Contracts and Ownership: Some New Thoughts. *American Economic Review* 97 182–186.

[84] Hart, O., A. Shleifer and R.W Vishny (1997) The proper scope of government: Theory and an application to prisons. *Quarterly Journal of Economics* 112 1127–1161.

[85] Hartley, K. (2002) The economics of military outsourcing. *Public Procurement Law Review* 5 287–297.

[86] Hartley, K. (2003a) Problems of using Partnering and Similar Private Sector Practice in the Public Sector Environment: The Example of PPPs/PFI. In *Public Procurement: The Continuing Revolution*, edited by S. Arrowsmith and M. Trybus, pp 187–195. Kluwer Law International.

[87] Hartley, K. (2003b) Defence industries. In *Industries in Europe: Competition, Trends and Policy Issues* edited by P. Johnson. Cheltenham: Edward Elgar.

[88] Hartley, K. (2004) The UK's 2004 Defence White Paper: An Economic Perspective. Unpublished manuscript, University of York.

[89] Hartley, K. (2011) The Economics of Defence Policy. London: Routledge.

[90] Hartley, K. and P. MacDonald (2010) Country Survey XXI: The United Kingdom. *Defence and Peace Economics* 21 (1) 43-63.

[91] Hartley, K. and T. Sandler (eds) (1990) *The Economics of Defence Spending: An international survey*. London: Routledge.

[92] Hausman, J. (1983) Simultaneous Equation Models. In *Handbook of Econometrics*, Vol. 1, edited by Z. Griliches and M. Intriligator. Amsterdam: North-Holland Publishing Company.

[93] Hayashi, F. (2000) *Econometrics*. Princeton, NJ: Princeton University Press.

[94] Healthcare Commission (2005) *Acute Hospital Portfolio Reviews: Pathology, Facilities Management, Therapy and Dietetics, Information and Records*. London: Commission for Healthcare Audit and Inspection.

[95] Hermalin, B.E., and M.L. Katz (1991) Moral hazard and Verifiability: The Effects of Renegotiation in Agency. *Econometrica* 59 1735–1753.

[96] HM Revenue and Customs (2010) Historic exchange rates available at `https://www.uktradeinfo.com/index.cfm?task=exchange` accessed January 2010.

[97] HM Treasury (2003) *Supplementary Green Book Guidance*. London: HM Treasury.

[98] HM Treasury (2009) PFI Signed Projects List. Available at `http://webarchive.nationalarchives.gov.uk/+/http://www.hm-treasury.gov.uk/ppp_pfi_stats.htm` accessed September 2009.

[99] Hobijn, B., P.H. Frances and M. Ooms (1998) *Generalizations of the KPSS-test for Stationarity*. Econometric Institute Report 9802/A, Econometric Institute, Erasmus University Rotterdam.

[100] Hodge, G.A. (2000) *Privatisation: An International Review of Performance*. Boulder, CO: Westview Press.

[101] Holmstrom, B., and P.R. Milgrom (1991) Multitask Principal-Agent Analysis: Incentive Contracts, Asset Ownership, and Job Design. *Journal of Law, Economics and Organization* 7 24–52.

[102] Hosek, J. and M. Totten (2004) The effect of deployment on first- and second-term reenlistment in the US active duty force. Defence and Peace Economics 15 (5), 433–451.

[103] House of Commons Committee of Public Accounts (2005) *Ministry of Defence: Major Projects Report 2005*. Fiftieth Report of Session 2005–2006. London: The Stationary Office.

[104] House of Commons Committee of Public Accounts (2005) *Improving patient care by reducing the risk of hospital acquired infection: a progress report*. Twenty-Fourth Report of Session 2004–2005. London: The Stationary Office.

[105] Howard (1998) The Canadian Alternative Service Delivery Program, Presentation to the *Commercial Support Program Benchmarking Seminar*, Canberra, 29–30 September. Cited in [100].

[106] Iossa, E., and D. Martimort (2008) *The Simple Micro-Economics of Public-Private Partnerships*. CEIS Tor Vergata Research Paper Series 6(12).

[107] Jaffry, S., Y. Ghulam and A. Apostolakis (2010) Analysing quits and separations from the Royal Navy. *Defence and Peace Economics* 21 (3) 207-28.

[108] Julius, D. (2008) *Public Services Industry Review: Understanding the Public Services Industry: How big, how good, where next?* London: Department for Business Enterprise & Regulatory Reform.

[109] Industry Commission (1996) *Competitive tendering and contracting by public sector agencies*. Report No. 48. Melbourne: Australian Government Publishing Service.

[110] Jensen, P., and R. Stonecash (2005) Incentives and the efficiency of public sector-outsourcing contracts. *Journal of Economic Surveys* 19 (5) 767–787.

[111] Katz, L. and K. Murphy (1992) Changes in relative wages, 1963–1987: Supply and demand factors. *Quarterly Journal of Economics* 107 35–78.

[112] Kennedy, P. (1998) *A Guide to Econometrics*, Fourth Edition. Oxford: Blackwell.

[113] Kirk, R. and T. DePalma (2005) *Performace-Based Logistics Contracts: A Basic Overview*. CRM D0012881.A2/Final. Alexandria, Virginia: Center for Naval Analyses.

[114] Kopač, E. (2006) Defense Industry Restructuring: Trends in European and U.S. Defense Companies. *Transition Studies Review* 13 (2) 283–296

[115] Kulmala, H., M. Ojala, L. Ahoniemi and E. Uusi-Rauva (2006) Unit cost behaviour in public sector outsourcing. *International Journal of Public Sector Management* 19 (2) 130–149.

[116] Laffont, J.-J., and J. Tirole (1993a) Privatization and Incentives. *Journal of Law, Economics, and Organization* 7 84–105.

[117] Laffont, J.-J., and J. Tirole (1993b) The Provision of Quality in Procurement. In *Incentives in procurement contracting*. Pew Studies in Economics and Security, edited by J. Leitzel and J. Tirole, pp 77–86. Boulder and Oxford: Westview Press.

[118] Laffont, J.-J., and J. Tirole (1993c) *A Theory of Incentives in Procurement and Regulation*. Cambridge, MA: MIT Press.

[119] Lafontaine, F., and M. Slade (2007) Vertical Integration and Firm Boundaries: The Evidence. *Journal of Economic Literature* 45 (3) 629-685.

[120] Leitzel, J., and J. Tirole (Eds.) (1993d) *Incentives in procurement contracting.* Pew Studies in Economics and Security. Boulder and Oxford:Westview Press.

[121] Lemer, J. (2009) MoD considers call for big rise in outsourcing. *Financial Times*. 18 November 2009.

[122] Levin, J., and S. Tadelis (2005) *Contracting for Government Services: Theory and Evidence from U.S. Cities*. Unpublished manuscript. University of California, Berkeley.

[123] Lewis, T.R., and D. Sappington (1993) Procurement and Quality Monitoring. In *Incentives in procurement contracting*. Pew Studies in Economics and Security, edited by Leitzel J., and J. Tirole, pp 61–70. Boulder and Oxford: Westview Press.

[124] Maddala, G. (2001) *Introduction to Econometrics*, Third Edition. Chichester: John Wiley and Sons Ltd..

[125] McDavid, J.C. (1985) The Canadian experience with privatizing residential solid waste collection services. *Public Administration Review* 45 (5) 602–608.

[126] MacDonald, P. (2006) Capital-labour substitution in UK armed forces. *Defence and Peace Economics* 17(2) 141–153.

[127] McGuire, M. (1982) U.S. Foreign assistance, Israeli resource allocation, and the arms race in the middle eat. *Journal of Conflict Resolution* 26(2) 199–235.

[128] McKinlay, R. (1989) *Third World military expenditure: determinants and implications*. London: Pinter.

[129] MacLeod, W.B. (2006) *Reputations, Relationships and the Enforcement of Incomplete Contracts*. CESifo Working Paper No. 1730.

[130] Martimort, D., and J. Pouyet (2008) To build or not to build: Normative and positive theories of public-private partnerships. *International Journal of Industrial Organisation* 26 393–411.

[131] Mikusheva, A. and B. Poi (2006) Tests and confidence sets with correct size when instruments are potentially weak. *Stata Journal* 6 335–347.

[132] Murray, M. (2006) Avoiding Invalid Instruments and Coping with Weak Instruments. *Journal of Economic Perspectives*, 20(4) 111–132.

[133] Ministry of Defence (1991) *Options for Change*. London: HMSO.

[134] Ministry of Defence (1994) *Front Line First: The Defence Costs Study*. London: HMSO.

[135] Ministry of Defence (2004) Delivering Security in a Changing World: Future Capabilities. Cm 6269.

[136] Ministry of Defence (2005) *Defence Industrial Strategy: Defence White Paper*. Cm 6697. London: TSO.

[137] Mobley, B.W. (2004) Outsourcing Post-conflict Operations: Designing a System for Contract Management and Oversight. *Journal of Public and International Affairs* 15 21–37.

[138] Molas Gallart, J., and P. Tang (2006) Ownership Matters: Intellectual Property, Privatization and Innovation. *Research Policy* 35 200–212.

[139] National Audit Office (2004) *Battlefield Helicopters*. London:The Stationary Office, HC 486, Session 2003–2004. London: TSO.

[140] National Audit Office (2006) *Ministry of Defence Major Projects Report 2006*, HC 23–I, Session 2006–2007. London: TSO.

[141] National Audit Office (2007a) *Transforming logistics support for fast jets*. HC 825, Session 2006–2007. London: TSO.

[142] National Audit Office (2007b) *Improving the PFI tendering process*. HC 149, Session 2006–7. London: TSO.

[143] National Audit Office (2008) *Allocation and management of risk in Ministry of Defence PFI projects*. HC 343, Session 2007–2008. London: TSO.

[144] National Audit Office (2009) *Support to High Intensity Operations*. HC 508, Session 2008–2009. London: TSO.

[145] National Audit Office (2010) *Ministry of Defence: Delivering multi-role tanker aircraft capability*. HC 433, Session 2009–2010. London: TSO.

[146] National Board for Prices and Incomes (1970–71) *Report No. 142*. Cm. 4291, London: HMSO.

[147] Nelson, C.R. and R. Startz (1990) The distribution of the Instrumental Variables estimator and its t-ratio when the instrument is a poor one. *Journal of Business*, 63(1) S125–40.

[148] Ng, S. and P. Perron (1995) Unit root tests in ARMA models with data-dependent methods for the selection of the truncation lag. *Journal of the American Statistical Association* 90 268–281.

[149] Nikolaidou, E. (2008). The demand for military expenditure: Evidence from the EU15 (1961–2005). *Defence and Peace Economics* 19 (4) 273–292.

[150] Niskanen, W. (1971) *Bureaucracy and Representative Government*. Chicago: Aldine-Atherton.

[151] Niskanen, W. (2001) Bureaucracy. In W. Shugart II and L. Razzolini (eds) *The Elgar Companion to Public Choice*. Cheltenham: Edward Elgar.

[152] Ohlsson, H. (2003) Ownership and production costs: Choosing between public production and contracting-out in the case of Swedish refuse collection. *Fiscal Studies* 24(4) 451–476.

[153] Owen, N. (1994) How many men do armed forces need? An international comparison. *Defence and Peace Economics* 5 (4), 269–288.

[154] PartnershipsUK (2009) Project Database. Available at `http://www.partnershipsuk.org.uk/PUK-Projects-Database.aspx` accessed December 2009.

[155] Pagan, A. and D. Hall (1983) Diagnostic tests as residual analysis. *Econometric Reviews* 2(2) 159–218.

[156] Parker, D., and K. Hartley (1997) The economics of partnership outsourcing versus adversarial competition: a critique. *European journal of Purchasing and Supply Management* 3 (2) 115–125.

[157] Parker, D., and K. Hartley (2003) Transaction costs, relational contracting and public private partnerships: a case study of UK defence. *Journal of Purchasing & Supply Management* 12 97–108.

[158] Peltz, E., M. Robbins, K. Girardini, R. Eden, J. Halliday and J. Angers (2009) *Sustainment of Army Forces in Operation Iraqi Freedom Major Findings and Recommendations*. Santa Monica: RAND.

[159] Perlo-Freeman, S., and E. Sköns (2008) The Private Military Services Industry. *SIPRI Insights on Peace and Security* No. 2008/1.

[160] Perron, P. (1989) The Great Crash, the Oil Price Shock, and the Unit Root Hypothesis. *Econometrica* 57(6) 1361–1401.

[161] Perron, P. (2006) Dealing with structural breaks. In *The Palgrave Handbook of Econometrics, Vol. 1: Econometric Theory*, edited by T. Mills and K. Patterson. Basingstoke: Palgrave Macmillan.

[162] Pesaran, M.H. and Pesaran, B. (1997). *Working with Microfit 4.0: Interactive Econometric Analysis*. Oxford: Oxford University Press.

[163] Pesaran, M.H., Y. Shin and R. Smith (2001). Bounds testing approaches to the analysis of level relationships. *Journal of Applied Econometrics* 16(3) 289–326.

[164] Pesaran, M.H. and L. Taylor (1999) Diagnostics for IV regressions. *Oxford Bulletin of Economics and Statistics* 61(2) 255–281.

[165] Phillips, P. (1983) Exact small sample theory in the simultaneous equations model. In *Handbook of Econometrics*, Vol. 1, edited by Z. Griliches and M. Intriligator. Amsterdam: North-Holland Publishing Company.

[166] Pier, W.J., R.B. Vernon and J.H. Wicks (1974) An empirical comparison of government and private production efficiency. *National Tax Journal* 27 (4) 653–656

[167] Propper, C., S. Burgess and D. Gassage (2008) Competition and Quality: Evidence from the NHS internal market 1991–1999. *Economic Journal* 118 138–170.

[168] Rajan, R. and L. Zingales (1998) Power in a theory of the firm. *Quarterly Journal of Economics* 113 387–432.

[169] Ridge, M. and R. Smith (1991) UK military manpower and substitutability. *Defence Economics* 2(4) 283–293.

[170] Rogerson, W.P. (1994) Economic Incentives and the Defense Procurement Process. *Journal of Economic Perspectives* 8 65–90.

[171] Rogerson, W.P. (1995) Incentive Models of the Defense Procurement Process. In *Handbook of Defense Economics, Volume 1*. Handbooks in Economics, Vol. 12. Edited by K. Hartley and T. Sandler, pp 309–346. Amsterdam; New York and Oxford: Elsevier.

[172] Rosenkranz, S., and P.W. Schmitz (1999) Know-how disclosure and incomplete contracts. *Economics Letters* 63 181–185.

[173] Sandler, T. and K. Hartley (1995) *The Economics of Defence*. Cambridge Surveys of Economic Literature. Cambridge: Cambridge University Press.

[174] Sappington, D. and J. Stiglitz (1987) Privatization, Information and Incentives. *Journal of Policy Analysis and Management* 6 567–582.

[175] Sargan, J. (1988) Testing for misspecification after estimation using instrumental variables. In *Contributions to econometrics: John Denis Sargan*, Vol. 1, edited by E. Maasoumi. Cambridge: Cambridge Universtiy Press. Previously a 1976 unpublished paper from the London School of Economics.

[176] Schwartz, M. (2009) *Department of defence Contractors in Iraq and Afhanistan: Background and Analysis.* Congressional Research Service Report R40764.

[177] Schmidt, K.M. (1996a) The Costs and Benefits of Privatization: An Incomplete Contracts Approach. *Journal of Law, Economics, and Organization* 12 1–24.

[178] Schmidt, K.M. (1996b) Incomplete contracts and privatization. *European Economic Review* 40 569–579.

[179] Schmitz, P.W. (2001) The hold-up problem and incomplete contracts: A survey of recent topics in contract theory. *Bulletin of Economic Research* 53 1–17.

[180] Schwert, G.W. (1989) Tests for Unit Roots: A Monte Carlo Investigation. *Journal of Business and Economic Statistics* 7(4) 147–159.

[181] Shapiro, C., and R.D. Willig (1990) Economic rationales for the scope of privatization. In *The political economy of public sector reform and privatisation*, edited by E. Suleiman and J. Waterbury, pp 55–87. Boulder, CO: Westview Press.

[182] Simons, D. (2004) Occupation for hire: Private military companies and their role in Iraq. *RUSI Journal* 149 68-71.

[183] SIPRI (Various years 1977–2007) *SIPRI Yearbook*. Oxford; Oxford University Press.

[184] Smith, R. (1980) The demand for military expenditure. *The Economic Journal* 90 811–820.

[185] Smith, R. (1989) Models of military expenditure. *Journal of Applied Econometrics* 4 345–359.

[186] Smith, R. (1995) The demand for military expenditure. In *Handbook of Defence Economics*, edited by K. Hartley and T. Sandler. Amsterdam: North-Holland.

[187] Smith, R., A. Hummel and J. Fontanel (1987) Capital-labour substitution in defence provision. In *Defence, Security and Development*, edited by S. Deger and R. West. London: Frances Pinter.

[188] Smoker, R. (1979) Cost as a criterion in determining the mix of military and civilian personnel. In *Defence Manpower Policy: Presentations from the 1976 Rand conference on defence manpower*, edited by R. Cooper. R-2396-ARPA. Santa Monica: RAND.

[189] Snyder, C.M., R.P. Trost and R.D. Trunkey (2001a) Bidding Behavior in the Department of Defense's Commercial Activities Competitions. *Journal of Policy Analysis and Management* 20 21–42.

[190] Snyder, C.M., R.P. Trost and R.D. Trunkey (2001b) Reducing Government Spending with Privatization Competitions: A Study of the Department of Defense Experience. *Review of Economics and Statistics* 83 108–117.

[191] Sodexo (2009) Fiscal 2009 Reference Document. Paris: Sodexho.

[192] Solomon, B. (2005) The demand for Canadian defence expenditures. *Defence and Peace Economics* 16(3) 171–189.

[193] Staiger, D. and J. Stock (1997) Instrumental Variables Regression with Weak Instruments. Econometrica 65(3) 557–586.

[194] Statement on the Defence Estimates, Vol. II (1980, 1982, 1987). London: HMSO.

[195] Stevens, B.J. (1978) Scale, market structure and the cost of refuse collection. *The Review of Economics and Statistics* 60 (3) 438–448.

[196] Stock, J. and M. Watson (2007) *Introduction to Econometrics, Second Edition*. Boston, MA: Pearson Education.

[197] Stock, J. and M. Yago (2005) Testing for Weak Instruments in Linear IV Regression. In *Identification and Inference for Econometric Models: Essays in Honour of Thomas Rothenberg*. Edited by D. Andrews and J. Stock. Cambridge: Cambridge University Press.

[198] Szymanski, S. (1996) The impact of compulsory competitive tendering on refuse collection services. *Fiscal Studies* 17 (3) 1–19.

[199] Szymanski, S., and S. Wilkins (1993) Cheap rubbish? Competitive tendering and contracting out in refuse collection. *Fiscal Studies* 14(3) 109–130

[200] Tadelis, S. (2002) Complexity, Flexibility, and the Make-or-Buy Decision. *American Economic Review Papers and Proceedings* 92(2) 433–437.

[201] Tadelis, S., and O. Williamson (2007) *Transaction Cost Economics.* Unpublished manuscript. University of California, Berkeley.

[202] Thomas, K. (2009) The future of defence outsourcing. Presentation given on 10th September 2009 available at `http://investors.babcock.co.uk/~/media/Files/B/Babcock-Investors/pdfs/results-and-presentations/2009/Babcock-The-Future-of-Defence-Outsourcing-10-09-09.pdf` accessed June 2010.

[203] Tirole, J. (1999) Incomplete contracts: where do we stand? *Econometrica* 67 741–781.

[204] Trimble, S., (2010) US Military Maintenance: Service Conflict. *Flight International*. 18–24 May, pp32–34.

[205] Tullock, G. (1965) *The Politics of Bureaucracy*. Washington, D.C.: Public Affairs Press.

[206] *UK Defence Statistics* (various years 1992–2006). London: HMSO/TSO.

[207] United States Government Accountability Office (1994) *Public-Private Mix: Extent of Contracting Out for Real Property Management Services in GSA.* Briefing report to the Ranking Minority Member, Subcommittee on Investigations and Oversight, Committee on Public Works and Transportation, House of Representatives, May, GAO/GGD-94-126BR, Washington, D.C..

[208] United States Government Accountability Office (2003) *Contract Management: High-Level Attention to Transform DoD Services Acquisition*. GAO-03-935, Washington, D.C..

[209] United States Government Accountability Office (2003b) *Military Operations: Contractors Provide Vital Services to Deployed Forces but Are Not Adequately Addressed in DoD Plans*. GAO-03-695, Washington, D.C..

[210] United States Government Accountability Office (2005) *Defence Management: DoD Needs to Demonstrate That Performance-Based Logistics Contracts Are Achieving Expected Benefits*. GAO-05-966, Washington, D.C..

[211] United States Government Accountability Office (2006) *Military Operations: High level DoD action needed to address long-standing problems with management and oversight of contractors supporting deployed forces.* GAO-07-145, Washington, D.C..

[212] United States Government Accountability Office (2007) *Defense Acquisitions: DoD Needs to exert management and oversight to better control acquisition of services*. GAO-07-359T, Washington, D.C..

[213] United States Government Accountability Office (2008) *Military Operations: DoD Needs to Address Contract Oversight and Quality Assurance Issues for Contracts Used to Support Contingency Operations*. GAO-08-1087, Washington, D.C..

[214] United States Government Accountability Office (2009) *Defense Management: Actions Needed to Overcome Long-standing Challenges with Weapon Systems Acquisition and Service Contract Management*. GAO-09-392T, Washington, D.C..

[215] Uttley, M. (1993) Contracting-Out and Market-Testing in the UK Defence Sector: Theory, Evidence and Issues. *Public Money & Management* 13 55–60.

[216] Uttley, M. (2005) *Contractor on Deployed Military Operations: United Kingdom Policy and Doctrine*. Carlisle, Pennsylvania: Strategic Studies Institute.

[217] Valila, T. (2005) How Expensive are Cost Savings? On the Economics of Public-Private Partnerships. *European Investment Bank Papers* 10(1) 94–119.

[218] Warner, J. and B. Asch (1995) The Economics of Military Manpower. In *Handbook of Defense Economics*, Vol. 1, edited by K. Hartley and T. Sandler. Amsterdam: Elsevier Science B.V..

[219] Warner, J. and S. Negrusa (2006) The Economic Case for All-Volunteer Forces. In *Service to Country: Personnel Policy and the Transformation of Western Militaries*, edited by C. Gilroy and C. Williams. Cambridge, MA: The MIT Press.

[220] White, G. (2009) VT Group's war chest should boost future profits. *The Daily Telegraph*. 23 November. Available at `http://www.telegraph.co.uk/finance/markets/questor/6637625/VT-Groups-war-chest-should-boost-future-profits.html` accessed January 2010.

[221] Williamson, O.E. (1975) *Markets and Hierarchies: Analysis and Antitrust Implications: A study in the Economics of Internal Organisation*. New York:Free Press.

[222] Williamson, O.E. (1979) Transaction-Cost Economics: The Governance of Contractual Relations. *Journal of Law and Economics* 22 233–261.

[223] Williamson, O.E. (1985) *The Economic Institutions of Capitalism*. New York:The Free Press.

[224] Williamson, O.E. (1999) Public and Private Bureaucracies: A transactions cost economics perspective. Journal of Law, Economics & Organisation 15, 306–342.

[225] Williamson, O.E. (2000) The New Institutional Economics: Taking Stock, Looking Ahead. *Journal of Economic Literature* 38 595–613.

[226] Williamson, O.E. (2008) Transaction cost economics: an introduction. In *The Edward Elgar Companion to Transaction Cost Economics*, edited by P. Klein and M. Sykuta Cheltenham: Edgar Elgar. Preprinted in Economics Discussion Papers, No 2007–3. http://www.economics-ejournal.org/economics/discussionpapers/2007–3.

[227] Wilson, J.Q. (1989). *Bureaucracy: What Government Agencies Do and Why they Do It*. New York:Basic Books.

[228] Wooldridge, J. (1990) A note on the Lagrange multiplier and F-statistics for two stage least squares regressions. *Economics Letters* 34 (2), 151–155.

[229] Wooldridge, J. (2002) *Econometric Analysis of Cross Section and Panel Data*. Cambridge, MA: MIT Press.

[230] Wooldridge, J. (2003) *Introductory Econometrics: A Modern Approach*. Mason, Ohio: Thompson South-Western.

[231] Yildirim, J., and S. Sezgin (2003) Military expenditure and employment in Turkey. *Defence and Peace Economics* 14(2) 129–139.