Preliminary Draft, June 2005, Prepared for
Handbook of Defence Economics, Volume 2 Chapter 10.
We would welcome comments and suggestions

THE ECONOMETRICS OF MILITARY ARMS RACES[1]

J. Paul Dunne
University of the West of England
John2.Dunne@uwe.ac.uk

Ron P. Smith
Birkbeck College, London
R.Smith@bbk.ac.uk

Abstract

This chapter reviews the econometric issues involved in estimating models of the competitive acquisition of military capability by hostile powers. Econometrics involves the synthesis of theory, data and statistical methods thus in reviewing the econometrics of arms races, we will pay as much attention to theory and data as we do to statistical methods. After discussing the choice of data and the theoretical issues in specification, we then examine four types of model: time-series estimation of classical Richardson type action-reaction models, using India and Pakistan as an example; Markov switching estimation of game-theory type models, using Greece and Turkey as an example; cross-section models and panel models. Our general conclusion is that the theory does not lead us to expect that arms race interactions will have parameters which are constant over time and the empirical literature largely confirms this.

Keywords
Arms Race, econometric methods, ..

1. Introduction

Arms races - enduring rivalries between pairs of hostile powers, which prompt competitive acquisition of military capability - appear to be a pervasive phenomenon. From the Cold War competition between the US and the USSR to the regional antagonisms between Greece and Turkey or India and Pakistan, they are a matter of continuing concern. The concern arises both because they consume scarce resources

[1] We are grateful for comments on an earlier version from the editors Keith Hartley and Todd Sandler, our co-authors Maria Garcia-Alonso and Paul Levine and Jurgen Brauer.

and because of the danger that they may increase the probability of war. Gibler, et al. (2005) is a recent contribution to the empirical literature on whether arms races lead to war.

The question we address in this paper is how does one use data on military capabilities to estimate and test arms-race models. There are different questions, about the effects of arms races for instance, that take arms races for granted and provide criteria to define them. For instance, Gibler et al. (2005) define them as cases where both countries have increased their military spending or personnel by 8% or more in every year of a three year period and there was qualititative historical evidence that this build-up involved a competitive dynamic between the rivals. This is sensible given their purpose, but other definitions are possible.

The theory of arms races comes in two main forms, a two-agent two-choice game, where each country has a discrete choice to arm or not to arm; and a Richardson type of dynamic action reaction process determining continuous measures of military capability. The theory of arms races is well surveyed by Gleditsch and Njolstad (1990) Brito and Intrilligator (1995) and Intriligator and Brito (2000). Smith (1995) discusses estimation of demand functions for military expenditure, but largely treats foreign military expenditures as exogenous, whereas the central feature of arms race models is the mutual determination of stocks of weapons. Murdoch (1995) deals with alliance issues which overlap with arms races. In this volume, Levine and Garcia-Alonso (chapter 11) and Brauer (chapter 12) examine how arms races drive the arms trade and arms production. Smith et al. (2000) provide more detail on some of the technical time-series issues raised below. The literature is vast and our survey of contributions is far from comprehensive. There are also a number of statistical techniques that we do not discuss, e.g. Andreou et al. (2003) use genetically evolved fuzzy cognitive maps to study the Greek-Turkish arms race.

In evolutionary biology the idea of intra-species or inter-species arms races driving co-evolution has been empirically very fruitful. If you put ‘arms race’ into Google-Scholar, most of the hits are in biology. In contrast, the econometric evidence for military arms races has been less compelling. Sandler and Hartley (1995 ch4) in their review of the empirical analysis of arms races comment: ‘To date, the empirical results can be best described as disappointing.’ Brauer (2002) in a review of the literature on Greece and Turkey comments: ‘The literature, as presently constituted, has reached the point of rapidly declining returns. Running more single or simultaneous regression equations, even when incorporating all the latest quirks of mathematical statistics, is unlikely to much advance our substantive knowledge.’ Subsequent literature has not taken his advice. As teachers of ‘the latest quirks in mathematical statistics’, we would make a distinction between those who use arms race data to illustrate that they can implement the latest statistical model and those who ask which available statistical model best describes the data. The fact that we will be sceptical about the ability of the models to describe the data, should not be taken as scepticism about the need for statistical analysis of the data.

Econometrics involves the synthesis of theory, data and statistical methods[2] thus in reviewing the econometrics of arms races we will pay as much attention to theory and data as we do to statistical methods. Section 2 discusses the data and section 3 the theoretical issues in specification. Section 4 examines time-series estimation of action-reaction models, using India and Pakistan as an example. Section

[2] ‘Econometrics may be defined as the quantitative analysis of actual economic phenomena based on the concurrent development of theory and observation, related by the appropriate methods of inference’ Samuelson, Koopmans and Stone (1954).

5 examines Markov switching estimation of game-theory type models, using Greece and Turkey as an example. Section 6 examines cross-section models, section 7 panel models and section 8 concludes.

2. Data

Probably the most important, though often least discussed, modelling issue is the choice of the measure of arms. Brauer (2002) discusses the issues in the case of Greece and Turkey and reviews the various choices in that literature. In some cases the obvious measure is the number of a certain sort of weapons. As Intriligator and Brito (2000 p46) note that until the East West arms race of the Cold War, most arms races were naval: while armies were labour intensive navies were capital intensive. Before World War I, Britain had a policy of matching competitors in numbers of major warships, giving rise to a typical Richardson type reaction function. This is a quantitative (numbers of battleships) symmetric (both acquired the same weapons) arms race. Craft (2000) examines the effect of the Washington Naval Agreement of 1922 on naval expenditure, suggesting that it caused diversion of resources to more advanced and expensive systems. McGuire (1977) analyses another quantitative symmetric arms race, between the US and Soviet Union, using various measures of nuclear arsenals. Desai and Blake (1981) criticise his dynamic specification and McGuire responds. This debate anticipates many of the issues that are central to the subsequent literature.

A qualitative and asymmetric arms race was that between fortifications and siege trains in the late Mediaeval period, where there were evolutionary improvements in the technology of both, until the advent of gunpowder gave the besiegers an advantage the fortifiers could not counter. The current increasing returns to scale in military technology, emphasised by Intriligator and Brito (2000), leads to a virtual monopoly of military force by the US which gives its opponents incentives to choose asymmetric warfare and an asymmetric arms race. Estimating qualitative asymmetric arms-races is inevitably difficult and we will confine out attention to quantitative symmetric ones. This is a significant restriction, because what may be currently the most important arms race, between governments and terrorist opponents, is inherently qualitative and asymmetric. Data on preparations by terrorist groups, as distinct from actual attacks, are hard to come by. The 9/11 Commission estimated that the cost of preparing the attack by the four planes was between $400,000 and $500,000; tiny compared to typical military expenditure numbers. Dunne et al. (2005) examine asymmetric conflict, Enders in this volume discusses estimated models of the process by which terrorists substitute between different types of target in response to investments in defence.

For a country involved in an enduring hostility the focus must be on military capability: the probability of prevailing in a conflict. This will be a function of its levels of forces, measured by military capital (troops, stocks of weapons etc.), relative to those of its opponent. However, capability will also depend on how well those forces are used; a matter of strategy, tactics, training and leadership. Using a large sample of battles, Rotte and Schmidt (2003) show that relative force size is a poor predictor of victory. Instead much quantitative analysis suggests that the strategy adopted by each side is the most important determinant of victory in war, Reiter (1999).

Measuring military capability ex ante, before an actual conflict, is inherently problematic. Measuring forces is easier, but the long list of elements which go to make up a force structure cannot be well summarised by a single number. In particular, adding up spending on the different categories (personnel, R&D, procurement, etc.) may not be the appropriate form of aggregation. Dunne et al. (2004) suggest an aggregate of the form

$$K_i = N^{\nu}\left[\sum_{j=1}^{N}(u_{ij}m_{ij})^{\alpha}\right]^{1/\alpha}$$

where m_{ij} is the quantity of weapons system of type j that country i fields, u_{ij} the quality of the system and N is the total number of systems. Thus there are trade-offs between the quality, quantity and number of systems in the inventory. While useful theoretically, empirical implementation is difficult, though much military operational research does construct such force aggregates in the context of particular scenarios.

The level of forces reflects depreciated past stocks plus investment paid for by military expenditure. Payment for troops represents investment in human military capital. Again the conversion of military expenditure into effective forces is not a straightforward process, reflecting the efficiency of the arms production industries and personnel policies (e.g. the use of volunteers or conscripts). There are many cases where high military expenditures have not produced capable forces. Despite this, rather than using stocks of particular types of weapons, most empirical studies use measures based on military expenditures, either as levels, M_t, shares of GDP, M_t/Y_t, growth rates $(M_t - M_{t-1})/M_{t-1}$ or stocks calculated as $K_t = M_t + (1-\delta)K_{t-1}$ where the rate of depreciation, δ, and initial stock are either estimated or assumed. In principle one can test between the alternative measures, but this is not often done. Since the military expenditures of the two countries have to be in common units, there are issues about appropriate exchange rates and price indexes, and measures of the levels of military expenditure can be very sensitive to these choices. For instance, SIPRI (2005) estimate that Chinese military expenditure was $35 billion at market exchange rates and $161 billion at purchasing power parity exchange rates. This is not an issue for shares, since they are pure ratios, but shares may be sensitive to the measurement of GDP. Most researchers use data from either WMEAT or SIPRI and there are often substantial differences between them, Lebovic (1999). In addition, the figures are often revised, Brauer (2002, p90) notes the big differences between the military expenditure series used in two studies of the Greek-Turkish arms race although both were based on the same SIPRI data source. Murdoch in Chapter 3 of this volume discusses world military expenditures.

Governments determine the level of military expenditure by first making a strategic assessment of the threat and of the effectiveness of military spending in countering the threat. It then balances the strategic assessment against the opportunity costs of the military spending, given available national output. The outcome of that political-economy calculation is the choice of a share of output to devote to military spending. The military may get a smaller share because the threat is thought to be less; or because military expenditure appears less effective at meeting the threat than alternative measures such as confidence building initiatives; or because the opportunity costs appear greater. While the share of military expenditures in output is clearly a measure of priorities not capabilities, arms races should be reflected in priorities, thus shares may be an interesting measure in certain contexts.

The 1975 revision of CIA estimates of Soviet figures illustrates the issue. The CIA calculated Soviet military spending by first estimating the number of goods and services purchased - number of troops, tanks, ships, soldiers etc – from intelligence sources. It then estimated what these would have cost the USA to get a dollar figure. This was then multiplied by an estimated rouble/dollar exchange rate, to get a rouble figure, which could then be expressed as a share of CIA estimates of Soviet GNP. In 1975, the CIA decided that the Soviet military industry was much less efficient than previously thought and altered the exchange rate to reflect this, raising the estimated share of military expenditure from 6-8% to 11-13%. Although this did not change their estimate of Soviet forces or the dollar figure for Soviet military expenditure, the revision to the estimated Soviet share of military expenditure was widely interpreted in the US as indicating an increased Soviet threat. With hindsight, we know the Soviet economy was even less efficient than the CIA thought, and the actual share was probably well over 20%.

3. Theoretical issues in specification

The natural theoretical starting point for analysis of arms races is the Prisoner's dilemma or similar two-agent two-choice games and this is where Sandler and Hartley (1995 section 4.1) start. However, game theory has had very little influence on empirical work and Brauer (2002, p102) comments that 'the extant economics literature never addresses the disagreements and conflict from a game theoretic point of view'. Within the theoretical literature the arms-race game tends to be treated either as one stage of a more complex game, e.g. followed by decisions about war or preceded by negotiation (as in Baliga & Sjostrom (2004) who provide a review of the literature) or as a repeated game. Intriligator and Brito (2000) note that the existence of stocks means that arms races are not Markov processes, depending just on the state in the previous period. They suggest that the sufficient assumptions needed to transform this dynamic economic model into a repeated game are that the choice set is discrete, there are no income effects, and investment in weapons is reversible. We return to game theory models in section 5.

Most empirical work has started from the model of Richardson (1960). This suggested that for some continuous measure of military preparedness of country $i = 1, 2$ at time t, $m_i(t)$, the arms race dynamics could be described by a pair of differential equations in continuous time;

$$\frac{dm_1(t)}{dt} = a_1 + b_1 m_2(t) - c_1 m_1(t)$$

$$\frac{dm_2(t)}{dt} = a_2 + b_2 m_1(t) - c_2 m_2(t)$$

Richardson interpreted the a_i as exogenous 'grievance' terms, b_i as 'reaction' or defence terms and c_i as 'fatigue' terms. The system can be written in discrete time as:

$$\Delta m_{1t} = a_1 + b_1 m_{2t} + c_1 m_{1,t-1}$$

$$\Delta m_{2t} = a_2 + b_2 m_{1t} + c_2 m_{2,t-1}$$

We return to the econometrics of this in the next section. The Richardson system is descriptive, in that it does not have explicit behavioural foundations based on optimisation or other explicit forms of decision-making and it does not include a budget constraint. There are many ways of providing those foundations, Intrilligator (1975) is an early example. We will examine a model based on Levine and Smith (1995).

Consider two countries, $i, j = 1, 2$, each has a budget constraint in terms of national income, Y_{it}, consumption, C_{it}, investment in quantity of military goods and services, M_{it}, at price, P_{it}:

$$Y_{it} = C_{it} + P_{it} M_{it}.$$

The stock of arms accumulate according to

$$K_{it+1} = (1-\delta_i) K_{it} + M_{it}$$

where δ_i is the rate of depreciation. Welfare in any period, W_{it} is a CES function of security S_{it} and consumption,

$$W_{it} = \left[(1-a_i) C_{it}^{-\rho_i} + a_i S_{it}^{-\rho_i} \right]^{-1/\rho_i}$$

where $\sigma_i = 1/(1+\rho_i)$ is the elasticity of substitution between security and consumption. Security is given by the difference between the forces the other country would need to make a successful attack, $\beta_i K_{it} + \alpha_i$, and the actual level of the other country's forces:

$$S_{it} = \beta_i K_{it} + \alpha_i - K_{jt}.$$

Since welfare in each period is a function of stocks, we have a dynamic game, which requires an inter-temporal utility function;

$$U_i = \sum_{t=0}^{\infty} \mu_i^t W_{it}$$

with discount rate μ. Solving this set of equations yields the open-loop Nash equilibrium as functions of the known path for the exogenous variables, incomes and prices. In zero growth steady state the security consumption ratio is given by

$$\frac{S_i}{C_i} = \left\{ \frac{\beta_i a_i \mu_i}{(1-a_i)[1-\mu_i(1-\delta)]P_i} \right\}^{\sigma_i} = \phi_i$$

and the steady state equilibrium reaction functions by

$$K_i = (\beta_i + \phi_i P_i \delta_i)^{-1}(-\alpha_i + K_j + \phi_i Y_i)$$

This is a standard Richardson type long-run relation, making the stock held by one country a function of the stock held by the other; with the difference that price and income appear explicitly, and the long-run coefficients to be estimated are complicated functions of the underlying parameters and price.

This is a full information model, in practice we have to allow for expectations. To introduce the issues, consider the case where only one of the countries is responding, say country one, while country two sets their military expenditure autonomously. Suppose the loss function for country one is forward looking with discount factor μ and penalises both the difference between the two countries military expenditure (this can be easily generalised as above) and changes in its own military expenditure, because of adjustment costs:

$$L_t = \sum_{i=0}^{\infty} \mu^i \left[\frac{1}{2}(m_{1t+i} - m_{2t+i})^2 + \frac{\theta}{2} \Delta m_{1t+i}^2 \right].$$

The Euler equation takes the form:

$$\mu m_{1t+1} - (1+\mu+\theta^{-1}) m_{1t} + m_{1t-1} = -m_{2t}\theta^{-1}.$$

Solving the Euler equation requires finding the two roots, $\nu_1 < 1 < \nu_2$ which are the solution to the characteristic equation

$$\delta\nu^2 - (1+\delta+\theta^{-1})\nu + 1 = 0.$$

Calling the stable root (i.e. the one that is less than unity) ν , the optimal policy is then given by the partial adjustment equation, of the same form as the Richardson equation above:

$$\Delta m_{1t} = (1-\nu)(\widehat{m}_{2t} - m_{1t-1})$$

but with a forward looking target:

$$\widehat{m}_{2t} = (1-\mu\nu) \quad E_t\left[\sum_{i=0}^{\infty}(\mu\nu)^i m_{2t+i}\right].$$

If a time-series process is specified for m_{2t} an estimate of the expected value can be obtained in terms of observed values of m_{2t} and the parameters of the time-series process determining it. For instance, suppose

$$\Delta m_{2t} = g(1-\rho) + \rho\Delta m_{2t-1}$$

where g is the long run growth rate, then:

$$\Delta m_{1t} = \frac{\mu\nu(1-\nu)(1-\rho)g}{(1-\mu\nu)(1-\rho\mu\nu)} + \frac{1-\nu}{1-\rho\mu\nu}\Delta m_{2t} + (1-\nu)(m_{2t-1} - m_{1t-1}).$$

This is an Error Correction Model, ECM, of the form commonly estimated. Notice that it differs from the Richardson equation in including m_{2t-1} in addition to m_{1t-1}. The estimated coefficients are functions not merely of the preference and adjustment parameters μ and ν, but of the parameters of the expectations process g and ρ. There is then the familiar Lucas Critique problem, if the expected behaviour of country two changes, i.e. g or ρ change, then the parameters of country one's reaction function will change.

Given that strategic circumstances do show quite marked changes, this might make us sceptical of the likely structural stability of arms race models. Murdoch (1995) discusses the effect of the move from the doctrine of massive retaliation to flexible response on the parameters of NATO alliance models. Smith (1989) estimates structurally stable models of the UK and French shares of military expenditure as functions of US and Soviet shares, 1951-1987; but the equations, like the Soviet Union, collapsed a year after publication. Another form of structural instability could result from substitution between the weapons being raced. For instance, if a country is losing the conventional forces race, it may switch to building nuclear weapons or supporting terrorism. Such substitutions may not show up particularly well in aggregate military expenditure data.

The leader-follower model above is particularly simple. If both countries react and have rational expectations the problem becomes more complex, because there may not be a unique solution. There may be an infinite number of rational expectations equilibria, in each of which both countries expectations prove self-fullfilling, all associated with different growth rates of military expenditure. Furthermore there may be transient 'bubbles', like those in markets for financial assets: periods when countries compete explosively driven by the expectation that the others are arming; expectations which are confirmed by events. The form of the dynamics will reflect both expectations and adjustment lags and the theory gives little guidance on this, leaving it to be largely an empirical choice.

The models above are linear, primarily for mathematical convenience and theory gives relatively little guidance about the appropriate functional forms. Non-linear reaction functions may intersect more than once, giving multiple equilibria and zones of stability and instability, Brito and Intrilligator (1995). One may then observe, jumps from one equilibrium to another, which the Richardson type models do not catch and provide a role for arms control in stabilising at a low equilibrium.

Probably the central question raised by this very brief theoretical review is the likely structural stability of the underlying parameters. Some may be very stable deep parameters, e.g. the elasticity of substitution between security and consumption or the discount rate, but many others may not be stable. Changes in military technology and tactical improvements will change the rate of depreciation, δ_i, and the parameters determining the forces country j needs to successfully attack country i: $\alpha_i + \beta_i K_i$. Changes in the strategic environment will change the country seen as a threat. Similar concerns arise with the parameters of the adjustment and expectations processes. As we shall see, structural instability is a pervasive feature of the econometric estimates.

4. Action-reaction models

The 'structural form' of the Richardson system introduced in the last section written in discrete time with the addition of a stochastic error term is:

$$\Delta m_{1t} = a_1 + b_1 m_{2t} + c_1 m_{1,t-1} + \varepsilon_{1t}$$

$$\Delta m_{2t} = a_2 + b_2 m_{1t} + c_2 m_{2,t-1} + \varepsilon_{2t}$$

where it is usually assumed that:

$$E(\varepsilon_{it}) = 0;\ E(\varepsilon_{it}{}^2) = \sigma_{ii};\ E(\varepsilon_{it}\varepsilon_{jt}) = \sigma_{ij};\ E(\varepsilon_{it}\varepsilon_{j,t-s}) = 0, s \neq 0;\ i, j = 1, 2.$$

These structural shocks, ε_{it}, will be driven partly by idiosyncratic factors (e.g. events in the Balkans for Greece and the conflict with the Kurds for Turkey) and partly by common factors (events in Cyprus or NATO for both). So, we would not expect the structural shocks to be independent of each other. They may also not be independent of the regressors and it is possible that systems like this may indicate an arms race between the two countries, when in fact the military expenditures of both are responding to common unobserved shocks.

This system is just-identified and can be estimated by, for instance, two stage least squares. However, the plausibility of the just identifying restrictions ($m_{1,t-1}$ does not appear in the equation for Δm_{2t} and vice versa) may be questioned. For instance, the restriction does not hold in the ECM model derived above, which allowed for expectations. The interpretation of the coefficients is not straightforward. For instance, Richardson interpreted the c_i as fatigue terms, economists tend to interpret them as adjustment coefficients, political scientists as measures of bureaucratic inertia. We will follow the economic convention and write the model as:

$$\Delta m_{1t} = \lambda_1(\alpha_1 + \beta_1 m_{2t} - m_{1t-1}) + \varepsilon_{1t}$$

$$\Delta m_{2t} = \lambda_2(\alpha_2 + \beta_2 m_{2t} - m_{2t-1}) + \varepsilon_{2t}$$

The model is stable as long as the adjustment parameters satisfy $-1 < \lambda_i < 1$, in which case there are two long-run reaction functions, $m_i = \alpha_i + \beta_i m_j$, with static long run equilibria $m_i = (1 - \beta_i \beta_j)^{-1}(\alpha_i + \beta_i \alpha_j)$.

This form is structural in that current values of the endogenous variables appear on the right hand side of the equation. Sets of equations with this general form can be derived from a variety of different theories. The structural coefficients will be functions of the underlying parameters of the system. For instance in the theoretical model above, they are functions of prices and income, the depreciation rates for weapons, strategic parameters which describe the nature of conflict, discount rates, and the elasticity of substitution between security and consumption. Even if one can estimate the structural coefficients consistently, this may not allow one to recover the underlying parameters implied by the theory.

The reduced form of the system, in terms of predetermined, lagged, variables can be written in the VECM (Vector Error Correction Model) form of a first order VAR (Vector Autoregression) as:

$$\Delta m_{1t} = \frac{a_1 + b_1 a_2}{1 - b_1 b_2} + \frac{c_1 + b_1 b_2}{1 - b_1 b_2} m_{1t-1} + \frac{b_1(1 + c_2)}{1 - b_1 b_2} m_{2t-1} + \frac{\varepsilon_{1t} + b_1 \varepsilon_{2t}}{1 - b_1 b_2}$$

$$\Delta m_{2t} = \frac{a_2 + b_2 a_1}{1 - b_1 b_2} + \frac{c_2 + b_1 b_2}{1 - b_1 b_2} m_{1t-1} + \frac{b_2(1 + c_1)}{1 - b_1 b_2} m_{2t-1} + \frac{b_2 \varepsilon_{1t} + \varepsilon_{2t}}{1 - b_1 b_2}$$

or

$$\Delta m_{1t} = \pi_{10} + \pi_{11} m_{1t-1} + \pi_{12} m_{2t-1} + u_{1t}$$

$$\Delta m_{2t} = \pi_{20} + \pi_{21} m_{1t-1} + \pi_{22} m_{2t-1} + u_{2t}$$

with $E(u_{it}) = 0$, $E(u_{it}{}^2) = \omega_{ii}$, $E(u_{it} u_{jt}) = \omega_{ij}$ $E(u_{it} u_{j,t-s}) = 0, s \neq 0$, $i, j = 1, 2$. Each equation of the reduced form can be estimated consistently by least squares. There is Granger causality from m_1 to m_2 if $\pi_{21} \neq 0$ and from m_2 to m_1 if $\pi_{12} \neq 0$, however Granger causality rarely corresponds to economic causality. For instance, suppose $b_1 = 0$, country one is not arms-racing; m_1 will still Granger cause (help predict) m_2, as long as $c_2 \neq 0$. In addition, expectations mean that in economics effects can precede causes. Unless one can identify the model, the reduced form is uninformative about contemporaneous interactions.

If the variables are I(1)[3], or equivalently contain a stochastic trend, then there is a danger of spurious regression. In a regression of one I(1) variable on another, the R^2 tends to unity with the sample size and the t ratio to a non zero value, even if the two series are unrelated. The requirement for the regression not to be spurious is that the two variables cointegrate. If this is the case then the process can be represented as a restricted form of the VECM above. If a long run relationship of the form $m_{1t} = \beta m_{2t}$ exists, then the disequilibrium or error correction term is measured by $z_t = m_{1t} - \beta m_{2t}$, which will be I(0). The VECM then takes the form:

$$\Delta \mathrm{m}_{1t} = \delta_{10} + \alpha_1 \mathrm{z}_{t-1} + \mathrm{v}_{1t}$$

$$\Delta \mathrm{m}_{2t} = \delta_{20} + \alpha_2 \mathrm{z}_{t-1} + \mathrm{v}_{2t}$$

where the feedbacks are stabilising if $\alpha_1 \leq 0, \alpha_2 \geq 0$, and at least one is non-zero. In the I(0) case, discussed above, there were two long-run reaction functions and both military expenditures tended to constants. In the I(1) case, both military expenditures are driven by a single stochastic trend, so they can behave like random walks and go anywhere, and there is a single long-run relation, the cointegrating vector, so they do not diverge too much. Estimation and testing of the cointegrating vectors can be done in a number of ways, including within the maximum likelihood framework suggested by Johansen (1988).

[3] A variable is said to be I(d), integrated of order d, if it must be differenced d times to become covariance stationary. A variable is said to be covariance stationary if its expected value, variances and auto-covariances are all constant, perhaps after the removal of a deterministic trend. An I(0) variable is thus stationary. If there is a linear combination of two I(1) variables which is I(0), the two variables are said to cointegrate.

Unit root tests, VARs and cointegration have been widely adopted in the defence economics literature and in the study of arms races. However, there are a number of problems with the techniques. Both the tests for unit roots (used to determine the order of integration) and the tests for cointegration tend to have low power, so determining the order of integration and cointegration is not straightforward. The tests are also sensitive to the choice of lag order, the treatment of the deterministic elements, the presence of structural breaks and various other factors. There are also questions of interpretation, since the order of integration is not a structural property of the series but a description of the time-series properties of a sample. Series which appear I(1) on short spans of data often appear I(0) on long spans, where span refers to the length in time of the series not the number of observations. Over centuries of data, the UK share of military expenditure is clearly I(0), over shorter spans it appears to be I(1). While cointegration allows us to estimate the long-run equilibrium, it does not help in identifying the short-run structural interaction.

As noted above, the Richardson model lacks a budget constraint, but this can be dealt with, as in the model of section 3 by including GDP. Care needs to be taken in including extra variables within a VAR. The number of parameters grows very rapidly with the number of variables and large VARs can have very bad small sample properties. If there are m variables in the system and r cointegrating vectors, there are $m-r$ stochastic trends. When there are r cointegrating vectors r just identifying restrictions are required on each cointegrating vector to interpret them. When $r=1$, the restriction is just a normalisation, choosing the dependent variable; when $r>1$, it may be difficult to find just identifying restrictions that allow the results to be interpreted.

Many papers have estimated VARs and VECMs for the Greek-Turkish arms race and the evidence for an action-reaction type arms race is, at best, mixed. Dunne et al. (2003) describe an unsuccessful specification search. However, there is more evidence for an arms race between India and Pakistan. Deger and Sen (1990) obtained well defined results and Ocal (2003) using data 1949-1999 and smooth transition regression, which allows for non-linearities in the responses of growth rates in military expenditure, also finds well defined feedback.

Dunne et al. (2003) found that for SIPRI data, 1960-1996, various tests suggested a second order VAR in Indian and Pakistan real military expenditures (rather than logarithms or shares) and that GDP was Granger non-causal for military expenditures. Working with a VECM in the two military expenditure series, using unrestricted intercept and no trend, they found one cointegrating vector by both trace and eigenvalue tests, irrespective of the treatment of the deterministic elements. This was: $Z_t = I_t - 2.0P_t$. Indian military expenditure tends to a long-run relationship where it is twice Pakistan military expenditure (plus a constant). The VECM showed significant feedbacks in both directions and interesting dynamics. The equations passed a variety of specification tests and in particular Cusum and Cusum-squared tests indicated that the equations were structurally stable. The data were updated to 1960-2003 using revised SIPRI data. Re-estimating on 1962-1996, again gave one cointegrating vector, though it was slightly different: $Z_t = I_t - 2.51P_t$, perhaps as a consequence of rebasing and revision. The estimated VECM coefficients and the fit were very similar.

When the data were extended to 1962-2003, there was much less evidence for cointegration. With no intercept or trend, a very unrealistic assumption, both tests indicated two cointegrating vectors. The trace test for restricted intercept no trend indicated one cointegrating vector, all the other tests indicated no cointegrating vectors. Both Schwartz and Akaike criteria indicated that unrestricted intercept and no trend was still the appropriate assumption. The extra seven years did not make a large difference to the estimated cointegration coefficient, but increased its standard error. The coefficients (standard errors) were –2.51 (0.09) for 1962-1996 and –3.04 (0.30) for 1962-2002. However the dynamics changed substantially. The feedback coefficient on the error correction term in the Indian equation went from being large and significant 1962-1996 to being small and insignificant 1962-2003. None of the variables in the Indian equation are significant and it looks like a random walk. The feedback coefficient in the Pakistan equation went from 0.11 (0.04) to 0.05 (0.02) remaining significant. For the Indian equation, the Chow forecast accuracy test (p=0.02) and the parameter equality test (p=0.009) rejected structural stability after 1996. The equation estimated using data up to 1996 substantially under-predicted subsequent Indian military expenditure. The Pakistan equation also under-predicted subsequent expenditure, but not by as much, and the forecast accuracy test did not reject structural stability (p=0.0.08), though the parameter equality test did (p=0.03).

The model was re-estimated using the logarithms of real military expenditure, LMI and LMP. Likelihood criteria which had indicated the linear was better for both countries for 1962-1996 now indicated that logarithmic was better for India and linear for Pakistan. A VAR2 with unrestricted intercept and restricted trend, the preferred specification, showed clear evidence of cointegration over 1962-2003, in fact both trace and eigenvalue tests indicated two cointegrating vectors, which would indicate that both variables are I(0). While the levels of military expenditure are clearly I(1), unit root tests suggest that the logarithm of Indian military expenditure is I(0) and Pakistan I(1). This is an example of conflict between tests.

Choosing one cointegrating vector gives estimates (standard errors)

$$Z_t = LMI_t - \underset{(0.099)}{0.309} LMP_t - \underset{(0.005)}{0.028}\, t$$

with strong and significant feedbacks for both countries, which were quite similar to the earlier linear results. The VECM estimates, (standard errors) for 1962-2003 are:

$$\Delta LMI_t = \underset{(0.69)}{3.28} + \underset{(0.13)}{0.60}\, \Delta LMI_{t-1} - \underset{(0.14)}{0.04}\, \Delta LMP_{t-1} - \underset{(0.12)}{0.57}\, Z_{t-1}$$

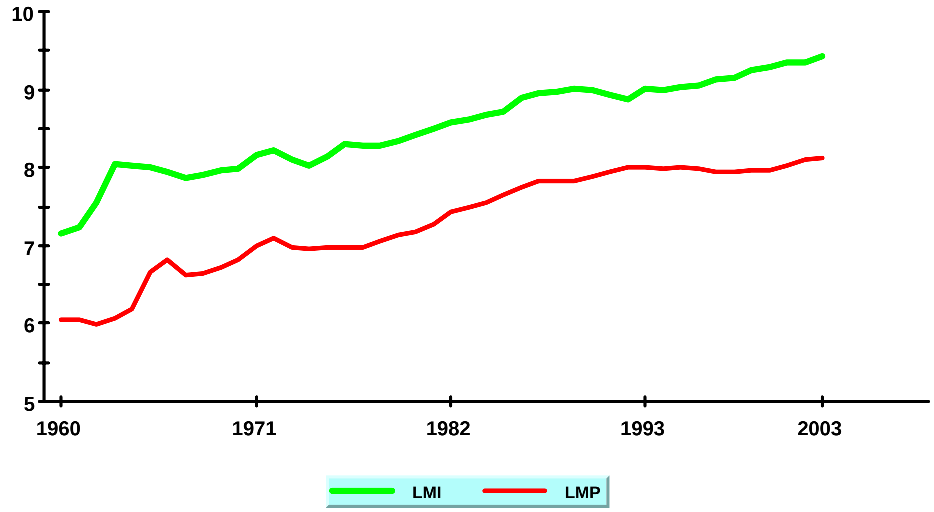
$R^2 = 0.47$; SER=0.079

$$\Delta LMP_t = \underset{(0.74)}{-2.29} - \underset{(0.14)}{0.10}\, \Delta LMI_{t-1} + \underset{(0.14)}{0.11}\, \Delta LMP_{t-1} + \underset{(0.13)}{0.41}\, Z_{t-1}$$

$R^2 = 0.27$; SER=0.086

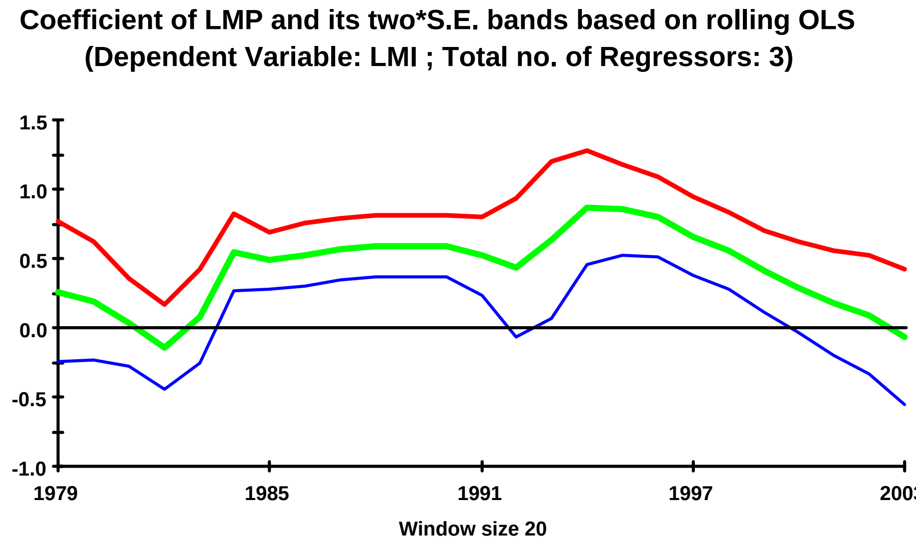
Unlike the linear equation estimated up to 1996, both equations failed a large number of misspecification tests, including Cusum and Cusum-squared tests for structural stability. There is marked serial correlation in the Pakistani equation, which can be removed by increasing the lag length to 3, but this then causes serial correlation in the Indian equation. Plots of the logarithms of the series are shown in Figure 1. The greater variance at the beginning of the sample, leads to the failure on heteroscedasticity. The linear model may have been preferred over the shorter period because it stabilised the variance of the residuals.

Figure 1. India and Pakistan log military expenditure.



To investigate how stable the parameters were, we ran a simple static rolling regression of log Indian military expenditure on log Pakistani military expenditure and a trend, using a window size of 20. Figure 2 clearly shows a changing elasticity of Indian to Pakistani military expenditure. Using the 20 years ending in 2003, gives a coefficient almost exactly zero. In 2004, Indian military expenditure sharply increased, while Pakistan expenditure did not, so the relationship is likely to get worse.

Figure 2, time varying elasticity and standard errors.



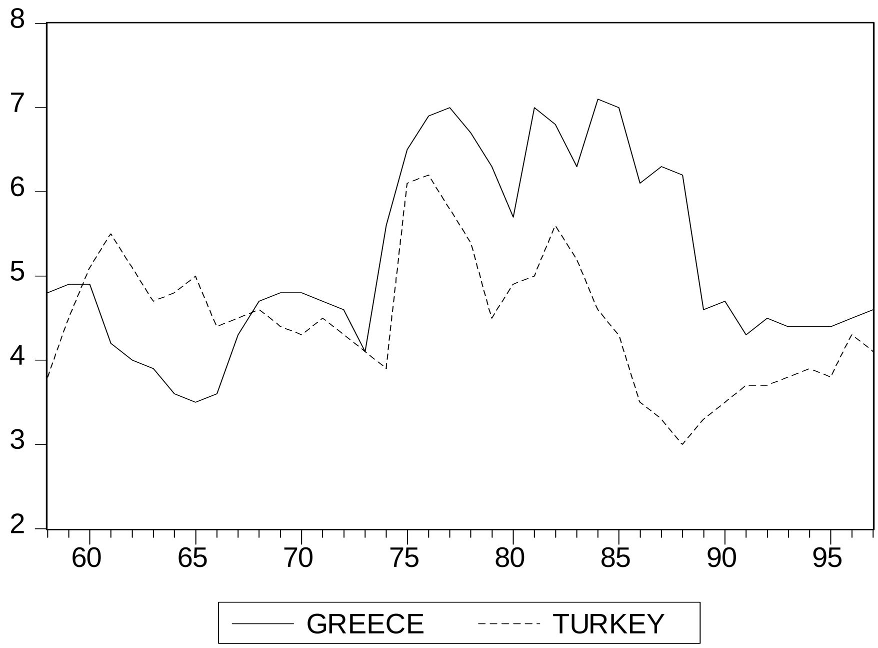
The VECM approach described above is standard (though not unquestioned) in time-series methodology and flows naturally from the Richardson model. However, our impression is that it does not work well on arms race data for a variety of reasons. The most important reason seems likely to be that structural stability may not be an appropriate assumption. The India-Pakistan example may be unusual in that their strategic relationship has been bureaucratised, perhaps because of a shared decision rules acquired from a their common colonial heritage. But even here the evidence for a structurally stable arms race is mixed. There is also the danger that India may be responding to China and the apparent arms race is produced by common shocks as noted above. The quality of Chinese military expenditure data is such that it would be difficult to test this.

5. Game Theory Models

As noted above, the natural theoretical starting point for analysis of arms races is the Prisoner's dilemma, and similar two-agent two-choice games. However, game theory has had very little influence on empirical work because it is quite difficult to translate the game theory models into statistical specifications. One paper which tries to do this is Smith Sola and Spagnolo (2000) which analyses the Greek Turkish arms-race. The paper models each country choosing either high or low military spending, with payoffs which depend on the choice of the other. This gives four possible states, with associated payoffs. Given the payoffs, the countries choose a strategy. This is a repeated or iterated game, played every year. One can imagine that during the annual budgetary cycle each country chooses its strategy, high or low military spending for the next year, knowing what its opponent chose for this year, but not what the opponent will do next year. The strategy they choose for the next year will be conditional on the state this year. A familiar example of a conditional strategy is Tit for Tat: do in period $t+1$ what your opponent did in t. This is a pure strategy, specifying either high or low for the next period conditional on the state in this period. Countries could also follow mixed strategies, choosing high with some probability p and low with some probability $(1-p)$. Mixed strategies are optimal for quite a wide range of games. The paper assumes countries play conditional mixed strategies, choosing a probability of being high or low next year depending on the current state, and tries to estimate these strategies to see what light they shed on the interaction between Greek and Turkish military expenditures. This is the reverse of the usual approach in game theory, which specifies the payoffs and then determine the optimal strategy. Here the payoffs are not observed; but the strategies the players adopt are observed: whether they choose high or low. This allows inference about the nature of the game.

This game approach differs from the Richardson approach in a number of ways. Firstly, it naturally handles non-linearities and structural change - jumps from high to low - which the Richardson approach does not. Secondly, all the other factors, which are treated as deterministic influences in the regression approach are treated stochastically: reflected in the conditional probabilities of choosing high or low. Thirdly, military expenditures are now discrete, taking two values high and low, rather than continuous as in the Richardson models. Obviously, this approach can only be applied where it is sensible to treat the outcomes as dichotomous. Therefore, it could not be applied to military expenditures which are trended upwards and there is no natural classification into high and low. However, it can be applied to shares of military expenditure or growth rates. It is clear that there have been marked changes both in the level of the Greek and Turkish shares and the relationship between them. As Fig 1 shows, to a first approximation the series appear to be well described by variations around distinct high and low levels, so modelling them in terms of a simple high-low choice plus some random errors may not be too unrealistic.

Fig 3. Greek and Turkish shares of military expenditure in GDP, percent.



Estimation of the strategies in a simple two by two game in which each side choses high or low can be done using the bivariate Hamilton (1989) discrete state switching model. There are four states or regimes: (1) both high, (2) Greece high, Turkey low, (3) Greece low, Turkey high and (4) both low. The parameters to be estimated are those for the two states, high and low, and the transition matrix. The state parameters are the high and low mean for each country, the variance for each country and the covariance between the shocks for each country, 7 in total. The evolution of the series is described by a four by four transition matrix, Π, the elements of which, π_{ij} give the probability of moving from state i in period t to state j in period $t+1$, $i, j = 1,2,3,4$. So π_{11} gives the probability of staying in state 1, both countries having high shares of military expenditure in the next period, given they both have high shares this period. Because the system must move to one of the four states in the next period, each of the four columns of the transition matrix sums to unity. Thus the unrestricted transition matrix has 12 free probabilities, which together with the 7 state parameters gives 19 free parameters in total. The parameter estimates are obtained by maximising the likelihood function. On the basis of the estimates, one can calculate the state probabilities for each year. This gives the probability of being in each of the four states, given the history of the process, for instance $P(s_t = 1)$ gives the probability of both countries being high in a particular year. If the four state model is a good description of the process, these probabilities should be close to one or zero, indicating a clear separation between the four regimes.

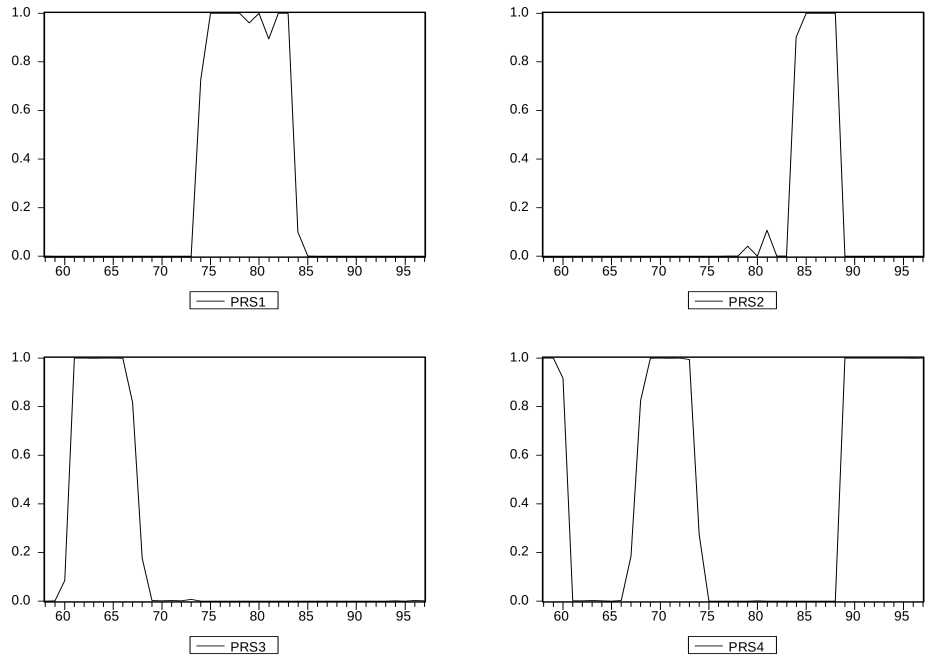
The unrestricted case allows unlimited dependence in the decisions of the two countries. The paper also considers three restricted transition matrices. The first assumes that each country determines its share of military expenditure independently. If two events A and B are independent, the probability of both happening is the product of the probabilities of each of them happening: $P(A \cap B) = P(A)P(B)$. Similarly, the probability of staying in state 1, both high, is just the product of the probability of Greece staying in the high state and the probability of Turkey staying in the high state: $\pi_{11} = \pi_{GH}\pi_{TH}$. With independence, there are only four free probabilities: that of Greece staying in a high state, π_{GH}; Turkey staying in the high state, π_{TH}; Greece staying in a low state, π_{GL} and Turkey staying in the low state, π_{TL}.

The second and third restricted versions assume that one or other country plays tit-for-tat and follows the leader. That is that Greece (Turkey) is always in the state that Turkey (Greece) was one period ago. In this case there are only two free probabilities, that of the leader staying in the high state or of the leader staying in the low state. Suppose Greece leads and Turkey plays tit-for-tat, then Turkey will always be in the state Greece was one period ago. There are then only two free probabilities, that represent Greek strategy, π_{GH} and π_{GL}. If Greece is high this period, Turkey will be high next period, so any state with Turkey low has zero probability. What drives the system is the probabilities of Greece staying in the high or low states. Similarly if Turkey leads Greece.

This technique was applied to SIPRI annual data on Greek and Turkish shares of military expenditure 1958-1997. First, the unrestricted Markov structure is estimated, which allows for twelve free probabilities in the transition matrix. The Greece Low state has a mean military expenditure of 4.37% of GDP, the Greece High state has a mean 2.16 percentage points of GDP higher, and this shift is very significant, with a t ratio over 15. The Turkey Low state has a mean share of military expenditure of 3.86% with the Turkey High state 1.44 percentage points higher, again the difference between the high and low means is very significant. Turkey shows a higher variance around its means than does Greece: the standard deviations of the shocks are 0.43 for Greece and 0.55 for Turkey. Thus the variations in the share of military expenditure around the means are about a half a percent of GDP. There is a positive covariance between the shocks, that is on the margin of significance at the 5% level.

The probabilities of being in each of the four regimes from the unrestricted, general, model are shown in Figure 2. They show a very clear separation into regimes, the probabilities are close to zero or one, and match the impression of the time series. The series starts with both being low (state 4); switches in 1961 (about the time of a Turkish military coup) to Greece low, Turkey high (state 3); stays in that state till 1967 (the year of the Greek military coup), when it switches back to both low (state 4) until 1973. With 1974 (the invasion of Cyprus) it switches to both high (state 1) until 1983 (the end of a period of military rule in Turkey). From 1984-1988 Greece is High, Turkey Low (state 2). From 1989 (the end of the Cold War) to 1997 shares of military expenditure are in state 4 (both low). Although there is some correspondence, the division into states does not match up neatly with military rule: 1967 to 1974 for Greece and 1960-61, 1971-73 and 1980-1983 for Turkey.

Fig 4. State Probabilities for the four regimes.



The first restriction tested is that the probability of switching between states in each country is independent. Each country has its own probabilities of switching, which does not depend on the state the other country is in. This reduces the number of free parameters in the transition matrix from 12 to 4, two for each country. This would be the case if each was responding to its own domestic political economy or strategic concerns, which determined the probability of staying in the high state or the low state, without regard for whether the other was in a high or low state. Imposing independence reduces the maximised log-likelihood from 15.2 to 10.6, which is not a significant reduction at any conventional level. The Likelihood ratio test would give LR=9.2 whereas the 5% critical value for $\chi^2(8)$ is 15.51. The estimates of the coefficients and the assignment of the sample into the four states are hardly changed by the restriction of independence. Furthermore, the estimates of the transition probablities appear to be persistent and highly significant with values respectively given by $\pi_{GH} = 0.96,\ \pi_{GL} = 0.95,\ \pi_{TH} = 0.86,\ \pi_{TL} = 0.93$. If the two countries get into state 4 (low-low), there is a high probability of them staying there: $\pi_{44} = \pi_{TL}\pi_{GL} = 0.95 \times 0.93 = 0.88.$ Independence of the two countries transition probabilities does not imply that the two series are independent or that they cannot change regime simultaneously. Because of the positive covariance through the error terms: shocks (e.g. Cyprus) hit them both in the same way.

The second and third restrictions tested were that one country determined its state independently and the other played tit-for-tat, playing what the leader had played in the previous period. The Greece leads Turkey or Turkey leads Greece restrictions lead to a large reduction in the maximised log likelihood. Using Likelihood Ratio tests both restrictions are rejected at the 1% level.

The simplicity of the model allows us to test sharp hypotheses suggested by game theory, which is difficult to do in a regression framework. The independence restriction can be interpreted either as the switches are determined by other factors than the Greece-Turkey antagonism or that each country plays a mixed strategy that does not condition on the other side's behaviour. Neither interpretation provides much support for a traditional action-reaction type arms-race in which military expenditure is a response to an external threat represented by the other countries military spending. Instead the estimates of the transition probabilities are more consistent with an internal explanation in which political or bureaucratic inertia mean that once either country gets into a particular state, high or low, there is a high probability that it will stay there.

It is also interesting that a very simple theoretical model translates into quite a complicated empirical model. To apply the simple unrestricted two by two game to data on observed choices requires 19 free parameters. However, these parameters can be related to game theory strategies in a straightforward way, which is not the case with regression based arms race models. Theories about the possible strategies followed by the players, such as independence or tit-for-tat, can reduce the parameter space substantially. Such models, while useful in some cases, will only work where the observations can be divided into high and low, as in this case, and tend not to predict very well. Although it was not an issue in this case, the likelihood function in these models can be badly behaved, giving numerical problems in estimation, difficulty in obtaining convergence and sensitivity to choice of initial conditions.

6. Cross-section Models

Cross-section models of the demand for military expenditure, particularly for developing countries have been common for some time. These include neighbours military expenditure as determinants as well as other economic and political variables. Rosh (1988) introduced the concept of a “Security Web”, defined by neighbours and other countries (such as regional powers) that can affect a nation’s security. Rosh calculates the degree of militarisation of a nation’s Security Web by averaging the military burdens of those countries in the web, finding it to have a significant positive effect on a country’s military burden. More generally spillover effects have been attracting increasing attention, e.g Murdoch and Sandler (2002, 2004).

To illustrate the issues suppose that we have a cross-section of countries $i = 1, 2, .., N$, with the average value for some military measure m_i over some time-period and other determinants x_i. If one particular country i feels threatened by an alliance of j and k; the equations for these 3 (out of N) observations take the form

$$m_i = \alpha + \beta^e (m_j + m_k) + \gamma' x_i + \varepsilon_i$$

$$m_j = \alpha + \beta^e m_i + \beta^a m_k + \gamma' x_j + \varepsilon_j$$

$$m_k = \alpha + \beta^e m_i + \beta^a m_j + \gamma' x_k + \varepsilon_k$$

where β^e measures the arms race effect from an enemy and β^a the spillover effect from an ally. Given data on N countries and knowledge of threats and alliances, this model could be estimated by say OLS.

There are obvious practical problems; how to determine the pattern of threats and alliances, what time period to average over, which military measure to use, how to aggregate if adding allies expenditure is not appropriate, etc. There are also econometric issues. Having m_i appear as a dependent variable in one equation and as an independent variable in another is not a standard simultaneity issue, it is more like a lagged dependent variable issue and the conditions for consistent estimation by OLS are the same as that case: m_i can be treated as predetermined in the equation for m_j if $E(\varepsilon_i \varepsilon_i) = 0$. If there are correlated shocks to different countries or spatial serial correlation, this assumption will not hold.

Collier and Hoeffler (2004) use a system like this to estimate arms race multipliers. Consider a two country case

$$m_i = a_i + bm_j$$

an exogenous unit increase in a_i would cause an increase in the other countries military expenditure by bm_j feeding back on m_i and giving a total effect of $1/(1-b)$ on spending by country i. They distinguish own expenditure and neighbour expenditure arms race multipliers.

Using cross-section data, rather than time-series for individual countries increases the sample size considerably and allows one to measure the effect of variables which tend not to vary very much within countries. The cost of this benefit is the assumption that the arms race and alliance parameters are the same across all countries. This assumption can be weakened. Suppose we allow coefficients to differ across countries collect the independent variables in a vector z_i and write the model

$$m_i = \alpha_i + \beta_i' z_i + \varepsilon_i$$

$\alpha_i = \alpha + \nu_i$, $\beta_i = \beta + \eta_i$ where $\alpha = E(\alpha_i)$, $\beta = E(\beta_i)$ and η_i and ν_i are random, independent of the regressors, then OLS will provide consistent estimates of the average effect β. These assumptions are quite strong and with the availability of panel data models of the form

$$m_{it} = \alpha_i + \beta_i' z_{it} + \varepsilon_{it}$$

can be estimated and the homogeneity or independence assumptions tested.

Dunne and Perlo Freeman (2003a) estimate cross-section demand functions for developing countries using average data for Cold War (1981-88) and post Cold War (1990-97) periods. The dependent variable was the log of the share of military expenditure, significant explanatory variables were log population; log of the sum of its potential enemies military expenditure; log of the sum of the military expenditures of countries in its security web (including potential enemies); a democracy measure; civil war; external war and region dummies. There was little evidence of a change in the underlying cross-section relationship with the end of the Cold War, though as we shall see this is not a robust conclusion.

7. Panel Models

With panel data, one can employ a larger sample and allow for heterogeneity in the responses of different countries. Treatment and interpretation of heterogeneity in panels is a central issue. To focus on this, we will ignore the feedback issues discussed in section 4, to concentrate on a single equation. When one allows for panel VARS and VECMs, the situation becomes more complicated, and is discussed in Fuertes and Smith (2005). Smith and Tasiran (2005) compare alternative panel estimators in an application to the demand for arms imports. Given the number of different ways to treat heterogeneity and between group dependence, there are a large number of different panel estimators.

Suppose we have data on m_{it}, military expenditure, for a sample of countries $i = 1, 2, ..., N$ and years $t = 1, 2, ..., T$ and data on a vector of k exogenous variables x_{it}, which does not include unity for an intercept, but which includes other countries military expenditure. Suppose that we are interested in measuring the linear effect of x_{it} on m_{it}. This is a static model, in that x_{it} does not include lagged dependent variables; we discuss dynamics below.

A central issue in the choice of estimator is the relative size of N and T. The traditional panel literature deals with cases where N is large and T small, maybe only two or three time periods. Asymptotic analysis is done letting $N \to \infty$. The time-series literature deals with the case where T is large and N small and asymptotics let $T \to \infty$. Recently there has been interest in panel time-series where N and T are of the same orders of magnitude and asymptotics let both $N \to \infty$ and $T \to \infty$ in some way. What estimators are appropriate in the three cases differs. Define the country and overall means as

$$\overline{m}_i = T^{-1}\sum_{t=1}^{T} m_{it}; \quad \overline{m} = (NT)^{-1}\sum_{i=1}^{N}\sum_{t=1}^{T} m_{it}.$$

The total variation in the dependent variable is the sum of the within country variation and the between country variation

$$\sum_i\sum_t (m_{it} - \overline{m})^2 = \sum_i\sum_t (m_{it} - \overline{m}_i)^2 + T\sum_i (\overline{m}_i - \overline{m})^2,$$

similarly for the regressors. The main panel estimators differ in how they treat the within and between variation.

Pooled OLS gives the within and between variation equal weight, and uses least squares on:

$$m_{it} = \alpha + \beta' x_{it} + u_{it};$$

where β is a $k \times 1$ vector of slope coefficients. The within estimator (also known as the one way fixed effects, least squares dummy variables and a variety of other names) allows intercepts to differ across countries but constrains the slopes to be the same

$$m_{it} = \alpha_i + \beta' x_{it} + u_{it}$$

this only uses the within variation and is equivalent to OLS on

$$(y_{it} - \overline{y}_i) = \beta'(x_{it} - \overline{x}_i) + u_{it};$$

since $\alpha_i = \overline{y}_i - \beta'\overline{x}_i$. The between or cross-section estimator, only uses the cross-section variation in the country means:

$$\overline{m}_i = \alpha + \beta'\overline{x}_i + \overline{u}_i;$$

The two way fixed effects estimator constrains slopes to be the same, but allows intercepts to vary freely both over country and year:

$$m_{it} = \alpha_i + \alpha_t + \beta' x_{it} + u_{it}$$

allowing for a completely flexible trend, or unobserved common factor, which impacts each country by the same amount. The Swamy (1970) random coefficient model (RCM) allows all the parameters to differ over countries:

$$m_{it} = \alpha_i + \beta_i' x_{it} + u_{it}$$

and calculates weighted averages of the individual time-series estimates $\hat{\beta}_i$:

$$\widehat{\beta}_R = \sum_i W_i \widehat{\beta}_i,$$

the weights, W_i being based on the variances of the $\hat{\beta}_i$. Although we do not distinguish them in our notation: the β in the within equation is a different parameter from the β in the between equation and the cross-section may be measuring something quite different from the time-series.

There are a number of other estimators. The Zellner SURE estimator allows for non-zero between group covariances and possibly heterogeneous slopes. This is only feasible when T>N, and only appropriate when the factors causing between group dependence are not correlated with the regressors. There are other estimators which can deal with between group dependence when N>T, including the correlated common effects estimator which includes the means of the variables for each period. The TSCS (time-series cross-section) estimator allows for between group heteroskedasticity and within group serial correlation. This estimator is rarely used in economics, because it treats the symptoms of misspecification (heteroskedasticity and serial correlation) rather than their likely causes (parameter heterogeneity and dynamic misspecification), though it is popular elsewhere. The random effects estimator assumes that slopes are identical and intercepts are randomly distributed independently of the regressors. It calculates the optimal combination of within and between variation under these assumptions. There are a variety of tests to choose between the various estimators.

In static models, if the coefficients, α_i and β_i are randomly distributed, independently of the x_{it}, all these estimators will produce unbiased estimators of the expected values of the coefficients $E(\alpha_i)$ and $E(\beta_i)$. However, the independence assumption may not hold and the cross-section (between country) effects can be very different from the time-series (within country) effects. The between estimate β_B and the within estimate β_W will differ if the α_i are correlated with the $\overline{x}_i$ and the two estimates may even be of opposite sign.

A further issue arises with dynamic models since the within (fixed effect) estimator of

$$m_{it} = \alpha_i + \beta' x_{it} + \lambda m_{it-1} + u_{it}$$

is consistent for large T, but is not consistent for fixed T, large N. In this case the coefficient of the lagged dependent variable is biased downwards. This is the standard small T bias of the OLS estimator in models with lagged dependent variables. There are a variety of instrumental variable estimators for this case. However, if the true model is heterogeneous

$$m_{it} = \alpha + \beta_i' x_{it} + \lambda_i m_{it-1} + u_{it}$$

and homogeneity of the slopes is incorrectly imposed, the within estimator is not consistent even for large T. The coefficient of the lagged dependent variable is biased upwards towards unity (assuming the regressors are positively serially correlated as is usually the case). The RCM estimator is however consistent for large T, though it suffers the small T lagged dependent variable bias. Comparison of the various estimators, which are subject to different biases, can allow us to infer which biases are most important.

In the individual regressions,

$$m_{it} = \alpha_i + \beta_i' x_{it} + u_{it},$$

if the variables are integrated of order one, I(1) and are cointegrated (the error term u_{it} is I(0)) then the least squares estimate of $\hat{\beta}_i$ gives a super-consistent estimate of the long-run effect for large T. However, as noted above if the variables are I(1) but are not cointegrated (the error term is also I(1)), then $\hat{\beta}_i$ does not converge to β_i, as $T \to \infty$, but to a random variable. The regression is spurious. However, pooling or averaging over groups, can allow one to obtain a consistent estimate of an average long run effect from the levels regression. Thus the pooled or average estimates from static levels regressions may be of interest even if individual countries equations differ and do not cointegrate.

Dunne and Perlo-Freeman (2003b) estimate a very similar model to Dunne and Perlo Freeman (2003a) explaining the log of the share of military expenditure with the same explanatory variables. The difference is that rather than averaging the data they use it as an unbalanced panel of annual data for 98 developing countries 1981-1997. The static fixed effects estimates are quite similar to those found in cross-section. However, when dynamics are allowed for, through a lagged dependent variable the results are very different. The estimates are obtained by differencing the data to remove the fixed effect, then instrumenting the lagged change in the dependent variable, which becomes correlated with the error term from the differencing. In contrast to the cross-section results there is evidence of structural change, between Cold War and post Cold War periods in the dynamic panel model.

Collier and Hoefler (2004) use a pooled static panel of five-year averages explaining the share of military expenditure by measures of international war, civil war, external threat, international war, democratic government, neighbor's military expenditure, a post-Cold War shift after 1995, log population, log GDP per capita, aid to GDP and a dummy for Israel. They find the effect of neighbor's military expenditure quite large, meaning that increases in military expenditure are escalated among neighbors, making them a regional public bad. They also investigate the endogeneity of a number of variables and find that once instrumented, military expenditure does not deter rebellions.

Gadea et al. (2004) estimate a heterogeneous panel model for the military expenditures of NATO countries (a) allowing for endogenously determined structural breaks and (b) using average shares of military expenditure which controls for between group residual dependence and acts as a proxy for the threat. They find a lot of structural breaks. The results for the individual countries are difficult to interpret. This is a common finding in other areas. In panel studies, where the number of time-periods is large enough to estimate a separate time-series model for each country, the pattern tends to be that while average effects are sensible there is a very large amount of heterogeneity in the country specific estimates, which may not appear sensible.

8. Conclusions

Arms races are like bubbles in financial markets, most people agree they exist but it is remarkably difficult to define them unambiguously, estimate them or test for their existence. Our primary explanation for that difficulty is structural instability. The theory does not suggest that the underlying parameters of an arms race models would be constant over time and the econometrics confirms that. The one example of what looked like a stable action-reaction pattern, between India and Pakistan, has broken down since we last estimated it, though it could be rescued to some extent, by changing functional form. Of course, as this example shows, the apparent structural instability may be the product of misspecification and we would expect further research in that direction: experimenting with different specifications, functional forms, estimation methods etc. It is not clear that this the right route.

We are not claiming that there are no examples of equations where the military expenditure of antagonists, and allies, are significant in equations explaining the demand for military expenditure for particular countries over a particular sample. There are a lot of examples, where this is certainly the case. What we are claiming is that there are no examples of structurally stable dynamic action-reaction systems determining mutual military capabilities. Nonetheless, cross-section and panel estimates may be useful in that they allow estimation of average interaction effects which may be useful in calculating the costs of spill-over effects of increases in military expenditure in one country. Collier and Hoeffler (2004) draw a variety of policy conclusions from their estimates. Arms races have obvious implications for arms control, particularly where there are multiple equilibria, but the linear models common in the econometric literature do not allow for multiple equilibria.

The emphasis in this literature is on quantitative-symmetric arms races, because those are easier to estimate, but this emphasis may be misleading, qualitative-asymmetric arms races, particularly between governments and their non-governmental opponents, may be much more important. Arreguin-Toft (2001) comments “material power is useful for theory building because it is quantifiable and measurable in a way that courage, leadership and dumb luck are not. … The strategic interaction thesis makes clear, however, the limitations of relative material power by highlighting the conditions under which it matters more or less.” The argument of this paper, like that of Reiter (1999) is that strategy matters and when combatants fight in the same way, the stronger wins, but when they fight in different ways, the weak has an advantage. We may be missing important arms races, where the opponents are preparing to fight in different ways.

References

Andreou, A.S. N.H. Mateou and G.A. Zombanakis (2003) The Cyprus puzzle and the Greek-Turkish arms race: forecasting developments using genetically evolved fuzzy cognitive maps, Defence and Peace Economics, 14(4) 293-310.

Arreguin-Toft, I (2001) How the weak win wars: a theory of asymmetric conflict, International Security, 26(1) 93-128.

Baliga, Sandeep and Tomas Sjostrom (2004) Arms Races and Negotiations, Review of Economic Studies, 71(2) April p1-19.

Brauer, Jurgen, (2002) Survey and Review of the Defense Economics Literature on Greece and Turkey: What have we learned?, Defence and Peace Economics, 13 (2) 85-108.

Brito, D.L. and M.D. Intriligator (1995) Arms Races and Proliferation, Chapter 6 of Handbook of Defence Economics, ed. K. Hartley and T. Sandler, Elsevier Science.

Collier, P. and Hoeffler, A. (2004), Military Expenditure: Threats, Aid and Arms Races.

Craft, Cassady B. (2000) An analysis of the Washington Naval Agreement and the economic provisions of arms control theory, Defence and Peace Economics, 11(2) 127-148.

Deger, S. and Sen, S. (1990) Military Security and the Economy: Defence Expenditure in India and Pakistan, In (Hartley, K. and Sandler, T. eds) The Economics of Defence Spending, London: Routledge, pp189-227.

Desai, Meghnad and D Blake (1981) Modelling the Ultimate Absurdity: A Comment on 'A Quantitative Study of the Strategic Arms Race in the Missile Age', Review of Economics and Statistics, 63(4), November, pp629-32.
Dunne J. Paul, Eftychia Nikolaidou and Ron Smith (2003) Arms Race Models and Econometric Applications, in Arms Trade, Security and Conflict edited by Paul Levine and Ron Smith, Routledge p178-187.
Dunne, JP and Perlo Freeman, S (2003a) The Demand for Military Spending in Developing Countries, International Review of Applied Economics, Vol. 17, no. 1, 2003, pp. 23-48.
Dunne, JP and Perlo Freeman, S (2003b) The Demand for Military Spending in Developing Countries: A Dynamic Panel Analysis, Defence and Peace Economics, Vol 14, No. 6, 2003, pp. 461-474.
Dunne, JP M.D.C. Garcia-Alonso, P. Levine & R. Smith (2004) Military Procurement, Industry Structure and Regional Conflict. University of Kent Working Paper.
Dunne, JP M.D.C. Garcia-Alonso, P. Levine & R. Smith (2005) Managing Asymmetric Conflict. University of Kent Working Paper.
Fuertes A-M and R.P. Smith (2005) Panel Time Series, Cemmap course notes.
Gadea, M.D., E. Pardos and C. Perez-Forniez (2004) A long run analysis of defence spending in the NATO countries (1960-99) Defence and Peace Economics, 15(3) 231-250.
Gibler, D.M. T.J. Rider & M. Hutchison (2005) Taking Arms Against a Sea of Troubles: Conventional Arms Races During Periods of Rivalry, Journal of Peace Research, 42(2) p131-148.
Gleditsch, NP and Njolstad, O (eds) (1990) Arms Races, Technology and Political Dynamics. London: Sage Publications.
Hamilton, James D., 1989, `A new approach to the economic analysis of nonstationary time series and the business cycle', Econometrica 57, 357-384.
Hartley K and T Sandler (1995) editors Handbook of Defense Economics, Elsevier, Amsterdam.
Intriligator M.D. and D.L. Brito (2000) Arms Races, Defence and Peace Economics, 11 (1) 45-54.
Johansen, S., (1988) Statistical analysis of cointegrating vectors, Journal of Economic Dynamics and Control, 12, 231-254.
Lebovic, J.H. (1999) Using military expenditure data: the complexity of simple inference. Journal of Peace Research, 36(6), 681-697.
Levine P and R.P. Smith (1995) The arms trade and arms control, Economic Journal, 105, 471-484.
McGuire M (1977) A Quantitative Study of the Strategic Arms Race in the Missile Age, Review of Economics and Statistics; 59(3), August 328-39.
Murdoch, J.C. (1995) Military Alliances: Theory and Empirics, Chapter 5 of Handbook of Defence Economics, ed. K. Hartley and T. Sandler, Elsevier Science.
Murdoch, J.C. and Sandler, T (2002) Economic growth, civil wars, and spatial spillovers, Journal of Conflict Resolution, 46 91-110.
Murdoch, J.C. and Sandler, T (2004) Civil Wars and Economic Growth: Spatial Dispersion, American Journal of Political Science 48(1), 138-151.
Ocal, N (2003) Are the military expenditures of India and Pakistan external determinants for each other: An empirical investigation, Defence and Peace Economics, 14(2) p141-149.

Reiter, D. (1999) Military Strategy and the outbreak of international conflict: quantitative empirical tests, 1903-1992. Journal of Conflict Resolution, 43(3), 366-87.
Richardson, L. F. (1960) Arms and Insecurity: A Mathematical Study of Causes and Origins of War. Pittsburgh: Boxwood Press.
Rosh, R. M. (1988): ‘Third World Militarisation: Security Webs and the States they Ensnare’, Journal of Conflict Resolution 32(4), 671-698.
Rotte, R. & C.M. Schmidt (2003) On the production of victory: empirical determinants of battlefield success in modern war, Defence and Peace Economics, 14, 175-93.
Samuelson, P.A. T.C. Koopmans and J. R. N. Stone (1954) Report of an evaluative committee on Econometrica, Econometrica 22, 141-6.
Sandler, T. and Hartley, K. (1995) The Economics of Defense, Cambridge, Cambridge University Press.
SIPRI (1995) Armaments, Disarmament and International Security, Stockholm International Peace Research Institute, Oxford University Press.
Smith, R.P. (1989) Models of Military Expenditure, Journal of Applied Econometrics, 4, 345-359.
Smith, R. P. (1995): ‘The Demand for Military Expenditure’, In (Hartley, K & Sandler, T. eds), Handbook of Defense Economics. Amsterdam: Elsevier pp 71-87.
Smith, R. P . Sola, M and Spagnolo, F (2000) “The Prisoner’s Dilemma and Regime-Switching in the Greek-Turkish Arms Race”, Journal of Peace Research, 37, 6 pp 737-750.
Smith, R. P., J. P. Dunne and E. Nikolaidou (2000), `The Econometrics of Arms Races', Defence and Peace Economics, 11,1, 31-44.
Smith, R.P. and A. Tasiran (2005) The demand for arms imports, Journal of Peace Research, 42(2), 167-182.