BIROn - Birkbeck Institutional Research Online

Cavatorta, Elisa and Smith, Ron P. (2017) Factor models in panels with cross-sectional dependence: an application to the extended SIPRI military expenditure data. Defence and Peace Economics 28 (4), pp. 437-456. ISSN 1024-2694.

Downloaded from: https://eprints.bbk.ac.uk/id/eprint/19781/

Usage Guidelines:
Please refer to usage guidelines at https://eprints.bbk.ac.uk/policies.html or alternatively contact lib-eprints@bbk.ac.uk.

Factor models in panels with cross-sectional dependence: an application to the extended SIPRI military expenditure data

Elisa Cavatorta*and Ron P. Smith†

Received 4 April 2016
In final form: 9 November 2016

Abstract

Strategic interactions between countries, such as arms races, alliances and wider economic and political shocks, can induce strong cross-sectional dependence in panel data models of military expenditure. If the assumption of cross-sectional independence fails, standard panel estimators such as fixed or random effects can lead to misleading inference. This paper shows how to improve estimation of dynamic, heterogenous, panel models of the demand for military expenditure allowing for cross-sectional dependence in errors using two approaches: Principal Components and Common Correlated Effect estimators. Our results show that it is crucial to allow for cross-sectional dependence, that the bulk of the effect is regional and there are large gains in fit by allowing for both dynamics and between country heterogeneity in models of the demand for military expenditures.

JEL Category: C33, C82, H56
Keywords, Military Expenditures, Panel Data, Factor models.

*Elisa Cavatorta (corresponding author): Department of Political Economy, King's College London, Strand, London WC2R 2LS; e-mail: elisa.cavatorta@kcl.ac.uk.
†Ron P. Smith: Department of Economics, Mathematics and Statistics, Birkbeck College, Malet Street Bloomsbury, London WC1E 7HX; e-mail: r.smith@bbk.ac.uk.

1 Introduction

Strategic interactions between countries can induce cross-sectional dependence in models of military expenditures estimated using panel data, for countries $i = 1, 2, ..., N$ and time periods $t = 1, 2, ..., T$. These strategic interactions can arise for a variety of reasons: arms races between hostile countries; burden sharing within alliances; security-web interactions within regional networks; and economic and political shocks, like terrorist attacks, affecting the general perception of threat for all units. The cross-sectional dependence appears as correlations between the errors in different countries. This dependence may be local, interactions among neighbours for instance; or global, unobserved common factors that drive the military expenditures of all countries. If the unobserved strategic factors that cause the cross-sectional dependence are correlated with regressors of interest, such as income, then the estimated coefficients of those regressors will be biased and inconsistent.

While the importance of cross-sectional dependence has been widely recognised in empirical arms race and alliance models, the issue has usually been addressed in the context of quite restrictive models. For instance, one approach, used in Dunne et al. (2008), relies on qualitative judgements about a country's security-web to allow the military expenditures of a country's neighbours to be aggregated. This requires identifying allies and enemies and choosing judgement-based weights in the aggregation procedure. A second approach uses spatial econometric models, which specify a distance or contiguity matrix to characterise the spillovers. A third approach, recently adopted in the panel literature, assumes that the cross-sectional dependence can be characterised by a finite number of unobserved common factors affecting all units with different intensities. This approach can be implemented empirically if there exist proxies for the common factors. One way to obtain proxies is by extracting cross-sectional commonalities in military expenditures using Principal Component Analysis (PCA). PCA estimates the linear combinations of military expenditures and factor loadings that account for most of the variation in the data. A number of these linear combinations can be included in a demand model to control for cross-sectional dependence. A second way to obtain proxies, known as Common Correlated Effect (Pesaran, 2006), consists of approximating the unobserved common factor using cross-section averages of the dependent and independent variables. One advantage of the latter approach is that is easily implementable, it yields consistent estimates under a variety of conditions – e.g. serial correlation in errors, contemporaneous correlation between regressors and unobserved factors, spatial and temporal correlations as shown by Coakley et al. (2006); Kapetanios et al. (2011); Pesaran and Tosetti (2011)– and the estimates can be easily interpreted. However, this approach requires panel data with both large N and large T.

SIPRI provides the most reliable and widely used series of military expenditures data in academic research on defence economics, but until recently had the disadvantage that the time-series started in 1988. The recent release of the extended SIPRI military expenditure data since 1945 enables the application of large T panel techniques and the implementation of factor models that allow for cross-sectional dependence in errors. The US State Department *World Military Expenditures and Arm Trade*, only provides 11 years of data.

In this paper, we use the SIPRI data to develop a dynamic model of the demand for military expenditure where there is cross-sectional dependence between the errors of different units due to unobserved common factors generated by strategic interactions. This analysis develops the approach used in Cavatorta (2010) for the MENA region and builds on recent panel time-series procedures surveyed by Chudik and Pesaran (2015*b*) and given a text-book treatment in Pesaran (2015).

We discuss the estimation of dynamic, heterogeneous models of military expenditures with cross-sectional dependence in errors using Principal Components (PCs) and Common Correlated Effect (CCE) estimation procedures and compare the estimates with standard estimation techniques which ignore the issue. Unlike the standard fixed effect approach, we allow for heterogeneity between countries, both in regression coefficients and in the impact of the unobserved strategic factors. We also examine differences between the Cold War period and the post-Cold War period and between different regions. Our results show that it is crucial to allow for cross-sectional dependence: doing so, whether by PCs or CCE methods, substantially improves the fit. The bulk of the dependence appears to be within regions. The results also show that it is crucial to allow for both dynamics and between country heterogeneity, both of which improve the fit. Allowing for a structural break at the end of the Cold War in the dynamic heterogeneous model seems less important. Our conclusion is that it is important for researchers using the extended SIPRI dataset to take account of dynamics, heterogeneity and cross-sectional dependence when choosing their specification. In particular, mean group estimation of error correction models augmented by cross-section averages seems to provide an effective modelling framework.

The paper proceeds as follows: Section 2 discusses the data issues and provides summary statistics. Section 3 sets out the basic theoretical framework of factor models. Section 4 uses the PCs and CCE procedures to estimate static factor models that attempt to determine the relative influence of economic and strategic factors on the shares of military expenditure. Section 5 allows for dynamics both in PCs and CCE models. Section 6 provides some conclusions and suggestions for further research.

2 Data Issues

The SIPRI military expenditure series are the most widely used measures in academic research on military expenditures. They have the disadvantage that, up to now, SIPRI has only provided consistent data from 1988. This is quite a short post Cold War sample and researchers have tried to extend the data either by splicing to earlier, unapproved, SIPRI series or to Correlates of War (COW) series, neither of which are quite consistent with the later SIPRI authorised series. Brauner (2015) discusses the combination of SIPRI and COW series. The availability of a database with a longer version of the authorised SIPRI series thus enables more interesting explorations of the data.

SIPRI provides three series. D_{it} a domestic measure of military expenditure in local currency at current prices; M_{it} real military expenditure in constant US prices and exchange rates of a base year; and $S_{it} = M_{it}/Y_{it}$ the share of military expenditure as a percentage of Gross Domestic Product (GDP), also called military burden. The GDP series that SIPRI uses to construct the share, Y_{it}, is also available in the database.

The focus of discussions of data quality is usually on the problems of measurement of military expenditure and in their discussion of sources and methods. SIPRI comment on the limitations of the data in terms of reliability, validity and comparability. However, it should be recognised that there are substantial revisions in measured GDP and this is also a source of measurement error and data revisions.

We chose to use a balanced panel which gave a large T for a large set of N countries. The 50 year period 1965-2014 with some interpolation, gave a sample of 70 countries.[1] There is a trade-off between obtaining the maximum T or N. For explicit spatial models it is important that there are data for neighbours, thus increasing N at the expense of T is sensible. We are interested in the dynamic heterogeneous models so increasing T at the expense of N has advantages. This is not a random selection of countries. Having data for the whole of the period 1965-2014 excludes important countries such as China, which has a lot of missing data on military expenditure, and the Soviet Union and the 15 Soviet successor states including Russia.

We give summary statistics, for military expenditure, GDP and the share of military expenditure for the full sample and for the balanced panel in Table 1.

In terms of the share of military expenditure the full and balanced samples were very similar: means of 2.96% in the full sample and 2.89% in the balanced sample. The full sample, with a standard deviation of 3.5%, was more dispersed than the balanced sample, with a standard deviation of 2.9. In the balanced sample 80% of the variance came from the between-country cross-section dimension. Both distributions are highly skewed. While there is a minimum of zero to the left, there is no maximum to the right, military expenditure can be over 100% of GDP. The maximum in the full sample was 117%, Kuwait, in the balanced sample 30%. The distributions of military expenditure and GDP are also skewed. Military expenditure in the full sample, is about 60% of the balanced sample. However, GDP in the full sample is twice that of the balanced sample. Notice that the mean of the ratios of military expenditure to GDP is not the same as the ratio of the means.

[1]The countries are Algeria, Argentina, Australia, Austria, Belgium, Benin, Bolivia, Brazil, Burkino Faso, Burundi, Canada, Chile, Colombia, Denmark, Dominican Republic, Ecuador, Egypt, El Salvador, Ethiopia, Finalnd, France, Germany, Ghana, Greece, Guatamala, Honduras, India, Iraq, Ireland, Israel, Italy, Japan, Jordan, Kenya, Korea South, Liberia, Libya, Luxemburg, Madagascar, Malawi, Malaysia, Mali, Mauritius, Mexico, Morocco, Myamar, Netherlands, New Zealand, Nigeria, Norway Pakistan Paraguay Peru Phillipines Portugal, Saudi Arabia, Sierra Leone, South Africa, Spain Sri Lanka, Sweden Switzerland, Thailand, Tunisia, Turkey, Uganda, UK, USA Venuezela Zimbabwe.

The bottom panel of Table 1 gives summary statistics for the shares of military expenditure for the five regions. Four of the regions, Africa, Americas, Asia, and Europe, have very similar mean shares of military expenditure between 2-3%, the Middle East is much higher at 8.7%.

3 Modelling Framework

3.1 Strategic Interactions

The classic quantitative representation of the determination of military expenditures is the Richardson Arms Race Model. This describes the dynamic interaction of military expenditures of countries 1 and 2, $m_1(t)$, $m_2(t)$. There is a positive stimulus from the rival and negative fatigue effects from one's own spending:

$$\begin{aligned}\frac{dm_1(t)}{dt} &= a_1 + b_1 m_2(t) - c_1 m_1(t),\\ \frac{dm_2(t)}{dt} &= a_2 + b_2 m_1(t) - c_2 m_2(t).\end{aligned}$$

This implies dependence between the military expenditures of the different countries. This model has been widely applied, with relatively limited success, partly because interactions are rarely that mechanical or purely bilateral. Even in classic arms races, like those between Greece and Turkey or India and Pakistan, the actors are responding to other threats than from their antagonist; the Soviet Union in the case of Greece and Turkey and China in the case of India and Pakistan. The expenditures of possible allies would also matter, like the US in the case of Greece and Turkey. There is also a budget constraint limiting expenditures.

The model that we use to provide a framework, which is very standard in the literature, generalises the Richardson model and is reviewed in Dunne and Smith (2007). It assumes that military expenditures are determined by both economic and strategic factors. The economic factors are typically measured by GDP to approximate the budget constraint and the strategic factors are usually measured by the military expenditures of other countries, allies or potential enemies, which represent the threat or fear factor. Although we will also estimate a dynamic model, to clarify the exposition we start from a simple static model determining m_{it}, the logarithm of real military expenditures of country $i = 1, 2, ..., N$ in year $t = 1, 2, ..., T$, by the logarithm of their real GDP, y_{it}, and other countries military expenditures, m_{jt}, for $j \neq i$,. Then the model can be written as a system of N equations of the form:

$$m_{it} = \alpha_i + \eta_i y_{it} + \sum_{j \neq i} \gamma_{ij} m_{jt} + u_{it}. \tag{1}$$

The income elasticity of demand for military expenditure is η_i, the feedback from other countries military expenditures is given by γ_{ij}. Smith (1995) discusses how equations of this sort can be derived from optimising a social welfare function, which depends on security and consumption, subject to a budget constraint. Other variables could be added such as indicators of political regime as in Brauner (2015); indicators of internal or external conflict as in Dunne et al. (2008) or other sources of income, such as aid, as in Collier and Hoeffler (2007). But given our focus is on cross-sectional dependence we will just use a simple model relating military expenditures to GDP.

Clearly it is not possible to freely estimate the $N-1$ feedback coefficients.[2] Various ways to deal with this curse of dimensionality have been adopted in the literature. Spatial models for the $N \times 1$ vector of dependent variables, military expenditures in our case, $\mathbf{y}$, determined by a $N \times K$ matrix of covariates, $\mathbf{X}$, allow spillovers mediated through a known $N \times N$ non-negative weight matrix $\mathbf{W}$, with zeros on the diagonal. This could represent distance between the units, whether the units share a common border, or some other measure of closeness. The general nesting spatial econometric model is:

$$\begin{aligned}\mathbf{y} &= \mathbf{X}\beta_0 + \mathbf{W}\mathbf{X}\beta_1 + \gamma\mathbf{W}\mathbf{y} + \mathbf{e} \\ \mathbf{e} &= \rho\mathbf{W}\mathbf{e}\end{aligned}$$

There are endogeneity problem in cases where either of the scalars γ or ρ are not equal to zero and other estimators than OLS are available in the literature. If $\mathbf{W}$ is known this reduces the curse of dimensionality, but it may not be known. More importantly for our purposes, the term of most interest is the spatial autoregressive term, $\gamma\mathbf{W}\mathbf{y}$, which captures the influence of other countries military expenditures on the expenditures of the focus country. But this only has a single parameter γ which cannot capture the different signs of the interactions with allies and enemies.

In (1), one might expect that for enemies $\gamma_{ij} > 0$, reflecting arms races; for allies $\gamma_{ij} < 0$, since their military expenditures can substitute for yours; and for uninvolved pairs $\gamma_{ij} = 0$. In the literature on alliances, surveyed in Murdoch (1995), considerable attention is paid to the technology by which the military expenditures of allies is aggregated. This imposes a particular structure on the γ_{ij}. The technology may make the strength of the alliance depend on the simple sum, the best shot or the weakest link. Another common procedure to reduce the curse of dimensionality is to use *ad-hoc* weights to sum the military expenditures of potential allies to give a measure of friends spending and sum that of potential adversaries to give a measure of foes spending. These *ad-hoc* weights can be based on qualitative judgements about the security web, the nature of the linkages with the other countries, as in Dunne et al. (2008).

3.2 Principal Component Approach

If a set of allies are all responding to a common threat, they are likely to all move their military expenditures together generating a positive correlation between them as well as between them and their potential enemies. Such positive correlations

[2] If the dependence was on own and others lagged military expenditure, (1) would correspond to the infinite VAR discussed by Chudik and Pesaran (2011).

between allies are common in the literature. This positive correlation among all the military expenditures of a group of interacting nations can be represented by a common unobserved threat factor driving the military expenditures of the interacting countries. Among a large group of countries, there are likely to be more than one strategic interaction, so the military expenditures may be driven by more than one threat factor. Assume that there are K such interactions with K such unobserved latent factors, f_{kt}^*, $k = 1, 2, .., K$. Then we can write the model as

$$m_{it} = \alpha_i + \eta_i y_{it} + \sum_{k=1}^{K} \lambda_{ik} f_{kt}^* + e_{it} \tag{2}$$

Where the weights λ_{ik} are non-zero if country i is involved in interaction k. The f_{kt} can be estimated by the method of principal components, PCs, as linear combinations of the military expenditures:

$$f_{kt}^* = \sum_{j=1}^{N} a_{kj} m_{jt} \qquad k = 1, 2, .., K \tag{3}$$

and one might hope that a few PCs would account for a lot of the variation in military expenditures. One would expect K to be much smaller than N, so estimating (2) will be much easier than estimating (1) . One can recover the coefficients on other countries military expenditure as

$$\gamma_{ij} = \lambda_{ik} a_{kj}.$$

If we define the logarithm of the share of military expenditure in GDP as $s_{it} = m_{it} - y_{it}$, then we can write the model in log shares as

$$s_{it} = \alpha_i + \beta_i y_{it} + \sum_{k=1}^{K} \lambda_{ik} f_{kt} + e_{it} \tag{4}$$

where $\beta_i = \eta_i - 1$. If the income elasticity of demand for military expenditures is unity, as is often assumed, $\beta_i = 0$ and log GDP drops out of the equations. Shares may also be a better indicator of threat perceptions than military expenditures, not being dominated by size. So one could estimate the factors as the PCs of the shares of military expenditure:

$$f_{kt} = \sum_{i=1}^{N} a_{ki} s_{it} \qquad k = 1, 2, .., K. \tag{5}$$

Estimating the factors using (5) rather than (3) implies that the threat from enemies or support from allies is represented not by the level of their military expenditures but by their share of military expenditure in GDP, perhaps as an indication of commitment or intent. This is not implausible given the importance attached by NATO to the commitment made at the 2014 Cardiff summit to spend at least 2% of GDP on defence. The fit of the shares model can be compared with the levels model using some model selection criterion like the BIC,[3] which can be used even though the models are not nested. To nest model (4) in model (2) , one would need to add the log GDP of the other countries to (2).

The strategic factors may not be global but regional so the model could be applied not over all countries N, but over the number of countries in the region N_r, for $r = 1, 2, ...R$. One could also allow for the US being a dominant unit, as discussed in Chudik and Pesaran (2013) that appears as an explanatory variable in every region. The Principal Components allow us to measure how much of the variance of the shares or military expenditure is accounted for by these strategic factors and how they influence each country.

Model (4) is heterogeneous, the coefficients differ across countries and we can report the mean group estimator of Pesaran and Smith (1995) reporting the average and standard error of the coefficients. A special case of model (4) is the two way fixed effect model, which imposes slope homogeneity, $\beta_i = \beta$, and the factors having the same effect on each country $\sum_{k=1}^{K} \lambda_{ik} f_{kt} = \alpha_t$. The model is then:

$$s_{it} = \alpha_i + \alpha_t + \beta y_{it} + e_{it} \tag{6}$$

A different model which is intermediate between the mean group and two way fixed effects is the interactive fixed effects model of Bai (2009), which assumes slope homogeneity ($\beta_i = \beta$) but allows the effect of the factors to differ over countries.

Model (4) is a very parsimonious model. Clearly, there are many other variables that one might think are omitted from the model. These might include measures of conflict and quality of the institutions in the country, and many other economic and political variables. Using a parsimonious model has the advantage that we can use the maximum number of observations for which we have SIPRI military share data, not losing data because of missing observations on other variables. It also allows us to focus on the role of cross-sectional dependence in a simple case. Denote these omitted variables by the vector z_{it}, so that the correct model is

$$s_{it} = \alpha_i^* + \beta_i^* y_{it} + \sum_{k=1}^{K} \lambda_{ik}^* f_{kt} + \phi_i' z_{it} + e_{it}.$$

Notice that β_i^* is measuring a different parameter of interest from β_i. The parameter β_i^* measures the effect of a change in income holding z_{it} constant, while β_i measures the effect of a change in income allowing z_{it} to adjust as it does in the sample. The effect of this omission depends on the correlation between income, the global factors, f_{kt}, and the country specific omitted variables z_{it}. Consider the case of democracy. Democracy seems a relevant variable since democracies spend less than autocracies on the military and are richer. If the country is a democracy throughout the period, $z_{it} = z_i$, as it does not vary over time. The effect of any time invariant variable is picked up by the intercept α_i. After the Cold War many country undertook a process of democratisation. This time-varying factor will be correlated across countries and

[3] The BIC seems more appropriate than the AIC because it is more parsimonious and with large data sets it is easy for parameters to proliferate.

hence with the global factors, which pick up the reduction in the share of military expenditure during the democratisation process. In sum, to the extent that the country specific variables z_{it} are correlated with GDP or the factors, these variables will pick up the effects of the z_{it} and may be a parsimonious representation of many influences that lack consistent data across the sample. If, however, there are time-varying unobserved country specific variables that vary in a way that is uncorrelated with income or the factors, then this will increase the unexplained variance. How big an improvement in fit results from including these possible omitted variables is a subject for future research.

3.3 Common Correlated Effect Approach

Above we assumed that the unobserved factors were estimated by principal components, but they can also be allowed for using the correlated common effect, CCE, estimator. Reparameterise (2) in terms of shares, so $\beta_i = \eta_i - 1$, and for exposition initially assume that there is only a single threat factor, though in estimation we allow for a multi-factor model (e.g. there there may be separate factors for enemies and allies). Then

$$s_{it} = \alpha_i + \beta_i y_{it} + \lambda_i f_t + e_{it}. \tag{7}$$

Average (7) over the countries to give

$$\overline{s}_t = \overline{\alpha} + \overline{\beta}\overline{y}_t + \overline{\lambda} f_t + \overline{e_t} + \overline{\eta}_t \tag{8}$$

where

$$\begin{array}{lll} \overline{s}_t = \sum_{i=1}^N s_{it}/N, & \overline{\lambda} = \sum_{i=1}^N \lambda_i/N, & \overline{e}_t = \sum_{i=1}^N e_{it}/N, \\ \overline{\beta} = \sum_{i=1}^N \beta_i/N, & \overline{y_t} = \sum_{i=1}^N y_{it}/N, & \overline{\eta}_t = \sum_{i=1}^N \left(\beta_i - \overline{\beta}\right) y_{it}/N. \end{array}$$

Assuming $\overline{\lambda} \neq 0$, we can write (8) as

$$f_t = \overline{\lambda}^{-1} \left(\overline{s}_t - \overline{\alpha} - \overline{\beta}\overline{y}_t - \overline{e_t} - \overline{\eta}_t\right)$$

thus we can approximately filter out the effect of the factor by including $\overline{s}_t$ and $\overline{y}_t$ in (7) instead of the factor

$$\begin{aligned} s_{it} &= \alpha_i + \beta_i y_{it} + \lambda_i \left[\overline{\lambda}^{-1} \left(\overline{s}_t - \overline{\alpha} - \overline{\beta}\overline{y}_t - \overline{e_t} - \overline{\eta}_t\right)\right] + e_{it} \\ s_{it} &= a_i + \beta_i y_{it} + \delta_{1i}\overline{s}_t + \delta_{2i}\overline{y}_t + u_{it} \end{aligned} \tag{9}$$

where

$$\begin{aligned} u_{it} &= \lambda_i \overline{\lambda}^{-1}(-\overline{e_t} - \overline{\eta}_t) + e_{it}, \\ a_i &= \alpha_i - \lambda_i \overline{\lambda}^{-1}\overline{\alpha}. \end{aligned}$$

Pesaran (2006) provides more details and the generalisation to the dynamic case, which we use below, is provided in Chudik and Pesaran (2015*a*). Note that the covariance of $\overline{s}_t$ with u_{it} declines with N, so for N large we can treat $\overline{s}_t$ as exogenous.

We can compare (7) , using two factors, estimating the f_t by the principal components of the shares, with (9) and see which fits better. If we use two PCs both the CCE and PC equations will have the same degrees of freedom, so their fit is directly comparable.

Both (7) and (9) assume heterogeneous relationships, different for every country. We could also see whether there is any evidence of homogeneity by comparing the fit of the heterogeneous models with the fit of the two way fixed effect model (6) , above which imposes slope homogeneity $\beta_i = \beta$ and that the factor has the same effect everywhere so $\lambda_i f_t = \lambda f_t = \alpha_t$.

3.4 Testing for error cross section dependence

This section is based on Pesaran (2015) section 29.7, which considers the issues in more depth. The tests for error correlation use the estimated correlations between the equation residuals:

$$\hat{\rho}_{ij} = \sum_{t=1}^{T} \hat{u}_{it}\hat{u}_{jt} / \left(\sum_{t=1}^{T} \hat{u}_{it}^2 \right)^{1/2} \left(\sum_{t=1}^{T} \hat{u}_{jt}^2 \right)^{1/2}.$$

Under the null of no cross-sectional dependence for finite N and large T, $\sqrt{T}\widehat{\rho}_{ij} \sim N(0,1)$ and $T\widehat{\rho}_{ij}^2$ is χ_1^2.

The Breusch-Pagan LM test is for the hypothesis $H_0 : \rho_{ij} = 0$, all $i \neq j$, and uses the $N(N-1)/2$ squared correlations, $\widehat{\rho}_{ij}^2$. This test tends to be highly over-sized, rejecting too often, when N is large, as in our case with $N = 70$. The hypothesis, $\rho_{ij} = 0$, for all $i \neq j$, is sensible when N is small and fixed as $T \to \infty$. However, when N is large, requiring all correlations, 2415 in our case, to be zero is overly restrictive. A few non-zero correlations are not going to adversely affect the properties of the regression parameters of interest. In addition, unless T is very large, coincident outliers in pairs of countries can generate high correlations.

To avoid the problem that the LM test is not correctly centred, Pesaran (2004) proposed a cross-sectional dependence test ($CSD_{Pesaran}$) based on the average correlations, which has mean zero for fixed values of N and T. This is the test that we shall use and is given by:

$$CSD_{Pesaran} = \sqrt{\frac{2T}{N(N-1)}} \left(\sum_{i=1}^{N-1} \sum_{j=i+1}^{N} \widehat{\rho}_{ij} \right) \sim N(0,1).$$

The implicit null hypothesis of this test is weak cross-sectional dependence. However, even this test is likely to over-reject in the case of models with weakly exogenous regressors if N is much larger than T, e.g. lagged dependent variables. Because this test is based on the average correlation, positive and negative correlations cancel out. We also report the average absolute correlation to take this into account.

4 Estimation of Static Factor Models

4.1 Estimation of the Common Factors (PCA)

We begin by estimating the common factors as the principal components, PCs, of the shares of military expenditure, the logarithms of military expenditures, the logarithms of GDP and the logarithms of the shares for the 70 countries in our balanced panel over the whole period, 1965-2014; for the two sub-periods 1965-88 and 1989-2014; and for the five regions. The cumulative share of the variance of military burden explained by the first 5 Principal Components is given in Table 2 for each case.

For all countries and the whole period, the first PC explains 50% of the variance of the share, the second 17% and the third 7%, so the first three explain 74% of the variance and the first 5, 82%. Clearly, there are strong common factors that drive these series. In the Cold War the first PC explains less of the variation than in the post Cold War periods, though the total explained by the first 5 is similar. The strength of the common factors differ across regions. For Africa and the Americas, the first PC explains a much smaller part of the variance than it does in the other regions and the second a larger part relative to the other regions. The first PC explains 37% in Africa and 43% in the Americas, compared to 60% or more in the other regions. It seems plausible that more idiosyncratic factors drive the shares of military expenditures in Africa and the Americas. Notice that the shares are not weighted, so in the measured variance of the shares in the Americas, the US gets equal weight to any other country. Because of this, including the US in the other regions did not change the results very much, since it has a small weight. In Asia, Europe and the Middle East the first PC explains a larger proportion of the variance than in the full sample, in Europe a striking 79% of the variance is explained by the first PC. Europe was in the front line of the Cold War and it is plausible that European shares of military expenditure were driven largely by the Cold War factor.

For all the countries and the whole period, the first PC is plotted against the second PC in Figure 1. The bulk of the observations lie in a vertical column, with a value for PC1 of just over 0.1. This roughly corresponds to the mean share, giving an an equal positive weight to most countries. There are a group of countries that have negative values for PC1, these are Algeria, Burundi, Colombia, Congo, Ecuador, Japan, Liberia, Sierra, Leone, Sri Lanka. Libya, Mexico and Uganda have very small positive values. All these countries except Japan and, perhaps, Mexico have seen substantial civil wars. It is interesting that this PCA procedure identifies these as outliers even though it was not designed to do so. There is also a pattern for countries in the main group with a high or low value for PC2, to have a value of PC1 closer to zero, introducing some curvature.

When one looks at it by sub-period, the post-Cold War period figure looks very similar to Figure 1, with most countries having a positive weighting. However, for the Cold War period the pattern is very different. Figure 2 shows a large block of countries having a positive weighting, and a large block having a negative weighting.

The curvature apparent in Figure 1 is now apparent in both blocks. It looks almost like a circle.

Figure 3 plots PC1 and PC2. PC1 is roughly constant till about 1985 and then trends steadily downwards. It reflects the high shares of the cold war, then the downward trend. PC2 trends upwards to the early 1980s then trends downwards. It is not so clear what the interpretation of PC2 is. This is a limitation of PCs, it is often difficult to interpet them.

Table 3 gives the proportion explained by the first five PCs for the log share, the log of military spending and the log of GDP. The proportions explained were higher for the log of share than share, the first PC explaining 55% of log share as compared to 50% of share. Not surprisingly the common factors in log military expenditure and log GDP are much higher than for log shares because the level variables have more variance, and log GDP has a strong trend. For log GDP the first PC explains 95% of the variance. Clearly a much stronger factor drives GDP than the shares of military expenditure.

4.2 CCE and PC static model estimates

We report the Pesaran and Smith (1995) mean group estimates for these models. Estimates are given for the full sample, two sub-periods 1965-88 and 1989-2014 and for the five regional groupings. The tables give the Wald test for the hypothesis that the means of the three slope coefficients are zero. It is a test on the averages of the coefficients. The Sum LL is the sum of the maximised log likelihood over the individual regressions for the N countries. It is not the log likelihood for the system, since it ignores the covariances. For model selection, we also use an approximate Bayesian information criterion $BIC = \sum_{i=1}^{N} LL_i - 0.5K \ln NT$, where K is the total number of prameters estimated. This is approximate both because the Sum LL ignores the covariances and because it is not clear that N and T should be treated symmetrically. The literature on model selection in panel data is relatively small. The notes to the tables also give the $CSD_{Pesaran}$ test statistic.

The mean group estimates using PCs MG-PC are given in Table 4, together with the estimates with no factors in the bottom panel. Adding the factors clearly improves the fit and changes the estimate of the coefficient of GDP. It also reduces cross-sectional dependence the $CSD_{Pesaran}$ test statistic, which is $N(0, 1)$ under the null of no correlation, falls from 62.5 to 0.405 when the PCs are added. There is clear evidence of a structural break. Splitting the data into two sub-periods improves the fit substantially, increasing the log-likelihood for the model including PCs from 1980 to 3301. The pattern of coefficients is quite different in the whole period, where log GDP is insignificant and the two factors significant, and the two sub-periods, where GDP is significant and, with the exception of the first factor in the pre-Cold War period, the factors are not significant on average. Notice that while the average of the coefficients of the factors may not be significant, they may be significant in individual countries: some responding positively and some negatively to the factors. Although the coefficients of GDP are significant in the sub-periods they are not large, indicating that the common assumption that the income elasticity of demand for military expenditure is close to one, is not unreasonable. Disaggregation by regions has a much smaller effect, raising the LL from 1980 to 2138. In the regions, the first factor is significant and positive, but the second factor is insignificant.

The mean group estimates using cross-section means MG-CCE are given in Table 5. Including the means reduces the $CSD_{Pesaran}$ test statistic from 62.5 to 5.8, but not by as much as the PCs. As with the PC estimates, splitting the sample into sub-periods improves the fit substantially increasing LL from 1757 to 3023. It is also clear that including two PCs works better than including the means of the share and log GDP. As in the PC case, the coefficient of log GDP is insignificant in the whole period but significant in the two sub-periods. Whereas the pattern of GDP only being significant when the sample is split into two sub-periods is the same in the PC and CCE cases, the size of the coefficient are not. The CCE estimates suggest that military expenditure is very inelastic, with an elasticity of 0.49 in the first period compared with the PC estimate of 0.79. In the second period the difference is smaller, with the CCE estimate of 0.77 and the PC of 0.86. The fact that the coefficient of GDP is larger in the CCE estimator suggests that it is a countries GDP relative to the world average that matters. In all cases the mean of the coefficients of log GDP, $\bar{\beta} = N^{-1} \sum \beta_i$ is of opposite sign to the mean of the coefficients of GDP average, $\bar{\delta}_2 = N^{-1} \sum \delta_{1i}$. In the two sub-periods when $\bar{\beta}$, the coefficient of log GDP, is significant that on average GDP, $\bar{\delta}_2$, is also significant. The PC estimate based on the shares cannot capture this relative feature whereas the CCE estimator does.

It is also noticeable that in almost all the cases the mean of the coefficients of the mean share, $\bar{\delta}_1 = N^{-1} \sum \delta_{1i}$ from (9) is close to one and is very significant. In the two way fixed effect estimator the coefficient of the mean of the dependent variable, δ_1, would be one and the coefficient on the mean of the independent variable, δ_2 would be equal and opposite the coefficient of the independent variable, since (6) can be written as

$$s_{it} - \bar{s}_t = \alpha_i + \beta(y_{it} - \bar{y}_t) + \varepsilon_{it}.$$

Figure 4 plots the histogram of the coefficients of income for both procedures, PCA and CCE. The CCE estimates seem more dispersed, though in both cases the range is quite wide (income elasticities between -0.5 and almost 2): there is clearly considerable heterogeneity in the income coefficients. Figure 5 plots the scatter diagram between the two sets of income coefficients. There is a positive though not very strong relationship, the fitted regression line being dominated by an outlier, where both estimates are close to -1.5. The average income coefficients differ between the two methods, primarily because of the high variance of the CCE estimates.

Table 6 shows the fit statistics for the 10 models. Allowing for heterogeneity, cross-section dependence and a break at the end of the Cold War all improve the fit. The approximate BIC would choose the MG-PC split into sub-samples as the best model and an AIC calculated in the same way would also choose this. However, these are all static models and in time-series dynamics are likely to be important.

In addition, the variables may be I(1) and possibly cointegrated.[4]

5 Dynamic Factor Models

5.1 PCs and CCE dynamic models

We now present the dynamic models. The dynamic CCE equation is:

$$\begin{aligned} s_{it} &= \alpha_{0i} + \alpha_{1i}s_{i,t-1} + \beta_{0i}y_{it} + \beta_{1i}y_{i,t-1} + \delta_{10,i}\overline{s}_t + \delta_{11,i}\overline{s}_{t-1} \\ &\quad +\delta_{20,i}\overline{y}_t + \delta_{21,i}\overline{y}_{t-1} + u_{it}. \end{aligned} \tag{10}$$

This can also be written in error correction form

$$\begin{aligned} \Delta s_{it} &= a_{0i} + a_{1i}s_{i,t-1} + b_{0i}\Delta y_{it} + b_{1i}y_{i,t-1} + d_{10,i}\Delta\overline{s}_t + d_{11,i}\overline{s}_{t-1} \\ &\quad +d_{20,i}\Delta\overline{y}_t + d_{21,i}\overline{y}_{t-1} + u_{it}. \end{aligned} \tag{11}$$

The error correction form is useful to capture both long and short-term dynamics in a single model. The dynamic PCs model specifications are similar, substituting $\Delta\bar{f}1$ for $\Delta\overline{s}_t$ and $\bar{f}1$ for $\overline{s}_{t-1}$ and $\Delta\bar{f}2$ for $\Delta\overline{y}_t$ and $\bar{f}2$ for $\overline{y}_{t-1}$. The number of estimated parameters remains the same, so the PC and CCE models are directly comparable using log-likelihoods.

5.2 PCs and CCE dynamic estimates

The mean-group estimates of the dynamic model (11) are reported in Table 7. Adding dynamics substantially improves the fit, increasing the BIC considerably, the best static model has a BIC of 994, the best dynamic model, MG, has a BIC of 1834. Table 8 presents the fit statistics for the ten dynamic models models. Unlike in the static case, neither splitting the sample or allowing for cross-sectional dependence improves the fit according to the BIC. However, on the basis of the $CSD_{Pesaran}$ statistics, including the world means reduces cross-sectional dependence considerably. The dynamic MG with no correction for cross-section dependence has a CSD statistic of 18.31, with an average correlation of 0.053 and an average absolute correlation of 0.126. The MG-CCE which includes means has a CSD statistic of -2.12, average correlation of -0.006 and an average absolute correlation of 0.126, so reducing the extent of cross-sectional dependence as one might expect. Unlike the static case the one way fixed effect has a BIC that is similar to MG-CCE. Moving to two way fixed effect does not improve the BIC, indicating that allowing for a common factor does not improve fit in the homogeneous model.

Based on the log-likelihood or BIC, the model using PCs performs slightly better than the model using CCE, but the difference is small. However, the magnitude of the factors' coefficients are difficult to interpret in the PC case as they do not

[4] Breitung and Pesaran (2008) discuss unit roots and cointegration in panels. Kapetanios et al. (2011) discuss panels with non-stationary multifactor error structures.

have clear-cut units of measurement, whereas the coefficients of the mean have a straightforward interpretation. In the CCE models, the mean group estimate of the long-run effect of a common increase in the level of threat, represented by the global average share of military expenditure, can be calculated as $-\bar{d}_{11}/\bar{a}_1$. In this case, the two coefficients have the same units, in PCs models the two coefficients have different units. We show the heterogeneity of long-run effects of military expenditures to common threat estimated using CCE in Figure 6. The distribution of long-run effects using PCs model is almost uniformly centered at zero because the estimated coefficient of the first factor, d_{11}, is almost zero in any country. The US is an extreme outlier because the coefficient a_{1i} is almost zero making the ratio very large. The heterogeneous estimator, on which the mean-group estimates are based, also perform much better than the fixed effect, FE, estimator which imposes homogeneity. Thus we do not report the FE results. The coefficient on $\bar{s}_{t-1}$ in the FE model (0.0965) and that of the mean-group estimator (0.206) are different, suggesting that country heterogeneity is important.

Splitting the CCE model by pre- and post-1988 period does not improve the fit as measured by the BIC. This is because there is a trade-off between fit and number of estimated parameters. The time-series for the sub-periods are quite short and the increase in estimated parameters imposes a large penalty on the BIC. There is a small T downward bias in the coefficient of the lagged dependent variable and it is noticeable that in both sub-periods the coefficient of the lagged dependent variable is more negative than in the whole period. The long-run effect of the world share, possibly a measure of the effect of common threats, is slightly larger in the Cold War period than in the post Cold War period. This suggests that using a simple Cold War dummy variable for pre- and post-1988 is not sufficient to capture that variation. In addition, there is little correlation between the estimated coefficient on y_{t-1} in the Cold War and post-Cold War period (these correlations are -0.09 and 0.03 using the CCE and PCs models, respectively.

5.3 Disaggregating the factors

Up to now we have assumed a single global factor, now we want to allow for separate influences from global and regional factors and the impact of the US, as a dominant unit. We keep world GDP since a global economic cycle is plausible and the PCA analysis above indicated that the first factor accounted for the bulk of the variation in GDP. The results are shown in Table 9 where in addition to the change and lagged level or the world share, the change and lagged level of the regional and US share are included, where the region is the geographical region of country i defined by SIPRI (i.e., Africa, Americas, Asia & Oceania, Europe and Middle East). The world share and US shares are not significant but both the change and lag in the regional shares are very significant: regional factors seem to be important. The pattern seems to be very similar across the two sub-periods, with the exception of the greater downward bias on the lagged dependent variable as a result of the reduction in T. The short-run effects of regional shares are almost identical to that for the whole period at

0.85, the long-run effects are also very similar 0.92 for the whole period, 0.87 for the Cold War and 0.95 for the post Cold War. These coefficients are slightly larger than the long-run effect of the world share in Table 7, which was 0.6.

The growth rate and lagged level of both own and world log GDP are significant. The coefficients of lagged own and lagged world GDP are roughly equal and opposite, which suggests that it may be own GDP relative to the world average which has a positive effect. The coefficients are small so the long-run income elasticity is close to unity, as one would expect. The short-run effect of GDP growth is negative: this may be because military expenditure adjusts slowly, so transitory increases in the growth rate are associated with a falling share, as military expenditure does not keep up.

There is still some cross-section dependence, the CSD test statistics is -3.33 with an average correlation of -0.010 and an average absolute correlation of 0.131. Examining the individual correlations suggests that they are random rather than reflecting omitted strategic interactions. The three largest correlations are: Mauritius-New Zealand (+0.567), Thailand-Bolivia(-0.565), Turkey-Saudi Arabia, (-0.525) and none of the top ten seemed sensible and the correlations may reflect coincidence of outliers.[5] Cases where one might expect arms races showed low correlations: Greece-Turkey -0.291 and India-Pakistan 0.142.

Since the world shares and the US shares were not individually significant, they were dropped and only the change and lagged levels of regional shares and world GDP were retained. The results are in Table 9. The reduction in the log-llikelihood from 3854 to 3534 in the whole period model is very small given that we are saving 4×69 parameters. The CSD results are very similar and, as before, splitting into sub-periods does not improve the fit. The effect of regional shares is slightly smaller, but apart from that, removing world and US shares has not changed the results very much.

6 Conclusions

In this paper, we consider a model of the demand for military expenditure where there is cross-sectional dependence in the errors due to unobserved common factors generated by strategic interactions. If such omitted factors are correlated with the regressors, as they well may be, they will cause the coefficients of those regressors to be biased. The procedures to allow for cross-section dependence, which we describe in detail, are relatively straightforward to implement: add cross-section means or principal components. Either cross-section means or principal components are significant and reduce the degree of cross-section dependence. These are acting as proxies for the unobserved factors causing the cross-section dependence. Clearly there are strong strategic factors driving the shares of military expenditure, so it is

[5] The other seven were Portugal-Burkina Faso; Burkina Faso-Austria; Pakistan-Japan; Belgium-Austria; Colombia-Mali; Colombia-Paraguay; Libya-Malaysia; all with correlations between 0.5 and 0.45.

important that one allows for these factors. These factors seem to be region specific, rather than global and the US does not seem to be acting as a dominant unit. There is also evidence of substantial heterogeneity across countries, so that assuming slope homogeneity as it is done in fixed effect models may be misleading. There is little evidence of a significant structural break at the end of the Cold War.

There are a range of natural extensions. Firstly, to emphasise the role of cross-section dependence and to obtain the largest possible sample, we have used a very simple model with income as the only independent variable. There are many other economic and political measures that have been used to explain military expenditures in the literature and their role could be investigated. Secondly, we have assumed that income is exogenous despite the fact that there is a large literature investigating the effect of military expenditure on growth. This issue could be investigated by estimating a VAR in military expenditures and GDP and testing for the pattern of Granger causality. Thirdly, our results show that there is considerable heterogeneity and there is scope to examine the factors that determine that heterogeneity: are there characteristics of the countries that explain the differences in coefficients? This could be linked to case studies of the individual countries. The extended SIPRI data set opens up the possibility of a large range of quantitative and qualitative studies.

Acknowledgements

This paper is part of a larger SIPRI project supported by Sweden's Riksbankens Jubileumsfond. The authors would like to thank participants at the SIPRI Expert Workshop on Military Expenditure Data (Stockholm) and the Econometrics workshop on Cross-sectional, Spatial Dependence and Heterogeneity in Panel Data Models (Queen Mary, University of London) for stimulating discussions. We are grateful to Hashem Pesaran, Sam Perlo-Freeman, Paul Dunne, Christos Kollias and an anonymous referee for comments. The responsibility for any errors or omissions is our own.

Disclosure Statement

No potential conflict of interest was reported by the authors.

References

Bai, Jushan (2009), 'Panel data models with interactive fixed effects', *Econometrica* **77**(4), 1229–1279.

Brauner, Jennifer (2015), 'Military spending and democracy', *Defence and Peace Economics* **26**(4), 409–423.

Breitung, Jörg and M Hashem Pesaran (2008), Unit roots and cointegration in panels, *in* L.Matyas and P.Sevestre, eds, 'The Econometrics of Panel Data', Springer-Verlag, Berlin.

Cavatorta, Elisa (2010), 'Unobserved common factors in military expenditure interactions across mena countries', *Defence and Peace Economics* **21**(4), 301–316.

Chudik, Alexander and M Hashem Pesaran (2011), 'Infinite-dimensional vars and factor models', *Journal of Econometrics* **163**(1), 4–22.

Chudik, Alexander and M Hashem Pesaran (2013), 'Econometric analysis of high dimensional vars featuring a dominant unit', *Econometric Reviews* **32**(5-6), 592–649.

Chudik, Alexander and M Hashem Pesaran (2015*a*), 'Common correlated effects estimation of heterogeneous dynamic panel data models with weakly exogenous regressors', *Journal of Econometrics* **188**(2), 393–420.

Chudik, Alexander and M Hashem Pesaran (2015*b*), Large panel data models with cross-sectional dependence: a survey, *in* B.Baltagi, ed., 'The Oxford Handbook on Panel Data', Oxford University Press, Oxford, chapter 1.

Coakley, Jerry, Ana-Maria Fuertes and Ron Smith (2006), 'Unobserved heterogeneity in panel time series models', *Computational Statistics & Data Analysis* **50**(9), 2361–2380.

Collier, Paul and Anke Hoeffler (2007), 'Unintended consequences: Does aid promote arms races?', *Oxford Bulletin of Economics and Statistics* **69**(1), 1–27.

Dunne, J Paul and Ron P Smith (2007), The econometrics of military arms races, *in* T.Sandler and K.Hartley, eds, 'Handbook of Defense Economics', Vol. 2, Elsevier, pp. 913–940.

Dunne, J Paul, Sam Perlo-Freeman and Ron P Smith (2008), 'The demand for military expenditure in developing countries: hostility versus capability', *Defence and Peace Economics* **19**(4), 293–302.

Kapetanios, George, M Hashem Pesaran and Takashi Yamagata (2011), 'Panels with non-stationary multifactor error structures', *Journal of Econometrics* **160**(2), 326–348.

Murdoch, James C (1995), 'Military alliances: Theory and empirics', *Handbook of Defense Economics* **1**, 89–108.

Pesaran, M Hashem (2004), 'General diagnostic tests for cross section dependence in panels', *CESifo working paper series* .

Pesaran, M Hashem (2006), 'Estimation and inference in large heterogeneous panels with a multifactor error structure', *Econometrica* **74**(4), 967–1012.

Pesaran, M Hashem (2015), *Time-series and panel data econometrics*, Oxford University Press, Oxford.

Pesaran, M Hashem and Elisa Tosetti (2011), 'Large panels with common factors and spatial correlation', *Journal of Econometrics* **161**(2), 182–202.

Pesaran, M Hashem and P. Ron Smith (1995), 'Estimating long-run relationships from dynamic heterogeneous panels', *Journal of econometrics* **68**(1), 79–113.

Smith, Ron (1995), The demand for military expenditure, *in* K.Hartley and T.Sandler, eds, 'Handbook of defense economics', Vol. 1, Elsevier, Amsterdam, pp. 69–87.

Table 1: Descriptive Statistics

	N group	Obs	Mean	Std.Dev.	Min	Max
Entire panel (1965-2014)						
burden	172	6,420	0.029608	0.035163	0	1.173498
milex (billion)	172	6,372	9.108008	45.84714	0	720.2188
GDP (billion)	172	8,600	32652.5	511791.7	0	2.72E+07
Balanced panel						
burden	70	3,500	0.028917	0.029275	0.001399	0.304638
milex (billion)	70	3,412	14.75106	60.63815	0.001655	720.2188
GDP (billion)	70	3,500	15802.79	92983.97	1.55E-11	150286
Military burden by region						
Africa	19	950	0.023637	0.022778	0.001399	0.267347
Americas	16	800	0.020704	0.013485	0.003482	0.090634
Asia	11	550	0.027417	0.015566	0.005447	0.069917
Europe	17	850	0.022885	0.011169	0.004746	0.059623
Middle East	6	300	0.087121	0.057432	0.015873	0.304638

Table 2: Cumulative Proportions of Military Burden explained by Principal Components

	All	Cold War period	Post-Cold War	Africa	Americas	Asia	Europe	Mid East
N	72	72	72	20	16	11	17	7
T	65-14	65-88	89-14	65-14	65-14	65-14	65-14	65-14
PC1	0.4982	0.4195	0.5711	0.374	0.4293	0.6654	0.7929	0.5961
PC2	0.672	0.5976	0.7068	0.6008	0.6414	0.783	0.8827	0.8021
PC3	0.7397	0.7323	0.7993	0.6854	0.7479	0.8475	0.9389	0.8963
PC4	0.7899	0.7928	0.8432	0.7531	0.8134	0.8974	0.9569	0.9518
PC5	0.8241	0.8398	0.8659	0.8102	0.863	0.9282	0.9711	0.9776

Table 3: Cumulative Proportions of (log) Military Burden/Expenditure/GDP explained by Principal Components

	log(burden)	log(milex)	log(GDP)
N	70		
T	1965-2014		
PC1	0.548	0.6141	0.9491
PC2	0.7146	0.7782	0.9717
PC3	0.7775	0.849	0.9913
PC4	0.8263	0.8897	0.995
PC5	0.861	0.9206	0.9972

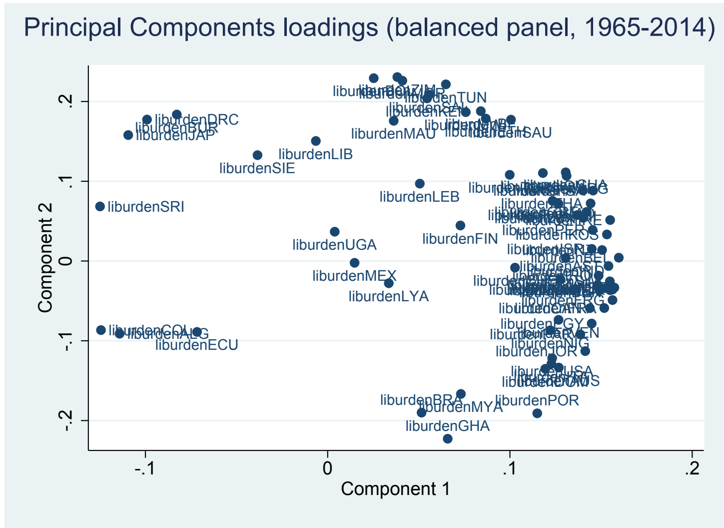

Figure 1: Principal Components (balanced panel, 1965-2014)

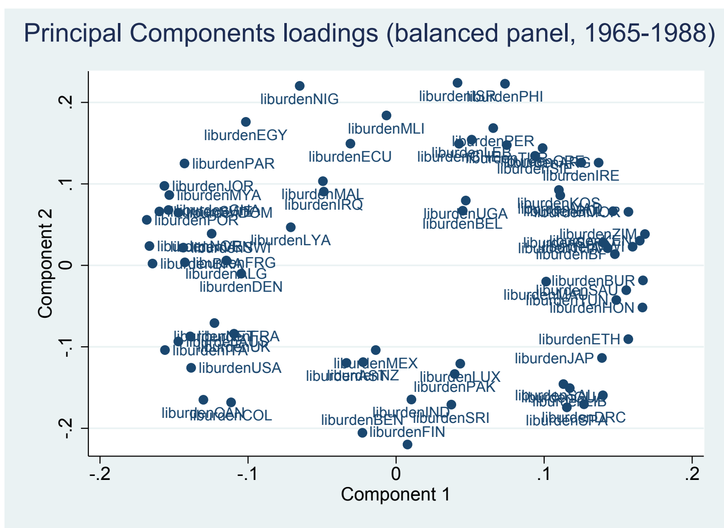

Figure 2: Principal Components pre-Cold War period (balanced panel, 1965-1988)

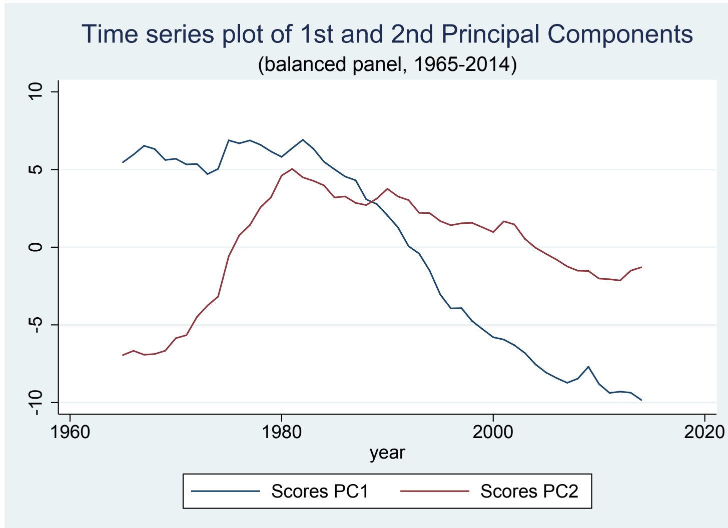

Figure 3: Times series of Principal Components (balanced panel, 1965-2014)

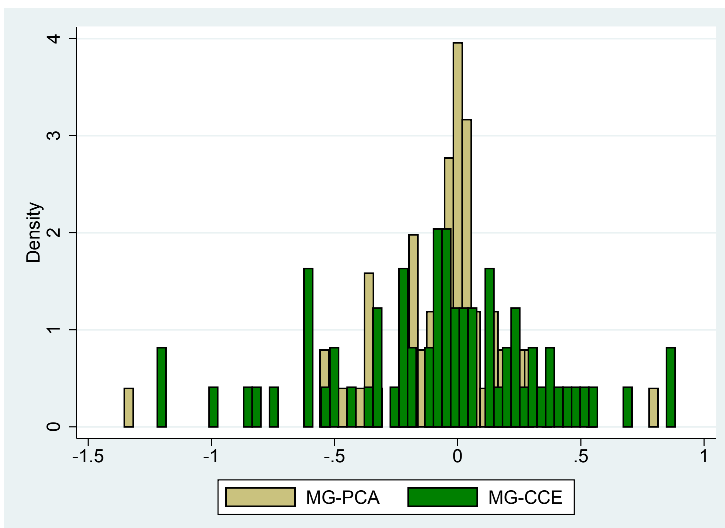

Figure 4: Density of income coefficients by methods (PCA and CCE)

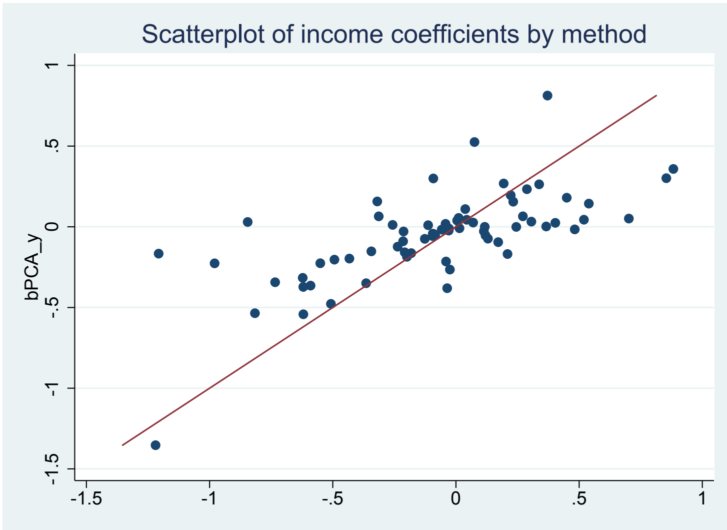

Figure 5: Scatterplot of income coefficients by methods (PCA on y-axis and CCE on x-axis)

Table 6: Goodness-of-fit statistics for static models, entire period and sub-periods

	MG-CCE	MG-PC	MG	FE2	FE1
	Estimation for the entire period				
SumLL	1755	1907	277	-1379	-1707
NT	3430	3430	3430	3430	3430
k	280	280	140	119	71
BIC	615	767	-293	-1863	-1996
	Estimation by sub-period				
SumLL	3017	3273	1942	-613	-766
NT	3430	3430	3430	3430	3430
k	560	560	280	189	142
BIC	738	**994**	802	-1382	-1343

Table 8: Goodness-of-fit statistics for dynamic models, entire period and sub-periods

	MG-CCE	MG-PC	MG	FE2	FE1
	Estimation for the entire period				
MLL	3511	3576	2974	1606	1535
NT	3430	3430	3430	3430	3430
k	560	560	280	121	73
BIC	1232	1297	**1834**	1114	1238
	Estimation by sub-period				
MLL	4492	4458	3631	1690	1642
NT	3430	3430	3430	3430	3430
k	1120	1120	560	193	146
BIC	-67	-100	1352	904	1048

Table 4: Mean-group PCA

	All: 1966-2014	Cold War: 1966-1988	Post Cold War: 1989-2014	Africa	Americas	Asia	Europe	Middle East
log(GDP)	-0.0454	-0.2401***	-0.1416**	-0.0813	0.03811	0.0131	-0.0569	-0.0487
	(0.033)	(0.045)	(0.055)	(0.059)	(0.031)	(0.027)	(0.036)	(0.051)
Factor 1	0.0257***	0.0382***	0.0135*	0.0634***	0.0849***	0.0643*	0.0607***	0.1557***
	(0.008)	(0.008)	(0.006)	(0.020)	(0.025)	(0.037)	(0.007)	(0.024)
Factor 2	0.01722**	-0.0074	0.0021	0.0351	-0.0016	0.0186	0.01871	-0.0382
	(0.008)	(0.005)	(0.005)	(0.033)	(0.028)	(0.037)	(0.024)	(0.025)
constant	1.7092**	6.637***	4.2782***	2.2638*	-0.3996	0.3825	2.195**	3.0331***
	(0.823)	(1.039)	(1.493)	(1.261)	(0.760)	(0.782)	(0.970)	(1.191)
N	3430	1610	1820	931	784	539	833	294
N groups	70	70	70	19	16	11	17	6
Periods	49	23	26	49	49	49	49	49
Wald	42.78	31.91	31.95	12.51	15.12	9.28	139.09	101.43
Sum LL	1907	1461	1812	119	286	470	1107	159

Baseline mean-group estimator without factors

	All: 1966-2014	Cold War: 1966-1988	Post Cold War: 1989-2014	Africa	Americas	Asia	Europe	Middle East
log(GDP)	-0.124173***	0.0489	-0.2897***	-0.0967	-0.0748***	-0.0864*	-0.2396***	-0.1516**
	(0.022)	(0.034)	(0.038)	(0.059)	(0.024)	(0.044)	(0.036)	(0.058)
constant	3.887287***	-0.1286	8.2048***	2.7751**	2.4401***	3.134**	6.983***	5.5318***
	(0.530)	(0.801)	(0.961)	(1.296)	(0.683)	(1.235)	(0.988)	(1.340)
N	3500	1610	1820	931	784	539	833	294
N groups	70	70	70	19	16	11	17	6
Periods	50	23	26	49	49	49	49	49
Wald	32.988	1.96	56.965	2.55	9.48	3.80	44.16	6.84
Sum LL	277	770	1172	-375	10	188	429	46

Notes: Factor 1 and Factor 2 are the first and second principal component scores obtained from the military expenditure shares within each group. The Pesaran (2004) tests for cross-sectional dependence (and their p-values) on the PCs models are as follows. Entire period: 0.405 (0.685); Cold War period 1966-88: 14.859 (0.000); post Cold War 1989-2014: 3.750 (0.000) . Without correcting for CSD these, the CSD test (and p-values) are: Entire period: 62.468 (0.000); Cold War period 1966-88: 9.146 (0.000); post Cold War 1989-2014: 8.240 (0.000).

Table 5: MEAN-GROUP CCE

	All: 1966-2014	Cold War: 1966-1988	Post Cold War: 1989-2014	Africa	Americas	Asia	Europe	Middle East
log(GDP)	-0.0614	-0.4890***	-0.2342**	-0.1478	-0.0745*	0.0342	-0.1942	-0.1293
	(0.051)	(0.077)	(0.098)	(0.175)	(0.046)	(0.116)	(0.118)	(0.087)
log(burden) avg	0.9944***	1.3084***	0.9175***	0.9896***	1.0653**	1.110***	0.9557***	0.938505***
	(0.188)	(0.302)	(0.333)	(0.255)	(0.474)	(0.359)	(0.162)	(0.208)
log(GDP) avg	0.0123	0.3339***	0.2150**	0.0553	0.0416	-0.0275	0.1756	0.082387**
	(0.041)	(0.061)	(0.094)	(0.174)	(0.039)	(0.131)	(0.116)	(0.033)
constant	0.9658	3.811***	0.2487	1.873	0.9248	-0.1804	0.5358	1.269784
	(0.682)	(1.030)	(2.241)	(1.310)	(1.090)	(1.444)	(1.505)	(2.058)
N	3430	1610	1820	931	784	539	833	294
N groups	70	70	70	19	16	11	17	6
Periods	49	23	26	49	49	49	49	49
Wald	43.35	39.63	16.05	12.51	9.437	14.815	97.338	73.763
Sum LL	1755	1344	1673	-133	195	398	1065	154

Notes: The Pesaran (2004) tests for cross-sectional dependence (and their p-values) on the CCE models are as follows. Entire period: 5.760 (0.000); Cold War period 1966-88: 1.101 (0.271) ; post Cold War 1989-2014: 4.513 (0.000).

Table 7: Dynamic models correcting for cross-sectional dependence, Mean-Group estimator

	Common Correlated Effect				Principal Components		
	All period: 1966-2014	Cold-War: 1966-1988	Post Cold-War:1989-2014		All period:1966-2014	Cold-War: 1966-1988	Post Cold-War:1989-2014
	(1)	(2)	(3)		(4)	(5)	(6)
$shares_{t-1}$	-0.346182***	-0.537727***	-0.535553***	$shares_{t-1}$	-0.360873***	-0.600078***	-0.567210***
	(0.020)	(0.036)	(0.034)		(0.021)	(0.033)	(0.035)
Δy_t	-0.575561***	-0.773903***	-0.609458***	Δy_t	-0.453167***	-0.743175***	-0.483785***
	(0.062)	(0.102)	(0.074)		(0.048)	(0.078)	(0.067)
y_{t-1}	0.041528*	-0.053314	-0.033627	y_{t-1}	-0.012754	-0.122054***	-0.048003
	(0.024)	(0.085)	(0.057)		(0.014)	(0.033)	(0.041)
$\Delta s\bar{h}ares_t$	0.841572***	1.098419***	0.883816***	Δf_1	0.025485***	0.011801	0.024136***
	(0.157)	(0.222)	(0.254)		(0.006)	(0.010)	(0.007)
$s\bar{h}ares_{t-1}$	0.206372**	0.729032***	0.596947***	$f1, t-1$	0.010922***	0.021640***	0.008144
	(0.081)	(0.229)	(0.231)		(0.004)	(0.006)	(0.005)
$\Delta\bar{y}$	0.183700***	0.477572***	0.301562**	Δf_2	0.018661**	0.007614	0.008434*
	(0.068)	(0.131)	(0.141)		(0.007)	(0.005)	(0.005)
$\bar{y}_{t-1}$	-0.032691*	0.009545	0.069088	$f2, t-1$	0.005316*	0.000372	0.000927
	(0.019)	(0.066)	(0.066)		(0.003)	(0.004)	(0.003)
constant	-0.180541	1.324848	-1.044626	constant	0.634145*	3.651121***	1.764053
	(0.313)	(0.879)	(1.501)		(0.345)	(0.784)	(1.106)
Observations	3,430	1,610	1,820		3,430	1,610	1,750
LL	3511	1991	2501		3576	1978	2480
LL no correction for CS dependence $*$	2974	1545	2086		2974	1545	2086
LL ECM Fixed Effect	1606	739	951		1591	723	917
Periods	49	23	26		49	23♮	25♮
CSD-test (Pesaran, 2004) p-value	0.034	0.153	0.327		0.447	0.000	0.639
Abs(corr)	0.126	0.180	0.167		0.127	0.207	0.185

Notes: Mean-Group estimator, standard errors of coefficients in parentheses. ♮– we loose one observation in each sub-period since we need the first difference in the factors, Δf_1 and Δf_2, estimated by PCA by Cold War period. In CCE models, the first difference of the cross-section averages in the period 1989-2014 is always available. The statistics marked with $*$ indicates log-likelihood from a dynamic model without correction for cross-sectional dependence. The Pesaran (2004) tests for cross-sectional dependence (and their p-values) on the dynamic models not correcting for CSD are: all period 1966-2014: 18.310 (0.000); Cold War period 1966-88: 11.163 (0.000); post Cold War 1989-2014: 4.77 (0.000). The average absolute correlations are: all period 1966-2014, 0.137; Cold War period 1966-88: 0.192; post Cold War 1989-2014: 0.173.

Table 9: Dynamic Models with regional share averages (ECM form)

	USA as dominant unit			Restricted models		
	All period: 1966-2014	Cold-War: 1966-1988	Post Cold-War:1989-2014	All period:1966-2014	Cold-War: 1966-1988	Post Cold-War:1989-2014
	(1)	(2)	(3)	(4)	(5)	(6)
$shares_{t-1}$	-0.434***	-0.750***	-0.715***	-0.344***	-0.570***	-0.514***
	(0.023)	(0.043)	(0.038)	(0.020)	(0.036)	(0.032)
$\Delta \bar{shares}^{r}_{t}$	0.853***	0.860***	0.847***	0.872***	0.891***	0.812***
	(0.109)	(0.170)	(0.145)	(0.095)	(0.124)	(0.131)
$\bar{shares}^{r}_{t-1}$	0.400***	0.653***	0.679***	0.274***	0.554***	0.330***
	(0.082)	(0.189)	(0.178)	(0.052)	(0.109)	(0.100)
$\Delta \bar{shares}^{w}_{t}$	0.048	0.152	0.021			
	(0.170)	(0.274)	(0.243)			
$\bar{shares}^{w}_{t-1}$	-0.029	0.223	-0.004			
	(0.123)	(0.381)	(0.315)			
$\Delta shares^{USA}_{t}$	-0.004	-0.013	-0.011			
	(0.036)	(0.067)	(0.074)			
$shares^{USA}_{t-1}$	-0.017	-0.089	-0.048			
	(0.025)	(0.084)	(0.056)			
Δy_t	-0.431***	-0.677***	-0.471***	-0.454***	-0.682***	-0.542***
	(0.052)	(0.116)	(0.071)	(0.052)	(0.104)	(0.058)
y_{t-1}	0.055**	-0.120	-0.043	0.068***	-0.038	-0.007
	(0.026)	(0.102)	(0.076)	(0.021)	(0.074)	(0.055)
$\Delta \bar{y}$	0.162**	0.294*	0.243**	0.149**	0.320**	0.254***
	(0.072)	(0.153)	(0.115)	(0.065)	(0.130)	(0.099)
$\bar{y}_{t-1}$	-0.049**	0.026	0.041	-0.055***	0.036	0.001
	(0.021)	(0.083)	(0.079)	(0.017)	(0.065)	(0.057)
constant	-0.171	2.706**	0.348	-0.347	0.229	0.229
	(0.385)	(1.322)	(1.665)	(0.244)	(0.833)	(1.027)
Observations	3381	1587	1794	3381	1587	1794
N groups	69	69	69	69	69	69
Periods	49	23	26	49	23	26
LL	3854	2462	2897	3534	1994	2479
CSD-test (Pesaran, 2004)	-3.333	-1.23	-2.041	-3.052	-0.304	-1.811
CSD-test (Pesaran, 2004) p-value	0.0009	0.219	0.041	0.002	0.761	0.070
Abs(corr)	0.131	0.212	0.195	0.125	0.188	0.170
Benchmark test (Pesaran, 2004)⋆	18.11	11.269	4.742	18.11	11.269	4.742
Benchmark test (Pesaran, 2004) p-value	0.000	0.000	0.000	0.000	0.000	0.000
Abs(corr) no CSD correction	0.137	0.192	0.171	0.137	0.192	0.171

Notes: Mean-Group estimator, standard errors of coefficients in parentheses. $\bar{shares}^{r}$ and $\Delta \bar{shares}^{r}$ indicate the average *regional* shares of military expenditures, where the relevant region is the region of country i, and its change, respectively. The statistics marked with $*$ indicate the Pesaran (2004)'s test for CSD for cross-sectional dependence (and their p-values) on the dynamic models not correcting for CSD. These tests are computed on the sample excluding USA ($N = 69$) since the dynamics models with dominant unit have $N = 69$ groups.

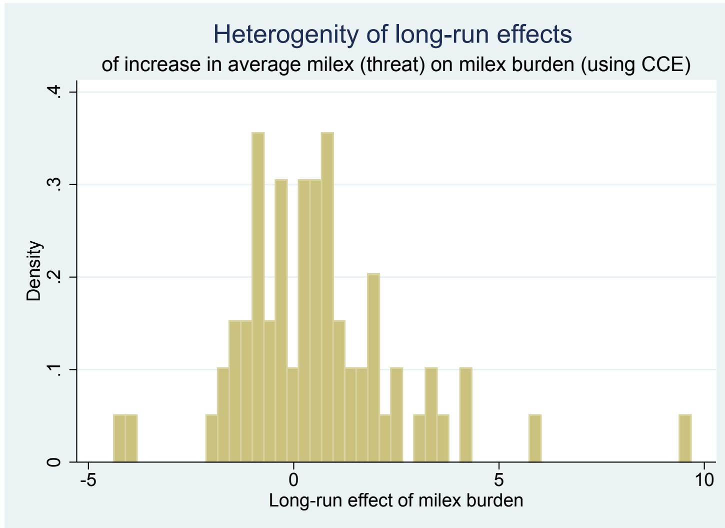

Figure 6: Histogram of log-run burden effect