Bayesian Estimation of Risk-Premia in an APT Context

Theofanis Darsinos and Stephen E. Satchell

June 2003

DAE Working Paper No. 0329

Bayesian Estimation of Risk-Premia in an APT Context

Theofanis Darsinos and Stephen E. Satchell*
Faculty of Economics, University of Cambridge, UK.

May 28, 2003

Abstract

Recognizing the problems of estimation error in computing risk premia via arbitrage pricing, this paper provides a Bayesian methodology for estimating factor risk premia and hence equity risk premia for both traded and non-traded factors. Some illustrative calculations based on UK equity are also provided.

1 Introduction

The calculation of factor-risk premia is one of the major contributions of Arbitrage Pricing Theory as espoused by Ross (1976) and Ingersoll (1987). In this literature, two cases are considered; when the factors are traded portfolios and when they are not (see for example Cambpell, Lo and Mackinlay (1997)). Whilst practitioner oriented models focus on the former, the academic literature is more concerned with the latter (see Burmeister and McElroy (1988)). There are considerable problems in estimating factor-risk premia, as discussed in Pitsillis and Satchell (2001), Pitsillis (2002), and elsewhere. To alleviate some of the estimation problems, we consider a Bayesian approach to estimation, so that prior information can be utilised to improve accuracy. Many authors (see for example Polson and Tew (2000), Ericsson and Karlsson (2002)) have shown that Bayesian approaches to linear factor models and portfolio theory have been successful in reducing some of the excess variability in the data.

Regarding the choice of factors, interest rates, returns on broadbased portfolios one of which, typically, approximates the market portfolio, growth in consumption as tabulated for example from inflationary data, production and other macroeconomic variables that measure the state of the economy are potential risks that are rewarded in the stock market and could be included in a

*We thank Marios Pitsillis for providing us with the data for the empirical application in the paper.

factor model. Furthermore, variables that signal changes in the future, such as term premiums, credit spreads, etc. are also reasonable to include (see Ericsson and Karlsson (2002). Our focus is on the case of non-traded factors, so that our factors are macroeconomic ones. Several authors have used macroeconomic variables as factors (see for example Jagannathan and Wang (1996), Reyfman (1997)).Chen et al. (1986) test whether innovations in macroeconomic variables are risks that are rewarded in the stock market. Included variables are: the spread between long and short interest rates, expected and unexpected inflation, industrial production, the spread between high and low-grade bonds, market portfolio, aggregate consumption and oil prices. Other macro-economic variables have also been considered.

Fama and French (1992,1993,1996) advocate a model with the market return, the return of small less big stocks (SMB) and the spread between high and low book-to-market stocks (HML) as factors. However, although empirically very successful, the nondiversifiable risk that is proxied by the returns of the HML and SMB is not clear.

The structure of the paper is as follows. In section 2 we briefly introduce the general APT framework. Section 2.1 outlines the excess return generating process when factors are traded portfolios and suggests how a Bayesian estimation framework can be utilised in this case. Section 2.2 considers the case of non-traded or macroeconomic factors and section 3 derives the prior and posterior estimates for the (non-traded) factor risk-premia. In section 4 we provide an empirical application to illustrate how the methodology developed in the paper could be utilised in practice. Concluding comments follow in section 5.

2 The General APT Framework

2.1 The Excess Return Generating Process (when factors are traded portfolios)

We have N assets. Then for each asset i, its excess return x^i is generated by:

$$x^i = \sum_{j=1}^{K} \beta_{ij} f_j + \varepsilon^i \tag{1}$$

where

$f = (f_1, ..., f_K)'$
$b_i = (\beta_{i1}, ..., \beta_{iK})'$
$E(\varepsilon^i) = 0$
$E(f\varepsilon^i) = 0$

The f_j's are the factors, K factors in total. The β_{ij} are the betas or factor loadings, and the ε^i 's are white noise errors. Exact factor pricing implies that in the case of traded factors the intercept term in the factor model (i.e. equation 1) is zero. In this case the risk premia on the factors λ_k can be estimated directly from the sample means of the excess returns on the traded portfolios that constitute the factors. Thus

$$\widehat{\lambda}_k = \overline{x}_k$$

x_k represents expected excess return from the portfolio that mimics perfectly factor k. Note that as the sample size n increases we expect

$$\overline{x}_k \sim N(\mu_k, \frac{\sigma_k^2}{n})$$

via the usual central limit theorems.

$N(\cdot)$ denotes the Normal distribution. μ_k and σ_k^2 are the mean and variance respectively of the factor mimicking portfolio for factor k. For this model, unlike the case where factors are non-traded, the APT restriction $\mu^i = \sum \beta_{ij}\lambda_j$ holds exactly upon taking the expected value of equation (1).

2.1.1 A Bayesian Framework

When portfolios are factors it is quite straightforward to introduce a Bayesian framework in the estimation of risk premia. This case strongly resembles the popular Black-Litterman model used in tackling asset management problems (see Black and Litterman (1991), (1992)).

We wish to obtain the posterior probability density function (*pdf*) for μ_k which will then gives us directly a posterior estimate for the risk premium λ_k. We assume that observations on x_k are drawn from a normal population with unknown mean μ_k and known variance σ_k^2. As regards a prior *pdf* for μ_k we assume using standard Bayesian methodology (see for example Zellner (1971), Bauwens, Lubrano, and Richard (1999) that

$$\mu_{\Pr ior} = \begin{pmatrix} \mu_1 \\ \vdots \\ \mu_K \end{pmatrix} \sim N\left[\begin{pmatrix} \widetilde{\mu}_1 \\ \vdots \\ \widetilde{\mu}_K \end{pmatrix}, \begin{pmatrix} \widetilde{\sigma}_1^2 & \cdots & 0 \\ \vdots & \ddots & 0 \\ 0 & 0 & \widetilde{\sigma}_K^2 \end{pmatrix}\right]$$

The values of these parameters are assigned by the investigator on the basis of his/her initial information. Prior values of risk premia are simply in this case

$$\lambda_{\Pr ior} = \begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_K \end{pmatrix} = \begin{pmatrix} \widetilde{\mu}_1 \\ \vdots \\ \widetilde{\mu}_K \end{pmatrix}$$

Then using Bayes's theorem to combine the likelihood function with the prior *pdf* to obtain the posterior *pdf* for μ_k $(k = 1, ..., K)$ we have that

$$\mu_k^{Posterior} \sim N\left(\frac{\overline{x}_k(\sigma_k^2/n)^{-1} + \widetilde{\mu}_k(\widetilde{\sigma}_k^2)^{-1}}{(\sigma_k^2/n)^{-1} + (\widetilde{\sigma}_k^2)^{-1}}, \frac{1}{(\sigma_k^2/n)^{-1} + (\widetilde{\sigma}_k^2)^{-1}}\right)$$

Hence a posterior estimate of the risk premium, λ_k, for factor k is given by the mean of $\mu_k^{Posterior}$.

Extensions for the cases of correlated factors and/or factors where the variance-covariance matrix of the factor mimicking portfolios is unknown and stochastic are straightforward to derive. For example a diffuse prior (i.e. $pdf(\mu, \Sigma) \propto |\Sigma|^{-(k+1)/2}$) or a Normal-Inverted Wishart prior (i.e. $pdf(\mu|\Sigma)$ ~Normal and $pdf(\Sigma)$ ~Inverted Wishart) both lead to matrixvariate t distributions for the posterior (i.e. $pdf(\mu|\Sigma)$ ~matrixvariate t). See for example Satchell and Scowcroft (2000) for an illustration on incorporating stochastic volatility in Black-Litterman type models.

2.2 The Excess Return Generating Process (when factors are macroeconomic variables or non-traded portfolios)

This time for each asset i, the excess return x^i is generated by:

$$\begin{aligned} x^i &= \mu_i + \sum_{j=1}^{K} \beta_{ij} f_j + \varepsilon^i \\ &= \mu_i + b_i' f + \varepsilon^i \\ &= a_i' h + \varepsilon^i \end{aligned} \tag{2}$$

where $h = (f_1, ..., f_K, 1)'$, $a_i = (\beta_{i1}, ..., \beta_{iK}, \mu_i)'$, $E(\varepsilon^i) = 0$, $E(f\varepsilon^i) = 0$

The factors f_j's are now constructed so that they have zero-mean (they are in fact factor deviations from their mean) and μ_i denotes the expected excess return for asset i.(i.e. $\mu_i = E(x^i)$). To generalize the above setting for our N assets we write:

$$\begin{bmatrix} x^1 \\ \vdots \\ x^N \end{bmatrix} = \begin{bmatrix} \beta_{1,1} & \cdots & \beta_{K,1} & \mu_1 \\ \vdots & \ddots & \vdots & \vdots \\ \beta_{1,N} & \cdots & \beta_{K,N} & \mu_N \end{bmatrix} \begin{bmatrix} f_1 \\ \vdots \\ f_K \\ 1 \end{bmatrix} + \begin{bmatrix} \varepsilon^1 \\ \vdots \\ \varepsilon^N \end{bmatrix}$$

$$\Rightarrow$$

$$x = A'h + \varepsilon \tag{3}$$

where

$$\varepsilon \backsim NID_N(0, \Psi), \tag{4}$$

NID : Independent Normal, Ψ is the positive-definite $N \times N$ residual variance-covariance matrix and A is the regression coefficients matrix:

$$A = \begin{bmatrix} B \\ \mu' \end{bmatrix} = \begin{bmatrix} \beta_{1,1} & \cdots & \beta_{1,N} \\ \vdots & \ddots & \vdots \\ \beta_{K,1} & \cdots & \beta_{K,N} \\ \mu_1 & \cdots & \mu_N \end{bmatrix} \tag{5}$$

The matrix version of the above multivariate regression for T observations is:

$$\begin{aligned} \underset{T\times N}{X} &= \underset{T\times(K+1)}{H} \times \underset{(K+1)\times N}{A} + \underset{T\times N}{E} \\ &= [F, i_T]\begin{bmatrix} B \\ \mu' \end{bmatrix} + E \end{aligned} \tag{6}$$

where $E \backsim MN_{T\times N}(0, \Psi \otimes I_T)$ and $X, H,$ and E are obtained by stacking the row vectors x', h', and ε' respectively. MN denotes a matrixvariate normal distribution[1]. F is obtained by stacking together $f'$and is the factor data matrix.

2.3 Obtaining the $(K \times 1)$ Vector of Risk Premia λ

The APT pricing relationship is

$$E(x^i) = \mu_i = \lambda_1\beta_{i1} + ... + \lambda_K\beta_{iK} = b_i'\lambda \tag{7}$$

where $\lambda = (\lambda_1, ..., \lambda_K)'$; i.e. a $K \times 1$ vector of factor risk-premia. We write this as an equality although it is only approximately equal in the case of non-traded factors. This then implies that

$$\mu = B'\lambda \tag{8}$$

μ is a $(N \times 1)$ vector of expected excess returns for our N assets and B is a $(K \times N)$ matrix of betas. Then the best (cross-sectional) linear predictor for the column vector of excess returns μ, is obtained when

$$E[B(\mu - B'\lambda)] = 0 \tag{9}$$

This implies that

$$\begin{aligned} E[B\mu] - E[BB']\lambda &= 0 \\ \lambda &= [E(BB')]^{-1}E[B\mu] \end{aligned} \tag{10}$$

where

$$E(BB') = \begin{bmatrix} \sum_{i=1}^{N}\left[Var(\beta_{1,i}) + [E(\beta_{1,i})]^2\right] & \cdots & \sum_{i=1}^{N}\left[Cov(\beta_{1,i}, \beta_{K,i}) + E(\beta_{1,i})E(\beta_{K,i})\right] \\ \vdots & \ddots & \vdots \\ \sum_{i=1}^{N}\left[Cov(\beta_{1,i}, \beta_{K,i}) + E(\beta_{1,i})E(\beta_{K,i})\right] & \cdots & \sum_{i=1}^{N}\left[Var(\beta_{K,i}) + [E(\beta_{K,i})]^2\right] \end{bmatrix}$$

$\Rightarrow$

$$E(BB') = \begin{bmatrix} tr\Omega_{11} + \sum_{i=1}^{N}[E(\beta_{1,i})]^2 & \cdots & tr\Omega_{1K} + \sum_{i=1}^{N}E(\beta_{1,i})E(\beta_{K,i}) \\ \vdots & \ddots & \vdots \\ tr\Omega_{1K} + \sum_{i=1}^{N}E(\beta_{1,i})E(\beta_{K,i}) & \cdots & tr\Omega_{KK} + \sum_{i=1}^{N}[E(\beta_{K,i})]^2 \end{bmatrix}$$

[1]See the appendix for a definition of the Matrixariate Normal distribution. For more information we refer the reader to Bauwens, Lubrano and Richard (1999).

and

$$E[B\mu] = \begin{bmatrix} \sum_{i=1}^{N} \left[Cov(\beta_{1,i}, \mu_i) + E(\beta_{1,i})E(\mu_i)\right] \\ \vdots \\ \sum_{i=1}^{N} \left[Cov(\beta_{K,i}, \mu_i) + E(\beta_{K,i})E(\mu_i)\right] \end{bmatrix} \tag{11}$$

$$= \begin{bmatrix} tr\Omega_{1\mu} + \sum_{i=1}^{N} E(\beta_{1,i})E(\mu_i) \\ \vdots \\ tr\Omega_{K\mu} + \sum_{i=1}^{N} E(\beta_{K,i})E(\mu_i) \end{bmatrix}$$

Note that the Ω's are $N \times N$ matrices obtained from the $(K+1)N \times (K+1)N$ variance-covariance matrix Σ_A of the regression coefficients matrix A:

$$\Sigma_A = \begin{bmatrix} \Omega_{11} & \cdots & \Omega_{1K} & \Omega_{1\mu} \\ \vdots & \ddots & \vdots & \vdots \\ \Omega_{K1} & \cdots & \Omega_{KK} & \Omega_{K\mu} \\ \Omega_{\mu 1} & \cdots & \Omega_{\mu K} & \Omega_{\mu\mu} \end{bmatrix} \tag{12}$$

3 Introducing a Bayesian Framework using a Minnesota Prior (Litterman's Prior)

We start by assuming that the residual variance-covariance matrix Ψ defined in equation (13) is fixed and diagonal:

$$\Psi = \begin{bmatrix} \psi_{11} & 0 & \cdots & 0 \\ 0 & \psi_{22} & \cdots & 0 \\ \vdots & \cdots & \ddots & \vdots \\ 0 & 0 & \cdots & \psi_{NN} \end{bmatrix} \tag{13}$$

This of course implies that the diagonal elements of Ψ, need to be specified. We shall follow the empirical Bayes approach where the ψ_{ii} are replaced by s_i^2, the sample residual variance estimates.

It is possible to relax the assumption that Ψ is fixed and diagonal and work with Ψ non-diagonal. This is the approach followed by Chamberlain and Rothschild (1983), Ingersoll (1984) and Connor and Korajczyk (1993). These authors allow for non-diagonality in the variance-covariance matrix which will disappear as N gets large. The strict factor model we assume is more in line with practitioner factor models where the emphasis is on making clear the distinction between systematic risk (through common factors) and idiosyncratic risk (stock-specific risk). In any case we shall outline the appropriate extensions to our results when Ψ is non-diagonal.

3.1 Prior Estimates of the Risk Premia

In Bayesian statistics, parameters are treated as random variables and are assigned probability distributions. We assume that the prior distribution of the regression parameters is

$$A \backsim MN\left[Vec\widetilde{A}, \widetilde{\Sigma}\right] \tag{14}$$

where MN denotes the matrixvariate Normal distribution (see the appendix for a definition). $Vec\widetilde{A}$ denotes the prior mean of A and $\widetilde{\Sigma}$ the prior variance-covariance matrix of the regression parameters. In partitioned form we have

$$\begin{bmatrix} B \\ \mu' \end{bmatrix} \backsim MN\left[\begin{pmatrix} Vec\widetilde{B} \\ \widetilde{\mu} \end{pmatrix}, \begin{bmatrix} \Theta_{BB} & \Theta_{B\mu} \\ \Theta_{\mu B} & \theta_{\mu\mu} \end{bmatrix} \otimes I_N\right] \tag{15}$$

where we assume that

$$\Theta_{BB} = \begin{bmatrix} \theta_{11} & \cdots & \theta_{1K} \\ \vdots & \ddots & \vdots \\ \theta_{K1} & \cdots & \theta_{KK} \end{bmatrix}$$

i.e. a $(K \times K)$ matrix and

$$\Theta_{B\mu} = \begin{bmatrix} \theta_{1\mu} \\ \vdots \\ \theta_{K\mu} \end{bmatrix}$$

i..e a $(K \times 1)$ column vector and $\theta_{\mu\mu}$ is a scalar. This assumption implies that equation (12) above is specialised so that $\Omega_{11} = \theta_{11} I_N$, $\Omega_{12} = \theta_{12} I_N$,...etc. To clarify what this means, we shall consider a simple example. Suppose we have a classic Fama and French three factor model (i.e.$K = 3$: Market portfolio, Size ranking, and Book to Market ratio. Then for each stock we have betas for each of these 3 factors. Assuming that $\Omega_{11} = \theta_{11} I_N$ is tantamount to assuming that all the betas with respect to the market are drawn from a common population with variance θ_{11}. Likewise $\Omega_{12} = \theta_{12} I_N$ means that the betas from the market and the size factors have a common covariance θ_{12}. Such an assumption is both helpful for interpretation and also leads to empirical Bayes analysis. Similar assumptions are made for the means. In particular we assume that the betas of all assets with respect to each factor k are drawn from a common population with mean $\widetilde{\beta}_k$. Thus

$$\begin{aligned} Vec\widetilde{B} &= \begin{bmatrix} \widetilde{\beta}_1 & \cdots & \widetilde{\beta}_1 \ , \cdots\cdots , & \widetilde{\beta}_K & \cdots & \widetilde{\beta}_K \end{bmatrix}' \\ &= \begin{bmatrix} \widetilde{\beta}_1 & \cdots & \widetilde{\beta}_K \end{bmatrix}' \otimes i_N \end{aligned}$$

and finally

$$\widetilde{\mu} = \begin{bmatrix} \mu_0 & \cdots & \mu_0 \end{bmatrix}' = \mu_0 \otimes i_N$$

We stress that the above assumptions are standard in Bayesian finance and particularly helpful in understanding cross sectional analyses. So changes in θ_{11} would represent a change in the volatility exposure of factor 1.

3.1.1 Standard Bayesian Case: (Prior independence between B and μ; i.e. $\Theta_{B\mu} = 0$)

Here we assume that $\Theta_{B\mu} = 0$. This prior assumption would be held by an investor who was sceptical that the β_{ij}'s influenced μ^i and expresses a disbelief in linear factor modelling and the APT. Subcases where a subset of $\Theta_{B\mu}$ is set to zero could also be considered.

The prior estimate for the risk premium vector λ is then obtained from equation (10), where

$$E(BB') = \begin{bmatrix} N\theta_{11} + \sum_{i=1}^{N} [E(\beta_{1,i})]^2 & \cdots & N\theta_{1K} + \sum_{i=1}^{N} E(\beta_{1,i})E(\beta_{K,i}) \\ \vdots & \ddots & \vdots \\ N\theta_{1K} + \sum_{i=1}^{N} E(\beta_{1,i})E(\beta_{K,i}) & \cdots & N\theta_{KK} + \sum_{i=1}^{N} [E(\beta_{K,i})]^2 \end{bmatrix}$$

and

$$E[B\mu] = \begin{bmatrix} \sum_{i=1}^{N} E(\beta_{1,i})E(\mu_i) \\ \vdots \\ \sum_{i=1}^{N} E(\beta_{K,i})E(\mu_i) \end{bmatrix}$$

We therefore obtain:

$$\lambda_{prior} = \begin{bmatrix} \theta_{11} + \widetilde{\beta}_1^2 & \cdots & \theta_{1K} + \widetilde{\beta}_1 \widetilde{\beta}_K \\ \vdots & \ddots & \vdots \\ \theta_{1K} + \widetilde{\beta}_1 \widetilde{\beta}_K & \cdots & \theta_{KK} + \widetilde{\beta}_K^2 \end{bmatrix}^{-1} \begin{bmatrix} \widetilde{\beta}_1 \mu_0 \\ \vdots \\ \widetilde{\beta}_K \mu_0 \end{bmatrix} \tag{16}$$

This can be expanded as

$$\begin{aligned} \lambda_{prior} &= \left[\Theta_{BB} + Vec(\widetilde{B})Vec(\widetilde{B})'\right]^{-1} Vec(\widetilde{B})\mu_0 \qquad (17) \\ &= \Theta_{BB}^{-1} Vec(\widetilde{B})\mu_0 - \frac{\Theta_{BB}^{-1} Vec(\widetilde{B})Vec(\widetilde{B})'\Theta_{BB}^{-1} Vec(\widetilde{B})\mu_0}{1 + Vec(\widetilde{B})'\Theta_{BB}^{-1} Vec(\widetilde{B})} \\ &= \frac{1}{1 + Vec(\widetilde{B})'\Theta_{BB}^{-1} Vec(\widetilde{B})} \Theta_{BB}^{-1} Vec(\widetilde{B})\mu_0 \end{aligned}$$

Thus comparative statics are easy to calculate. We can for example compute $\frac{\partial\lambda}{\partial\mu_0}, \frac{\partial\lambda}{\partial\Theta_{BB}}, \frac{\partial\lambda}{\partial Vec(\widetilde{B})}$.

For example for the CAPM case where the only factor is the market portfolio, we have that the prior estimate of the risk premium is

$$\lambda_1 = \frac{\widetilde{\beta}_1\mu_0}{\theta_{11} + \widetilde{\beta}_1^2}$$

$\widetilde{\beta}_1$is the prior estimate of the market beta, μ_0 is the prior mean of the excess return assigned by the investigator to be common for all assets, θ_{11} is the prior variance of $\widetilde{\beta}_1$. Also

$$\frac{\partial\lambda}{\partial\mu_0} = \frac{\widetilde{\beta}_1}{\theta_{11} + \widetilde{\beta}_1^2}$$

When we have 2 factors we get:

$$\lambda_{prior} = \begin{bmatrix} \lambda_1 \\ \lambda_2 \end{bmatrix} = \begin{bmatrix} \frac{(\theta_{22}+\widetilde{\beta}_2^2)(\mu_0\widetilde{\beta}_1)-(\theta_{12}+\widetilde{\beta}_1\widetilde{\beta}_2)(\mu_0\widetilde{\beta}_2)}{(\theta_{11}+\widetilde{\beta}_1^2)(\theta_{22}+\widetilde{\beta}_2^2)-(\theta_{12}+\widetilde{\beta}_1\widetilde{\beta}_2)^2} \\ \\ \frac{(\theta_{11}+\widetilde{\beta}_1^2)(\mu_0\widetilde{\beta}_2)-(\theta_{12}+\widetilde{\beta}_1\widetilde{\beta}_2)(\mu_0\widetilde{\beta}_1)}{(\theta_{11}+\widetilde{\beta}_1^2)(\theta_{22}+\widetilde{\beta}_2^2)-(\theta_{12}+\widetilde{\beta}_1\widetilde{\beta}_2)^2} \end{bmatrix}$$

We can proceed similarly for 3 or more factors.

3.1.2 General Case: (Prior dependence between B and μ; i.e. $\Theta_{B\mu} \neq 0$)

That there is no cross-sectional dependence, a priori, between a given beta and the mean excess return seems highly unlikely given the nature of many practitioner processes and their possible knowledge of asset-pricing theory. Typical processes involve sorting by factor exposure or variable, so fund managers will believe that high growth leads to high return. Here we generalize the above setting to allow for the more realistic case of prior dependence between B and μ. We therefore assume that $\Theta_{B\mu} = [\theta_{1\mu}, ..., \theta_{K\mu}]' \neq 0$. The general formula for the prior estimate for the factor risk-premium now becomes

$$\lambda_{prior} = \frac{1}{1 + Vec(\widetilde{B})'\Theta_{BB}^{-1}Vec(\widetilde{B})}\Theta_{BB}^{-1}(Vec(\widetilde{B})\mu_0 + \Theta_{B\mu})$$

From this we can now write very easily the analytic formulae for the prior estimates of the one factor case (i.e. CAPM), two factors etc. For example for The risk premium for the one factor CAPM case now is

$$\lambda_1 = \frac{\theta_{1\mu} + \widetilde{\beta}_1\mu_0}{\theta_{11} + \widetilde{\beta}_1^2}$$

with $\theta_{1\mu}$ being the prior covariance between $\widetilde{\beta}_1$ and μ_0. Similarly the risk premium for the 2-factor case with prior dependence between B and μ is obtained from

$$\lambda_{prior} = \left[\begin{array}{c} \lambda_1 \\ \lambda_2 \end{array} \right] = \left[\begin{array}{c} \frac{(\theta_{22}+\widetilde{\beta}_2^2)(\theta_{1\mu}+\mu_0\widetilde{\beta}_1)-(\theta_{12}+\widetilde{\beta}_1\widetilde{\beta}_2)(\theta_{2\mu}+\mu_0\widetilde{\beta}_2)}{(\theta_{11}+\widetilde{\beta}_1^2)(\theta_{22}+\widetilde{\beta}_2^2)-(\theta_{12}+\widetilde{\beta}_1\widetilde{\beta}_2)^2} \\ \\ \frac{(\theta_{11}+\widetilde{\beta}_1^2)(\theta_{2\mu}+\mu_0\widetilde{\beta}_2)-(\theta_{12}+\widetilde{\beta}_1\widetilde{\beta}_2)(\theta_{1\mu}+\mu_0\widetilde{\beta}_1)}{(\theta_{11}+\widetilde{\beta}_1^2)(\theta_{22}+\widetilde{\beta}_2^2)-(\theta_{12}+\widetilde{\beta}_1\widetilde{\beta}_2)^2} \end{array} \right]$$

Extensions for 3 or more factors are straightforward to obtain.

3.2 Posterior Estimates of the Risk Premia

We now turn to deriving the posterior estimates of the risk-premia. If the prior distribution of the regression coefficients matrix A is as in (14), then the posterior distribution of A is (see for example Kadiyala and Karlsson (1997)):

$$A \backsim MN\left[VecA_p, \Sigma_p\right]$$

where

$$\begin{aligned} \Sigma_p &= \left[(\widetilde{\Sigma}^{-1} + ((H'H)^{-1} \otimes \Psi)^{-1})\right]^{-1} \qquad (18) \\ &= \left(\widetilde{\Sigma}^{-1} + ((H'H) \otimes \Psi^{-1})\right)^{-1} \\ &= (\widetilde{\Sigma}^{-1} + \widehat{\Sigma}^{-1})^{-1} \end{aligned}$$

(Note that$\widehat{}$ denotes a sample estimate and $\widetilde{}$denotes a prior estimate) and

$$\begin{aligned} VecA_p &= \Sigma_p \left((\widetilde{\Sigma}^{-1} \times Vec\widetilde{A}) + ((H'H)^{-1} \otimes \Psi)^{-1} \times Vec[(H'H)^{-1}H'X]) \right) \\ &= \Sigma_p \left((\widetilde{\Sigma}^{-1} \times Vec\widetilde{A}) + (\widehat{\Sigma}^{-1} \times Vec\widehat{A}) \right) \end{aligned}$$

Now to write Σ_p and $VecA_p$ in partitioned form note first that

$$\begin{aligned} H &= [F, i_T] \Rightarrow \\ H'H &= \left[\begin{array}{cc} F'F & F'i_T \\ i_T'F & i_T'i_T \end{array} \right] \\ &= \left[\begin{array}{cc} F'F & 0 \\ 0 & T \end{array} \right] \end{aligned}$$

where we have assumed by construction that the factors have zero sample mean (i.e. $F'i_T$ and $i_T'F = 0$). Therefore

$$\Sigma_p = \left(\left[\left[\begin{array}{cc} \Theta_{BB} & \Theta_{B\mu} \\ \Theta_{\mu B} & \theta_{\mu\mu} \end{array} \right] \otimes I_N \right]^{-1} + \left[\begin{array}{cc} F'F & 0 \\ 0 & T \end{array} \right] \otimes \Psi^{-1} \right)^{-1} \qquad (19)$$

3.2.1 Standard Bayesian Case: (Prior independence between B and μ; i.e. $\Theta_{B\mu} = 0$)

Turning to the posterior estimates of the mean and variance, with prior independence we get the following simplifications (for detailed calculations see the appendix):

$$\Sigma_p = \begin{bmatrix} [\Theta_{BB}^{-1} \otimes I_N + F'F \otimes \Psi^{-1}]^{-1} & 0 \\ 0 & [\theta_{\mu\mu}^{-1} \times I_N + T \times \Psi^{-1}]^{-1} \end{bmatrix} \tag{20}$$

and

$$VecA_p = \begin{bmatrix} VecB_p \\ \mu_p \end{bmatrix}$$

$$= \begin{bmatrix} [\Theta_{BB}^{-1} \otimes I_N + F'F \otimes \Psi^{-1}]^{-1} \times \{(\Theta_{BB}^{-1} \otimes I_N)Vec\widetilde{B} + [F'F \otimes \Psi^{-1}]Vec\widehat{B}\} \\ [\theta_{\mu\mu}^{-1} \times I_N + T \times \Psi^{-1}]^{-1} \times \{(\theta_{\mu\mu}^{-1} \times I_N)\widetilde{\mu} + T\Psi^{-1}\widehat{\mu}\} \end{bmatrix}$$

The estimates for the risk-premium vector λ now follow straightforwardly from the procedure outlined in section (2.3).

3.2.2 General Case: (Prior dependence between B and μ; i.e. $\Theta_{B\mu} \neq 0$)

When there is prior dependence between B and μ things get a bit more complicated. We know that

$$\Sigma_p = \left(\left[\begin{bmatrix} \Theta_{BB} & \Theta_{B\mu} \\ \Theta_{\mu B} & \theta_{\mu\mu} \end{bmatrix} \otimes I_N \right]^{-1} + \begin{bmatrix} F'F & 0 \\ 0 & T \end{bmatrix} \otimes \Psi^{-1} \right)^{-1}$$

This can now be written as

$$\Sigma_p = \left(\begin{bmatrix} \Theta_{BB}^{-1}(I + \Theta_{B\mu}\Xi_2\Theta_{\mu B}\Theta_{BB}^{-1}) & -\Theta_{BB}^{-1}\Theta_{B\mu}\Xi_2 \\ -\Xi_2\Theta_{\mu B}\Theta_{BB}^{-1} & \Xi_2 \end{bmatrix} \otimes I_N + \begin{bmatrix} F'F & 0 \\ 0 & T \end{bmatrix} \otimes \Psi^{-1} \right)^{-1}$$

where

$$\Xi_2 = (\theta_{\mu\mu} - \Theta_{\mu B}\Theta_{BB}^{-1}\Theta_{B\mu})^{-1}$$

From this we have (see the appendix) that the posterior variance-covariance in partitioned form is:

$$\Sigma_p = \begin{bmatrix} \Sigma_{p11} & \Sigma_{p12} \\ \Sigma_{p21} & \Sigma_{p22} \end{bmatrix}$$

where

$$\begin{aligned} \Sigma_{p11} &= V_{11}^{-1}(I + \Upsilon_2(V_{22} - \Upsilon_2 V_{11}^{-1}\Upsilon_2)^{-1}\Upsilon_2' V_{11}^{-1}) \\ \Sigma_{p12} &= -V_{11}^{-1}\Upsilon_2(V_{22} - \Upsilon_2 V_{11}^{-1}\Upsilon_2)^{-1} \\ \Sigma_{p21} &= \Sigma_{p12}' \\ \Sigma_{p22} &= (V_{22} - \Upsilon_2 V_{11}^{-1}\Upsilon_2)^{-1} \end{aligned}$$

with

$$\begin{aligned}
\Upsilon_1 &= [\Theta_{BB}^{-1}(I + \Theta_{B\mu}\Xi_2\Theta_{\mu B}\Theta_{BB}^{-1})] \otimes I_N \\
\Upsilon_2 &= [-\Theta_{BB}^{-1}\Theta_{B\mu}\Xi_2] \otimes I_N \\
\Upsilon_3 &= \Xi_2 \times I_N
\end{aligned}$$

and

$$\begin{aligned}
V_{11} &= \Upsilon_1 + (F'F \otimes \Psi^{-1}) \\
V_{22} &= \Upsilon_3 + T\Psi^{-1}
\end{aligned}$$

Similarly the posterior estimates for the regression coefficients are given by (see the appendix)

$$\begin{aligned}
VecA_p &= \Sigma_p \left((\widetilde{\Sigma}^{-1} \times VecA) + (\widehat{\Sigma}^{-1} \times Vec\widehat{A}) \right) \\
&= \begin{bmatrix} VecB_p \\ \mu_p \end{bmatrix}
\end{aligned}$$

where

$$\begin{aligned}
VecB_p = \; & (\Sigma_{p11}\Upsilon_1 + \Sigma_{p12}\Upsilon_2')VecB + \Sigma_{p11}[F'F \otimes \Psi^{-1}]Vec\widehat{B} \\
& + (\Sigma_{p12}\Upsilon_3 - \Sigma_{p11}\Upsilon_2)\underline{\mu} + \Sigma_{p12}T\Psi^{-1}\widehat{\mu}
\end{aligned}$$

and

$$\begin{aligned}
\mu_p = \; & (\Sigma_{p21}\Upsilon_1 + \Sigma_{p22}\Upsilon_2')VecB + \Sigma_{p21}[F'F \otimes \Psi^{-1}]Vec\widehat{B} \\
& + (\Sigma_{p22}\Upsilon_3 - \Sigma_{p21}\Upsilon_2)\underline{\mu} + \Sigma_{p22}T\Psi^{-1}\widehat{\mu}
\end{aligned}$$

Now that we have obtained the posterior estimates for Σ_p and $VecA_p$ in partitioned form it is straightforward to obtain the risk premium vector λ following the procedure of section (2.3).

3.2.3 Extensions of the Minnesota prior

The Bayesian framework introduced above used the Minnesota prior. It is possible to generalize this framework by allowing for a non-diagonal variance-covariance matrix and/or by taking Ψ to be unknown. Possibilities include using a Normal-Wishart prior, a Normal-Diffuse prior (introduced by Zellner (1971)) or an extended Natural Conjugate Prior (see Dreze and Richard (1983)). For the latter two priors no closed form solution for the posterior moments exist, and numerical methods such as importance sampling or Gibbs sampling are required. (see Kadiyala and Karlsson (1997)). An additional possibility is to maintain the assumption of a diagonal Ψ matrix while taking the diagonal elements to be unknown. Independent inverse gamma priors on the diagonal elements then lead to marginal multivariate t priors and posteriors for the parameters of each equation.

4 An Empirical Application

We now present an empirical application of our methodology, using non-traded factors. Our use of several observed macroeconomic risk factors to explain asset returns can be justified by the newest generation of empirical research, as summarized for example in Cochrane (2001). One of the earliest examples of applying macroeconomic risks in the APT is the paper by Chen, Roll and Ross (1986) analyzing the pricing of such factors in the US market. Recognizing the ability of investors to diversify and the co-movements of asset prices, the authors suggest the presence of pervasive or systematic influences as the likely source of investment risk. In particular Chen, Roll and Ross find that 1) unanticipated changes in the expected level of production, 2) unanticipated shifts in the shape of the term structure, 3) changes in default premiums and 4) unexpected inflation are risks that are significantly priced in the US market. By contrast, risks stemming from unanticipated changes in the market portfolio, aggregate consumption and oil prices were found not to be priced by the authors.

4.1 Data

The choice of candidate macroeconomic factors in this paper is largely inspired by Chen et al.(1986) and is a subset of the factors presented in Antoniou et al. (1998). All data for measuring the macroeconomic factors are obtained from Datastream. In addition, data on total monthly logarithmic returns for UK stocks are also obtained from Datastream. Our sample spans a period of five years, starting from the end of November 1993 until the end of September 1998. The sample comprises 66 stocks from the FTSE 100 Index on which data are available throughout the sample period. Thus there is bound to be some survivorship bias The overall sample mean of excess returns for the 66 stocks over our sample period is calculated to be 16.7% (annualized).

Apart from spanning the space of returns, the most important property required of appropriate factor measures is that they cannot be predictable from their own past. To avoid problems caused by the potential presence of autocorrelation in the variables, simple ARIMA models were fitted to pre-whiten the series. It has to be noted, however, that although this procedure is designed to avoid spurious correlation, it carries a danger of possible misspecification. Given finite samples, the fitted ARIMA models can only be approximatinos to the true data generating process. The measurement of the risk factors used in our empirical application is explained below.

4.1.1 Industrial Production (Ind. Prod.)

In line with Chen, Roll and Ross, we use the monthly growth rate in industrial production. This is defined as:

$$MP_t = \ln IP_t - \ln IP_{t-1}$$

where IP denotes industrial production. We use the UKINPRODG Datastream series defined as the "UK industrial production - total production vol.". An AR(1) model was used to derive the innovations in industrial production.

4.1.2 Inflation (Infl.)

We use the difference in the logarithm of the consumer price index (CPI) to capture the effect of the inflation factor, as follows:

$$IR_t = \ln CPI_t - \ln CPI_{t-1}$$

We use the series UKRP...F defined as the "UK Retail Price Index NADJ". An ARMA model was used to derive the unexpected component of this series.

4.1.3 Market risk premium (Market)

To capture the effect of the market risk premium factor, we use the difference in the returns on the equity market (EM) and the government bond market (BM) in line with the definition of the risk premium in Datastream:

$$RP_t = [\ln EM_t - \ln EM_{t-1}] - [\ln BM_t - \ln BM_{t-1}$$

We use the FTALLSH(RI) series defined as the "FTSE All share - Total Return Index" and the series FTAGOVT(RI) defined as the "FTA Government All Stocks - Total return Index". An AR(1) model was used to derive the innovations in this series.

4.1.4 Term structure (Term Str.)

To capture the effect of unanticipated shifts in the term structure, in line with the approach followed in the literature we use the spread between long-term (LTR) and short-term interest rates (STR):

$$TS_t = LTR_t - STR_t$$

We use the first difference in the logarithm of the series BMUK30Y(RI) defined as the "UK Benchmark 30 Years DS Government Index - Total Return Index" and the series LDNT3BM defined as the "UK Treasury Bill Discount 3 month - Middle Rate".

4.2 Results

4.2.1 Empirical Bayesian approach as prior

We start by employing an empirical Bayesian approach in specifying our prior (hyper)parameters of section (3.1). We first estimate equation (2):

$$x^i = \mu_i + \sum_{j=1}^{K} \beta_{ij} f_j + \varepsilon^i$$

using OLS regression for each of the 66 stocks in our portfolio. Next, the prior parameter μ_0 is calculated as the mean of the intercepts from our estimated regressions

$$\mu_0 = \frac{1}{N}\sum_{i=1}^{N}\widehat{\mu}_i$$

(The symbol$\widehat{}$denotes an OLS estimate). Similarly the prior parameters $\begin{bmatrix} \widetilde{\beta}_1 & \cdots & \widetilde{\beta}_K \end{bmatrix}$ are calculated as the means of the estimated factor β's. Thus

$$\widetilde{\beta}_k = \frac{1}{N}\sum_{i=1}^{N}\widehat{\beta}_{ik}\forall k \in (1, ..., K)$$

Finally Θ_{BB}, $\theta_{\mu\mu}$, and $\Theta_{B\mu}$ are assigned values taken from the cross-sectional variance-covariance matrix of the estimated μ's (intercepts) and β's. (see for example Table (2) in the appendix where we report the estimated cross-sectional correlation matrix of μ's and β's). Prior and posterior estimates for the risk premia are next derived using the two subcases presented in sections (3.1) and (3.2) (see Table 4 in the appendix). Note that for the subcase where by construction we assume prior independence between B and μ, (see sections (3.1.1) and (3.2.1)) $\Theta_{B\mu} = 0$.

4.2.2 CAPM as prior

We now make the assumption that we are CAPM advocates and we want to incorporate this in our prior information. This effectively means that we construct our priors so that we now have $\widetilde{\beta}_{\text{Market}} = 1$ and $\widetilde{\beta}_1, ..., \widetilde{\beta}_K = 0$, $\forall\ k \neq$Market. For the remaining hyperperameters we continue to use the empirical Bayes approach arising from the procedure of estimating unconstrained regressions outlined in the previous section. However we note that since in this case our prior information implies the exclusion of certain regressors another possibility would be to compute the remaining hyperparameters by estimating constrained least squares regressions.

We estimate risk premia both with and without the assumption of prior independence between μ and B. Our results are exhibited in the appendix. In particular Table 1 reveals the correlation structure between our factors; notable are the high correlation between Industrial Production and Inflation (positive) and Industrial Production and the Term Structure (negative). Table 2 reports the pattern of cross-sectional OLS beta's and mean excess returns for our 66 stocks. Table 3 reports our beta results. Imposing the restriction $\Theta_{\mu B} = 0$ has little impact on the results. Using the Empirical Bayes prior leads to very little Bayesian updating; our prior and posterior estimates are essentially equal. However imposition of the CAPM prior leads to posterior estimates of betas that are smaller in magnitude for inflation than before, whilst (unsurprisingly) increasing substantially the role of the market. Industrial Production has changed little whilst term structure has more or less disappeared. Turning to Table 4, the market risk premia with a CAPM prior are now positive except for industrial production and are dominated by the market risk premium. Finally in Table 5 we compute the mean overall prior and posterior return premium arising from the APT relationship $\mu = B'\lambda$. Purists might make a case for transforming the posterior joint distributions to derive the distribution of the posterior risk premium and then calculating our mean. However, since our results are essentially the same whether we impose the prior that $\Theta_{\mu B} = 0$ or $\Theta_{\mu B} \neq 0$, we thought this would add unnecessary complications. Our final result of a CAPM posterior of 16.6% ($\Theta_{\mu B} \neq 0$) or 16.9% ($\Theta_{\mu B} = 0$) can be compared with the overall annualized sample mean excess return of stocks in the period of 16.7%. This compares favorably with the empirical Bayes posterior estimates of about 8% and also with the CAPM estimate of about 18%.

5 Conclusion

Our paper has set out to illustrate how to use Bayesian methods to compute factor risk premia in an APT framework for both cases of traded and non-traded factors. Using a sample of UK stocks from 1993 to 1998 we found evidence that a CAPM prior seemed to produce more data consistent results than an empirical Bayes approach. However since our CAPM prior still retains some empirical Bayes hyperparameters based on the APT, a role of importance for the CAPM versus the APT, is by no means conclusive. In fact, our results suggest that a Bayesian mixture of CAPM as prior and APT as the data generating process outperforms both classical cases of CAPM or APT alone.

One case not considered in this paper, or indeed elsewhere in the literature, is the important hybrid case where some factors are traded whilst some are not. We hope to address this problem in a later paper.

6 Appendix

6.1 Definition of the Matrixvariate Normal Distribution

Let X and $VecX$ denote a $p \times q$ random matrix and its $pq-$dimensional column expansion respectively. X is said to have a matrixvariate normal distribution with parameters $M \in \Re^{p \times q}$, $P \in C_p$, and $Q \in C_q$ (C_n denotes the set of $n \times n$, positive definite symmetric matrices)

$$X \sim MN_{p \times q}(VecM, Q \otimes P)$$

if and only if

$$VecX \sim N_{pq}(VecM, Q \otimes P)$$

(this denotes a multivariate normal distribution).

Therefore its density function is given by

$$f_{MN}^{p \times q}(X|M, Q \otimes P) = C_{MN}^{-1}(P, Q; p, q) \times \exp\{-\frac{1}{2} tr[Q^{-1}(X-M)' P^{-1}(X-M)]\}$$

where

$$C_{MN}(P, Q; p, q) = [(2\pi)^{pq} |P|^q |Q|^p]^{1/2}$$

6.2 Posterior risk premia calculations

6.2.1 Prior independence (i.e $\Theta_{\mu B} = 0$)

$$\begin{aligned}
\Sigma_p &= \left(\begin{bmatrix} \Theta_{BB}^{-1} \otimes I_N + F'F \otimes \Psi^{-1} & 0 \\ 0 & \theta_{\mu\mu}^{-1} \times I_N + T \times \Psi^{-1} \end{bmatrix} \right)^{-1} \\
&\Rightarrow \\
\Sigma_p &= \begin{bmatrix} \Sigma_{p11} & 0 \\ 0 & \Sigma_{p22} \end{bmatrix} = \begin{bmatrix} [\Theta_{BB}^{-1} \otimes I_N + F'F \otimes \Psi^{-1}]^{-1} & 0 \\ 0 & [\theta_{\mu\mu}^{-1} \times I_N + T \times \Psi^{-1}]^{-1} \end{bmatrix}
\end{aligned}$$

and

$$\begin{aligned}
VecA_p &= \begin{bmatrix} VecB_p \\ \mu_p \end{bmatrix} = \begin{bmatrix} \Sigma_{p11}\{(\Theta_{BB}^{-1} \otimes I_N)Vec\widetilde{B} + [F'F \otimes \Psi^{-1}]Vec\widehat{B}\} \\ \Sigma_{p22}\{(\theta_{\mu\mu}^{-1} \times I_N)\widetilde{\mu} + T\Psi^{-1}\widehat{\mu}\} \end{bmatrix} \\
&= \begin{bmatrix} [\Theta_{BB}^{-1} \otimes I_N + F'F \otimes \Psi^{-1}]^{-1} \times \{(\Theta_{BB}^{-1} \otimes I_N)Vec\widetilde{B} + [F'F \otimes \Psi^{-1}]Vec\widehat{B}\} \\ [\theta_{\mu\mu}^{-1} \times I_N + T \times \Psi^{-1}]^{-1} \times \{(\theta_{\mu\mu}^{-1} \times I_N)\widetilde{\mu} + T\Psi^{-1}\widehat{\mu}\} \end{bmatrix}
\end{aligned}$$

6.2.2 Prior dependence (i.e. $\Theta_{\mu B} \neq 0$)

$$\begin{aligned}
\Sigma_p &= \left(\begin{bmatrix} [\Theta_{BB}^{-1}(I + \Theta_{B\mu}\Xi_2\Theta_{\mu B}\Theta_{BB}^{-1})] \otimes I_N & [-\Theta_{BB}^{-1}\Theta_{B\mu}\Xi_2] \otimes I_N \\ [-\Xi_2\Theta_{\mu B}\Theta_{BB}^{-1}] \otimes I_N & \Xi_2 \times I_N \end{bmatrix} + \begin{bmatrix} F'F \otimes \Psi^{-1} & 0 \\ 0 & T \times \Psi^{-1} \end{bmatrix} \right) \\
&= \begin{bmatrix} \Upsilon_1 + (F'F \otimes \Psi^{-1}) & \Upsilon_2 \\ \Upsilon_2' & \Upsilon_3 + T\Psi^{-1} \end{bmatrix}^{-1}
\end{aligned}$$

where

$$\begin{aligned}
\Upsilon_1 &= [\Theta_{BB}^{-1}(I + \Theta_{B\mu}\Xi_2\Theta_{\mu B}\Theta_{BB}^{-1})] \otimes I_N \\
\Upsilon_2 &= [-\Theta_{BB}^{-1}\Theta_{B\mu}\Xi_2] \otimes I_N \\
\Upsilon_3 &= \Xi_2 \times I_N
\end{aligned}$$

Now let

$$\begin{aligned}
V_{11} &= \Upsilon_1 + (F'F \otimes \Psi^{-1}) \\
V_{22} &= \Upsilon_3 + T\Psi^{-1}
\end{aligned}$$

The posterior variance-covariance in partitioned form is therefore:

$$\begin{aligned}
\Sigma_p &= \begin{bmatrix} V_{11} & \Upsilon_2 \\ \Upsilon_2' & V_{22} \end{bmatrix}^{-1} \\
&= \begin{bmatrix} V_{11}^{-1}(I + \Upsilon_2(V_{22} - \Upsilon_2 V_{11}^{-1}\Upsilon_2)^{-1}\Upsilon_2' V_{11}^{-1}) & -V_{11}^{-1}\Upsilon_2(V_{22} - \Upsilon_2 V_{11}^{-1}\Upsilon_2)^{-1} \\ -(V_{22} - \Upsilon_2 V_{11}^{-1}\Upsilon_2)^{-1}\Upsilon_2' V_{11}^{-1} & (V_{22} - \Upsilon_2 V_{11}^{-1}\Upsilon_2)^{-1} \end{bmatrix} \\
&= \begin{bmatrix} \Sigma_{p11} & \Sigma_{p12} \\ \Sigma_{p21} & \Sigma_{p22} \end{bmatrix}
\end{aligned}$$

where

$$\begin{aligned}
\Sigma_{p11} &= V_{11}^{-1}(I + \Upsilon_2(V_{22} - \Upsilon_2 V_{11}^{-1}\Upsilon_2)^{-1}\Upsilon_2' V_{11}^{-1}) \\
\Sigma_{p12} &= -V_{11}^{-1}\Upsilon_2(V_{22} - \Upsilon_2 V_{11}^{-1}\Upsilon_2)^{-1} \\
\Sigma_{p21} &= \Sigma_{p12}' \\
\Sigma_{p22} &= (V_{22} - \Upsilon_2 V_{11}^{-1}\Upsilon_2)^{-1}
\end{aligned}$$

Similarly

$$\begin{aligned}
VecA_p &= \Sigma_p\left((\widetilde{\Sigma}^{-1} \times VecA) + (\widehat{\Sigma}^{-1} \times Vec\widehat{A})\right) \\
&= \Sigma_p(\begin{bmatrix} \Upsilon_1 & \Upsilon_2 \\ \Upsilon_2' & \Upsilon_3 \end{bmatrix} \times \begin{bmatrix} VecB \\ \underline{\mu} \end{bmatrix} + \begin{bmatrix} F'F \otimes \Psi^{-1} & 0 \\ 0 & T \times \Psi^{-1} \end{bmatrix} \times \begin{bmatrix} Vec\widehat{B} \\ \widehat{\mu} \end{bmatrix}) \\
&= \Sigma_p(\begin{bmatrix} \Upsilon_1 VecB - \Upsilon_2\underline{\mu} \\ \Upsilon_2' VecB + \Upsilon_3\underline{\mu} \end{bmatrix} + \begin{bmatrix} [F'F \otimes \Psi^{-1}]Vec\widehat{B} \\ T\Psi^{-1}\widehat{\mu} \end{bmatrix}) \\
&= \begin{bmatrix} \Sigma_{p11} & \Sigma_{p12} \\ \Sigma_{p21} & \Sigma_{p22} \end{bmatrix} \times \begin{bmatrix} \Upsilon_1 VecB - \Upsilon_2\underline{\mu} + [F'F \otimes \Psi^{-1}]Vec\widehat{B} \\ \Upsilon_2' VecB + \Upsilon_3\underline{\mu} + T\Psi^{-1}\widehat{\mu} \end{bmatrix} \\
&= \begin{bmatrix} VecB_p \\ \mu_p \end{bmatrix}
\end{aligned}$$

where

$$\begin{aligned}
VecB_p &= \Sigma_{p11} \times \{\Upsilon_1 VecB - \Upsilon_2 \underline{\mu} + [F'F \otimes \Psi^{-1}]Vec\widehat{B}\} \\
&\quad + \Sigma_{p12} \times \{\Upsilon_2' VecB + \Upsilon_3 \underline{\mu} + T\Psi^{-1}\widehat{\mu}\} \\
&= \Sigma_{p11}\Upsilon_1 VecB - \Sigma_{p11}\Upsilon_2\underline{\mu} + \Sigma_{p11}[F'F \otimes \Psi^{-1}]Vec\widehat{B} \\
&\quad + \Sigma_{p12}\Upsilon_2' VecB + \Sigma_{p12}\Upsilon_3\underline{\mu} + \Sigma_{p12}T\Psi^{-1}\widehat{\mu} \\
&\Rightarrow \\
VecB_p &= (\Sigma_{p11}\Upsilon_1 + \Sigma_{p12}\Upsilon_2')VecB + \Sigma_{p11}[F'F \otimes \Psi^{-1}]Vec\widehat{B} \\
&\quad + (\Sigma_{p12}\Upsilon_3 - \Sigma_{p11}\Upsilon_2)\underline{\mu} + \Sigma_{p12}T\Psi^{-1}\widehat{\mu}
\end{aligned}$$

and

$$\begin{aligned}
\mu_p &= \Sigma_{p21} \times \{\Upsilon_1 VecB - \Upsilon_2 \underline{\mu} + [F'F \otimes \Psi^{-1}]Vec\widehat{B}\} \\
&\quad + \Sigma_{p22} \times \{\Upsilon_2' VecB + \Upsilon_3 \underline{\mu} + T\Psi^{-1}\widehat{\mu}\} \\
&= \Sigma_{p21}\Upsilon_1 VecB - \Sigma_{p21}\Upsilon_2\underline{\mu} + \Sigma_{p21}[F'F \otimes \Psi^{-1}]Vec\widehat{B} \\
&\quad + \Sigma_{p22}\Upsilon_2' VecB + \Sigma_{p22}\Upsilon_3\underline{\mu} + \Sigma_{p22}T\Psi^{-1}\widehat{\mu} \\
&\Rightarrow \\
\mu_p &= (\Sigma_{p21}\Upsilon_1 + \Sigma_{p22}\Upsilon_2')VecB + \Sigma_{p21}[F'F \otimes \Psi^{-1}]Vec\widehat{B} \\
&\quad + (\Sigma_{p22}\Upsilon_3 - \Sigma_{p21}\Upsilon_2)\underline{\mu} + \Sigma_{p22}T\Psi^{-1}\widehat{\mu}
\end{aligned}$$

6.3 Tables

Table1: Correlation structure of macroeconomic factors

Correlation	*Ind.Prod.*	*Infl.*	*Market.*	*Term Str.*
Ind.Prod.	1.000	0.275	0.075	-0.207
Infl.	0.275	1.000	0.090	-0.066
Market	0.075	0.075	1.000	-0.067
TermStr.	-0.207	-0.066	-0.067	1.000

Table 2: Cross-sectional correlation of means and beta's

Correlation	$\widehat{\mu}$	$\widehat{\beta}_{\text{Ind.Prod}}$	$\widehat{\beta}_{\text{Infl.}}$	$\widehat{\beta}_{\text{Market}}$	$\widehat{\beta}_{\text{Term.Str}}$
$\widehat{\mu}$	1.000	-0.247	-0.124	0.054	-0.027
$\widehat{\beta}_{\text{Ind.Prod}}$	-0.247	1.000	-0.355	0.050	0.297
$\widehat{\beta}_{\text{Infl.}}$	-0.124	-0.355	1.000	-0.178	0.008
$\widehat{\beta}_{\text{Market}}$	0.054	0.050	-0.178	1.000	-0.264
$\widehat{\beta}_{\text{Term.Str}}$	-0.027	0.297	0.008	-0.264	1.000

Table 3: Mean Prior and Posterior Values of regression coefficients

	Emp.Bayes prior ($\Theta_{\mu B} \neq 0$)	Emp.Bayes prior ($\Theta_{\mu B} = 0$)	CAPM prior ($\Theta_{\mu B} \neq 0$)	CAPM prior ($\Theta_{\mu B} = 0$)
Prior Values				
$\widetilde{\mu}$ (intercept)	19.2%	19.2%	19.2%	19.2%
$\widetilde{\beta}_{\text{Ind.Prod}}$	0.079	0.079	0	0
$\widetilde{\beta}_{\text{Infl.}}$	-2.295	-2.295	0	0
$\widetilde{\beta}_{\text{Market}}$	0.030	0.030	1	1
$\widetilde{\beta}_{\text{Term.Str}}$	-0.185	-0.185	0	0
Posterior Estimates (Means)				
$\overline{\mu}_p$ (intercept)	18.8%	18.8%	18.3%	18.8%
$\overline{\beta}_{p\text{Ind.Prod}}$	0.095	0.094	-0.101	-0.101
$\overline{\beta}_{p\text{Infl.}}$	-2.217	-2.221	-0.563	-0.563
$\overline{\beta}_{p\text{Market}}$	0.033	0.033	0.664	0.664
$\overline{\beta}_{p\text{Term.Str}}$	-0.189	-0.189	-0.005	-0.005

Table 4: Prior and Posterior estimates of the Risk Premia (annualised)

	Emp.Bayes prior ($\Theta_{\mu B} \neq 0$)	Emp.Bayes prior ($\Theta_{\mu B} = 0$)	CAPM prior ($\Theta_{\mu B} \neq 0$)	CAPM prior ($\Theta_{\mu B} = 0$)
Prior Estimates				
$\widetilde{\lambda}_{\text{Ind.Prod}}$	-0.022	+0.005	-0.002	-0.002
$\widetilde{\lambda}_{\text{Infl.}}$	-0.017	-0.011	+0.001	+0.001
$\widetilde{\lambda}_{\text{Mkt.risk.prem}}$	-0.032	-0.060	+0.182	0.180
$\widetilde{\lambda}_{\text{Term.str}}$	-0.182	-0.228	+0.043	+0.042
Posterior Estimates				
$\lambda_{p\text{Ind.Prod}}$	-0.016	+0.002	-0.008	-0.008
$\lambda_{p\text{Infl.}}$	-0.017	-0.013	+0.0005	+0.0005
$\lambda_{p\text{Mkt.risk.prem}}$	-0.033	-0.053	+0.249	+0.253
$\lambda_{p\text{Term.str}}$	-0.236	-0.268	+0.049	+0.049

Table 5: Overall Prior and Posterior excess return premiums ($\mu = B'\lambda$) (*annualised*)

$\mu = B'\lambda$	Prior	Posterior
Emp. Bayes prior ($\Theta_{\mu B} \neq 0$)	6.9%	8.1%
Emp. Bayes prior ($\Theta_{\mu B} = 0$)	6.5%	7.9%
CAPM prior ($\Theta_{\mu B} \neq 0$)	18.2%	16.9%
CAPM prior ($\Theta_{\mu B} = 0$)	18.0%	16.6%

Overall annualised SAMPLE mean excess return of stocks over the period 11/93-9/98:
16.7%

References

[1] Antoniou, A., Garrett, I., and Priestley, R. (1998). Macroeconomic variables as common pervasive risk factors and the empirical content of the arbitrage pricing theory. *Journal of Empirical Finance*, 221-240.

[2] Bauwens, L., M. Lubrano, and J-F. Richard (1999). *Bayesian inference in dynamic econometric models*. Advanced texts in econometrics, Oxford University Press.

[3] Black. F. and R. Litterman (1991). Global asset allocation with equities, bonds and currencies. Goldman Sachs and Co., October.

[4] Black, F. and R. Litterman (1992). Global portfolio optimization, *Financial Analysts Journal*, Sept-Oct, 28-43.

[5] Burmeister, E. and M. McElroy (1988). Joint estimation of factor sensitivities and risk premia for the Arbitrage Pricing Theory. *Journal of Finance*, 43, 721-733.

[6] Campbell, J., A. Lo, and C. MacKinlay (1997). *The Econometrics of Financial Markets*. Princeton University Press, Princeton, New Jersey.

[7] Chamberlain, G. and M. Rothschild (1983). Arbitrage, factor structure, and mean-variance analysis on large asset markets, *Econometrica*, 51, 1281-1304.

[8] Chen, N., R. Roll, and S. Ross (1986). Economic forces and the stock market, *Journal of Business*, 59, 383-403.

[9] Cochrane, J. (2001). *Asset Pricing*. Princeton University Press, Princeton, NJ.

[10] Connor, G. and R. Korajczyk (1993). A test for the number of factors in an approximate factor model, *Journal of Finance*, 48, 1263-1291.

[11] Dreze, J. and J.-F. Richard (1983). Bayesian analysis of simultaneous equations. In: Griliches, Z. and M. Intriligator (Eds.), *Handbook of Econometrics, Vol. I*, Chap. 9. Amsterdam: North-Holland.

[12] Ericsson, J. and S. Karlsson (2002). Selection of Factors in a Multifactor Pricing Model: A Bayesian Model Selection Approach. mimeo.

[13] Fama, E. and K. French (1992). The cross-section of expected stock returns, *Journal of Finance*, 47, 427-465.

[14] Fama, E. and K. French (1993). Common risk factors in the return of stocks and bonds, *Journal of Financial Economics*, 33, 3-56.

[15] Fama, E. and K. French (1996). Multifactor explanations for asset pricing anomalies, *Journal of Finance*, 51, 55-94.

[16] Ingersoll, J. (1984). Some results in the theory of arbitrage pricing, *Journal of Finance*, 39, 1021-1039.

[17] Ingersoll, J. (1987). *Theory of Financial Decision Making*. Rowman & Littlefield, New Jersey.

[18] Jagannathan, R. and Z. Wang (1996). The conditional CAPM and the cross-section of expected returns, *Journal of Finance*, 51, 3-53.

[19] Kadiyala, K.R. and S. Karlsson (1997). Numerical methods for estimation and inference in Bayesian VAR-models, *Journal of Applied Econometrics*, 12, 99-132.

[20] Pitsillis, M. (2002). *Econometric Analysis of National and International Risk Premia*. PhD thesis, University of Cambridge, UK.

[21] Pitsillis, M. and S. Satchell (2001). Improving the estimates of risk premia - application in the UK financial market. DAE working paper, University of Cambridge.

[22] Polson, N. G. and B. V. Tew (2000). Bayesian portfolio selection: An empirical analysis of the S&P 500 index 1970-1996. *Journal of Business and Economic Statistics*, 18, 164-173.

[23] Reyfman, A. (1997). *Labor market risk and expected asset returns*, PhD thesis, University of Chicago.

[24] Ross, S. A. (1976). The Arbitrage Theory of Capital Asset Pricing. *Journal of Economic Theory*, 13, 341-360

[25] Satchell, S. and A. Scowcroft (2000). A demystification of the Black-Litterman model: Managing quantitative and traditional portfolio construction, *Journal of Asset Management*, 1, 2, 138-150.

[26] Zellner, A. (1971). *An introduction to Bayesian inference in econometrics*, Wiley, New York.